Advances in VIRUS RESEARCH

VOLUME 47

ACADEMIC PRESS

Advances in

VIRUS RESEARCH

VOLUME 47

ADVISORY BOARD

Advances in

VIRUS RESEARCH

Edited by

KARL MARAMOROSCH
Department of Entomology
Rutgers University
New Brunswick, New Jersey

FREDERICK A. MURPHY
School of Veterinary Medicine
University of California, Davis
Davis, California

AARON J. SHATKIN
Center for Advanced Biotechnology and Medicine
Piscataway, New Jersey

VOLUME 47

ACADEMIC PRESS
San Diego London Boston
New York Sydney Tokyo Toronto

This book is printed on acid-free paper. ♾

Academic Press, Inc.
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.apnet.com

Academic Press Limited
24-28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

International Standard Serial Number: 0065-3527

International Standard Book Number: 0-12-039847-8

PRINTED IN THE UNITED STATES OF AMERICA
96 97 98 99 00 01 BC 9 8 7 6 5 4 3 2 1

CONTENTS

Geographic and Molecular Epidemiology of Primate T Lymphotropic Retroviruses: HTLV-I, HTLV-II, STLV-I, STLV-PP, and PTLV-L

ANTOINE GESSAIN AND GUY DE THÉ

This Page Intentionally Left Blank

ADVANCES IN VIRUS RESEARCH, VOL. 47

MARBURG AND EBOLA VIRUSES

Heinz Feldmann and Hans-Dieter Klenk

Institute of Virology
Philipps University
35037 Marburg, Germany

I. Introduction

Filoviruses are among the most pathogenic of human viruses. They are classified as "Biological Level 4" agents (WHO; Risk Group 4) based on their high mortality rate, person-to-person transmission, potential aerosol infectivity, and absence of vaccines and chemotherapy. Maximum containment is required for all laboratory work with infectious material (1, 2). Yet, we are only beginning to understand the interactions of these viruses with their host, and our knowledge on genetics, pathogenicity, and natural history is still limited. Even though outbreaks among human and nonhuman primates to date have always been self-limited, it is because of our ignorance about the natural

reservoir, the potential of these viruses to be transmitted by aerosol, and the lack of immunoprophylactic and chemotherapeutic measures that these infections are of great concern to biomedical scientists. Imported monkeys and international travel, especially rapid travel within the incubation time, are considerable risk factors for introduction of filovirus infections into nonendemic countries. Limited knowledge of the epidemiology and clinical picture of filoviral hemorrhagic fever (HF) and inexperience in diagnosing cases and in case management magnify the danger of an introduction.

Filoviruses, like other RNA viruses, presumably have a potential for rapid evolution due to an inherently high error rate of the virus-encoded polymerase and a lack of repair mechanisms (3). The consequence may be a spectrum of genetic variants that are selected by the host for different transmissibility, virulence, and other biological properties. Changes in socioeconomic structure, such as an increase in human population, increase in speed, variety, and frequency of travel, and disruption of social structures may augment the development of mutant virus populations and the probability of a filovirus emerging as a truly serious public health problem (4).

II. Epidemiology of Filoviruses

A. Outbreaks

1. Marburg Hemorrhagic Fever

Hemorrhagic fever caused by the Marburg virus first emerged in 1967 (Table I). The epidemic started in mid-August with three laboratory workers in a pharmaceutical factory in Marburg, Federal Republic of Germany, who became ill with a hemorrhagic disease after being engaged in processing organs from African green monkeys (*Cercopithecus aelhiops*). In the course of the epidemic 17 more patients were admitted, and two medical staff members became infected while attending the patients in the hospital. In November, the last patient was admitted who apparently had been infected by her husband during the convalescent period (5, 6). Six more cases, including two people with secondary infections, occurred in Frankfurt, Federal Republic of Germany, who apparently developed the disease at the same time (7, 8). Additional cases occurred in September in Belgrade, Yugoslavia: a veterinarian, who became infected while performing an autopsy on dead monkeys, and his wife, who nursed him during the first days of the illness (9). In all, there were 31 cases, including six secondary cases,

and there were seven deaths (Tables I and II) (10). Serologic data obtained some years after the epidemic suggest an additional primary case in Marburg during the 1967 outbreak (W. Slenczka, unpublished data) (Table II). The infective agent was introduced by infected monkeys imported from Uganda, among which a few originally infected animals were probably responsible for the whole episode (Fig. 1). Numbers on hemorrhagic disease and death among the monkeys from the single shipment from Uganda have never been published, but all African green monkeys experimentally inoculated with the virus died (11). Aerosol transmission during the epidemic is very unlikely and infection from monkey to man occurred by direct contact with blood or organs of the animals, including tissue culture (Table II).

Marburg virus (MBG) remained an obscure medical curiosity until 1975, when three cases of Marburg hemorrhagic fever were reported from Johannesburg, South Africa (Table I; Fig. 1) (12). The index patient was a man who had traveled in Zimbabwe shortly before becoming ill. Seven days after onset of his illness, his traveling companion became ill, followed by a nurse who came down with the symptoms 7 days after contact with the second patient. The index case patient died 12 days after onset of the disease, whereas both patients secondarily infected survived. The last two episodes occurred in 1980 and 1987 in Kenya (Table I; Fig. 1). The index patient in 1980 became ill in western Kenya and died in Nairobi. An attending physician became infected but survived. Further spread was prevented, presumably by use of barrier nursing procedures (13). In 1987, a single fatal Marburg case was reported in western Kenya, near where the index patient in the 1980 episode had become infected (E. D. Johnson, unpublished).

2. Ebola Hemorrhagic Fever

Hemorrhagic fever caused by Ebola virus (EBO), another filovirus, emerged in 1976, when two epidemics occurred simultaneously in Zaire and Sudan (Table I). The agent was isolated from patients in both countries and named after a small river in northwestern Zaire. In June and July the first cases were reported from Nzara in western Equatoria Province of southern Sudan, a small town bordering the African rain forest zone (Fig. 1). The outbreak was strongly associated with index cases in a single cotton factory in town, and spread was to close relatives (67 cases). The epidemic was intensified by the spread of cases to neighboring areas, Maridi, Tembura, and Juba. High levels of transmission occurred in the hospital of Maridi, a teaching center for student nurses (213 cases). Despite the similarities of the clinical diseases and mortality rates, the epidemic in Nzara differed from the

TABLE I

OUTBREAKS OF FILOVIRAL HEMORRHAGIC FEVER[a]

Location	Year	Virus/subtype[b]	Cases (mortality)	Epidemiology
Germany/Yugoslavia	1967	Marburg	32 (23%)[c]	Imported monkeys from Uganda source of most human infections
Zimbabwe	1975	Marburg	3 (33%)[d]	Unknown origin; index case infected in Zimbabwe; secondary cases were infected in South Africa
Southern Sudan	1976	Ebola/Sudan	284 (53%)[e]	Unknown origin; spread mainly by close contact; nosocomial transmission and infection of medical care personnel
Northern Zaire	1976	Ebola/Zaire	318 (88%)[f]	Unknown origin; spread by close contact and by use of contaminated needles and syringes in hospitals
Tandala, Zaire	1977	Ebola/Zaire	1 (100%)[g]	Unknown origin; single case in missionary hospital; other cases may have occurred nearby
Southern Sudan	1979	Ebola/Sudan	34 (65%)[h]	Unknown origin; recurrent outbreak at the same site as the 1976 outbreak
Kenya	1980	Marburg	2 (50%)[i]	Unknown origin; index case infected in western Kenya died; a physician secondarily infected survived
Kenya	1987	Marburg	1 (100%)[j]	Unknown origin; expatriate traveling in western Kenya

USA	1989	Ebola/Reston	4 (0%)[k]	Introduction of virus with imported monkeys from the Philippines; four humans asymptomatically infected
Italy	1992	Ebola/Reston	0 (0%)[l]	Introduction of virus with imported monkeys from the Philippines; no human infections associated
Ivory Coast	1994	Ebola/(Ivory Coast?)	1 (0%)[m]	Contact with chimpanzee; single case
Kikwit, Zaire	1995	Ebola/Zaire	315 (77%)[n]	Unknown origin; course of outbreak as in 1976
Gabon	1995/96	Ebola	37 (57%)[o]	Contact with chimpanzee

[a] Beside the well-documented episodes listed in this table two more suspected fatal and two nonfatal cases of Ebola hemorrhagic fever including a laboratory infection have been reported (165, 167, 168).

[b] Subtypes of Marburg are not classified.

[c] (127), numbers include a primary case which was diagnosed some years after the epidemic (W. Slenczka, unpublished data).

[d] (12).

[e] (22).

[f] (23)

[g] (166).

[h] (25).

[i] (13).

[j] E. D. Johnson, unpublished data.

[k] (29, 30).

[l] (40).

[m] (26).

[n] (27, 28).

[o] (169).

TABLE II

DISTRIBUTION OF CASES BY TRANSMISSION ROUTE DURING FOUR MAJOR EPIDEMICS

Transmission type	Germany/ Yugoslavia, 1967[a]	Sudan, 1976[b]	Zaire, 1976[c]	Sudan, 1979[d]
Nosocomial	4 (12.5%)	6 (2.1%)	85 (26.7%)	4 (11.8%)
Person-to-person	1 (3.1%)	231 (81.3%)	149 (46.9%)	27 (79.4%)
Nosocomial or person-to-person	1 (3.1%)	18 (6.3%)	43 (13.5%)	—[e]
Contact with infected monkeys	26 (81.3%)[f]	—	—	—
Neonatal	—	—	11 (3.5%)	—
Unknown	—	29 (10.2%)	30 (9.4%)	3 (8.8%)
Total number of cases	32 (100%)	284 (100%)	318 (100%)	34 (100%)
Total number of deaths (mortality)	7 (21.9%)	151 (53.2%)	280 (88.1%)	22 (64.7%)

[a] Based on (127).
[b] Based on (22).
[c] Based on (23).
[d] Based on (25).
[e] —, no data available.
[f] Including one primary case which was diagnosed some years after the epidemic (W. Slenczka, unpublished data).

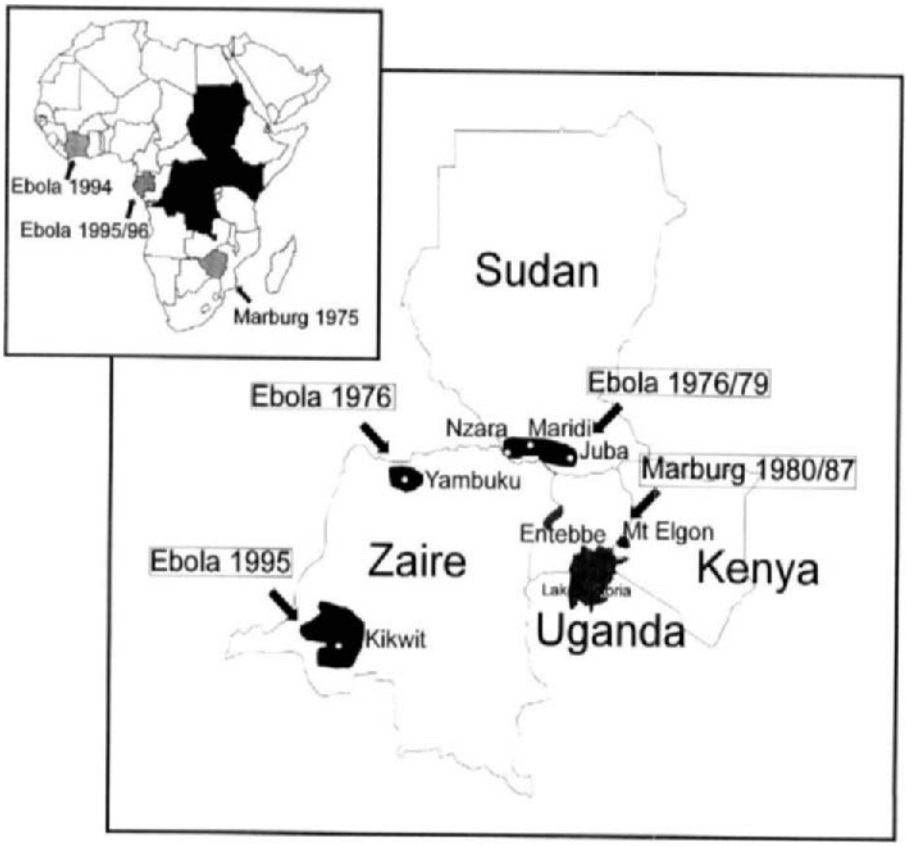


FIG. 1. Epicenters of hemorrhagic fever caused by filoviruses. Shown are the countries Kenya, Sudan, Uganda, and Zaire. The centers of the three major outbreaks of hemorrhagic fever caused by Ebola viruses are marked by black zones (Nzara, Maridi, and Juba, 1976 [22, 24, 25]; Yambuku, 1976 [23]; Kikwit, 1995 [27, 28]). Another black zone (Mount Elgon) indicates the region in which two index cases of Marburg hemorrhagic fever became infected (1980, 1987) (13; E. D. Johnson, unpublished data). A central holding station for wild-caught monkeys in the late 1960s was located near Entebbe at Lake Victoria. Infected vervet monkeys were shipped from here to Germany and Yugoslavia in 1967 (127). Three more episodes of filoviral hemorrhagic fever are indicated in the inset (gray): Marburg (Zimbabwe) 1975 (12), Ebola (Ivory Coast) 1994 (26), and Ebola (Gabon) 1995/96 (169).

one in Maridi. The Nzara outbreak involved factory workers and their close relatives whereas in Maridi the hospital served as both focus and amplifier of the infection. The outbreak lasted until November, during which time approximately 15 generations of person-to-person transmissions occurred (Table II). Transmission of the disease required close

contact with an acute case and was usually associated with nursing patients. The overall secondary attack rate was 12% and documented the relatively slow rate of spread into the community once out of the hospital. *In toto* there were 284 probable and confirmed cases involved with 151 deaths (53%) (Tables I and II).

By the end of August a second epidemic started in equatorial rain forest areas of northern Zaire (Fig. 1). A direct link between the two epidemics has always been discussed but never been verified. In total there were 318 probable or confirmed cases and 280 deaths (88%) (Tables I and II). The presumed index case came to Yambuku Mission Hospital for treatment of acute malaria, where he received an injection of chloroquine. It remains unclear whether this man was the source of the epidemic or became infected by the injection. Most persons acquired the disease following contact with patients, but for more than 25% the only risk factor elucidated was receipt of injections at Yambuku Mission Hospital (Table II). Nearly all survivors were infected by person-to-person contact. All ages and both sexes were affected, but the highest incidence was in women aged 15–29 years, who were frequently patients attending antenatal and outpatient clinics at the hospital. Although transmission was focused in the outpatient clinics of the hospital, there was subsequent dissemination in surrounding villages to people caring for sick relatives, attending childbirth, or having other forms of close contact. The secondary attack rate was approximately 5% overall but about 20% in close relatives of a patient. The epidemic, which lasted from the end of August until the end of October, spread relatively slowly in the epidemic area, and all infected villages (55; population <5000) were within 60 km of Yambuku.

Establishment of strict barrier nursing and classic public health principles, identification and isolation, was successful in controlling both epidemics. The spread by contaminated syringes and needles in Yambuku almost completely terminated when the hospital closed. The episode in Nzara died out spontaneously (14–23).

In 1979, Ebola hemorrhagic fever reemerged in Nzara and Yambio, which are located in the remote savanna of southern Sudan, near the border with Zaire (Table I; Fig. 1). The index case, a 45-year-old man, was admitted to the Nzara hospital with fever, vomiting, and diarrhea. He developed gastrointestinal bleeding and died 3 days postadmission. The index case worked in the same textile factory cited as the source of the 1976 outbreak in Sudan. The outbreak (July 31 to October 6, 1979) started from the hospital, where the index case patient was responsible for four nosocomial infections, which in turn led to disease in 5 families. Thirty-three cases could be traced to a human source of infection,

with 22 fatalities (65% mortality) (Tables I and II). Seven generations of virus transmission were estimated and mortality changed from 89% in the first four generations to 38% in the last three. Studies within the families confirmed reports from previous outbreaks (Sudan and Zaire, 1976) suggesting that Ebola virus is not easily transmitted and contact with body fluids of a patient is needed. Again, the hospital appeared to be the important focal point for dissemination (24, 25).

3. *New Emerging and Reemerging of Ebola Hemorrhagic Fever in Africa*

Two disease episodes of mortality were noticed among a group of chimpanzees in 1992 (8 deaths) and 1994 (12 deaths). The chimpanzees were objects of a 15-year observation by ethologists in the Tai National Park in western Ivory Coast. Several of the dead animals showed signs of hemorrhages, and one of the animals was autopsied in the field. A 34-year-old woman developed a dengue-like syndrome 8 days after performing the autopsy. She was admitted to the hospital in Abidjan 2 days later with continuing fever resistant to antimalarial treatment, diarrhea, and pruritic rush. Evacuation to Switzerland followed 5 days later when she developed a syndrome similar to that described for surviving Ebola-infected patients; she recovered without sequelae. An infection with an Ebola virus was confirmed by isolate-specific IgM and IgG antibodies, Ebola-Zaire-specific IgG antibodies, antigen-capture ELISA, reactivity to an Ebola serotype-specific monoclonal antibody, and virus isolation. Contact with infectious blood and tissues during the necropsy was considered to be the most likely source of the human infection. Organs of the dead chimpanzee were studied by immunohistochemistry and the findings were similar to those seen in material of the 1976 Ebola outbreaks and experimentally infected monkeys with Ebola virus. None of the persons in contact with either the case patient or the material of the chimpanzee tested antibody-positive (Table I; Fig. 1) (26).

Ebola hemorrhagic fever reemerged in Zaire in 1995. The first identified case related to the outbreak suffered an onset of illness on January 6, 1995 (Table I; Fig. 1). Until August 24, the official end of the epidemic, 315 cases had occurred, of which 244 died (77%). The center of the epidemic was Kikwit and the surrounding areas in Bandundu region in southwestern Zaire. The first case at Kikwit General Hospital was a male laboratory worker who had previously been admitted to Kikwit 2, a smaller second hospital in town. A laparotomy was performed after a differential diagnosis of typhoid fever with intestinal perforation. This was followed by a second laparotomy which showed

massive intraabdominal hemorrhage; the patient died 3 days later. Four days after the first laparotomy the first case among medical staff members occurred, with fever, headache, muscle aches, and hemorrhages. About three quarters of the first 70 patients were health care workers. The actual epidemic started within the hospitals. Prior to this time, cases had been sporadic. The major risk factors have been patient care in hospitals and households and preparation of bodies for burial. This is reflected by the fact that 26% of the cases with known professional occupation were medical staff members or students and 21% were housewives. During the course of surveillance, several chains of deaths have been identified which were traced as far back as late December 1994. The chain of the presumable index case, a charcoal worker, involves 7 out of 12 persons living in his household (27, 28).

An outbreak of Ebola hemorrhagic fever occurred in the village of Mayibout II, Makokou Health District, Ogooue-Ivindo Province, Gabon. It was linked to a chimpanzee found dead in the forest. A total number of 37 cases were diagnosed (mortality 56.8%) and 21 cases were directly exposed to the dead chimpanzee. A strain of Ebola virus was isolated from patient samples (169).

4. Reston Hemorrhagic Fever

In 1989, veterinary staff in a primate import quarantine facility in Reston, Virginia noted numerous deaths in cynomolgus monkeys in one animal room and suspected simian hemorrhagic fever (SHF) (Table I). Samples tested in the virus laboratory yielded SHF, but a filovirus closely related to Ebola was also isolated from those monkeys; the agent was called Ebola Reston (29–31). These cynomolgus monkeys (*Macaca fascicularis*) were imported from the Philippines. Shipments arrived either via Amsterdam or directly from the Philippines via the Pacific Ocean. No link to African or animals of other continents could be established on any route, so the presumption prevails that this new Ebola virus isolate is of Asian origin. The role of SHF in initiating or propagating the epidemic is unknown, but the filovirus was found to be pathogenic for monkeys under experimental conditions. Filoviral antigen and particles were found in tissues of naturally and experimentally infected monkeys in close anatomic relationship to the pathologic lesions (30–33). The epizootic occurring in monkeys spread through affected rooms by droplet contact with adjacent cages or to distant cages and different rooms by larger droplets and/or small-particle aerosols (34, 35). An airborne route of transmission was supported by the prominent respiratory involvement of the infected monkeys. Spread of the disease to other rooms of the facility lead to a decision to euthanize

all the monkeys in the building. Resumption of importation of monkeys led to new outbreaks of disease; subsequent investigation traced the source to a single source in the Philippines that was thought to have furnished all identified infected shipments, including monkeys sent to facilities in Texas and Pennsylvania (36). Four animal handlers at the quarantine facility became infected as judged by serological tests and, in one case, virus isolation. All four had high levels of daily exposure, but except for one, who cut himself while performing a necropsy, the mode of transmission is unclear (37–39). None of them had an unexplained febrile illness, suggesting that this virus may be less pathogenic for humans than previously known filoviruses, which have resulted in significant disease and mortality rates ranging from 22 to 88%. However, these observations should not be interpreted as assuring that this virus is not virulent for humans.

There is evidence that the outbreak in 1989 was not confined to the Reston facility, but occurred also in a branch in Münster, Germany, of the same company. However, since material has not been made available for laboratory diagnosis the nature of this agent remains uncertain.

In 1992, cynomolgus monkeys were imported into Italy from the same holding compound in the Philippines that exported the monkeys causing the 1989–1990 epizootic (Table I). Ebola Reston virus was isolated from three monkeys which died, and the remaining animals were sacrificed thereafter. No illness in associated humans has been reported (40). Reportedly, in 1991 monkeys were eliminated from the holding facility and cages disinfected. Thus, the virus either had persisted or was reintroduced by a similar mechanism as led to the 1989–1990 epizootic. Since the facilities in the Philippines were extensively cleaned following the first episode and since the isolated virus belonged to the EBO Reston subtype, it appears that reintroduction occurred from wild-caught animals of the same or a distinct endemic area.

Evidence for ongoing epizootic disease and transmission among captured monkeys at the export facility in the Philippines, which was the source of several shipments to the United States (1989–1990) and the shipment to Italy (1992) that contained infected monkeys, was demonstrated in 1990 and 1993. In 1990, filoviral antigen was detected by ELISA in 52.8% of dead monkeys from this facility in contrast to none in the dead monkeys from another facility in the Philippines. The investigation suggested that the type of holding cage was important in transmission, since being in a gang cage at the time of the initial serosurvey was a significant risk factor for subsequent infection (36). In the summer of 1993, high titered ELISA antibodies were present in monkeys held at that facility, but no evidence of viral antigen was found. Monkeys

imported from the facility into the United States in 1993 had stable IgG titers, suggesting infection in the recent past but not during quarantine (4). Even so, the original source remains unknown. It seems likely that naturally infected wild monkeys captured in the Philippines are the source of introduction.

B. Reservoir

Serological studies suggest that filoviruses are endemic in many countries of the Central African region (Fig. 2) (41–53). Recent serosurveys in other countries, such as Germany and the United States (54, 55), using several different techniques, suggest that filoviruses might also be endemic in those countries. Serological studies in relation to the EBO Reston outbreak indicated filovirus activity in the Philippines (Fig. 2). Although, as already mentioned, serological data based on IFA alone are of limited reliability, they at least suggest that subclinical infections caused by known or unknown filoviruses may occur and may be more common than expected. At this point, however, one has also to consider that filoviruses are members of the order *Mononegavirales*. This order includes many common human viruses that could be responsible for serological cross-reactivities.

MBG and subtypes Sudan and Zaire of EBO appear to be indigenous to the African continent, and both EBO subtypes have been isolated from human patients only in Africa. MBG has been isolated from human patients in Africa and Europe. The origin of the European cases could be traced back to foci in Uganda where vervet monkeys compounded in Entebbe (central holding station at Lake Victoria) were imported to Germany and Yugoslavia (Fig. 1). Complement-fixing antibodies were found in sera from some monkeys originally trapped near Lake Kyoga, the main area where vervet monkeys had been captured since the establishment of the trade in 1962. The finding of antibodies in three monkey trappers indicates that human infection may have occurred in Uganda during that time. However, all titers observed were weak, and an agent has never been isolated from the blood of a wild-trapped monkey or a monkey trapper (10, 56). Both index case patients of the episodes of MBG hemorrhagic fever in Kenya (1980/1987) had traveled in Mount Elgon region (Fig. 1). This region is not far from the shores of

FIG. 2. Prevalence of filovirus-reactive antibodies. *Key:* black, countries that have been subject to published serosurveys based mainly on IFA; gray, countries that have been subject to serosurveys based on different techniques. For references see Section II,B and reference list.

Lake Victoria and thus is close to the trapping place (Lake Kyoga, Uganda) and holding station (Entebbe, Uganda) of the monkeys that initiated the 1967 outbreak in Europe. One of the index cases had visited a cave (Kitum Cave) in that area shortly before becoming ill. Serological studies in this area, however, again failed to uncover the source of the virus. These studies included an extensive investigation of many animal species inhabiting the cave, including bats (E. D. Johnson, unpublished). Bats inhabiting buildings of a textile factory have also been considered as a potential source of the index cases of the 1976 and 1979 outbreaks of EBO hemorrhagic fever in southern Sudan. The geographical origin of both epidemics is less than 1000 km northwest of Lake Victoria and Mount Elgon (Fig. 1). A potential source for the 1976 Zairian outbreak has never been found, but a link to the simultaneous epidemic in Sudan has been discussed. These data strongly suggest an endemic focus for filoviruses in the equatorial rain forest areas of southern Sudan, northern Zaire, Uganda, and Kenya (Fig. 1).

The EBO Reston outbreak suggested for the first time the presence of a filovirus in Asia. Serological studies (IFA) among captive macaques in the Philippines indicated that the source of EBO Reston might be wild nonhuman primates. However, IFA-detected antibodies seem to be spurious, and latent infection in nonhuman primates has never been observed (57). Epidemiological data obtained in association with the 1994 Ivory Coast case suggested an Ebola epizootic among a group of chimpanzees as the cause of death. The pathogenicity of filoviruses, especially of EBO subtypes Sudan and Zaire and MBG, for nonhuman primates, however, does not support the concept of a reservoir in monkeys.

The reservoir of filoviruses remains a mystery. Many species have been discussed as possible natural hosts; however, no nonhuman vertebrate hosts or arthropod vectors have yet been identified. The high frequency of false-positives, especially when the EBO IFA is used, has contributed to the difficulties in finding the true reservoir for filoviruses. Filoviruses resemble "Old World" arenaviruses in several interesting biological properties, such as resistance to the antiviral effects of interferon, lack of *in vivo* neutralization, and lack of protection by convalescence sera. Arenaviruses cause chronic viremic infection in their rodent reservoirs. Thus, chronic infection of a mammal has to be considered as a mechanism that regulates survival of filoviruses in nature.

C. Transmission

Person-to-person transmission by physical contact with case patients is the main route of infection in human outbreaks (Table II). Activities such as nursing and preparing bodies for burial are espe-

cially associated with an increased risk of becoming infected. During the EBO outbreaks in 1976 and to some extent 1995, nosocomial transmission via contaminated syringes and needles was a major problem. Transmission does not seem to be efficient, as documented by secondary attack rates which on average rarely exceeded 12%. Thus, extreme care should be taken with blood, secretions, and excretions of infected patients. Sexual transmission has been described for MBG (58) and neonatal transmission has been reported for the 1976 outbreak in Zaire (16). Based on experience of the former episodes, isolation of patients and use of strict barrier nursing procedures (*e.g.*, protective clothing, respirator) are sufficient to interrupt transmission. Transmission by droplets and small-particle aerosols has been observed among experimentally infected (MBG) and quarantined imported monkeys (EBO RES, 1989–1990) (34, 59, 60). This is confirmed by identification of filovirus particles in alveoli of naturally and experimentally infected monkeys (33, 34, 61, 62) and human post mortem cases (63). However, the contribution of aerosol transmission to the course of human outbreaks is still unknown.

D. Molecular Evolution

The family *Filoviridae* has been constituted on the basis of unique morphologic, morphogenetic, physicochemical, and biological features of its members (64). Filoviruses can be separated into two types, which are clearly distinguished by the features listed in Table III. In general, the MEG viruses seem to be unique without known subtypes, but at least two different genetic lineages coexist (65). EBO, however, can be subdivided into at least three subtypes: Zaire, Sudan, and Reston (66–69). Molecular characterization of the Ivory Coast virus revealed a novel lineage, suggesting a fourth subtype of EBO (Fig. 1) (69). There is a lack of antigenic cross-reactivity between the types, but the subtypes of EBO share common epitopes (68, 70). Nucleotide sequence comparison among MBG and EBO shows only scattered similarities, which is in contrast to the similarities seen among amino acid sequences of structural proteins (71–73). This finding indicates that these agents may have diverged at some point in the distant past. A distinction within the EBO type is based on earlier peptide and oligonucleotide mapping (74, 75) and has been confirmed by recent sequence analysis of the glycoprotein genes. That study showed all four subtypes to differ from one another to a comparable extent: 37–41% nucleotide differences (69). This suggests that filoviruses have evolved into specific niches and may reflect a similar divergence in the natural hosts, assuming they have coevolved. Genetic variability seen among different virus isolates of one subtype seems to be much less than for some other RNA viruses.

TABLE III

CHARACTERISTICS OF MARBURG AND EBOLA TYPES OF FILOVIRUSES

Features	Type Marburg	Type Ebola
Serological cross-reactivity to other type	No	No
Subtypes	1[a]	3[b]
Glycoprotein (SDS-PAGE)[c]	About 170 kDa	About 140 kDa
Terminal sialylation of carbohydrates[d]	No	Yes
Nucleoprotein (SDS-PAGE)	About 95 kDa	About 105 kDa
Nonstructural proteins	No	1[e]
Editing	No	Yes[f]
Gene overlaps	1[g]	>1[h]
Overlapping ORF[i] in gene 2	Yes	No

[a] Type Marburg: subtype Marburg.
[b] Type Ebola: subtypes Zaire, Sudan, Reston.
[c] SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.
[d] By propagation of viruses in Vero E6 and MA-104 cells (monkey kidney cell lines).
[e] Small glycosylated protein encoded by gene 4.
[f] Gene 4.
[g] Between VP30 and VP24.
[h] Between VP35 and VP40, GP and VP30, and VP24 and L (Ebola, subtype Zaire).
[i] ORF, opening reading frame.

In particular, the close genetic relation of the two Zairian isolates from 1976 and 1995, with less than 1.6% difference in the GP gene (69), suggests that filovirus variants may not emerge as rapidly in nature.

Molecular analyses of the genomes clearly demonstrated that filoviruses are the closest relatives to *Rhabdoviridae* and *Paramyxoviridae* (Fig. 3). All nonsegmented negative-stranded (NNS) RNA viruses share a similar genome organization, with conserved regions at both ends encoding the core and L proteins surrounding a variable part in the middle encoding the envelope proteins (Fig. 3B). Filovirus genomes are more complex than those of lyssaviruses and vesiculoviruses and align organizationally more closely to members of the genera *Paramyxovirus* and *Morbillivirus*. This relationship is confirmed on the amino acid level, as demonstrated for the nucleoproteins and polymerases (L proteins) (67, 71). Transcription and replication of all NNS RNA viruses, including filoviruses, follow common principles, such as (i) helical ribonucleocapsid complex as the functional template for synthesis of replicative and messenger RNA, (ii) transcription of messenger RNAs by sequential interrupted synthesis from a single promoter at the 3′ genomic end, (iii) replication via a full-length anti-

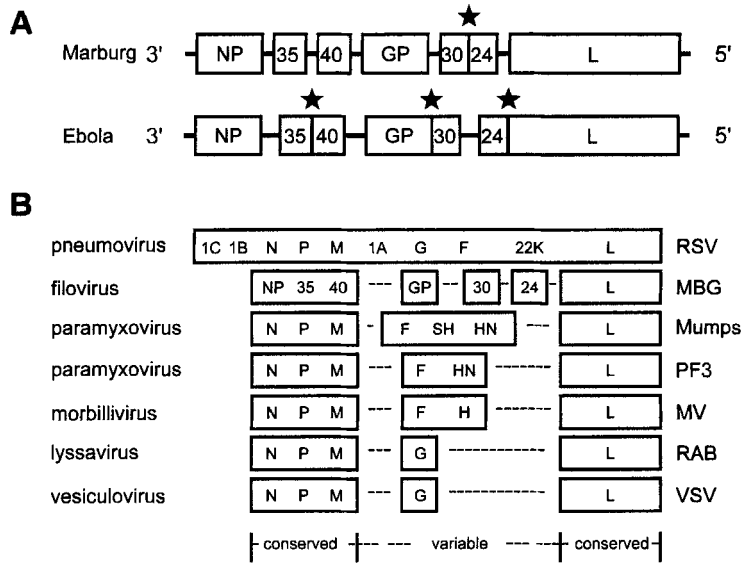


FIG. 3. (A) Genome organization of filoviruses. Filoviral genomes consist of a single, negative-stranded, linear RNA molecule. Differences in organization between Marburg and Ebola type viruses are indicated. Key: asterisk, position of gene overlap; GP, glycoprotein gene; L, polymerase (L) gene; NP, nucleoprotein gene; 24 / 30 / 35 / 40, virion structural protein (VP) genes. (B) Genome organizations of nonsegmented negative-stranded RNA viruses. Compared are genomes of viruses belonging to different genera of the three families Paramyxoviridae, Rhabdoviridae, and Filoviridae (Order *Mononegavirales*). Conserved and variable regions are identified. *Key:* RSV, human respiratory syncytial virus; MBG, Marburg virus; Mumps, mumps virus; PF3, human parainfluenza 3 virus; MV, measles virus; RAB, rabies virus; VSV, vesicular stomatitis virus; N and NP, nucleoprotein gene; P, phosphoprotein gene; 35, virion structural protein (VP) 35 gene, the putative P equivalent; M, matrix protein gene; 40, VP40 gene, the putative M equivalent; G, GP, F, H, HN, glycosylated membrane protein genes; SH, small hydrophobic protein gene; 22K, nonglycosylated membrane protein gene; 30, VP30 gene, unknown function; 24, VP24 gene, unknown function; L, polymerase gene.

genome of positive sense, (iv) transcription and replication in the cytoplasm, and (v) maturation by envelopment of independently assented ribonucleocapsid complexes at membrane sites containing viral proteins. This is reflected by common genomic features such as complementarity of the genome termini homologies in the 3′ leader regions, conservation of transcriptional signals, separation by intergenic sequences, and expression of virion-associated RNA-dependent RNA polymerases. In conclusion, all data available today support the concept of an order *Mononegavirales* comprising the three unique families *Paramyxoviridae, Rhabdoviridae,* and *Filoviridae* (76).

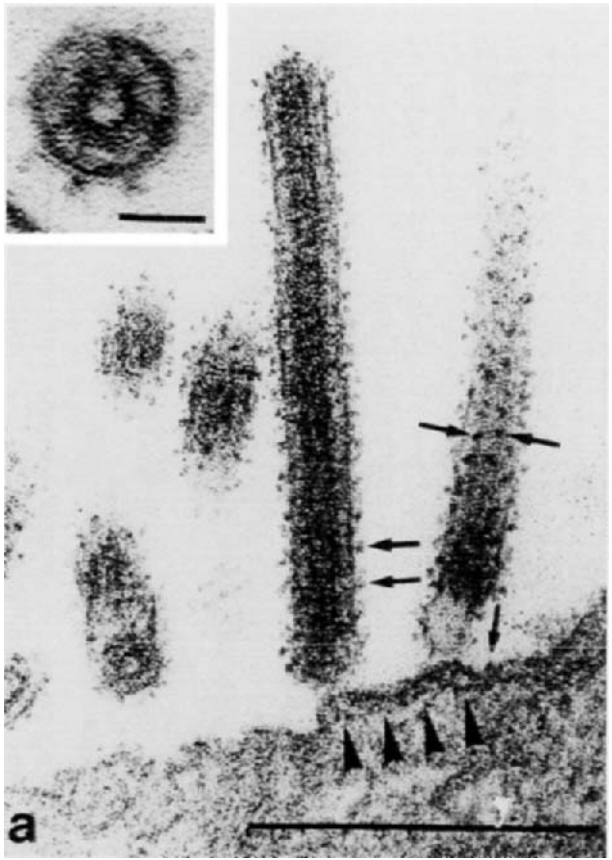
III. Biology of Filoviruses

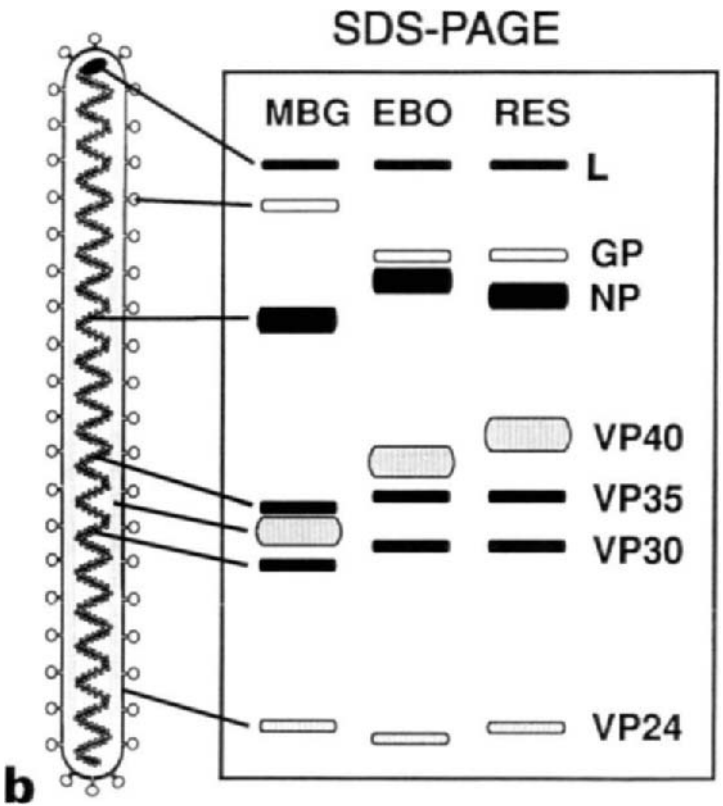
A. Morphology

The long filamentous shape of the particles is unique among viruses and has been decisive for classification. Particles appear in different shapes, such as branched, circular, or U- and 6-shaped forms. Virions vary greatly in length but show a uniform diameter of approximately 80 nm. Family members differ in length of virion particles but seem to be very similar in morphology. Peak infectivity has been associated with particles of 665 nm for MBG and 805 nm for EBO. Virions are composed of a central core formed by a ribonucleocapsid complex (RNP) which is surrounded by a lipid envelope derived from the host cell plasma membrane. Electron micrographs demonstrate an axial channel (10–15 nm in diameter) within the RNP. The channel is surrounded by a central dark layer (20 nm in diameter) and an outer helical layer (50 nm in diameter) with cross-striations at 5 nm intervals. Spikes approximately 7 nm in length and spaced at about 10 nm intervals form globular structures on the virion surface (Fig. 4A) (19, 58, 64, 77–79).

The RNP is composed of a single molecule of linear RNA and four of the seven virion structural proteins [nucleoprotein (NP), VP30, VP35, and the large (L) protein]. Genomic RNA has an M_r of 4.2×10^6 and constitutes 1.1% of the virion mass (80). The three remaining structural proteins are membrane-associated, with the glycoprotein (GP) as a type I transmembrane protein (81) and VP24 and VP40 probably located at the inner side of the membrane. Virus particles have an M_r of approximately $3–6 \times 10^8$ and a density in potassium tartrate of 1.14 g/cm^3 (Fig. 4B) (66).

Fig. 4. Morphology of filoviral particles. (A) Electron micrograph. Budding of Marburg virus particles from the plasma membrane of infected primary cultures of human endothelial cells. Particles consist of a nucleocapsid surrounded by a membrane in which spikes are inserted (*arrows*). The nucleocapsid contains a central channel (*inset*). The plasma membrane of infected cells is often thickened at locations where budding occurs (*arrowheads*). Ultrathin section—bar, 0.5 μm; bar inset, 50 nm (116). (B) Electrophoretic mobility patterns of filoviral structural proteins. The mobility patterns (SDS-PAGE) of structural proteins of Marburg (MBG, strain Musoke) and two Ebola subtypes, Zaire (EBO) and Reston (RES), are compared and differences illustrated. Four proteins are involved in nucleocapsid formation: polymerase or large (L) protein, nucleoprotein (NP), virion structural protein (VP) 30 and VP35 (*black*). The glycoprotein (GP) is a transmembrane protein and anchored with the carboxy-terminal part in the virion membrane (*white*). Homotrimers of GP form the spikes on the virion surface [*arrows* in (A)]. VP40 and VP24 are membrane-associated proteins (*gray*).


a


SDS-PAGE
MBG
EBO
RES
L
GP
NP
VP40
VP35
VP30
VP24
b

B. Genome

Genomes of filoviruses consist of a single negative-stranded linear RNA molecule (64, 80, 82). The RNA is noninfectious, not polyadenylated, and complementary to polyadenylated viral subgenomic RNA species (64, 72, 83). The nucleic acid sequences of two different isolates of MBG (83, 84) and the EBO Mayinga isolate (subtype Zaire) (72, 85) as well as parts of EBO Reston and EBO Maridi and Nzara isolates (subtype Sudan) (69; A. Sanchez, personal communication) have been elucidated. Filovirus genomes have a length of approximately 19 kb (19.1 for MBG and 18.9 kb for EBO) and are larger than all other negative-stranded RNA virus genomes. Genes have been identified by highly conserved transcriptional signals at their 3′ and 5′ ends. The following order is characteristic for filoviruses: 3′ leader–NP–VP35–VP40–GP–VP30–VP24–L–5′ trailer (Fig. 3A).

Genes are separated by intergenic regions varying in length and nucleotide composition. Some genes overlap but the positions and numbers of overlaps are different among filoviruses (Fig. 3A; Table III). Viruses belonging to the Zairian subtype of EBO possess three overlaps located between VP35 and VP40, GP and VP30, and VP24 and L, whereas MBG isolates have only one overlap between VP30 and VP24. The length of the overlaps is limited to five highly conserved nucleotides within the transcriptional signals (3′–UAAUU–5′) (Fig 5). Transcriptional start signals are conserved among filoviruses, and the sequence 3′–CUNCNUNUAAUU–5′ represents the consensus motif (Fig. 5). Transcriptional stop signals are identical for all genes (3′–UAAUUCUUUUU–5′) with the exception of the VP40 gene of MBG (C at position 2 instead of an A; genomic sense) (Fig. 5). Most genes tend to possess long noncoding sequences at their 3′ and/or 5′ ends which contribute to the increased length of the genome. Upstream of the N gene start site and downstream of the L gene stop site there are extragenic sequences, which are thought to be templates for small viral, nonpolyadenylated subgenomic RNAs synthesized during infection. The genomes are complementary at the very extreme ends, a feature known for all NNS RNA viruses (83, 85, 86).

C. Viral Proteins

1. NP—Nucleoprotein

The NP protein is encoded by gene 1 at the extreme 3′ end of the linear unsegmented RNA genome. NP proteins differ slightly in their electrophoretic mobility patterns, ranging from 95 kDa for MBG to 105

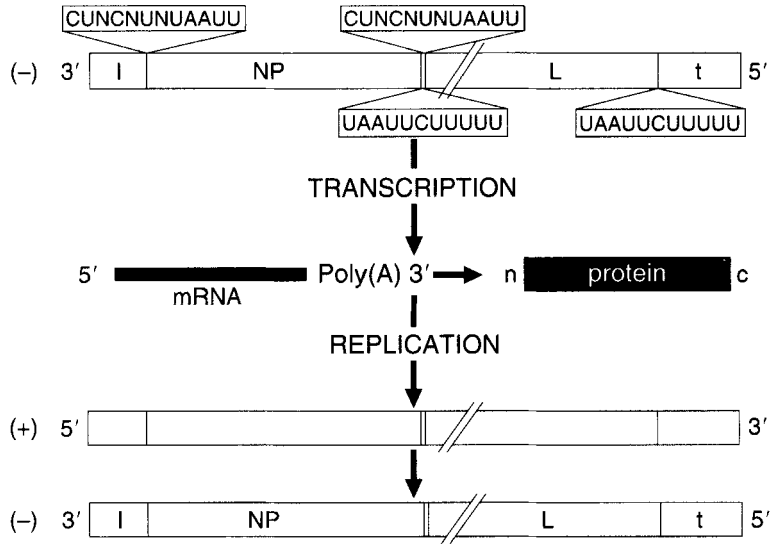


FIG. 5. Replication cycle of filoviruses. The model illustrates the mode of transcription and replication based on the data available to date. Each gene on the linear arranged nonsegmented (–)-sense genome is flanked by conserved transcriptional start (3′–CUNCNUN*UAAUU*–5′; indicated above) and termination signals (3′–UAAUUCUUUUU–5′; indicated beneath). Transcription starts at the 3′ end of the (–)-sense genome and leads to polyadenylated mRNA species. For replication a full-length (+)-sense antigenome is synthesized which serves as the template for the synthesis of progeny (–)-sense RNA anticomplementary to the parental RNA. *Key:* c, carboxy-terminal end of proteins; l, 3′ untranslated region (leader); n, amino-terminal end of proteins; Poly(A), polyadenylation of mRNA species; t, 5′ untranslated region (trailer); NP, nucleoprotein; L, viral RNA-dependent RNA polymerase.

kDa for EBO isolates. The molecular weights (MW) calculated from the deduced amino acid sequences of the corresponding genes of MBG (695 amino acids) and EBO (739 amino acids) are 78 kDa and 83 kDa, respectively, and the differences in lengths are related to the less-conserved C termini of the protein (71, 83, 84, 87). Thus, filovirus NP proteins possess an unusually high MW compared with other NNS-RNA virus nucleocapsid proteins, which range from 42 to 62 kDa. This suggests additional functions for the filovirus NP protein located in its unique C terminus. The NP protein is the major structural phosphoprotein, and only the phosphorylated form of the protein is incorporated into virions, as demonstrated for MBG (88–90). This finding may indicate that phosphorylation is needed to interact with genomic RNA and to form virion RNPs for progeny viruses. Sequence comparison of

NP proteins of MBG and EBO shows a high degree of homology within the first 400 predicted amino acids. The alignment shows that the region from position 130 to 392 of the MBG sequence has a very strong similarity and is highlighted by a run of 34 identical amino acids from position 296 to 329. The fact that two of the three cysteine residues of NP are conserved may indicate their role in proper folding of the molecule. A small region in the middle of the MBG and EBO NP sequences was found to contain a significant amino acid homology with paramyxoviruses and to a lesser extent with rhabdoviruses (71). The NP proteins of filoviruses and other NNS viruses also have the hydrophobicity of their N termini in common. A role of this region in either protein folding and/or RNA binding has been postulated for other NNS-RNA viruses (91, 92). The less-conserved C-terminal half of filovirus NP proteins, which is hydrophilic and very acidic, may function in the assembly process by interacting with the matrix proteins or the presumed second proposed nucleoprotein VP30 (66, 71, 88). Similar functions have been discussed before for the variable C termini of paramyxovirus NPs (91, 93). The NP protein is the major component of the RNP and is tightly bound within the complex. Although RNA binding has not yet been demonstrated, there is little doubt that this protein is the functional analogue of the nucleocapsid proteins of paramyxoviruses and rhabdoviruses (Fig. 4B; Table IV).

2. VP35—Polymerase Cofactor?

VP35 is encoded by gene 2. It has a length of 329 amino acids in MBG (73, 83); VP35 of EBO is 351 or 340 amino acids long (72, 73, respectively). RNP association of the protein is much weaker than that of NP and VP30, as demonstrated by nonionic detergent treatment of virion particles (66, 88). VP35 of EBO virions is not phosphorylated (88, 89). Expression studies of MBG VP35 in insect cells (SF9 cells), however, revealed weak phosphorylation of this protein, whereas VP35 expressed in HeLa cells using the vaccinia-virus-driven T7 polymerase system was not phosphorylated (H. Feldmann, unpublished data). Thus, VP35 may exist in a phosphorylated and an unphosphorylated form as has been demonstrated for NP. For NP, the unphosphorylated form seems to be incorporated into virion particles. Hydropathy plots of MBG and EBO VP35 showed similar profiles and a prominent common hydrophilic domain in close proximity to the N termini (MBG, positions 28 to 42; EBO, positions 57 to 76). This region may be involved in template binding; this is supported by the fact that VP35 unspecifically binds nucleic acids (H. Feldmann, unpublished data). In spite of the

TABLE IV

FILOVIRAL PROTEINS AND THEIR PROPOSED FUNCTION[a]

Designation	Virus type	Encoding gene	Localization	Proposed function
NP	MBG/EBO	1	Ribonucleocapsid complex	Encapsidation
VP35	MGB/EBO	2	Ribonucleocapsid complex	Phosphoprotein analogue
VP40	MBG/EBO	3	Membrane-association	Matrix protein
GP	MBG/EBO	4	Surface (transmembrane protein)	Receptor binding, fusion
VP30	MBG/EBO	5	Ribonucleocapsid complex	Encapsidation, RNA binding
VP24	MBG/EBO	6	Membrane-association	Unknown
L	MBG/EBO	7	Ribonucleocapsid complex	RNA-dependent RNA polymerase
sGP	EBO	4[b]	Nonstructural, secreted	Unknown

[a] NP, nucleoprotein; VP, virion structural protein; GP, glycoprotein; L, large protein (polymerase); sGP, small glycoprotein; MBG, type Marburg filoviruses; EBO, type Ebola filoviruses.

[b] Expressed by RNA editing (69, 112).

inconsistent data on phosphorylation among filoviruses and the lack of sequence homology, the genome position of the corresponding gene combined with the association in the RNP suggest that VP35 is functionally analogous to the P proteins of paramyxoviruses and rhabdoviruses. Further studies are needed to show if "P" protein would be an appropriate designation for this protein (Fig. 4B; Table IV).

3. VP40—Matrix Protein

VP40 of filoviruses is encoded by gene 3. It is 303 and 326 amino acids long in MBG and EBO, respectively (72, 73, 83). Differences in the electrophoretic mobilities of VP40 can be used to distinguish between MBG and EBO isolates and even among the three EBO subtypes (Fig. 4B) (68). Nitrocellulose-bound VP40 binds in a radiooverlay protein assay unspecifically to nucleic acids (H. Feldmann, unpublished data),

implicating a role in regulation of transcription/replication as has been described for the matrix protein of vesicular stomatitis virus (94, 95). VP40 is not associated with the RNP complex and behaves like a membrane-associated protein when analyzed following nonionic detergent treatment of virion particles (66, 88). This finding, together with a predominantly hydrophobic profile, the abundance in virion particles, and the genome localization of the corresponding gene, suggests that VP40 is the matrix protein analogue of filoviruses (Fig. 4B; Table IV).

4. GP—Glycoprotein

GP, encoded by gene 4 of the genome, is the only glycosylated structural protein of virions. GP of MBG and EBO is 681 and 676 amino acids in length, respectively (72, 81, 96, 97). Filovirus GPs are type I transmembrane proteins anchored via a C-terminal hydrophobic domain in the membrane (Fig. 6A). They are directed into the endoplasmic reticulum by an N-terminal hydrophobic domain which shows the structural requirements for signal peptides and can be cleaved by signal peptidases as shown directly for the MBG GP (81). Filovirus GPs contain *N*- and *O*-glycans that account for up to 50% of the MW of the mature proteins. Oligosaccharide side chains differ in their terminal sialylation patterns, which seem to be isolate- as well as cell-line-dependent (68, 98). Detailed structural analyses of filovirus carbohydrates are available for MBG only. The structures include oligomannosidic and hybrid-type *N*-glycans as well as bi-, tri-, and tetraantennary complex species and high amounts of neutral mucin-type *O*-glycans (99). Amino acid sequence comparison of filovirus GPs showed conservation at the N- and C-terminal ends of the proteins in which the two hydrophobic domains (signal peptide, membrane anchor) and most of the highly conserved cysteine residues are located. The middle region is variable and extremely hydrophilic and carries the bulk of the glycosylation sites for *N*- (EBO 17 sites; MBG 22 sites) and *O*-glycans (Fig. 6A). Recently it has been shown that the two cysteine residues at positions 671 and 673 of MBG GP are acylated (100). Acylation at the border between membrane anchor region and cytoplasmic tail has been shown for many viral type I transmembrane proteins. The special arrangement of all cysteine residues in the molecule favors an intermolecular cysteine bridge formation between the two external parts of the molecule, resulting in a stem region with a crown-like domain on the top carrying the mass of the carbohydrate side chains (Fig. 6B). For MBG it has been shown that the mature GP is inserted in the membrane as a homotrimer, and oligomerization seems to be mediated by intramolecular disulfide bridges, since complexes can be destroyed

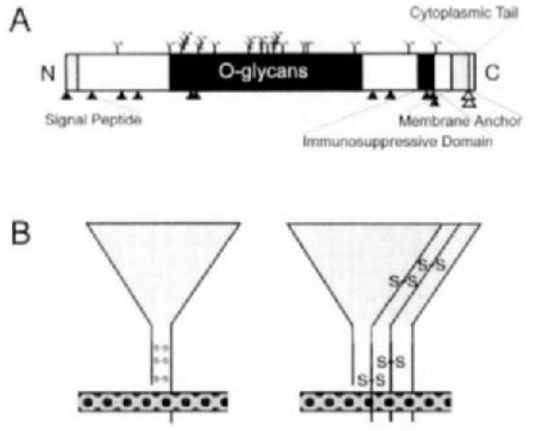


FIG. 6. (A) Schematic illustration of structural features of filoviral glycoproteins. The glycoprotein (GP) of filoviruses is encoded by gene 4. It is a type I transmembrane protein carrying two hydrophobic domains (*light gray*) at the amino-terminal (signal peptide) and carboxy-terminal ends (membrane anchor). An external domain in close proximity to the transmembrane region (*black*) shows significant homology to an immunosuppressive domain in envelope proteins of several retroviruses. The protein is highly glycosylated, carrying N- and O-linked oligosaccharides (γ) (here demonstrated for MBG strain Musoke). (B) Structural model of the surface spikes of filoviruses. The surface spikes of filoviral particles are formed by homotrimers of the glycoprotein (GP). Each monomer seems to form a stemlike structure with a crownlike domain on the top carrying the mass of the carbohydrate site chains. This formation may be mediated by intramolecular disulfide bridge formation (S–S) of cysteine residues located at the amino and carboxy termini of the protein [▲ in (A)]. Oligomerization seems to be due to intermolecular disulfide bridge formation (S–S) (98). *Key:* ▲, cysteine residues; △, acylated cysteine residues; γ, N-linked oligosaccharides.

under reducing conditions (Fig. 6B) (98). In general, filovirus GPs lack significant homologies with envelope proteins of other NNS RNA viruses. However, a region of 26 amino acids in the external domain in close proximity to the transmembrane region shows significant homology to an immunosuppressive domain in the envelope proteins of several retroviruses (Figs. 6A and 7) (72, 81, 96, 97). The immunosuppressive domain in retroviruses has been assumed to be involved in inhibition of blastogenesis of lymphocytes, decrease in monocyte chemotaxis and macrophage infiltration, inhibition of human natural killer-cell

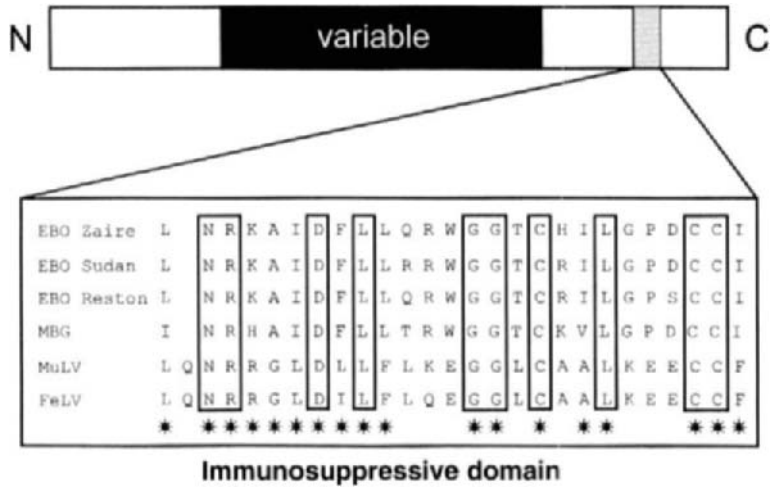


FIG. 7. Immunosuppressive motif located on filoviral glycoproteins. Shown is an alignment of the putative immunosuppressive motif (peptide of 26 amino acids) of the Zaire, Sudan, and Reston subtypes of Ebola viruses and Marburg virus (strain Musoke), compared with the known motif of two oncogenic retroviruses, murine leukemia virus (MuLV) and feline leukemia virus (FeLV). Identical amino acids are framed and functionally similar ones are identified beneath the alignment by asterisks. The motifs are located in close proximity to the transmembrane region on the conserved carboxy-terminal part of the molecules.

activity, and blocking of protein kinase C activity (101–103). Experimental data on GP function do not exist. However, the fact that GP is the only surface protein of the virions suggests a function in mediation of binding to cellular receptors and fusion with cellular membranes. Furthermore, GP is discussed as the major viral antigen and the main target for the host immune response (Figs. 4, 6, and 7; Table IV).

5. *VP30—Minor Nucleoprotein?*

VP30 is encoded by gene 5 and intimately associated with the RNP (66, 88). The protein has a length of 260 and 281 amino acids in EBO and MBG, respectively (72, 83, 84). The protein binds RNA under denaturing conditions in an RNA–protein overlay assay and forms complexes with the NP protein. Complexes of expressed recombinant VP30 and NP can be precipitated using either anti-VP30 or anti-NP antibodies (H. Feldmann, unpublished data). VP30 of EBO has been identified as the minor phosphoprotein of virions (88, 89). The available data indicate that VP30 may work as a functional unit in encapsidation of the

RNA genome (66), which could be achieved by binding to NP and/or binding to genomic RNA. It could also play a role as an additional cofactor of the transcriptase–replicase complex (Fig. 4; Table IV).

6. VP24—Membrane-Associated Protein of Unknown Function

VP24 is encoded by gene 6 of filoviruses. It is 253 and 251 amino acids in length in MBG and ERO, respectively (72, 83, 84). The protein is membrane-associated. Unlike VP40, it is not completely removed from the RNP under isotonic conditions (66, 88). VP24 presumably serves as a second matrix protein and may bind to the cytoplasmic tail of GP or may link the membrane proteins (VP40 and/or GP) to the RNP. There are minor differences in the SDS-PAGE migration profile of this protein among filovirus isolates (Fig. 4B). Such differences have also been observed between EBO wild type (subtype Zaire) and a highly pathogenic variant isolated after several passages from guinea pigs (104). To what extent these changes contribute to the higher pathogenic potential of the variant in guinea pigs is currently unknown (Fig. 4B, Table IV).

7. L—Large Protein

The L protein is encoded at the 5′ end of the linear genome and has a predicted MW of 267 kDa (2331 amino acids) for strains Musoke (105) and Popp (84) of MBG. Computer-assisted comparison revealed significant homologies to L proteins of other NNS RNA viruses. Homologies are mainly located in the N-terminal half of the protein and are concentrated within three common domains, named boxes A, B, and C (106). Other common features are a high content of leucine and isoleucine residues, a large positive net charge, clusters of basic amino acids, putative ATP binding sites, two neighboring cysteine residues located in the C-terminal half of the protein, and the genome localization of the encoding gene. A highly conserved peptide motif –GDNQ– located at the C-terminal end of domain B (positions 744–747) and flanked by hydrophobic amino acid residues seems to be correlated with enzymatic functions of the protein. Mutations in this domain, which is present in all NNS RNA virus L proteins, abolished activity of other NNS RNA virus L proteins (107). Furthermore, an –LDD– motif is present at positions 1095–1097. Similar motifs with alterations in the first amino acid have been described and discussed as active sites for some RNA-dependent RNA polymerases of plant, animals, and bacterial viruses (108–111). Even though transcriptase and replicase activities have not yet been demonstrated, the L protein is regarded as an RNA-dependent RNA polymerase (Fig 4B; Table IV).

8. *Nonstructural Proteins*

A nonstructural glycoprotein has recently been discovered in EBO (69, 112). This protein, designated sGP, is expressed from the glycoprotein gene by RNA editing (Fig. 8). sGP shares about 300 N-terminal amino acids with GP but has a different C terminus (~70 amino acids), which contains many charged residues as well as conserved cysteines. The protein is directed into the endoplasmic reticulum and becomes *N*- and *O*-glycosylated. sGP is secreted into culture medium in high quantities. The function of this nonstructural protein is unknown, but it could modulate the host immune response by binding antibodies. No similar protein is found in MBG (Table IV).

MBG isolates, on the other hand, possess a second small open reading frame (ORF) located within the ORF of gene 2 (65). Expression would result in an extremely basic protein of 60 amino acids. The putative product shares features with a recently identified nonstructural protein of vesicular stomatitis virus (VSV). This protein is also encoded by a second ORF of gene 2, which normally encodes for the P protein of VSV (113). The overlapping ORF in gene 2 is not found in EBO.

D. Virus Replication

1. *Virus Growth in Cell Cultures*

Vero cells, especially clone E6, are the most widely used line for virus isolation and propagation. Primary virus isolation has also been successful in MA-104 and SW13 cells (30, 114). Furthermore, a variety of other cell lines have been tested as substrates for filovirus replication (41, 114, 115). This includes a recently developed human microvascular endothelial cell line (HMEC-1) (H. Feldmann, unpublished data), primary cultures of human umbilical cord vein endothelial cells (HUVEC) (116), and human peripheral blood monocytes/macrophages (117).

MBG and EBO subtype Zaire cause lytic infections in cell culture with distinct cytopathogenic effects. Infection with EBO subtypes Sudan and Reston proceeds more slowly, and cytopathogenic effects are not as obvious. The course of filovirus infection in tissue culture can be controlled by indirect immunofluorescence assay (IFA) using antibodies directed against viral antigens. Plaque assays are performed in Vero E6 cells; however, the ratio of plaque-forming units to infectious particles is thought to be relatively low. The number of infectious particles may therefore be underestimated in this assay. When there are few or no cytopathogenic effects, reverse transcriptase–polymerase chain reaction (RT-PCR) on viral RNA isolated from infected cells and tissue

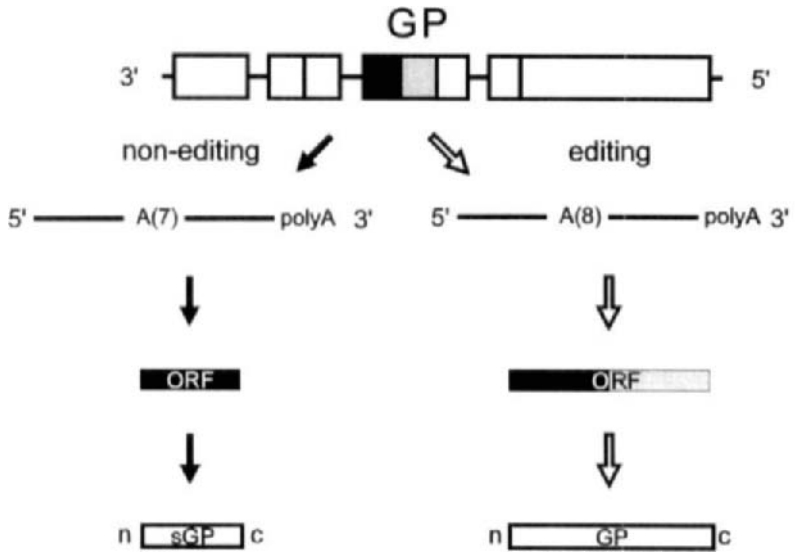


FIG. 8. Expression strategies of gene 4 of type Ebola viruses. Gene 4 of Ebola viruses is transcribed from two open reading frames. The primary gene product is a small glycoprotein (sGP). Full-length glycoprotein (GP) is expressed by RNA editing. *Key:* A, adenosine residue; c, carboxy-terminal end of proteins; n, amino-terminal end of proteins; GP, glycoprotein gene, 3′ and 5′, terminal ends of genomes and subgenomic RNAs.

culture supernatants can be helpful for quantification (116, 118). Furthermore, Northern blot hybridization and/or analysis of *in vivo* labeled viral subgenomic RNA have been used to study the course of infection (83, 119).

Viral RNA synthesis in tissue culture is detectable by at least 7 hr postinfection, reaches a maximum by 18 hr, and declines thereafter. Cytopathogenic effects are not seen before 48 hr postinfection. The first mRNA to be detected is NP-specific; it reaches levels sufficient to produce protein by 7 hr postinfection (119). This is in line with other NNS-RNA viruses, showing transcription to start at the most 3′-located gene of the linear genome (120, 121). All proteins are detectable by *in vitro* translation of polyadenylated RNA isolated 18 hr postinfection; thereafter the yield of translation products decreases (119). Recently, a PCR assay detected genomic RNA of MBG Strain Musoke (MUS) particles in clarified supernatants of infected cells as early as 12 hr postinfection, indicating that the replication cycle is approximately 12 hr (116). More sensitive technologies such as PCR have to be used in order to evaluate precisly the course of transcription and replication for filoviruses.

2. *Virus Entry*

Cell entry seems to be mediated by the GP as the only surface protein of virion particles. Studies on MBG MUS infections of hepatocytes have identified the asialoglycoprotein receptor present on these cells as a receptor candidate (122). However, one has to postulate additional receptors, since the asialoglycoprotein receptor is not expressed on many virus-susceptible cells. Furthermore, MBG GP is not generally lacking sialic acid, and sialylation has been shown to be cell-line-dependent (68). Whether the next step in virus entry involves a fusion process at the plasma membrane or fusion following endocytosis of virus particles is not known. The uncoating mechanism has not been studied either.

3. *Transcription, Translation, and Genome Replication*

Filovirus transcription and replication take place in the cytoplasm of infected cells. The data available so far suggest that the mechanisms involved resemble those observed for other NNS RNA viruses (Fig. 5). Transcription probably starts with a short (+)-leader sequence as in the case of some other NNS RNA viruses (123, 124). Subsequently, genomes are transcribed into monocistronic subgenomic RNA (mRNA) species which are complementary to viral genomic RNA and polyadenylated. Seven subgenomic RNA species have been detected in filovirus-infected cells by either Northern blot hybridization or RT-PCR amplification. There is no evidence for larger amounts of bi- or multicistronic subgenomic RNA species (72, 83, 119). Recent data on RNA amplification and direct sequencing of amplified products of four (NP, VP35, VP40, VP30) of the seven polyadenylated subgenomic RNA species of MBG MUS (125) revealed that the 5′ ends of the transcripts are two bases shorter than previously published (71, 72, 83). Thus, it appears that all start signals of filovirus genes possess the consensus sequence 3′–CUNCNUNUAAUU–5′ (Fig. 5). The 3′ ends of the transcripts carry a poly(A) tail generated by a stuttering mechanism of the viral polymerase at a run of five or six uridine residues located at the 5′ end of all transcription stop signals (72, 83). Therefore, the sequence 3′–UAAUUCUUUUU(U)–5′ serves as a transcription stop and polyadenylation signal (Fig. 5). As mentioned above, both signals carry the pentamer 3′–UAAUU–5′, which is characteristic for the family and a unique feature among NNS RNA viruses. The function of the pentamer is unknown, but it could serve as the recognition site for the polymerase complex. The surrounding semiconserved sequences may then mediate the exact initiation of transcription and termination–polyadenylation events.

Filovirus transcripts contain unusually long untranslated regions, especially at the 3′ ends. The 5′ end untranslated regions show a potential for formation of stable hairpin structures, which might play a role in transcript stability and ribosome binding (72). With the exception of the L protein, *in vitro* translation of purified subgenomic RNAs resulted in products comigrating on SDS-PAGE with the corresponding structural proteins. The finding that the L protein cannot be translated *in vitro* is probably the result of the low copy number of the L transcript, which has only been detected by RNA amplification (72, 83). The role of gene overlaps in regulation of transcription is unknown. Sanchez *et al.* (72) proposed that the downstream start site may allow the polymerase to recognize it after mRNA polyadenylation and possibly to reposition itself to initiate transcription. This "back up" mechanism is supported by the finding that attenuation of filovirus genes with start sites in overlaps does not occur to any greater degree than has been discussed for a much larger overlap found in the respiratory syncytial virus genome (126). Alternatively, the polymerase may occasionally terminate transcription at the overlap and initiate transcription of the downstream gene without polyadenylation of the upstream gene, but there is no evidence for detectable levels of transcripts lacking poly(A) tails.

With the EBO GP gene, transcription occurs from two open reading frames (Fig. 8). The primary gene product is a small nonstructural glycoprotein that is secreted from infected cells. Expression of full-length GP is achieved by transcriptional editing of a single nucleotide at a run of uridine residues (69, 112). MBG GP is expressed in a single frame and the gene does not contain sequences favoring RNA editing. A second ORF has been described for MBG MUS, but a corresponding gene product has not yet been identified (81). The difference in GP gene expression is one of several important distinctions between MBG and EBO (Table III).

The switch mechanism between transcription and replication is unknown. As with other NNS-RNA viruses, synthesis of the NP protein could be a key factor. Encapsidation and polymerase complex entry site are probably located on the 3′ leader sequence. The fact that the extremities of the genomes are complementary suggests a single identical encapsidation site on the genome and antigenome and an identical entry signal for the polymerase complex for both transcription and replication. Replication works via a full-length (+)-strand antigenome which serves as the template for synthesis of (–)-strand genome molecules. Encapsidated genomic RNA is incorporated into RNP. Those complexes may in part form inclusion bodies, which are prominent in infected cells (Fig. 5).

4. Virus Assembly and Exit

Virions usually bud at the plasma membrane, and the budding process is probably mediated at membrane locations where GP is incorporated. The cytoplasmic tail of GP is thought to interact with VP40 and/or VP24. VP40 may mediate the linkage between the RNP and the membrane proteins. Mature particles exit preferentially in a vertical mode, but budding via the longitudinal axis has also been observed. In macrophages, budding has also been observed at intracytoplasmic membranes surrounding vacuoles which form during infection (117). It is expected that the mechanism is similar to the situation at the plasma membrane.

IV. Clinical Virology

A. Clinical Syndrome

The onset of the disease is sudden, with fever, chills, headache, myalgia, and anorexia. This may be followed by symptoms such as abdominal pain, sore throat, nausea, vomiting, cough, arthralgia, diarrhea, and pharyngeal and conjunctival injection. Patients are dehydrated, apathetic, and disoriented and may develop a characteristic nonpruritic, maculopapular centripetal rash associated with varying degrees of erythema and then desquamate by day 5 or 7 of the illness. Hemorrhagic manifestations develop during the peak of the illness; they are of prognostic value for the disease. Bleeding into the gastrointestinal tract is most prominent along with petechiae and hemorrhages from puncture wounds and mucous membranes. Laboratory parameters are less characteristic, but the following findings are associated with the disease: leukopenia (as low as 1000/μl), left shift with atypical lymphocytes, thrombocytopenia (50,000–100,000/μl), markedly elevated serum transaminase levels (typically AST exceeding ALT), hyperproteinemia, and proteinuria. Prothrombin and partial thromboplastin times are prolonged and fibrin split products are detectable. In a later stage, secondary bacterial infection may lead to elevated white blood cell counts.

Nonfatal cases show fever for about 5–9 days; fatal cases develop clinical signs early during infection, and death commonly occurs between days 6 and 16 after the development of hemorrhage and hypovolemic shock (Fig. 9). Mortality is high for the African members of the family and varies between 22 and 88%, depending on the virus. The

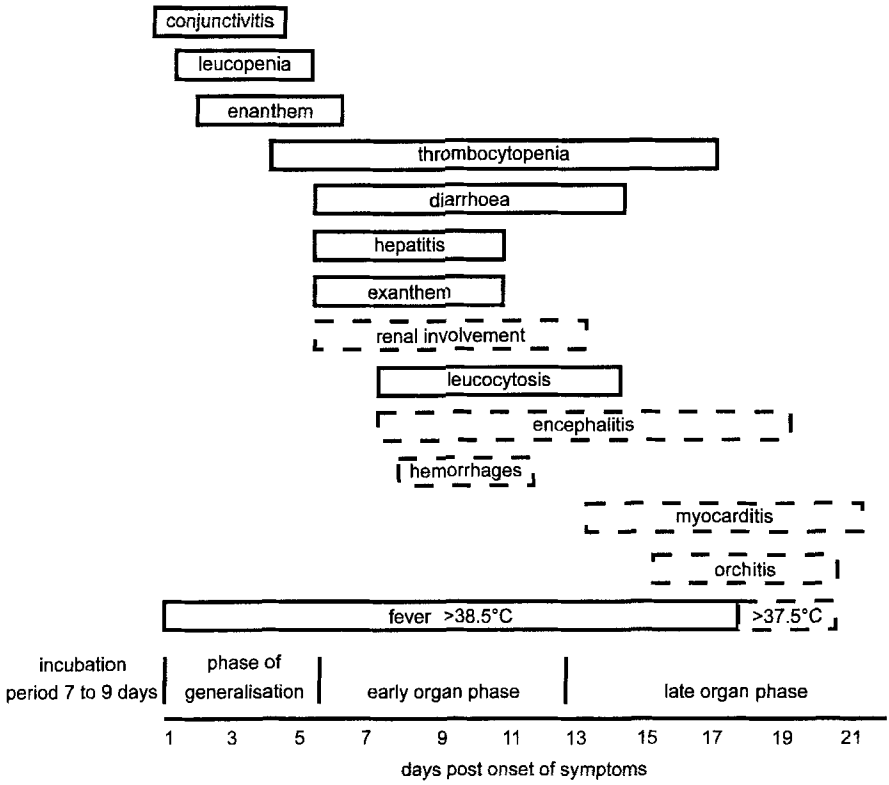


FIG. 9. Clinical course of hemorrhagic fever caused by Marburg virus. The drawing shows a summary of clinical observations during the 1967 outbreak of hemorrhagic fever caused by Marburg virus (adapted from W. Stille and E. Böhle [7]). *Key:* closed boxes, invariably present; open boxes, sometimes present.

highest rate has been reported for EBO Zaire. MBG infections are associated with the lowest mortality rates; however, most patients have been treated under European medical care standards, unlike in most of the EBO cases. The "Asian" filoviruses (EBO Reston) seem to possess a very low pathogenicity for humans or even to be apathogenic. This is interesting since genetic analyses have shown that EBO Reston seems to be most closely related to EBO Zaire (69).

Convalescence is prolonged and sometimes associated with myelitis, recurrent hepatitis, psychosis, or uveitis. An increased risk of abortion does exist for pregnant women, and clinical observations indicate a high death rate for children of infected mothers (20, 25, 127–131).

B. Diagnosis

Filoviruses cause acute infections in a variety of laboratory animals, although natural infections have only been reported in humans and nonhuman primates. In tropical settings, the identification of filoviral HF may be difficult since the most common causes of severe, acute febrile disease are malaria and typhoid fever. A wide range of infectious diseases has to be considered next, such as shigellosis, menigococcal septicemia, plague, leptospirosis, anthrax, relapsing fever, typhus, murine typhus, yellow fever, Chikungunya fever, Rift Valley fever, HF with renal syndrome, Crimean Congo HF, Lassa fever, and fulminant viral hepatitis. Travel, treatment in local hospitals, and contact with sick persons or wild and domestic monkeys are useful historical features in returning travelers, especially from Africa. Diagnosis of single cases is extremely difficult, but the occurrence of clusters of cases with prodromal fever followed by cases of hemorrhagic diatheses and person-to-person transmission are suggestive of viral HF, and containment procedures have to be initiated.

In filoviral HF, prostration, lethargy, wasting, and diarrhea seem to be more severe than observed in other viral HF patients. The rash is characteristic and extremely useful in differential diagnosis. Virologic diagnosis can be achieved during the febrile phase of the disease. Isolation attempts from serum and/or other clinical material should be performed using Vero or MA-104 cells (monkey kidney cells). However, most filoviruses do not cause extensive cytopathogenic effects on primary isolation. The most useful animal system after nonhuman primates is guinea pigs, which develop fever within 10 days of primary infection. Several passages, however, are necessary to produce a uniformly fatal disease. Often filoviruses do not kill newborn mice on primary isolation, suggesting that the most widely used animal system in laboratories may not be successful for virus isolation (30, 34, 128).

Laboratory diagnosis can be achieved in two different ways: measurement of the host-specific immunological response to the infection and detection of viral antigen and genomic RNA in the infected host (Table V). The most commonly used assay to detect antibodies to filoviruses is the indirect immunofluorescence assay (IFA) on acetone-fixed infected cells inactivated by gamma irradiation. The use of this assay, however, has been quite misleading, since a significant proportion of human and monkey sera will react with filovirus antigen without showing any symptoms of disease. Therefore, IFA results should be confirmed at least by an additional assay. Confirmatory tests include Western blot and enzyme-linked immunosorbent assays (ELISA). Direct IgG and

IgM ELISA are based on detergent-extracted infected cells adsorbed to plastic plates (55, 132, 133). In addition, an IgM capture assay has been developed which has correctly diagnosed acute infections with filoviruses in nonhuman primates, but still requires evaluation in humans (T. G. Ksiazek, personal communication). Radioimmune assays (RIA) are available but have not been evaluated for their use in diagnosis (70).

Direct detection of virus antigen, virus particles, and viral RNA can be achieved by several assays (Table V). Electron microscopy has been particularly useful in the diagnosis of filovirus infections (19, 30, 58, 134). Viral structures can be visualized by negative contrast electron microscopy after ultracentrifugation and fixation of initial passage cell culture supernatants. Thin-section microscopy can be performed on any infected material or infected cells which have been prepared by any standard fixation procedure. Immunohistochemisty on formalin-fixed material and paraffin-embedded tissues can be used for detection

TABLE V

LABORATORY DIAGNOSIS

Test	Target	Source	Remarks
Indirect immunofluorescence assay (IFA)	Antiviral antibodies	Serum	Simple to perform, but prone to nonspecific positives and *subjective interpretation*
Enzyme linked immunosorbent assay (ELISA)	Antiviral antibodies	Serum	Specific and sensitive, but initial response slower than IFA
Immuno blot	Antiviral antibodies	Serum	Protein-specific, but interpretation sometimes difficult
Antigen ELISA	*Viral antigen*	*Blood, serum,* tissues	*Rapid and sensitive, but* requires special equipment
Immunohistochemistry	Viral antigen	Tissues (*e.g.,* skin, liver)	Inactivated material, but requires time
Fluorescence assay (FA)	Viral antigen	Tissues (*e.g.,* liver)	Rapid and easy, but subjective information
Polymerase chain reaction (PCR)	Viral nucleic acid	Blood, serum, tissues	Rapid and sensitive, but requires expensive and special equipment
Electron microscopy	Viral particle	Blood, tissues	Unique morphology (immunostaining possible), but insensitive and requires expensive equipment
Virus isolation	Viral particle	Blood, tissues	Virus available for studies, but requires time

of filoviruses (30, 63), as can immunofluorescence on impression smears of tissues (135). Antigen detection ELISA (136) and reverse transcriptase–polymerase chain reaction (RT-PCR) (118) have been successfully used to detect filoviruses in clinical material. During the EBO Reston outbreak in 1989 both assays demonstrated their sensitivity and showed confirmation in nearly every case.

C. Patient Management and Prevention of Infection

A virus-specific treatment does not exist. Supportive therapy should be directed toward maintenance of effective blood volume and electrolyte balance. Management of shock, cerebral edema, renal failure, coagulation disorders, and secondary bacterial infection may be life-saving for patients. Heparin treatment should be considered only in cases with clear evidence of disseminated intravascular coagulopathy (DIC). Human interferon and human reconvalescence plasma have been used to treat patients in the past. Use of both therapies would be reasonable; however, there is no experimental data showing their efficacy. On the contrary, filoviruses are resistant to the antiviral effects of interferon, and interferon administration to monkeys has failed to increase survival rate or virus titer reduction. Ribavirin does not have any effect on filoviruses *in vitro* and thus is probably not of any clinical value, unlike in some other viral HFs. Isolation of patients is recommended, and protection of medical and nursing staff is required. This can be achieved by strict barrier nursing techniques and addition of HEPA filtered respirators for aerosol protection when feasible. For information regarding management of patients with suspected filoviral HF and approaches to minimize spread of virus in outbreak situations, especially in Africa, see published guidelines and recommendations (1, 34, 137–140).

Even though outbreaks of filovirus HF have been rare and mainly restricted to a small number of cases, vaccines would be of value for medical personnel in Africa as well as for laboratory personnel. Cross-protection among different EBO subtypes in experimental animal systems has been reported, suggesting the general value of vaccines (32, 141). Inactivated vaccines have been developed by formalin or heat treatment of cell-culture-propagated MBG and EBO subtypes Sudan and Zaire (142–144). Protection, however, has only been achieved by careful balance of the challenge dose and virulence. Because of the biohazardous nature of the agents, recombinant vaccines would be the way to go in the future. Immunizations of monkeys with purified NP and GP have demonstrated the induction of the humoral and cellular immune response and protected animals against challenge with lethal

doses (145). Thus, those two proteins and perhaps the sGP (EBO) may be candidates for recombinant vaccines. Recombinant GP-vaccinia- and baculoviruses have already been engineered for MBG MUS and EBO strain Mayinga (MAY), but have not yet been tested for a protective effect in animals (Centers for Disease Control and Prevention, Atlanta, Georgia, USA, unpublished data; Institute of Virology, Marburg, Germany, unpublished data; Institute of Molecular Biology, Koltsovo, Novosibirsk, Russian Federation).

Wild-caught monkeys are an important source for the introduction of filoviruses. This was clearly demonstrated in 1967 for MBG (127), in 1989, 1992, and 1996 for EBO Reston (30, 40, Centers for Disease Control and Prevention, personal communication), and in 1994 for the Ivory Coast EBO case (26). Quarantine of imported nonhuman primates and professional handling of animals will help prevent an introduction to humans. Guidelines for quarantine and proper handling of monkeys in medical research have been published (146).

Filovirus infectivity is quite stable at room temperature (20°C), but is destroyed in 30 min at 60°C. Infectivity is also destroyed by ultraviolet and gamma irradiation, formalin (1%), lipid solvents (deoxycholate, ether), β-propiolactone, and hypochloric and phenolic disinfectants (2, 138, 147).

V. Pathology and Immunology

A. Pathology in Experimental Animals

Monkeys, guinea pigs, suckling mice, and hamsters have been experimentally infected with filoviruses (30, 61, 114, 141, 148–150). MBG and EBO subtype Zaire are highly virulent for most of these species, and infection with low-passage virus stocks usually ends in death. Subtypes Sudan and Reston of EBO are less virulent, often causing a self-limited infection in guinea pigs and monkeys (32, 114, 151).

The incubation period for rhesus and African green monkeys inoculated with MBG and EBO subtype Zaire is 4–16 days. High titers of viruses can be detected in liver, spleen, lymph nodes, and lungs by onset of clinical symptoms. All of these organs, especially liver, show severe necrosis due to virus replication in parenchymal cells. Little inflammatory response at those sites is typical, which suggests that classical immunopathology may not be an important pathogenic consideration. Interstitial hemorrhage occurs and is most prominent in the gastrointestinal tract. In infected nonhuman primates, thrombocytopenia has been found accompanied by aggregation disorders of remain-

ing platelets in response to agonists such as ADP and collagen (61, 152). Histopathological damage of the target organs is at odds with serum transferase levels showing increase of ALT and AST with a ratio of AST:ALT of 7:1. This argues against hepatocellular dysfunction and raises the question of extrahepatic targets. Recent morphologic studies on EBO Reston-infected monkeys of the 1989 outbreak demonstrated extensive virus replication in tissue macrophages, interstitial fibroblasts of many organs, circulating monocytes/macrophages and, less frequently, in endothelial cells, hepatocytes, adrenal corticoid cells, and renal tubular epithelium (33). Similar results have been reported from experimentally infected MBG and EBO subtype Zaire monkeys (62). The pathophysiological basis for the hemorrhage and shock is still unknown. Prostaglandin-mediated dysfunction of endothelial cells and platelets and an unspecific immune response have been suggested to play an important role in the pathogenesis of the shock syndrome (152). Recent data, however, suggest that effects of viral infection on endothelial cells and/or virus-mediated release of mediators from infected monocytes/macrophages may be more important.

B. Pathology in Humans

In fatal cases, generalized hemorrhage is found macroscopically in most organ systems. Microscopic changes include focal necrosis in liver, lymphatic organs, kidneys, testes, and ovaries. The liver, while universally involved, with large eosinophilic intracytoplasmic inclusion bodies in hepatocytes and Councilman-like bodies within necrotic foci, is not the site of massive, potentially fatal necrosis. Generalized lymphoid necrosis is characteristic for the disease, and renal tubular necrosis is commonly found in agonal stages. A diffuse encephalitis as described for many viral infections has been observed in patients. Activation of the clotting system occurs and intravascular fibrin thrombi have been observed. Viral antigen can be detected by immunohistochemistry and electron microscopy in many organs, especially the liver, kidneys, spleen, and adrenal glands. Viral persistence has been demonstrated for MBG cases by isolation of virus from liver biopsy material and the anterior chamber of the eye after 4–5 weeks and semen after 12 weeks, despite an apparently normal immune response (20, 63, 127, 128, 134, 153).

C. Immunology

The mechanisms of recovery from filovirus infections in humans and wild as well as laboratory animals are unknown. *In vitro* neutralization has never been demonstrated by plaque reduction in cell culture

systems, and protection by convalescence sera has never been clearly shown. Fatal filovirus infections usually end with high viremia and no evidence of an immune response. In humans and monkeys they lead to extensive disruption of the parafollicular regions in the spleen and lymph nodes that contain the antigen-presenting dendritic cells (63). EBO Reston infection in monkeys is an exception in that a rise in nonproductive antibodies occurs shortly before death (128), Thus, cell-mediated immunity seems to mediate recovery from filovirus infections, although proof has not yet been presented.

GP is assumed to be the major antigenic molecule of virion particles. Its interaction with the host immune system may be modulated by the high content of carbohydrates. These sugars might cover antigenic epitopes, as has been demonstrated in other systems (154). For EBO, sGP production and secretion might interfere with the host immune response by neutralizing effective antibodies. As already mentioned filovirus GP molecules carry a presumably immunosuppressive domain close to the C terminus (Figs. 6A and 7). Peptides synthesized according to that 26-amino-acid-long region (Fig. 7) inhibited the blastogenesis of lymphocytes in response to mitogens, induced production of cytokines, and increased proliferation of mononuclear cells *in vitro.* Infected animals showed increased levels of mediators, in particular interferon and TNF. An activation of natural killer cells has been observed in the earlier stages of infection, whereas a complete lack of it has been observed in the later stages (155, 156). These findings are in line with the observation of immunosuppression in monkeys experimentally infected with filoviruses (MBG, EBO) and in humans and of proliferation of filoviruses in macrophages and monocytes *in vivo* and *in vitro* (33, 61–63, 117, 149). Monkeys that survived experimental MBG infection were susceptible to reinfection and showed shorter incubation periods and increased viremia (157). It is not yet known if the immunosuppressive domain on the GP is functional on mature molecules, but evidence for GP mediation of the above-mentioned effects has recently been reported (145).

D. Pathophysiology

The pathophysiological changes that make filovirus infections so devastating are just beginning to be unraveled. Pathogenesis in fatal infection in human and nonhuman primates is similar, suggesting the primate system as a model for studying filovirus HF (61, 62, 152, 158, 159). Clinical and biochemical findings support the anatomical observations of extensive liver involvement, renal damage, changes in vascular permeability including endothelial damage, and activation of the

clotting cascade. Visceral organ necrosis is a consequence of virus replication in parenchymal cells. However, no organ, not even the liver, shows sufficient damage to account for death. The role of disseminated intravascular coagulation (DIC) in pathogenesis is still controversial, since a laboratory confirmation of DIC in human infections has never been demonstrated. In nonhuman primates the intrinsic clotting pathway is most affected, whereas the extrinsic pathway is spared. The consequence is DIC in the final stages of the infection when parenchymal necrosis is extensive and leads to common terminal pathways, including DIC.

Laboratory parameters in the crucial early stage of filovirus HF, such as high AST:ALT ratio, normal bilirubin levels, and marked lymphopenia followed by a dramatic neutrophilia with left shift, suggest extrahepatic targets of infection. As with some other HFs [hemorrhagic fever with renal syndrome (HFRS), dengue HF, Lassa fever], fluid distribution problems and platelet abnormalities are dominant clinical manifestations indicating dysfunction or damage of endothelial cells and platelets. Post mortem there is little monocyte or macrophage infiltration in sites of parenchymal necrosis, suggesting that dysfunction of white blood cells also occurs. Morphological studies on EBO Reston-infected monkeys from the 1989 epizootic (33) and monkeys experimentally infected with EBO Zaire (62) showed that monocytes/macrophages and fibroblasts may be the preferred sites of virus replication in early stages, whereas other cell types may become involved as the disease progresses. Human monocytes/macrophages in culture are also sensitive to infection resulting in massive production of infectious virus and cell lysis (117). Although the studies on infected nonhuman primates did not identify endothelial cells as sites of massive virus replication, *in vitro* studies and post-mortem observations of human cases clearly demonstrated that endothelial cells of human origin are suitable targets for virus replication (63, 116). Here infection leads to cell lysis, indicating that damage of endothelial cells may be an important pathophysiological parameter during infection.

Besides evidence for direct vascular involvement in infected hosts, the role of active mediator molecules in the pathogenesis of the disorders has to be discussed. Although the source of these mediators during filovirus infections is still unknown, candidate cells exist. Besides the endothelium, which has yet to be examined for production and secretion of such mediators, the common denominator remains the monocyte/macrophage, a cell type shown to be a site of virus replication *in vivo* (33, 62, 63) and *in vitro* (117). Monocytes/macrophages are known as a pivotal source of different proteases, H_2O_2, and mediators such as tumor necrosis factor alpha (TNF-α), interleukins 6 and 8, and

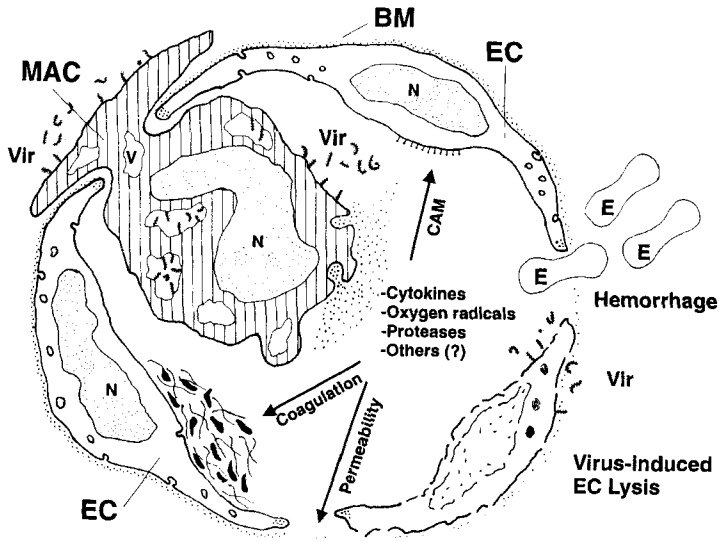


FIG. 10. Schematic drawing illustrating the possible role of macrophages and endothelial cells in the development of hemorrhagic fever caused by filoviruses. EC, endothelial cell; MAC, macrophage; Vir, virus particle; CAM, cell adhesion molecule; E, erythrocyte; BM, basement membrane; N, nucleus; V, vacuole.

platelet-derived growth factor (PDGF). TNF-α can result in secondary activation of mediators with important protective as well as deleterious effects. Recently it has been demonstrated that supernatants of filovirus-infected monocyte/macrophage cultures are capable of increasing paraendothelial permeability in an *in vitro* model (117). Examination for mediators in those supernatants revealed increased levels of secreted TNF-α, the prototype cytokine of macrophages. These data support the concept of a mediator-induced vascular instability and thus increased permeability as a key mechanism for the development of the shock syndrome seen in severe and fatal cases. Thus, the syndrome may be comparable to symptomatic shock in response to various endogenous and exogenous mediators (160–163). The bleeding tendency could be due to endothelial damage caused directly by virus replication as well as indirectly by cytokine-mediated processes. The onset of the bleeding tendency is supported by the loss of the integrity of the endothelium as demonstrated in tissue and organ culture (116) and infected animals (152). The hemorrhage occurs later in infection and could be due to extended damage which cannot be repaired by wound-healing mechanisms (164). The bleeding tendency is reinforced by a decrease in the bloodstream as a common consequence of shock.

The combination of viral replication in endothelial cells and virus-induced cytokine release from monocytes/macrophages may also promote a distinct proinflammatory endothelial phenotype that then triggers the coagulation cascade. A model summarizing the above discussed pathophysiological events is illustrated in Fig. 10.

NOTE

GenBank/EMBL data library accession numbers for nucleotide and amino acid sequences are as follows: MBG strain Musoke (EMBL data library, Z12132); MBG strain Popp (EMBL data library—X64405, X64406, X68493, X68494, X68495, Z29337); Ebola subtype Zaire, strain Mayinga (GenBank, L11365).

ACKNOWLEDGMENTS

The authors gratefully acknowledge suggestions, recommendations discussions, and the communication of unpublished data by (1) A. Sanchez, C. J. Peters, T. G. Ksiazek, P. E. Rollin of the Centers for Disease Control and Prevention, National Centers for Infectious Diseases, Division of Viral and Rickettsial Diseases, Special Pathogens Branch, Atlanta, GA 30333, USA; (2) W. Slenczka, V. E. Volchkov, E. Mühlberger, S. Becker of the Institute of Virology, Philipps-University, Marburg, Germany; (3) S. V. Netesov, H. Ryabchikova of the NPO "Vector," Institute of Molecular Biology, 633159 Koltsovo, Novosibirsk Region, Russian Federation; (4) M. P. Kiley of the Federal Laboratories for Health Canada and Agriculture and Agri-Food Canada, Winnipeg, Manitoba, Canada; (5) H.-J. Schnittler of the Institute of Anatomy, Julius-Maximilians-University, Würzburg, Germany. Heinz Feldmann had a National Research Council (NRC) fellowship from 1992 to 1994 at the Centers for Disease Control and Prevention, National Center for Infectious Diseases, Division of Viral and Rickettsial Diseases, Special Pathogens Branch, Atlanta, GA 30333, USA. Research on filoviruses at the Institute of Virology, Philipps-University of Marburg, was supported by grants from the Deutsche Forschungsgemeinschaft (Forschergruppe Kl 238/1-1, Sonderforschungsbereich 286, Sachbeihilfe Fe 286/4-1).

REFERENCES

1. World Health Organization (1985). "Arthropod-Borne and Rodent-Borne Viral Diseases." WHO Technical Report Series 719.
2. Centers for Disease Control and Prevention (1993). "Biosafety in Microbiology and Biomedical Laboratories." US Department of Health and Human Services (HHS), Publication No. (CDC) 93-8395. US Government Printing Office, Washington, DC.
3. Holland, J. J. (1993). Genetic diversity of RNA viruses *Curr. Top. Microbiol. Immunol.* **176**, 1–226.
4. Peters, C. J., Sanchez, A., Feldmann, H., Rollin, P. E., Nichol, S. T., and Ksiazek, T. G. (1994). Filoviruses as emerging pathogens. *Semin. in Virol.* **5**, 147–154.
5. Martini, G. A., Knauff, H. G. Schmidt, H. A., Mayer, G., and Baltzer, G. (1968). Über eine bisher unbekannte, von Affen eingeschleppte Infektionskrankheit: Marburg-Virus-Krankheit. *Dtsch. Med. Wochenschr. 93*, 559–571.

6. Martini, G. A. (1971). Clinical Syndrome. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 1–9. Springer-Verlag, New York.
7. Stille, W., and Böhle, E. (1971). Clinical course and prognosis of Marburg virus ("green monkey") disease. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 10–18. Springer-Verlag, New York.
8. Stille, W., Böhle, E., Helm, E., vanRey, W., and Siede, W. (1968). Über eine durch *Cercopithecus aethiops* übertragene Infektionskrankheit. *Dtsch. Med. Wochenschr.* **93**, 572–582.
9. Stojkovic, L. J., Bordjoski, M., Gligic, A., and Stefanovic, Z. (1971). Two cases of Cercopithecus-monkeys-associated hemorrhagic fever. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 24–33. Springer-Verlag, New York.
10. Hennessen, W. (1971). Epidemiology of "Marburg virus" disease. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 161–165. Springer-Verlag, New York.
11. Haas, R., and Maass, G. (1971). Experimental infection of monkeys with the Marburg virus. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 136–143. Springer-Verlag, New York.
12. Gear, J. S. S., Cassel, G. A., Gear, A. J., Trappler, B., Clausen, L., Meyers, A. M., Kew, M. C., Bothwell, T. H., Sher, R., Miller, G. B., Schneider, J., Koornhoff, H. J., Comperts, E. D., Isaäcson, M., and Gear, J. H. S. (1975). Outbreak of Marburg virus disease in Johannesburg. *Br. Med. J.* **4**, 489–493.
13. Smith, D. H., Johnson, B. K., Isaäcson, M., Swanepoel, R., Johnson, K. M., Kiley, M. P., Bagshawe, A., Siongok, T., and Keruga, W. K. (1982). Marburg-virus disease in Kenya. *Lancet* **1**, 816–820.
14. Bowen, E. T. W., Lloyd, G., Harris, W. J., Platt, G. S., Baskerville, A., and Vella, E. E. (1977). Viral haemorrhagic fever in southern Sudan and Northern Zaire. *Lancet* **1**, 571–573.
15. Johnson, K. M., Webb, P. A., Lange, J. V., and Murphy, F. A. (1977). Isolation and partial characterization of a new virus causing acute haemorrhagic fever in Zaire. *Lancet* **1**, 569–571.
16. Babiker, Mohd el Tahir (1978). The haemorrhagic fever outbreak in Maridi, western Equatoria, southern Sudan. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 125–127. Elsevier/North-Holland, Amsterdam.
17. Breman, J. G., Piot, P., Johnson, K. M., White, M. K., Mbuyi, M., Sureau, P., Heymann, D. L., van Nieuwenhove, S., McCormick, J. B., Ruppol, J. P., Kintoki, V., Isaäcson, M., van der Groen, G., Webb, P. A., and Ngvete, K. (1978). The epidemiology of Ebola haemorrhagic fever in Zaire, *1976*. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 103–124. Elsevier/North-Holland, Amsterdam.
18. Francis, D. P., Smith, D. H., Highton, R. B., Simpson, D. I. H., Pacifiko Lolik, Isiaih Mayom Deng, Anthony Lagu Gillo, Ali Ahmed Idrtis, and Babiker El Tahir (1978). Ebola fever in the Sudan, 1976: Epidemiological aspects of the disease. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 129–135. Elsevier/North-Holland, Amsterdam.
19. Murphy, F. A., van der Groen, G., Whitfield, S. G., and Lange, J. V. (1978). Ebola and Marburg virus morphology and taxonomy. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 61–84. Elsevier/North-Holland, Amsterdam.
20. Pattyn, S. R., ed. (1978). "Ebola Virus Hemorrhagic Fever," 1st Ed., pp. 1–436. Elsevier/North-Holland, Amsterdam.
21. Smith, D. H., Francis, D., Simpson, D. I. H., and Highton, R. B. (1978). The Nzara outbreak of haemorrhagic fever. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn,

ed.), 1st Ed., pp. 137–141. Elsevier/North-Holland, Amsterdam.
22. World Health Organization (1978). Ebola hemorrhagic fever in Sudan, 1976. *Bull. WHO* **56**, 247–270.
23. World Health Organization (1978). Ebola hemorrhagic fever in Zaire, 1976. *Bull. WHO* **56**, 271–293.
24. World Health Organization (1979). Viral hemorrhagic fever surveillance. *Weekly Epidemiol. Rec.* **54**, 342–343.
25. Baron, R. C., McCormick, J. B., and Zubeir, O. A. (1983). Ebola hemorrhagic fever in southern Sudan: Hospital dissemination and intrafamilial spread. *Bull. WHO* **61**, 997–1003.
26. LeGuenno, B., Formentry, P., Wyers, M., Gounon, P., Walker, F., and Boesch, C. (1995). Isolation and partial characterization of a new strain of Ebola virus. *Lancet* **345**, 1271–1274.
27. World Health Organization (1995a). Ebola hemorrhagic fever. *Weekly Epidemiol. Rec.* **70**, 149–152.
28. World Health Organization (1995b). Ebola hemorrhagic fever. *Weekly Epidemiol. Rec.* **70**, 241–242.
29. Centers for Disease Control and Prevention (1989). Ebola virus infection in imported primates—Virginia, 1989. *MMWR* **38**, 831–832, 837–838.
30. Jahrling, P. B., Geisbert, T. W., Galgard, D. W., Johnson, E. D., Ksiazek, T. G., Hall, W. C., and Peters, C. J. (1990). Preliminary report: Isolation of Ebola virus from monkeys imported to USA. *Lancet* **335**, 502–505.
31. Dalgard, D W., Hardy, R. J., Pearson, S. L., Pucak, G. J., Quander, R. V., Zack, P. M., Peters, C. J., and Jahrling, J. B. (1992). Combined simian hemorrhagic fever and Ebola virus infection in cynomolgus monkeys. *Lab. Animal Sci.* **42**, 152–157.
32. Fisher-Hoch, S. P., Brammer, L., Trappier, S. G., Hutwagner, L. C., Farrar, B. B., Ruo, S. L., Brown, B. G., Hermann, L. M., Perez-Oronoz, G. I., Goldsmith, C. S., Hanes, M. A., and McCormick, J. B. (1992). Pathogenic potential of filoviruses: Role of geographic origin of primate host and virus strain. *J. Infect. Dis.* **166**, 753–763.
33. Geisbert, T. W., Jahrling, P. B., Hanes, M. A., and Zack, P. M. (1992). Association of Ebola-related Reston virus particles and antigen with tissue lesions of monkeys imported to the United Staes. *J. Comp. Path.* **106**, 137–152.
34. Peters, C. J., Johnson, E. D., and McKee, K. T. (1991). Filoviruses and management and viral hemorrhagic fevers. *In* "Textbook of Human Virology" (R. B. Belshe, ed.), pp. 699–712. Mosby Year Book, St. Louis.
35. Peters, C. J., Johnson, E. D., Jahrling, P. B., Ksiazek, T. G., Rollin, P. E., White, J., Hall, W., Trotter, R., and Jaax, N. (1993). Filoviruses. *In* "Emerging Viruses" (S. S. Morse, ed.), pp. 159–175. Oxford University Press, Oxford.
36. Hayes, C. G., Burans, J. P., Ksiazek, T. G., DelRosario, R. A., Miranda, M. E. G., Manaloto, C. R., Barrientos, A. B., Robles, C. G., Dayrit, M. M., and Peters, C. J. (1992). Outbreak of fatal illness among captive macaques in the Philippines caused by an Ebola-related filovirus. *Am. J. Trop. Med. Hyg.* **46**, 664–671.
37. Centers for Disease Control and Prevention (1990). Update: Filovirus infection in animal handlers. *MMWR* **39**, 221.
38. Centers for Disease Control and Prevention (1990). Update: Evidence of filovirus infection in an animal caretaker in a research/service facility. *MMWR* **39**, 296–297.
39. Centers for Disease Control and Prevention (1990). Update: Filovirus infection associated with contact with nonhuman primates or their tissues. *MMWR* **39**, 404–405.
40. World Health Organization (1992). Viral hemorrhagic fever in imported monkeys. *Weekly Epidemiol. Rec.* **67**, 142–143.

41. van der Groen, G., Johnson, K. M., Webb, P. A., Wulff, H., and Lange, J. (1978). Results of Ebola antibody survey in various population groups. In "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 203–208. El Elsevier/North-Holland, Amsterdam.
42. Ivanoff, B., Duquesnoy, P., Languillat, G., Saluzzo, J. F., Georges, A., Gonzales, J. P., and McCormick, J. B. (1982). Haemorrhagic fever in Gabon. 1. Incidence of Lassa, Ebola and Marburg virus in Haut-Ogooue. *Trans. Soc. Trop. Med. Hyg.* **76**, 719–720.
43. Blackburn, N. K., Searle, L., and Taylor, P. (1982). Viral haemorrhagic fever antibodies in Zimbabwe school children. *Trans. Soc. Trop. Med. Hyg.* **76**, 803–805.
44. Bauree, P., and Bergmann, J. F. (1983). Ebola virus infection in man: A serological and epidemiological survey in the Cameroon. *Am. J. Trop. Med. Hyg.* **32**, 1465–1466.
45. Slenczka, W., Rietschel, M., Hoffmann, C., and Sixl, W. (1984). Seroepidemiologische Untersuchungen über das Vorkommen von Antikörpern gegen Marburg- und Ebola-Virus in Afrika. *Mitt. Oesterr. Ges. Tropenmed. Parasitol.* **6**, 53–60.
46. Johnson, B. K., Ocheng, D., Oogo, S., Gitau, L. G., Wambui, C., Gichogo, A., Libondo, D., Tukei, P. M., and Johnson, E. D. (1986). Seasonal variation in antibodies against Ebola virus in Kenyan fever patients. *Lancet* **1**, 1160.
47. van der Waals, F. J., Pomerov, K. L., Goudsmit, J., Asher, D. M., and Gajdusek, D. C. (1986). Hemorrhagic fever virus infection in an isolated rainforest area of central Liberia. Limitations of the indirect immunofluorescence slide test for antibody screening in Africa. *Trop. Geogr. Med.* **38**, 209–214.
48. Tomori, O., Fabiyi, A., and Sorungbe, A. (1988). Viral hemorrhagic fever antibodies in Nigerian population. *Am. J. Trop. Med. Hyg.* **38**, 407–410.
49. Gonzales, J. P., Johnson, E. D., Josse, R., Merlin, M., Georges, A. J., Abandja, J., Danyod, M., Delaporte, E., Dupont, A., Ghogomu, A., Kouka-Bemba, D., Madelon, M. C., Simav A., and Meunier, D. M. Y. (1989). Antibody prevalence against hemorrhagic fever viruses in randomized representative Central African populations. *Res. Virol.* **140**, 319–331.
50. Mathiot, C. C., Fontenille, D., Georges, A. J., and Coulanges, P. (1989). Antibodies to haemorrhagic fever viruses in Madagascar. *Trans. Soc. Trop. Med. Hyg.* **83**, 407–409.
51. Johnson, E. D., Gonzales, J. P., and Georges, A. (1993). Haemorrhagic fever virus activity in equatorial Africa: Distribution and prevalence of filovirus reactive antibodies in the Central African Republic. *Trans. Soc. Trop. Med. Hyg.* **87**, 530–535.
52. Johnson, E. D., Gonzales, J. P., and Georges, A. (1993). Filovirus activity among selected ethnic groups inhabiting the tropical rain forest of equatorial Africa. *Trans. Soc. of Trop. Med. Hyg.* **87**, 536–538.
53. Jahrling, P. B. (1995). Filoviruses and arenaviruses. *In* "Manual of Clinical Microbiology" (P. R. Murray, E. J. Baron, M. A. Pfaller, F. C. Tenover, and R. M. Yolken, eds.), pp. 1068–1081. American Society for Microbiology, Washington, DC.
54. Centers for Disease Control and Prevention (1990). Update: Filovirus infection among persons with occupational exposure to nonhuman primates. *MMWR* **39**, 266–267, 273.
55. Becker, S., Feldmann, H., Will, C., and Slenczka, W. (1992). Evidence for occurrence of filovirus antibodies in humans and imported monkeys: Do subclinical filovirus infections occur worldwide? *Med. Microbiol. Immunol.* **181**, 43–55.
56. Henderson, B. E., Kissling, R. E., Williams, M. C., Kafuko, G. W., and Martin, M. (1971). Epidemiological studies in Uganda relating to the "Marburg" agent. *In* "Marburg Virus Disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 166–176. Springer-Verlag, New York.
57. Fisher-Hoch, S. P., Perez-Oronoz, G. I., Jackson, E. L., Hermann, L. M., and Brown,

B. G. (1992). Filovirus clearance in non-human primates. *Lancet* **340**, 451–453.
58. Siegert, R., Shu, H.-L., Slenczka, W., Peters, D., and Müller, G. (1967). Zur Äthiologie einer unbekannten von Affen ausgegangenen Infektionskrankheit. *Dtsch. Med. Wochenschr.* **92**, 2341–2343.
59. Pokhodyaeu, V. A., Gonchar, N. I., and Pshenichnov, V. A. (1991). Experimental study of Marburg virus contact transmission. *Voprosy Virusologii* **36**, 506–508.
60. Bazhutin, N. B., Belanov, E. F., Spiridonov, V. A., Voitenko, A. V., Krivenchuk, N. A., Krotovf, S. A., Omelchenko, N. I., Tereschenko, A. Y., and Khomichev, V. V. (1992). The influence of the methods of experimental infection with Marburg virus on the features of the disease process in green monkeys. *Voprosy Virusologii* **37**, 153–156.
61. Murphy, F. A., Simpson, D. I. H., Whitfield, S. G., Zlotnik, I., and Carter, G. B. (1971). Marburg virus infection in monkeys. *Lab. Invest.* **24**, 279–291.
62. Ryabchikova, E. I., Kolesnikova, L. V., Tkachev, V. K., Pereboeva, L. A., Baranova, S. G., and Rassadkin, J. N. (1994). Ebola infection in four monkey species. Ninth International Conference on Negative Strand RNA Viruses, p. 164. Estoril, Portugal. [Abstract]
63. Zaki, S. (1995). Ebola virus infection. European Conference on Tropical Medicine, p. 2 (A22). Hamburg, Germany. [Abstract]
64. Kiley, M. P., Bowen, E. T. W., Eddy, G. A., Isaäcson, M., Johnson, K. M., McCormick, J. B., Murphy, F. A., Pattyn, S. R., Peters, D., Prozesky, O. W., Regnery, R. L., Simpson, D. I. H., Slenczka, W., Sureau, P., van der Groen, G., Webb, P. A., and Wulff, H. (1982). Filoviridae: A taxonomic home for Marburg and Ebola viruses? *Intervirology* **18**, 24–32.
65. Ströher, U., Sanchez, A., Klenk, H.-D., and Feldmann, H. (1995). The Marburg group of filoviruses: Genetic variability and characterization of a second ORF encoding a potential nonstructural protein. First European Meeting of Virology, P2/42. Würzburg, Germany. [Abstract]
66. Kiley, M. P., Cox, N. J., Elliott, L. H., Sanchez, A., DeFries, R., Buchmeier, M. J., Richman, D. D., and McCormick, J. B. (1988). Physicochemical properties of Marburg virus: Evidence for three distinct virus strains and their relationship to Ebola virus. *J. Gen. Virol.* **69**, 1957–1967.
67. Feldmann, H., Klenk, H.-D., and Sanchez, A. (1993). Molecular biology and evolution of filoviruses. *Arch. Virol.*, Supplement **7**, 81–100.
68. Feldmann, H., Nichol, S. T., Klenk, H.-D., Peters, C. J., and Sanchez, A. (1994). Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. *Virology* **199**, 469–473.
69. Sanchez, A., Trappier, S., Mahy, B. W. J., Peters, C. J., and Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3602–3607.
70. Richman, D. D., Cleveland, P. H., McCormick, J. B., and Johnson, K. M. (1983). Antigenic analysis of strains of Ebola viruses: Identification of two Ebola virus subtypes. *J. Infect. Dis.* **147**, 268–271.
71. Sanchez, A., Kiley, M. P., Klenk, H.-D., and Feldmann, H. (1992). Sequence analysis of the Marburg virus nucleoprotein gene: Comparison to Ebola virus and other non-segmented negative-strand RNA viruses. *J. Gen. Virol.* **74**, 347–357.
72. Sanchez, A., Kiley, M. P., Holloway, B. P., and Auperin, D. D. (1993). Sequence analysis of the Ebola virus genome: Organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.* **29**, 215–240.
73. Bukreyev, A. A., Volchkov, V. E., Blinov, V. M., and Netesov, S. V. (1993). The VP35

and VP40 proteins of filoviruses: Homology between Marburg and Ebola viruses. *FEBS Lett.* **322**, 41–46.
74. Buchmeier, M. J., DeFries, R., McCormick, J. B., and Kiley, M. P. (1983). Comparative analysis of the structural polypeptides of Ebola virus from Sudan and Zaire. *J. Infect. Dis.* **147**, 276–281.
75. Cox, N. J., McCormick, J. B., Johnson, K. M., and Kiley, M. P. (1983). Evidence for two subtypes of Ebola virus based on oligonucleotide mapping of RNA. *J. Infect. Dis.* **147**, 272–275.
76. ICTV (1991). The order *Mononegavirales.* Paramyxovirus Study Group of the Vertebrate Subcommittee. Virology Division News. *Arch. Virol.* **117**, 137–140.
77. Smith, C. E. G., Simpson, D. I. H., and Bowen, E. T. W. (1967). Fatal human disease from vervet monkeys. *Lancet* **2**, 1119–1121.
78. Kissling, R. E., Robinson, R. Q., Murphy, F. A., and Whitfield, F. G. (1968). Agent of disease contracted from green monkeys. *Science* **160**, 888–890.
79. Peters, D., Müller, G., and Slenczka, W. (1971). Morphology, development, and classification of Marburg virus. *In* "Marburg virus disease" (G. A. Martini and R. Siegert, eds.), 1st Ed., pp. 68–83. Springer-Verlag, New York.
80. Regnery, R. L., Johnson, K. M., and Kiley, M. P. (1980). Virion nucleic acid of Ebola virus. *J. Virol.* **36**, 465–469.
81. Will, C., Mühlberger, E., Linder, D., Slenczka, W., Klenk, H.-D., and Feldmann, H. (1993). Marburg virus gene four encodes for the virion membrane protein, a type I transmembrane glycoprotein. *J. Virol.* **67**, 1203–1210.
82. Kiley, M. P., Regnery, R. L., and Johnson, K. M. (1980). Ebola virus: Identification of virion structural proteins. *J. Gen. Virol.* **49**, 333–341.
83. Feldmann, H., Mühlberger, E. Randolf, A., Will, C., Kiley, M. P., Sanchez, A., and Klenk, H.-D. (1992). Marburg virus, a filovirus: Messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* **24**, 1–19.
84. Bukreyev, A. A., Volchkov, V. E., Blinov, V. M., Dryga, S. A., and Netesov, S. V. (1995). The nucleotide sequence of the Popp (1967) strain of Marburg virus: A comparison with the Musoke (1980) strain. *Arch. Virol.* **140**, 1589–1600.
85. Volchkov, V. E., Blinov, V. M., Kotov, A. N., Chepurnov, A. A., and Netesov, S. V. (1993). The full-length nucleotide sequence of the Ebola virus. IXth International Congress of Virology, p. 299. Glasgow, Scotland. [Abstract]
86. Kiley, M. P., Wilusz, J., McCormick, J. B., and Keene, J. D. (1986). Conservation of the 3′ terminal nucleotide sequence of Ebola and Marburg viruses. *Virology* **149**, 251–254.
87. Sanchez, A., Kiley, M. P., Holloway, B. P., McCormick, J. B., and Auperin, D. D. (1989). The nucleoprotein gene of Ebola virus: Cloning, sequencing, and in vitro expression. *Virology* **170**, 81–91.
88. Elliott, L. H., Kiley, M. P., and McCormick, J. B. (1985). Descriptive analysis of Ebola virus proteins. *Virology* **147**, 169–176.
89. Elliott, L. H., Sanchez, A., Holloway, B. P., Kiley, M. P., and McCormick, J. B. (1993). Ebola protein analysis for the determination of genetic organization. *Arch. Virol.* **133**, 423–436.
90. Becker, S., Huppertz, S., Klenk, H.-D., and Feldmann, H. (1994). The nucleoprotein of Marburg virus is phosphorylated. *J. Gen. Virol.* **75**, 809–818.
91. Barr, J., Chamers, P., Pringle, C. R., and Easton, A. J. (1991). Sequence of the major nucleocapsid gene of pneumonia virus of mice: Sequence comparison suggests structural homology between nucleocapsid proteins of pneumoviruses, paramyxoviruses, rhabdoviruses and filoviruses. *J. Gen. Virol.* **72**, 677–685.

92. Morgan, E. M. (1991). Evolutionary relationships of paramyxoviruses nucleocapsid-associated proteins. *In* "The Paramyxoviruses" (D. W. Kingsbury, ed.), pp. 163–179. Plenum Press, New York.
93. Peeples, M. E. (1991). Paramyxovirus M proteins: Pulling it all together and putting it on the road. *In* "The Paramyxoviruses" (D. W. Kingsbury, ed.), pp. 427–456. Plenum Press, New York.
94. Banerjee, A. K. (1987). Transcription and replication of rhabdoviruses. *Microbiol. Rev.* **51**, 66–87.
95. Wagner, R. R. (1987). Rhabdovirus biology and infection: An overview. *In* "The Rhabdoviruses" (R. R. Wagner, ed.), pp. 9–74. Plenum Press, New York.
96. Volchkov, V. E., Blinov, V. M., and Netesov, S. V. (1992). The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. *FEBS Lett.* **305**, 181–184.
97. Bukreyev, A, A,, Volchkov V. E., Blinov, V. M., and Netesov, S. V. (1993). The GP-protein of Marburg virus contains the region similar to the "immunosuppressive domain" of oncogenic etroviruses P15E proteins. *FEBS Lett.* **323**, 183–187.
98. Feldmann, H., Will, C., Schikore, M., Slenczka, W., and Klenk, H.-D. (1991). Glycosylation and oligomerization of the spike protein of Marburg virus. *Virology* **182**, 353–356.
99. Geyer, H., Will, C., Feldmann, H., Klenk, H.-D., and Geyer, R. (1992). Carbohydrate structure of Marburg virus glycoprotein. *Glycobiology* **2**, 299–312.
100. Funke, C., Becker, S., Dartsch, H., Klenk, H.-D., and Mühlberger, E. (1995). Acylation of the Marburg virus glycoprotein. *Virology* **208**, 289–297.
101. Cianciolo, G. J., Copeland, T. J., Oroszlan, S., and Snyderman, R. (1985). Inhibition of lymphocyte proliferation by a synthetic peptide homologous to retroviral envelope proteins. *Science* **230**, 453–455.
102. Harris, D. T., Cianciolo, G. J., Snyderman, R., Argov, S., and Koren, H. R. (1987). Inhibition of human natural killer cell activity by a synthetic peptide homologous to a conserved region in the retroviral protein, p15E. *J. Immunol.* **138**, 889–894.
103. Kadota, J., Cianciolo, G. J., and Snyderman, R. (1991). A synthetic peptide homologous to retroviral transmembrane envelope proteins depressed protein kinase C-mediated lymphocyte proliferation and directly inactivated proteine kinase C: A potential mechanism for immunosuppression. *Microbiol. Immunol.* **35**, 443–459.
104. Volchkov, V. E., Chepurnov, A., Dryga, S., Becker, S., Blinov, V., Kotov, A., Ternovoj, V., Klenk, H.-D., and Netesov, S. V. (1994). Molecular characterization of a pathogenicity variant of Ebola virus. Ninth International Conference on Negative Strand RNA Viruses, p. 176. Estoril, Portugal. [Abstract]
105. Mühlberger, E., Sanchez, A., Randolf, A., Will, C., Kiley, M. P., Klenk, H.-D., and Feldmann, H. (1992). The nucleotide sequence of the L gene of Marburg virus, a filovirus: Homologies with paramyxoviruses and rhabdoviruses. *Virology* **187**, 534–547.
106. Barik, S. E., Rud, W., Luk, D., Banerjee, A. K., and Yong Kang, C. (1990). Nucleotide sequence analysis of the L gene of vesicular stomatitis virus (New Jersey serotype): Identification of conserved domains in L proteins of nonsegmented negative-strand RNA viruses. *Virology* **175**, 332–337.
107. Sleat, D. E., and Banerjee, A. K. (1993). Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. *J. Virol.* **67**, 1334–1339.
108. Kamer, G., and Argos, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res.* **12**, 7269–7282.

109. Wain-Hobson S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9–17.
110. Chiu I. M., Yaniv, A., Dahlberg, J. E, Gazit, A., Skunitz, S. F., Tronick, S. R., and Aaronson, S. A. (1985). Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. *Nature (London)* **317**, 366–368.
111. Kemdirin, S., Palefsky, J., and Briedis, D. J. (1986). Influenza B virus PB1 protein: Nucleotide sequence of the genome RNA segment predicts a high degree of structural homology with the corresponding influenza A virus polymerase proteins. *Virology* **152**, 126–135.
112. Volchkov, V. E., Becker S., Volchkova, V. A., Ternovoj, V. A., Kotov, A. N., Netesov, S. V., and Klenk, H.-D. (1995). Gp mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology* **214**, 421–430.
113. Spiropoulou, C. F., and Nichol, S. T. (1993). A small highly basic protein is encoded in overlapping frame within the P gene of vesicular stomatitis virus. *J. Virol.* **67**, 3103–3110.
114. McCormick, J. B., Bauer, S. P., Elliott, L. H., Webb, P. A., and Johnson, K. M. (1983). Biological differences between strains of Ebola virus from Zaire and Sudan. *J. Infect. Dis.* **147**, 264–267.
115. Peters, C. J., Jahrling, P. B., Ksiazek, T. G., and Lupton, H. (1992). Filovirus contamination of cell cultures. *Develop. Biol. Standard* **76**, 267–274.
116. Schnittler, H. J., Mahner, F., Drenckhahn, D., Klenk, H.-D., and Feldmann, H. (1993). Replication of Marburg virus in human endothelial cells. A possible mechanism for the development of viral hemorrhagic disease. *J. Clin. Invest.* **91**, 1301–1309.
117. Feldmann, H., Bugany, H., Mahner, F., Klenk, H.-D., Drenckhahn, D., and Schnittler, H.-J. (1996). Filovirus-induced endothelial leakage triggered by infected monocytes/macrophages. *J. Virol.* **70**, 2208–2214.
118. Sanchez, A., and Feldmann, H. (1996). Detection of Marburg and Ebola virus infections by polymerase chain reaction assays. *In* "Frontiers in Virology—Diagnosis of Human Viruses by Polymerase Chain Reaction Technology" (Y. Becker and G. Darai, eds.), 2nd Ed., pp. 411–418. Springer-Verlag, New York.
119. Sanchez, A., and Kiley, M. P. (1987). Identification and analysis of Ebola virus messenger RNAs. *Virology* **157**, 414–420.
120. Kingsbury, D. W. (1974). The molecular biology of paramyxoviruses. *Med. Microbiol. Immunol.* **160**, 73–83.
121. Kolakofsky, D., Vidal, S., and Curran, J. (1991). Paramyxovirus RNA synthesis and P gene expression. *In* "The Paramyxoviruses" (D. W. Kingsbury, ed.), pp. 215–233. Plenum Press, New York.
122. Becker, S., Spiess, M., and Klenk, H.-D. (1995). The asialoglycoprotein receptor is a potential liver-specific receptor for Marburg virus. *J. Gen. Virol.* **76**, 393–399.
123. Colonno, R. J., and Banerjee A. K. (1977). Mapping and initiation studies on the leader RNA of vesicular stomatitis virus. *Virology* **77**, 260–268.
124. Leppert, M., Rittenhouse, L., Perrault, J., Summers, D. F., and Kolakofsky, D. (1979). Plus and minus strand leader RNAs in negative strand virus-infected cells. *Cells* **18**, 735–747.
125. Mühlberger, E., Trommer, S., Funke, C., Klenk, H.-D. and Becker, S. (1996). Determination of the termini of genomic RNA and mRNA species of Marburg virus. *Virology*, in press.
126. Collins, P. L. (1991). The molecular biology of human respiratory syncytial virus (RSV) of the genus *Pneumovirus*. *In* "The Paramyxoviruses" (D. W. Kingsbury, ed.), pp. 103–162. Plenum Press, New York.

127. Martini, G. A., and Siegert, R. (1971). "Marburg Virus Disease." 1st Ed. pp. 1–237. Springer-Verlag, New York.
128. Peters, C. J., Sanchez, A., Rollin, P. E., Ksiazek, T. G., and Murphy, F. A. (1996). Filoviridae: Marburg and Ebola viruses. *In* "Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), 3rd Ed., pp. 1161–1176. Raven Press, New York.
129. Wulff, H., and Conrad, J. L. (1977). Marburg virus disease. *In* "Comparative Diagnosis of Viral Diseases" (E. Kurstak, ed.), Vol. 2, pp. 3–33. Academic Press, New York.
130. Piot, P., Sureau, P., Bremen, G., Heyman, D., Kintoki, V., Masamba, M., Mbuyi, M., Miatudila, M., Ruppol, F., van-Nieuwenhove, S., White, M. K., van der Groen, G., Webb, P. A., Wulff, H., and Johnson, K. M. (1978). Clinical aspects of Ebola virus infection in Yambuku area, Zaire, 1976. *In* "Ebola Virus Haemorrhagic Fever" (S. R. Pattyn, ed.), pp. 7–14. Elsevier/ North-Holland, Amsterdam.
131. Sureau, P. H. (1989). Firsthand clinical observations of hemorrhagic manifestations in Ebola hemorrhagic fever in Zaire. *Rev. Infect. Dis.* **11**, 790–793.
132. Elliott, L. H., Bauer, S. P., Perez-Oronoz, G., and Lloyd, E. S. (1993). Improved specificity of testing methods for filovirus antibodies. *J. Virol. Methods.* **43**, 85–100.
133. Ksiazek, T. G. (1991). Laboratory diagnosis of filovirus infections in non-human primates. *Lab. Animal* **20**, 34–46.
134. Dietrich, M., Schumacher, H. H., Peters, D., and Knobloch, J. (1978). Human pathology of Ebola virus infection in Sudan. *In* "Ebola Virus" (S. R. Pattyn, ed.), pp. 37–42. Elsevier/North Holland, Amsterdam.
135. Rollin, P. E., Ksiazek, T. G., Jahrling, P. B., Hanes, M., and Peters C. J. (1990). Detection of Ebola-like viruses by immunofluorescence. *Lancet* **336(8730)**, 1591.
136. Ksiazek, T. G., Rollin, P. E., Jahrling, P. B., Johnson, E., Dalgard, D. W., and Peters, C. J. (1992). Enzyme immunosorbent assay for Ebola virus antigens in tissues of infected primates. *J. Clin. Microbiol.* **30**, 947–950.
137. Mitchell, S. W., and McCormick, J. B. (1982). "Mobile Clinical Laboratory Manual." pp. 1–60. Centers for Disease Control, Atlanta.
138. Centers for Disease Control and Prevention (1988). Management of patients with suspected viral hemorrhagic fever. *MMWR* **37 (Suppl 3)**, 1–16.
139. Foberg, U., Fryden, A., Isaksson, B., Jahrling, P. B., Johnson, A., McKee, K., Niklasson, B., Norrnann, B., Peters, C. J., and Bengtsson, M. (1991). Viral hemorrhagic fever in Sweden: Experiences from management of a case. *Scand. J. Infect. Dis.* **23**, 143–151.
140. World Health Organization (1995). Viral haemorrhagic fever—management of suspected cases. *Weekly Epidemiol. Rec.* **70**, 249–256.
141. Bowen, E. T. W., Platt, G. S., Lloyd, G., Raymond, R. T., and Simpson, D. I. H. (1980). A comparative study of strains of Ebola virus isolated from southern Sudan and northern Zaire in 1976. *J. Med. Virol.* **6**, 129–138.
142. Lupton, H. W., Lambert, R. D., and Bumgardner, D. L. (1980). Inactivated vaccine for Ebola virus efficacious in guinea pig model. *Lancet* **2**, 1294–1295.
143. Ignatyev, G. M., Agafonov, A. P., Strelysova, M. A., Kuzymin, V. A., Mainagasheva, G. I., Spirin, G. V., and Chernyi, N. B. (1991). A comparative study of the immunological indices in guinea pigs administered an inactivated Marburg virus. *Voprosy Virusologii* **36**, 421–423.
144. Agafonov, A. P., Ignatyev, G. M., Kuzymin, V. A., Akimenko, Z. L., Kosareva, T. V., and Kashentseva, E. A. (1992). The immunogenic properties of Marburg virus proteins. *Voprosy Virusologii* **37**, 58–61.
145. Agafanov, A. P., Ignatyev, G. M., Akimenko, Z. A., and Volchkov, V. E. (1993). Study of immunogenic and protective properties of Marburg virus GP, NP and VP40 proteins. IXth International Congress of Virology, p. 300. Glasgow, Scotland. [Abstract]

146. Centers for Disease Control and Prevention (1990). Update: Ebola-related filovirus infection in nonhuman primates and interim guidelines for handling nonhuman primates during transit and quarantine. *MMWR* **39(2)**, 22–24, 29–30.
147. Elliott, L. H., McCormick, J. B., and Johnson, K. M. (1982). Inactivation of Lassa, Marburg, and Ebola viruses by gamma irradiation. *J. Clin. Microbiol.* **16**, 704–708.
148. Kissling, R. E., Murphy, F. A., and Henderson, B. E. (1970). Marburg virus. *Ann. N. Y. Acad. Sci.* **174**, 932–945.
149. Baskerville, A., Fisher Hoch, S. P., Neild, G. H., and Dowsett, A. B. (1985). Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. *J. Pathol.* **147**, 199–209.
150. Pereboeva, L. A., Tkachev, V. K., Kolesnikova, L. V., Krendeleva, L. I., and Ryabchikova, E. I. (1993). The ultrastructural changes in guinea pig organs during the serial passage of the Ebola virus. *Voprosy Virusologii* **38(4)**, 179–182.
151. Bowen, E. T. W., Lloyd, G., Harris, W. J., Platt, G. S., Baskerville, A., and Vella, E. E. (1977). Viral haemorrhagic fever in southern Sudan and northern Zaire. *Lancet* **1**, 571–573.
152. Fisher-Hoch, S. P., Platt, G. S., Neild, G. H., Southee, T., Baskerville, A., Raymond, R. T., Lloyd, G., and Simpson, D. I. H. (1985). Pathophysiology of shock and hemorrhage in a fulminating viral infection (Ebola). *J. Infect. Dis.* **152**, 887–894.
153. Murphy, F. A. (1978). Pathology of Ebola virus infection. *In* "Ebola Virus Haemorrhagic Fever" (S. R. Pattyn, ed.), pp. 37–42. Elsevier/North-Holland, Amsterdam.
154. Munk, K., Pritzer, E., Gutte, B., Garten, W., and Klenk, H.-D. (1992). Carbohydrate masking of an antigenic epitope of influenza virus haemagglutinin independent of the oligosaccharide size. *Glycobiology* **2**, 233–240.
155. Chepurnov, A. A., Ignatyev, G. M., and Volchkov, V. E. (1993). Immunologic and biochemical indices Ebola fever. IXth International Congress of Virology, p. 300. Glasgow, Scotland. [Abstract]
156. Ignatyev G. M., Streltsova, M. A., Agafonov, A. P., and Netesov, S. V. (1993). The study of immunological properties of inactivated Marburg virus. IXth International Congress of Virology, p. 300. Glasgow, Scotland. [Abstract]
157. Belanov, E. F., Bukreyev, A. A., Bazhutin, N. B., Spiridonov, V. A., Netesov, S. V., and Blinov, V. M. (1993). Monkeys *C. aethiops* are sensitive to repeated infection with Marburg virus. IXth International Congress of Virology, p. 299. Glasgow, Scotland. [Abstract]
158. Simpson, D. I. H., Zlotnik, I., and Rutter, D. A. (1968). Vervet monkey disease. Experimental infection of guinea pigs and monkeys with the causative agent. *Br. J. Exp. Pathol.* **49**, 458–464.
159. Ellis, D. S., Bowen, E. T. W., Simpson, D. I. H., *et al.* (1978). Ebola virus: A comparison, at ultrastructural level, of the behaviour of the Sudan and Zaire strains in monkeys. *Br. J. Exp. Pathol.* **59**, 584–593.
160. Majno, G., and Palade, G. (1961). Studies on inflammation I. The effect of histamine and serotonin on vascular permeability, an electron microscopic study. *J. Biophys. Biochem. Cytol.* **11**, 571–605.
161. Crone, C. (1987). The Malpighi lecture. From "porositates carnis" to cellular microcirculation. *Int. J. Microcic. Clin. Exp.* **6**, 101–122.
162. Michel, C. C. (1988). Capillary permeability and how it may change. *J. Physiol.* **404**, 1–29.
163. Schnittler, H.-J., Wilke, A., Gress, T., Sutton, N., and Drenckhahn, D. (1990). Role of actin and myosin in the control of paracellular permiability in pig, rat, and human vascular endothelium. *J. Physiol.* **431**, 379–401.
164. Wong, M. K. K., and Gotlieb, A. I. (1988). The reorganization of microfilaments

centrosomes and microtubules during in vitro small wound reendothelialization. *J. Cell. Biol.* **107**, 1777–1783.

165. Emond, R. T. D. (1978). Isolation, monitoring and treatment of a case of Ebola virus infection. *In* "Ebola Virus Hemorrhagic Fever" (S. R. Pattyn, ed.), 1st Ed., pp. 27–32. Elsevier/North-Holland, Amsterdam.
166. Heymann, D. L., Weisfeld, J. S., Webb, P. A., Johnson, K. M., Cairns, T., and Berquist, H. (1980). Ebola hemorrhagic fever: Tandala, Zaire, 1977–78. *J. Infect. Dis.* **142**, 372–376.
167. Johnson, K. M., Scribner, C. L., and McCormick, J. B. (1981). Ecology of Ebola virus: A first clue? *J. Infect. Dis.* **143**, 749–751.
168. Teepe, R. G. C., Johnson, B. K., Ocheng, D., Gichogo, A., Langatt, A., Ngindu, A., Kiley, M. P., Johnson, K. M., and McCormick, J. B. (1983). A probable case of Ebola virus hemorrhagic fever in Kenya. *East Afr. Med. J.* **60**, 718–722.
169. World Health Organization (1996). Outbreak of Ebola haemorrhagic fever in Gabon officially declared over. *Weekly Epidemiol. Rec.* **71**, 125–126.

ADVANCES IN VIRUS RESEARCH, VOL. 47

MOLECULAR CHARACTERIZATION OF PESTIVIRUSES

Gregor Meyers* and Heinz-Jürgen Thiel†

*Federal Research Center for Virus Diseases of Animals
72076 Tübingen, Germany
†Institut für Virologie (Veterinärmedizin)
Justus-Liebig-Universität Giessen
35392 Giessen, Germany

I. Introduction

Diseases caused by pestiviruses affect members of the Suidae (swine) and several species of ruminants, including cattle, sheep, goats, and wild ruminants. Classical swine fever (CSF) was first recognized in 1833 in Ohio, and it has been speculated that CSF was at that time a new disease of swine. Bovine viral diarrhea (BVD), mucosal disease (MD), and border disease of sheep (BD) were described in the 1940s and 1950s, respectively. The two former diseases both occur in cattle, show distinct clinical courses, and are actually caused by the same virus, termed BVD virus (BVDV). The finding that the causative agents of CSF and BVD are serologically related (Darbyshire, 1960) laid the ground for the establishment of pestiviruses as one group. The

respective genus *Pestivirus*, which currently comprises the three species BVDV, CSF virus (CSFV), and BD virus (BDV), was initially part of the *Togaviridae* (Westaway *et al.*, 1985).

Even though pestiviruses cause interesting and economically important diseases, research on basic aspects of pestiviruses made little progress for several years after discovery of their relationship. This was mainly due to difficulties inherent in the particular virus system: the titers of pestiviruses grown in tissue culture cells are generally low and attempts to obtain highly purified preparations of virions have been unsuccessful (Moennig and Plagemann, 1992). Thus, neither viral proteins nor the respective nucleic acids were available in sufficient quantity and purity. Major breakthroughs did not happen until modern techniques of molecular biology were applied to pestiviruses. Molecular cloning and sequencing of pestiviral genomes, as well as expression of defined parts of their polyproteins, allowed crucial conclusions concerning, in particular, the strategy of gene expression, genome organization, composition of virions, relationships of the three species at the molecular level, etc. At this point it appeared justified to reclassify pestiviruses as a new genus of the family *Flaviviridae* (Wengler *et al.*, 1995). This chapter provides an overview of the different aspects of pestiviruses, including diseases they cause and their molecular biology. Particular emphasis is placed on the molecular characterization of cytopathogenic pestiviruses which arise by nonhomologous RNA recombination.

II. Pestiviruses: Diseases, Genomes, and Proteins

A. Significance, Taxonomy, and General Properties of Pestiviruses

Pestiviruses represent causative agents of diseases that have a significant negative financial impact for the livestock industry worldwide. The genus *Pestivirus* currently comprises three species named according to the host from which they were isolated and where they cause disease: classical swine fever virus (CSFV) (also termed hog cholera virus or European swine fever virus), bovine viral diarrhea virus (BVDV), and border disease virus (BDV) of sheep (Wengler *et al.*, 1995). The serological relationship of CSFV and BVDV was first described in 1960 (Darbyshire, 1960) and led to the establishment of the genus *Pestivirus*; however, their family status was recently changed. Pestiviruses used to be classified as members of the *Togaviridae* (Westaway *et al.*, 1985), but results concerning strategy of gene expression and genome organization led to reclassification as an additional

genus in the *Flaviviridae*; previously, this family consisted only of the genus *Flavivirus*. Recently, the *Hepatitis C-like Viruses* (HCV) have also been included as a third genus of this family (Wengler *et al.*, 1995). Common characteristics of all family members are (i) single-stranded genomic RNA of positive sense with a size generally between 9.5 and 12.5 kb, (ii) presence of one large open reading frame (ORF) which is translated into a hypothetical polyprotein, the latter being processed co- and posttranslationally by host cell and virus encoded proteases, (iii) similar genomic localization and function of several structural and nonstructural proteins, (iv) size and composition of virions with 2–3 envelope-associated (glyco)proteins and one nucleocapsid protein. For more details about members of the genera *Flavivirus* and *Hepatitis C-like Viruses* the reader is referred to respective review articles (Chambers *et al.*, 1990; Rice, 1996; Houghton, 1996; Monath and Heinz, 1996).

One remarkable property of pestiviruses is the existence of two biotypes that were recognized according to morphological changes they cause during growth in tissue culture cells. Noncytopathogenic (noncp) pestiviruses replicate without clearly visible effects, while cytopathogenic (cp) viruses lead to lysis of appropriate target cells. The molecular basis for this distinction is the subject of current investigations and will be the major focus of this review.

B. Diseases Caused by Pestiviruses

1. General

All pestiviruses are able to cross the placenta and to infect the fetus (Thiel *et al.*, 1996; Moennig and Plagemann, 1992). In this context it should be noted that the placenta of ruminants and pigs does not allow transfer of antibodies from the pregnant animal to the fetus. The time point of gestation at which the intrauterine infection occurs largely determines the outcome for the fetus. Such intrauterine infections may lead to abortion, stillbirth, fetal malformation, or weak or apparently healthy offspring. The last may have been infected toward the end of pregnancy and mounted an effective immune response against the virus which cleared the infection. Of particular importance is the capability of noncytopathogenic pestiviruses to establish persistent infections in fetuses not yet immunocompetent. Such animals acquire an immunotolerance to the virus and they are persistently viremic for the rest of their lives. In ruminants, persistent infections may remain clinically inapparent and last for many years, while persistently infected pigs invariably become sick and die, at the latest several months

after birth. Persistently infected animals which continuously shed virus have a major impact on the epidemiology of pestivirus-induced diseases (Thiel *et al.*, 1996).

Postnatal infections of ruminants with BVDV or BDV are generally considered to be harmless. One exception is the recently described hemorrhagic syndrome of cattle (see the following). The infection of pigs with CSFV, however, is usually characterized by acute disease and sometimes high mortality rates.

2. Bovine Viral Diarrhea and Mucosal Disease

Seroepidemiological surveys have shown that most cattle contract a BVDV infection during their lifetime. Isolations from the field show that in the vast majority of cases the noncytopathogenic biotype induces these antibodies. Such so-called acute infections may be associated with transient leukopenia, fever, mild diarrhea, a respiratory disease complex, and immunosuppression. Alternatively the infection may proceed subclinically. The vast majority of BVDV infections apparently remain unnoticed (Thiel *et al.*, 1996). However, a severe clinical condition in calves associated with BVDV was recently described. The disease is characterized by severe thrombocytopenia and hemorrhages (Pellerin *et al.*, 1994; Ridpath *et al.*, 1994; Corapi *et al.*, 1989; Rebhun *et al.*, 1989). The symptoms actually resemble those of other hemorrhagic diseases like CSF. The causative agent has been described as noncytopathogenic and falls into a "novel" group of ruminant pestiviruses (see below).

Mucosal disease (MD) is a sporadic disease in cattle which generally occurs between the ages of 6 months and 2 years. The disease is characterized by high case fatality, with death occurring usually within 2 weeks after the onset of clinical signs (Baker, 1987). Extensive ulceration of the gastrointestinal tract is the most prominent lesion (Baker, 1987; Moennig and Plagemann, 1992).

Since its first description about 40 years ago, the sporadic incidence of MD has puzzled investigators. Importantly, MD occurs only in cattle that had been infected with a noncp BVDV during the first trimester of gestation (Liess *et al.*, 1974). Such animals acquire immunological tolerance with regard to the infecting BVDV strain and develop a persistent infection. The immunotolerance is restricted to the particular noncp virus, and there is no indication of either formation of antibodies or cytolytic T-cells against the persisting virus. This is in contrast to other virus systems, in which persistent infections with a so-called split immunotolerance have been observed, for example, the well-studied lymphocytic choriomeningitis virus (LCMV) system (Ahmed *et al.*, 1984).

A major breakthrough was the observation that both viral biotypes, noncp BVDV as well as cp BVDV, were consistently found in animals that had died of MD (McKercher *et al.*, 1968; McClurkin *et al.*, 1985; Wilhelmsen *et al.*, 1991). For one given animal with MD the noncp virus together with the cp virus is called a "virus pair." In contrast to the known variability of BVDV field isolates, the members of such a "pair" of noncp BVDV and cp BVDV have been shown in many cases to be antigenically very closely related. This finding led to the hypothesis that during pathogenesis of MD a cp virus develops from the noncp virus by undergoing some kind of mutation (Corapi *et al.*, 1988; Howard *et al.*, 1987). Moreover, MD can be reproduced experimentally by superinfection of persistently viremic animals with cp BVDV; there is some evidence that the cp virus used for superinfection has to be antigenically closely related to the "endogenous" noncp BVDV strain (Bolin *et al.*, 1985; Brownlie *et al.*, 1984; Moennig *et al.*, 1990).

Tissues from cattle that died of MD were analyzed with regard to replication of noncp and cp virus. It was striking that the cp BVD virus appeared to replicate particularly efficiently in those organs where dramatic damage was observed, for example, the lymphoid tissue of the gastrointestinal tract (Liebler *et al.*, 1991). In contrast, the noncp virus was distributed in basically all tissues of the affected animals.

Some BVDV infected cattle suffer from clinical illness for a prolonged period, a condition termed "chronic disease" or "chronic MD." The disease has not been reproduced experimentally, but from natural cases there is evidence that affected animals are also persistently viremic. It has been suggested that chronic disease might develop when viremic animals are superinfected with a cp BVDV sharing only partial homology with the "endogenous" noncp BVDV (Brownlie, 1991; Moennig *et al.*, 1993).

3. Border Disease

BDV infections of sheep resemble in many aspects those of cattle with BVDV, especially since congenital BDV infections are of particular importance. Acute, postnatal BDV infections lead to no or mild clinical symptoms (Shaw *et al.*, 1967; Vantsis *et al.*, 1979). Like other pestiviruses, BDV will cross the placenta of pregnant animals to cause fetal death, malformations, or persistent infection (Barlow, 1972). Animals persistently infected with BDV may appear healthy. Alternatively such animals may show the "hairy shaker" syndrome, which is characterized by hairy fleece, low birth weight, tremors, and ataxia (Nettleton *et al.*, 1992). In analogy to the BVDV system the occurrence of a mucosal disease-like syndrome in association with the recovery of cp BDV has been described (Nettleton *et al.*, 1992).

4. Classical Swine Fever

According to descriptions from the last century, CSF used to be a peracute to acute disease with a short incubation period and high mortality rates (Dunne, 1973). The course of the disease has long since changed dramatically. Today the peracute form has disappeared and the acute kind no longer reaches high mortality rates. In addition, chronic forms of the disease are common (van Oirschot, 1992; Wensvoort and Terpstra, 1985). The different clinical pictures can be attributed to CSFV strains of varying virulence. The disease is characterized by high fever, leukopenia, and hemorrhages of skin and inner organs; the latter lesions can be typical for classical swine fever (Thiel *et al.*, 1996). CSF viruses may be categorized into groups with high, moderate, low, and no apparent virulence. Nonvirulent strains are used as live attenuated vaccine viruses (Baker, 1946; Koprowski *et al.*, 1946). The molecular basis for attenuation of CSFV is not known.

C. The Pestiviral Polyprotein: Localization, Generation, Structure, and Function of Individual Proteins

The pestiviral genome consists of a 5′ noncoding region (NCR) of almost 400 nucleotides with multiple AUG codons, a single long open reading frame (ORF) encoding approximately 4000 amino acids, and a 3′ NCR of more than 200 nucleotides (Collett *et al.*, 1988a: Meyers *et al.*, 1989a; Moormann *et al.*, 1990). Similar features have been described for hepatitis C virus (HCV). The ORF of HCV encodes a polyprotein of about 3000 amino acids; the noncoding sequences for the 5′ NCR comprise about 340 nucleotides, again with multiple AUG codons and up to 42 nucleotides for the 3′ NCR (Rice, 1996). For members of the genus *Flavivirus* the size of the ORF, about 3400 codons, is similar to the one of members from the other two genera; however, for flaviviruses only 95–132 nucleotides represent the 5′ NCR and 114–624 nucleotides the 3′ NCR (Rice, 1996). In general the 3′ end of all members of the *Flaviviridae* is not polyadenylated. Similar to eukaryotic mRNAs, the 5′ end of the flaviviral genome is capped and initiation of translation presumably occurs by ribosome scanning. For HCV and pestiviruses, however, it is assumed that the genomes are translated in a cap-independent manner. Analyses of the long 5′ NCR of both groups suggest complex secondary structures in this region comparable to secondary structures found in picornaviruses. These findings, and especially the presence of multiple AUG triplets upstream of the actual initiation codon, indicate the existence of an internal ribosomal entry

site (IRES) for HCV and pestiviruses. More recent studies have demonstrated that translation of the HCV genomic RNA is indeed initiated by a cap-independent internal ribosome binding mechanism, and the same apparently applies for pestiviruses (Rice, 1996; Poole *et al.*, 1995).

The hypothetical polyprotein translated from the large ORF is processed co- and posttranslationally in an orderly fashion by host-cell and virus-encoded proteases to result in 11–12 mature proteins (Fig. 1). The first cleavage event in pestiviral protein biogenesis is due to an autoproteolytic activity of the N-terminal protease (N^{pro}) that is responsible for cleavage between N^{pro} and the nucleocapsid protein (C) (Stark *et al.*, 1993; Wiskerchen *et al.*, 1991). The respective cleavage site has been determined for CSFV by N-terminal sequencing of the capsid protein (Stark *et al.*, 1993). Cleavage occurs between Cys-168 and Ser-169. The presence of a highly conserved region preceding Ser-169 lets us assume that processing by N^{pro} occurs at this site for all pestiviruses (Fig. 2). In order to determine the type of protease, mutagenesis studies were performed. It was intended to show that N^{pro} is a serine-type proteinase. However, the experiments failed to demonstrate conclusively that N^{pro} belongs to this class of proteinases (Wiskerchen *et al.*, 1991). Interestingly, comparisons with cellular proteases and proteases encoded by other positive-stranded RNA viruses showed homology between N^{pro} and the group of papain-like cysteine proteases (PCPs) (Stark *et al.*, 1993). As the first protein of the pestiviral polyprotein N^{pro} fits into the group of accessory leader proteases identified in viruses from animals and plants (Gorbalenya *et al.*, 1991). According to the available data on polyprotein processing (see below), N^{pro} appears to be involved only by releasing itself from the nascent polypeptide. It will be interesting to determine the role of N^{pro} in replication of pestiviruses, especially since this nonstructural protein has no counterpart in HCV and other flaviviruses.

With regard to the polyprotein, the nucleocapsid protein C is followed by an internal signal sequence which mediates translocation of the glycoproteins (E proteins), namely E0, E1, and E2, in the order of their arrangement in the polyprotein. In the past, pestiviral glycoproteins have mostly been named according to their apparent molecular weights, and thus analogous proteins from different pestiviral strains and isolates have different names. A standardized nomenclature, especially for the E proteins of pestiviruses, would help avoid confusion when comparing published data. A nomenclature based on analogies for different genera of the *Flaviviridae* would facilitate comparisons within the family (see addendum).

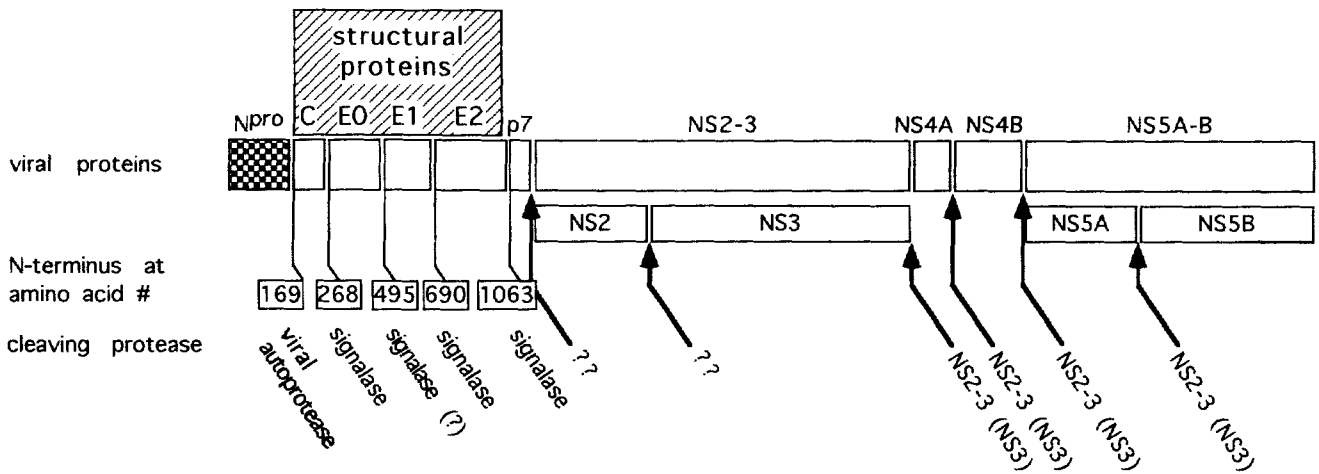


FIG. 1. Genome organization of pestiviruses. For designations of the individual structural and nonstructural viral proteins (upper part of the figure) see text and addendum. The N-terminal amino acids of the structural proteins and p7 are indicated in the second row; the numbers refer to the polyprotein of CSFV. The different proteases involved in processing are listed in the bottom row; ?, protease not known.

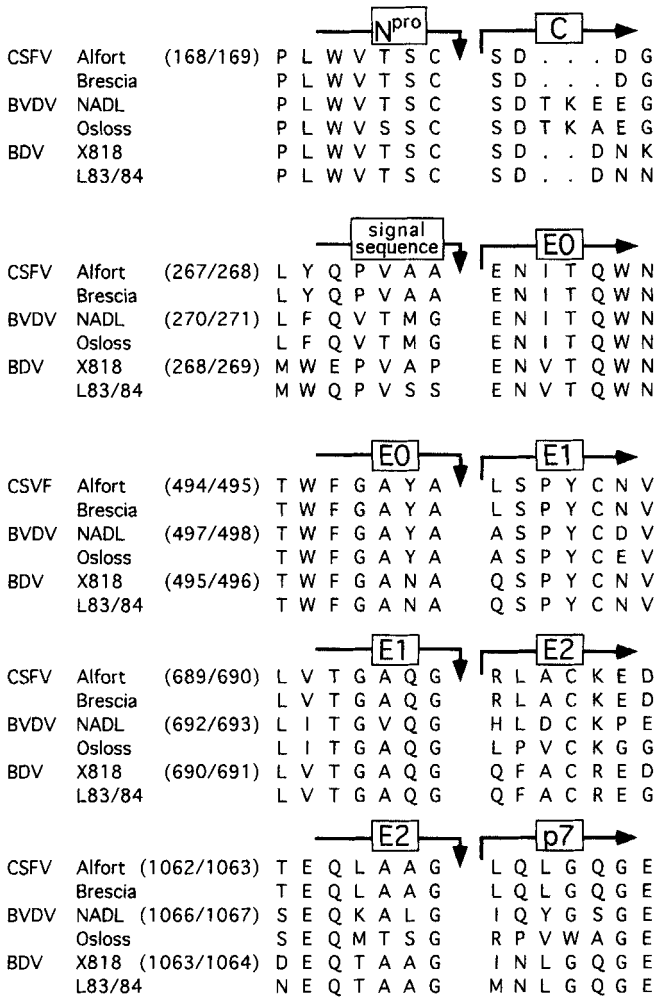


FIG. 2. Comparison of putative cleavage sites among pestiviruses which involve generation and/or release of structural proteins. The N termini of the structural proteins were determined for CSFV Alfort; the N terminus of p7 was determined for BVDV CP7 (not listed in the figure).

For E1 and E2, membrane anchors have been identified by different means (Rümenapf *et al.*, 1993; van Zijl *et al.*, 1991), whereas for E0 the mechanism of association with the virion remains unclear. All three glycoproteins are present as intermolecular disulfide-linked complexes, an E0 homodimer, an E1–E2 heterodimer, and an E2 homodimer, all of which can be detected in infected cells as well as in virions (Thiel *et al.*, 1991;

Weiland *et al.*, 1990). The cysteine residues of the glycoproteins are highly conserved among all pestiviruses sequenced so far (Weiland *et al.*, 1990). According to a recently proposed model, cysteines in the C-terminal half of E2 lead to intermolecular disulfide bonds with E1 or E2, whereas cysteines in the N-terminal half form intramolecular disulfide bonds (van Rijn *et al.*, 1994).

E0, which has no counterpart in other flaviviruses, shares regions of significant sequence similarity with certain fungal and plant ribonucleases (Hulst *et al.*, 1994; Schneider *et al.*, 1993). Although the sequence similarities are limited to two small regions of E0, they include all of the proposed catalytically important residues. Experiments conducted with purified E0 from CSFV-infected cells showed that E0 actually exhibits RNase activity (Schneider *et al.*, 1993). Biochemical characterization of E0 revealed a pronounced substrate specificity for uridine and inhibition by zinc ions (Schneider *et al.*, 1993). Using recombinant E0 expressed in insect cells it has been shown that neither the formation of homodimers nor the presence of carbohydrate moieties is a rerequisite of RNase activity (Windisch *et al.*, 1996). It is particularly interesting that monoclonal antibodies which effectively neutralized CSFV also exerted a high degree of E0 ribonuclease inhibition. This striking correlation suggests that the ribonuclease activity of CSFV E0 plays a role in the viral life cycle (Windisch *et al.*, 1996). The presence of a viral envelope protein with an intrinsic RNase activity is thus far unique for pestiviruses. Elucidation of the function of this enzymatic activity in the viral life cycle represents one of the most intriguing aspects of pestivirus research.

With respect to the polyprotein, E2 is followed by a hypothetical protein of 7 kDa (p7). Recently, p7 has been identified in virus-infected cells but not in virions (Elbers *et al.*, 1996). Interestingly, cleavage between E2 and p7 is not complete, leading to two forms of E2 with different C termini. This situation is reminiscent of HCV, where also two forms of E2 as well as a protein corresponding to p7 have been observed (Rice, 1996). For CSFV, the N termini of the structural proteins and p7 have been determined (Elbers *et al.*, 1996; Rümenapf *et al.*, 1993); the respective cleavage sites are apparently conserved for all pestiviruses (Fig. 2). The first step in the production of pestiviral glycoproteins is the translocation of the internal signal peptide downstream of C. The cleavage between C and the nascent E012 occurs rapidly upon translocation and is most likely effected by signalase (Rümenapf *et al.*, 1993). It is not known whether the signal sequence represents an integral part of C, rendering it an "anchored core" as described for flaviviruses. Cleavage between E1, E2, and p7 is pre-

sumably also due to signalase activity and is considered to occur instantly upon translocation. Processing between E0 and E1, however, apparently represents an exception, since it occurs in a slightly delayed fashion (Rümenapf *et al.*, 1993). Cleavage at this site actually takes place downstream of a sequence similar to those commonly employed by signalase but lacks an upstream hydrophobic domain. This could account for a delayed cleavage by signalase. Alternatively the sequence could be cleaved by a different protease (Rümenapf *et al.*, 1993).

The C-terminal two thirds of the ORF encode exclusively nonstructural proteins (Fig. 1); downstream of p7, the first cleavage product is represented by NS2–3, which for most pestiviruses is partially processed to yield NS2 and NS3 (Collett *et al.*, 1988b; Meyers *et al.*, 1991, 1992). There are, however, remarkable exceptions: after infection with noncp BVDV strains, only NS2–3 but no respective processing products can be detected (Donis and Dubovi, 1987c; Pocock *et al.*, 1987). In addition, certain cp BVDV strains express NS3 from a duplicated genomic region (see below) and their NS2–3 is not cleaved. If cleavage of NS2–3 occurs, there is apparently no precursor–product relationship between NS2–3 and NS3 (Collett *et al.*, 1991). Interestingly, the degree of cleavage of NS2–3 differs markedly among pestiviruses. In cells infected with cp BVDV the level of NS3 is generally at least as high as the level of NS2–3. After infection with either noncp BDV strains or noncp CSFV strains, however, only small amounts of NS3 can be detected when compared with NS2–3 (Thiel *et al.*, 1991; Becher *et al.*, 1994); the situation for a few characterized cp strains from these two pestiviral species is described below. It is noteworthy that the pestiviral NS2 represents a hydrophobic protein, whereas NS3 is rather hydrophilic. On this basis it has been suggested that the intracellular localization of NS2–3 and its cleavage products is different (Wiskerchen and Collett, 1991).

The pestiviral NS3 contains sequence motifs characteristic for the analogous protein of all members of the *Flaviviridae*. The predicted enzymatic activities, namely serine protease, NTPase, and helicase activity, have been directly demonstrated (Wiskerchen and Collett, 1991; Tamura *et al.*, 1993; Warrener and Collett, 1995). Accordingly, NS3 and NS2–3 are expected to exhibit these activities. The three predicted residues in the proteinase catalytic triad are located in the N-terminal domain of NS3; the NTPase and helicase motifs can be found approximately in the center of the protein. The serine protease is responsible for the cleavages which occur at its own C terminus and further downstream (Wiskerchen and Collett, 1991; Tautz *et al.*, in preparation). However, for one BVDV strain it has been claimed that

the NS3 proteinase cleaves also between NS2 and NS3 (Wiskerchen and Collett, 1991), but this does certainly not apply in other cases. For example, in cp BVDV strains with ubiquitin-coding sequences within NS2–3 (strains Osloss and CP 14, see Section III,A), a cellular protease is responsible for generation of the amino terminus of NS3. In many other cases, including cp BVDV, noncp BDV, and noncp CSFV strains, neither the protease(s) responsible for processing of NS2–3 nor the cleavage site(s) are known. Interestingly, cleavage at the NS2–3 site of the HCV polyprotein is catalyzed by a virus-encoded protease which is different from the NS3 protease of HCV and appears to be unique for members of the genus *Hepatitis C-like Viruses*. This so-called second HCV-encoded proteinase encompasses the C-terminal portion of NS2, the cleavage site of NS2–3, and the serine proteinase domain of NS3. It has been proposed that the enzyme represents a zinc-dependent metalloproteinase (Rice, 1996).

In the polyprotein the nonstructural proteins NS4 and NS5 are located downstream of NS3; like for HCV, both are further processed to yield NS4A, NS4B and NS5A, NS5B (Collett *et al.*, 1988a, 1991; Meyers *et al.*, 1992). It appears likely that the cleavage between NS3 and NS4A occurs only in cis. The sites downstream of NS4A, however, are probably cleaved in trans. For HCV and presumably also pestiviruses, NS4A serves as a cofactor, being required for certain proteolytic cleavages accomplished by NS3 (Rice, 1996; Tautz *et al.*, submitted). Studies are in progress to elucidate the role of NS4A in polyprotein processing of pestiviruses. The functional role of NS4B in virus replication is not known.

The products of NS5 cleavage exhibit different intracellular halflives. While NS5A represents a stable protein, NS5B has a short halflife (Collett *et al.*, 1991). The presence of characteristic sequence motifs, including the highly conserved tripeptide Gly–Asp–Asp, suggests that NS5B represents the viral RNA-dependent RNA polymerase (Meyers *et al.*, 1989a). Although the same holds true for NS5B of HCV, an uncleaved NS5 precursor has not been demonstrated in the latter virus system (Rice, 1996).

D. Properties of Virions

Pestiviruses grow to only modest titers in tissue culture cells, are in general released poorly from cells, and tend to remain within infected cells. Liberated virions associate with components originating from serum and cells, properties that hamper purification and electron microscopic examination (Laude, 1977). In addition, the low buoyant

density of pestiviruses makes it difficult to purify them by physical means, such as equilibrium centrifugation (Rümenapf *et al.*, 1991). In suspensions of concentrated virions, enveloped particles with diameters between 40 and 60 nm have been observed (Moennig and Plagemann, 1992). Purified pestivirions consist of RNA, C, and an envelope with three associated glycoproteins (E proteins) (Thiel *et al.*, 1991). Immunoelectron microscopy has revealed that E0 and E2 are located at the surface of the virions (Weiland *et al.*, 1992; Weiland *et al.*, in preparation). The envelope surrounds a probably hexagonally shaped, electron-dense inner core structure with a diameter of about 30 nm (Horzinek *et al.*, 1967). Virions probably mature at intracytoplasmatic membranes, and their liberation is most likely achieved by exocytosis of virus-containing membrane vesicles (Bielefeldt Ohmann and Bloch, 1981; Bielefeldt Ohmann, 1988a, 1988b; Gray and Nettleton, 1987).

E. Relationships among Pestiviruses

The antigenic relationship between BVDV and CSFV was demonstrated in early investigations by employing immunodiffusion with polyclonal antisera (Darbyshire, 1960; Dinter, 1963). Subsequently, serological reagents were used to detect not only relationships but differences between species and strains of pestiviruses (Cay *et al.*, 1989; Edwards *et al.*, 1989; Moennig and Plagemann, 1992; Peters *et al.*, 1986; Weiland *et al.*, 1992; Wensvoort *et al.*, 1989). For this, cross-neutralization assays with polyclonal sera and other assay systems, in particular in conjunction with monoclonal antibodies (MAbs), were applied. MAbs have also been used for the identification and characterization of individual proteins; the MAbs described so far are specific for E0, E2, or NS3 (Corapi *et al.*, 1990; Greiser-Wilke *et al.*, 1990, 1991; Paton *et al.*, 1991; Weiland *et al.*, 1990, 1992). In addition, sera against bacterial fusion proteins and synthetic peptides comprising defined regions of the pestiviral polyprotein have served to establish antigenic relationships between pestiviruses.

MAbs directed against NS3 generally recognize conserved epitopes and their reactivity pattern is considered panpestivirus-specific (Edwards *et al.*, 1989; Peters *et al.*, 1986). In contrast, MAbs against E0 and E2 have been used to discriminate between pestivirus species as well as between strains of one species (Cay *et al.*, 1989; Weiland *et al.*, 1992; Wensvoort *et al.*, 1989; Kosmidou *et al.*, 1994). MAbs against E2 have also been used to map respective epitopes by using competitive binding assays, antigen capture assays, virus neutralization assays, deletion mutants, and neutralization escape mutants (Moennig *et al.*,

1988; Paton *et al.*, 1992; van Rijn *et al.*, 1993; Wensvoort, 1989). Although E2 appears to represent the major target of virus neutralizing antibodies, MAbs against E0 also mediate virus neutralization (Bolin *et al.*, 1988; Donis *et al.*, 1988; Greiser-Wilke *et al.*, 1990; Weiland *et al.*, 1990, 1992; Wensvoort *et al.*, 1989). As observed for HCV, the E2 protein of pestiviruses is highly variable in its N-terminal portion, and the majority of neutralization epitopes is obviously located in this part of E2. Thus far cross-neutralization among pestiviral species by MAbs has not been demonstrated.

Nucleotide sequence data from whole pestiviral genomes are available for BVDV and CSFV; for BDV, only partial sequences have been determined (Becher *et al.*, 1994, 1995, 1996). The analyses of nucleotide and deduced amino acid sequences have led to the evaluation of (i) the overall similarity between BVDV and CSFV strains, (ii) the homology of defined regions, as well as (iii) the relationship among individual proteins (Collett *et al.*, 1988a; Deng and Brock, 1992; Meyers *et al.*, 1989a; Moormann *et al.*, 1990; De Moerlooze *et al.*, 1993). For example, there is a high degree of variability in the structural glycoprotein E2, with values of amino acid identities as low as 80% within one species and less than 60% between all three species (Becher *et al.*, 1994); these values are comparable to the ones obtained for NS2. There is only one less-conserved pestivirus-encoded protein, namely p7, for which the degree of amino acid identity between species can be as low as 43% (Elbers *et al.*, 1996). In contrast, NS3 represents the most conserved protein among pestiviruses (Meyers *et al.*, 1989a), with more than 90% amino acid identity among pestiviral species. The comparison of nucleotide sequences has shown that the highly conserved 5′ NTR is particularly suited to identify pestiviruses by PCR; moreover, nucleotide sequencing of this short region seems to provide sufficient information for their differentiation (Hofmann *et al.*, 1994; Harasawa and Tomiyama, 1994; De Moerlooze *et al.*, 1993). In addition, the comparison of E2 amino acid sequences has been used to demonstrate variation among pestiviruses (Becher *et al.*, 1994). The latter study led to the identification of two groups of ovine pestiviruses, namely BVDV-like strains and "true" BDV strains. Recently a third group of sheep-derived pestiviruses has been identified by PCR and nucleotide sequencing of the 5′ NTR as well as of the N^{pro} and C coding regions (Becher *et al.*, 1995). Interestingly, a "novel" group of BVDV strains, which was identified mainly by PCR and nucleotide sequencing of the 5′ NTR (Pellerin *et al.*, 1994; Ridpath *et al.*, 1994), shows a high degree of similarity with the third group of ovine pestiviruses (Becher *et al.*, 1995). Such bovine strains may cause a hemorrhagic syndrome (see above). These findings

show that the current taxonomy, classifying pestiviruses according to host and disease, is misleading, especially since cross-species infections occur quite frequently. It has therefore been suggested that nucleotide sequence homologies be used for differentiation of pestiviruses and that they be divided into types, namely pestivirus type 1 (classic BVDV strains), pestivirus type 2 (CSFV strains), pestivirus type 3 ("true" BDV strains), and pestivirus type 4 ("novel" group of isolates from cattle and sheep) (Fig. 3) (Becher *et al.*, 1995). It remains

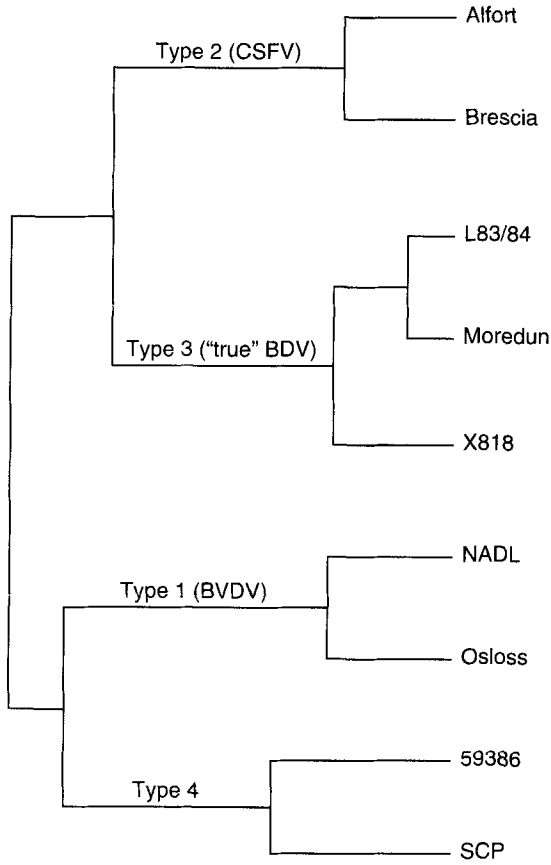


FIG. 3. Dendritic diagram comparing amino acid sequences of autoprotease N^{pro} from different pestiviruses. Four separate groups of pestiviruses can be distinguished. It has been suggested that these be called Type 1 (BVDV), Type 2 (CSFV), Type 3 ("true" BDV), and Type 4 (recently identified group of pestiviruses isolated from cattle and sheep); also see text.

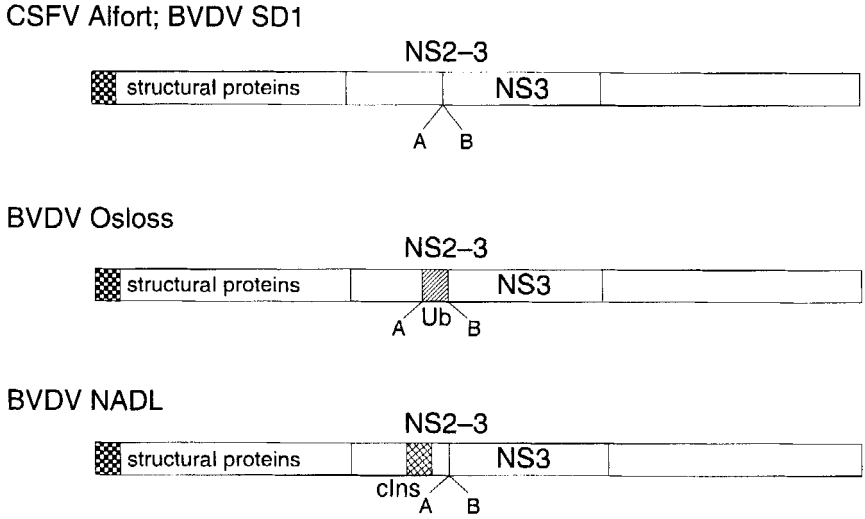


FIG. 4. Schematic presentation of the genomes of BVDV strains Osloss and NADL in comparison with CSFV Alfort and BVDV SD-1. The locations of the genes coding for the structural proteins and the nonstructural proteins NS2–3, NS3, and N^{pro} (checkered bar) are indicated. A and B designate the nucleotides preceding and following the ubiquitin-coding insertion in the BVDV Osloss genome; Ub, ubiquitin-coding insertion; cIns, cellular insertion of BVDV NADL.

to be seen whether additional pestivirus types will be identified. It would be no surprise if the molecular analysis of pestiviruses from wild ruminants leads to establishment of additional types.

III. RNA Recombination: The Molecular Basis for Cytopathogenicity of Pestiviruses

A. Gene Pool Parasites: BVD Viruses with Integrated Cellular Sequences

1. Identification of Cellular Sequences in the Genome of BVDV

Comparison of the genomic sequences of two laboratory strains of BVDV, namely Osloss (Renard *et al.*, 1987; de Moerlooze *et al.*, 1993) and NADL (Collett *et al.*, 1988a), with the one of CSFV Alfort (Meyers *et al.*, 1989a) resulted in identification of strain-specific insertions in the BVDV genomes (Meyers *et al.*, 1989b, 1990). The inserted sequence

```
BVDV  OSLOSS:    N L E H L G W I L K M Q I F V K T L T G K T I T L E V E P S D T
BVDV  NADL:      N L E H L G W I L R
Ubiquitin:                        M Q I F V K T L T G K T I T L E V E P S D T

BVDV  OSLOSS:    I E N V K A K I Q D K E G I P P D Q Q R L I F A G K Q L E D G R
Ubiquitin:       I E N V K A K I Q D K E G I P P D Q Q R L I F A G K Q L E D G R

BVDV  OSLOSS:    [S] L S D Y N I Q K E S T L H L V L R L R G [S] G P A V C K K
BVDV  NADL:                                                  G P A V C K K
Ubiquitin:       [T] L S D Y N I Q K E S T L H L V L R L R G [G]
```

FIG. 5. Amino acid sequence comparison of parts of the polyproteins encoded by the genomes of BVDV Osloss and NADL and the conserved animal ubiquitin. Amino acids differing between the animal ubiquitin and the protein encoded by the Osloss insertion are marked by boxes.

has a length of 228 nucleotides (BVDV Osloss) or 270 nucleotides (BVDV NADL). Both insertions are located in the region of the genome coding for NS2–3 but not at the same position (Fig. 4).

The overall nucleotide sequence homology between CSFV Alfort and BVDV NADL is about 66%, and at the protein level the two viruses exhibit 85% identical amino acids (Meyers *et al.*, 1989a). Disregarding the BVDV insertions, the encoded polyproteins of the three viruses differ in length by only one amino acid (CSFV Alfort 3898 residues, BVDV NADL minus insertion 3898 residues, BVDV Osloss minus insertion 3899 residues). Therefore, the discovery of dramatic differences between the genomes of BVDV and CSFV was highly unexpected, and it was interesting to question the significance of the inserted sequences.

The first clue was obtained for the BVDV Osloss insertion. A data library search revealed that the polypeptide encoded by the inserted sequence was almost identical with a well-known cellular protein, namely ubiquitin. The ubiquitin sequence of 76 amino acids is conserved for all animals for which the respective genes have been analyzed so far (Rechsteiner, 1987). The ubiquitin-like element from BVDV Osloss exhibits only two amino acid exchanges with respect to the animal ubiquitin sequence (Fig. 5) (Meyers *et al.*, 1989b, 1990). Because of the homology with the cellular ubiquitin sequence the BVDV Osloss insertion was designated "ub."

Equivalent analyses were then carried out for the BVDV NADL insertion. However, comparison of the respective sequence with the data available in different libraries did not result in identification of a

cellular counterpart. For further analysis, Northern hybridization experiments with insertion-specific probes were conducted. Three bands with sizes of about 2.5, 2.9, and 4.0 kb were detected in poly(A) RNA from bovine kidney cells (Meyers *et al.*, 1990). After cloning and sequencing, a cellular cDNA fragment was identified which contained a region almost identical to the NADL insertion (Fig. 6). Only two nucleotide exchanges were found, one of which leads to a change of the deduced amino acid sequence. Since the insertion in the genome of BVDV NADL apparently is also of cellular origin it was termed "cIns," for cellular insertion.

The findings outlined above indicated that cellular sequences had become integrated by some kind of recombination into the genomes of BVDV Osloss and NADL. Whereas the BVDV Osloss insertion is capable of encoding a complete ubiquitin protein, the cIns insertion probably represents an internal fragment of a bovine mRNA and codes for only part of a cellular protein. The extreme conservation of the cellular inserts suggests that, in terms of evolution, the recombination events occurred recently.

2. *A Concept for the Biological Significance of the Cellular Insertions*

As described above, cytopathogenic and noncytopathogenic isolates of BVDV are known. In addition to their different effects on tissue culture cells, the two biotypes can be distinguished by presence or

```
NADL:   agggt ATGTGCAGCCGATGCCAGGGAAAGCATAGGAGGTTTGAAATGGACCGGGAACCT   5046
            M  C  S  R  C  Q  G  K  H  R  R  F  E  M  D  R  E  P
cIns:   cgatg ATGTGCAGCCGATGCCAGGGAAAGCATAGGAGGTTTGAAATGGACCGGGAACCT

  5047  AAGAGTGCCAGATACTGTGCTGAGTGTAATAGGCTGCATCCTGCTGAGGAAGGTGACTTT   5106
        K  S  A  R  Y  C  A  E  C  N  R  L  H  P  A  E  E  G  D  F
        AAGAGTGCCAGATACTGTGCTGAGTGTAATAGGCTGCATCCTGCTGAGGAAGGTGACTTT
                          ▼
  5107  TGGGCAGAGTCGAGCATGTTGGGCCTCAAAATCACCTACTTTGCGCTGATGGATGGAAAG   5166
        W  A  E  S  S  M  L  G  L  K  I  T  Y  F  A  L  M  D  G  K
        TGGGCAGAGTCAAGCATGTTGGGCCTCAAAATCACCTACTTTGCGCTGATGGATGGAAAG
                    ▲
  5167  GTGTATGATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCAC   5226
        V  Y  D  I  T  E  W  A  G  C  Q  R  V  G  I  S  P  D  T  H
        GTGTATGATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCCCAGATACCCAC
                    ▼
  5227  AGAGTCCCTTGTCACATCTCATTTGGTTCACGGATG cctttcaggcaggaa           5283
        R  V  P  C  H  I  S  F  G  S  R  M
                 Y
        AGAGTCCCTTATCACATCTCATTTGGTTCACGGATG ccaggcaccagtggg
                 ▲
```

FIG. 6 Sequence comparison of the cIns insertion in the BVDV NADL genome and part of a bovine-cell-derived cDNA clone (cIns). The nucleotide sequence of the region identified as insertion in the viral RNA is given in capital letters and the deduced amino acid sequence is also shown. Exchanges between viral and cellular sequence are marked by triangles.

absence of NS3, the marker protein of cp BVD viruses. BVDV NADL and Osloss are both cp strains. Interestingly, the cellular insertions are located within the region of the NS2–3 gene where the cleavage generating NS3 is expected to occur. Thus, the inserted sequences might play a crucial role in the cp-specific generation of NS3. According to this theory, such insertions should only be present in the genomes of cp viruses. Indeed, the complete nucleotide sequence of the noncp BVDV strain SD-1 has been found to contain no insertion or other obvious difference with respect to the CSFV Alfort genome (Deng and Brock, 1992) (Fig. 4). In comparison with this sequence, the identified insertions are located between nucleotides 4992/4993 and 5152/5153 (BVDV NADL and Osloss, respectively).

In the process of MD development the generation of a cytopathogenic BVDV variant is regarded as causative for outbreak of the disease (Corapi *et al.*, 1988). Since MD occurs only in persistently infected animals, mostly at an age between 6 months and 2 years, an enormous number of replication cycles is apparently necessary to acquire a cytopathogenic phenotype. Therefore, a simple exchange of one or two nucleotides could hardly lead to cytopathogenicity. On this basis the following working hypotheses were put forward: (i) One possible mutation changing noncp BVDV into cp BVDV is a recombination between cellular and viral RNA. This process leads to integration of additional sequences into the NS2–3 gene. (ii) The inserted cellular element is responsible for generation of NS3. (iii) Fatal mucosal disease results from generation of such a cytopathogenic mutant.

3. First Experimental Evidence: Analysis of BVDV Pair #1

The working hypotheses were initially based on data obtained by comparing the genomes of two cp BVDV laboratory strains and one noncp CSFV isolate; at that time the analysis of noncp BVDV strains was missing. Support for the model was sought by analysis of cp and noncp viruses from defined BVDV pairs. BVDV NCP1 and CP1, a pair of viruses which had been isolated from an animal with MD (Corapi *et al.*, 1988), were chosen for these experiments. According to the hypotheses, one would expect a very close relationship between the cp and noncp viruses and a genomic change in the cytopathogenic isolate. After cDNA cloning and sequencing the region of the NCP1 genome coding for NS2–3 was found to be perfectly colinear with the CSFV Alfort sequence; no insertion or other discontinuity could be identified (Fig. 7) (Meyers *et al.*, 1991). In contrast, the CP1 genome contains a host-cell-derived insertion. As already described for BVDV Osloss, the inserted element is derived from a ubiquitin gene. The cellular in-

sertion of CP1 has a length of 366 nucleotides and codes for an amino-terminally truncated ubiquitin followed by one complete monomer of the same protein (Fig. 7). In contrast to BVDV Osloss, the amino acid sequence of this ubiquitin element is identical with the one of animal ubiquitin (Meyers *et al.*, 1991).

Remarkably, the host-cell-derived insert in the CP1 RNA is flanked by a duplication of 2384 b or bases of viral sequences encompassing the carboxy-terminal two-thirds of the NS2–3 gene. The duplication corresponds to nucleotides 5153 to 7537 of the published sequence of BVDV SD-1 (Deng and Brock, 1992) (Fig. 7, B and C_1, respectively). In total, the CP1 genomic RNA was found to be about 2.75 kb longer than that of NCP1. This size difference could also be demonstrated on Northern blots (Meyers *et al.*, 1991). Thus, the cytopathogenic virus CP1 has a genome organization differing from all pestiviruses analyzed up to that time.

According to the working hypotheses the genome of CP1 should not only exhibit a recombination-induced change with respect to the NCP1 RNA but also possess a high degree of sequence homology with the NCP1 sequence. Comparison of the virus-derived part of the determined CP1 and NCP1 sequences revealed 99.6% identity, which is at least 9% higher than the homology of this region with the one of other BVDV strains (Meyers *et al.*, 1991). Thus, CP1 certainly represents a mutant which was generated by recombination between NCP1 RNA and a cellular sequence.

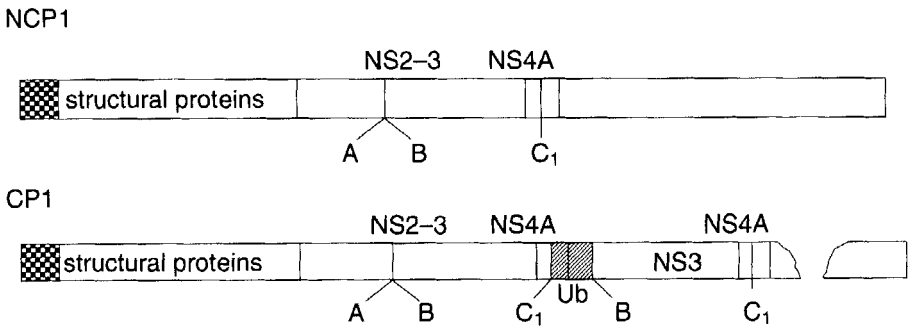


FIG. 7. Schematic presentation of the genome organization of BVDV NCP1 and CP1, the viruses from pair #1. The location of nonstructural genes NS2–3, NS3, and NS4A is indicated. C_1, last nucleotide of the viral sequence duplicated in the genome of CP1. The other elements are the same as in Fig. 4.

4. A Key Player: Ubiquitin

Analysis of pair #1 resulted in identification of the second cp BVDV isolate with a host-cell-derived insertion coding for ubiquitin. These two strains, however, represented only the starting point for detection of ubiquitin-coding sequences in BVDV genomes. In the meantime, using conventional cDNA cloning, RT-PCR, and Northern hybridization, a total of nine cp BVDV isolates have been found to contain ubiquitin-coding sequences in their genomes (Meyers *et al.*, 1989b, 1991, unpublished; Qi *et al.*, 1992; Tautz *et al.*, 1993). Among these viruses, eight belong to pairs of noncp and cp viruses isolated directly from diseased animals. For seven of these viruses the estimated genome size of more than 14 kb argues in favor of a genome organization with duplicated viral sequences corresponding to that of CP1. Besides BVDV Osloss and CP1, the genomes of four isolates were analyzed in more detail by partial cloning and sequencing of the genome. The characteristics of these viruses are presented in the following sections.

a. Pair #14 (Tautz *et al.*, 1993; Meyers, unpublished data). The genome of CP14, the cytopathogenic isolate of pair #14, was found to contain a ubiquitin-coding insertion of 498 nucleotides. This host-cell-derived sequence encodes one amino-terminally truncated and two complete monomers of ubiquitin (Fig. 8). The amino acid sequences of the virus-encoded ubiquitin proteins are identical with the animal ubiquitin sequence. The genome organization of CP14 is reminiscent of BVDV Osloss because the insertion is not flanked by duplicated viral sequences and is located at exactly the same position as found for the Osloss strain (Fig. 8, inserted between A and B). The homology between the CP14 sequence and corresponding regions of NCP14 is 99%.

b. Pair 190/VM (Qi *et al.*, 1992). The genome organization of BVDV 190, the cp virus of this pair, is very similar to that of BVDV CP1. A host-cell-derived insertion was found that codes for 16 carboxy-terminal amino acids of ubiquitin and one complete ubiquitin monomer. This cellular element is again flanked by a large duplication of viral sequences starting with nucleotide position 5153 and ending 2535 nucleotides downstream at position 7687 (Fig. 8, position B and C_2, respectively). Thus, the cellular insertion is 90 nucleotides shorter than that of CP1, while the duplication is about 150 residues longer. The ubiquitin sequence encoded by the genome of BVDV 190 exhibits three exchanges with respect to animal ubiquitin. Importantly, the 3′ recombination position with the 3′ end of a ubiquitin gene monomer fused to the viral sequence downstream of position 5153 is conserved with respect to BVDV Osloss, CP1, and CP14.

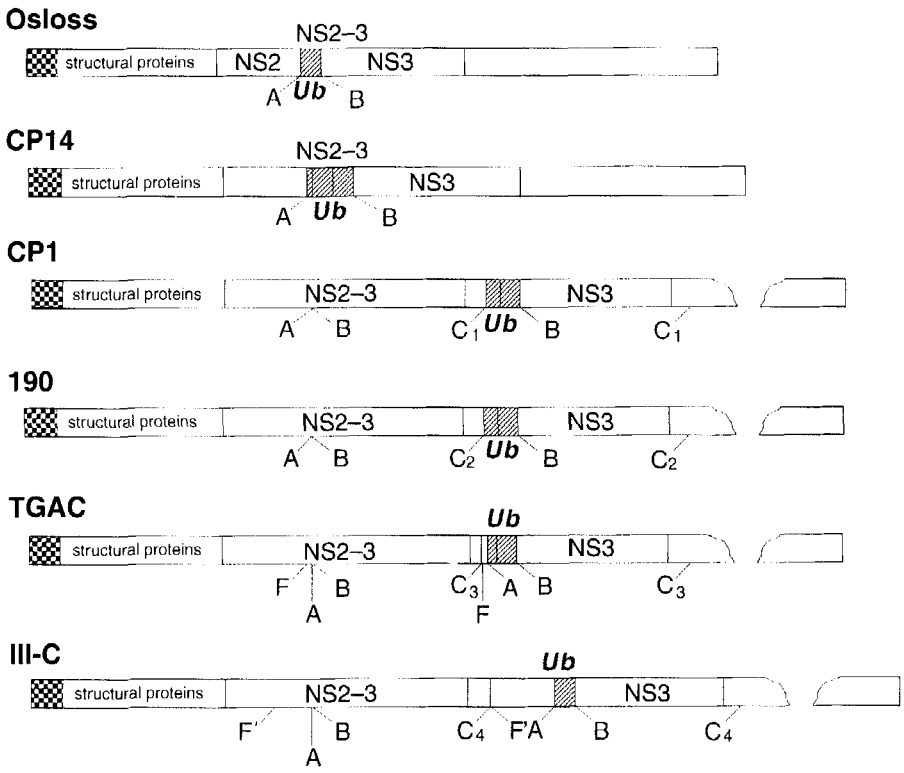


FIG. 8. Genome organization of different cp BVDV isolates with ubiquitin-coding insertions. C_1, C_2, C_3, C_4 indicate the last residues of the NS3 gene regions duplicated in the genomes of BVDV CP1, 190, TGAC, and III-C, respectively; F and F', first residues of the NS2-derived sequences duplicated in the genomes of BVDV TGAC and III-C. The other elements have already been introduced in Fig. 4. Note that the genomic positions A and B are the same as in Figs. 4 and 7.

c. Pair TGAC/TGAN (Qi *et al.*, 1992). In the case of the cp virus TGAC, again a ubiquitin-coding insertion flanked by duplicated viral sequences was found, and the 3′ recombination position is identical with those of other viruses with similar insertions (Fig. 8, position B). At the amino terminus, the inserted element is two codons shorter than that of BVDV 190. The duplication extends to position 7483 (C_3 in Fig. 8) and thus has a length of only 2331 nucleotides. According to our sequence comparison studies, the genome organization of BVDV TGAC shows one very interesting difference when compared with isolates CP1 and 190. Between the duplicated viral sequence located upstream

of the cellular insertion and the ubiquitin-coding element, a stretch of eight codons is present which is neither homologous to the viral genome downstream of position 7483 nor part of the cellular sequence (F to A in Fig. 8). In our analyses these eight codons were found to correspond to positions 5129–5152 of the viral genome. Since the viral sequence downstream of the cellular insertion starts again at position 5153 of a regular genome, the ubiquitin-coding element has actually been inserted between formerly neighboring nucleotides, which is similar to the situation found for BVDV Osloss and CP14 (Fig. 8). Because of the unusual genome organization it is likely that BVDV TGAC represents a secondary recombinant in which the 5′ half of the genome was derived from a noncp virus and the 3′ half starting eight codons before the cellular insertion was derived from a cp virus similar to BVDV Osloss or CP14 (see also Section III,F). Qi and co-workers identified only 89% homology between the first and the second version of the duplicated sequences in the TGAC genome and even lower similarity between the second version of the duplication and the respective sequence in the genome of TGAN, the ncp virus of this pair. This finding argues for our hypothesis outlined above and, moreover, indicates that the two partners for the secondary recombination probably represent different BVD viruses.

d. Pair III-C/III-NC (Qi *et al.*, 1992). The genome organization of cp BVDV III-C is very similar to that of BVDV 190. The 5′ half of the genome until position 7530 (position C_4 in Fig. 8) is typical for a noncp virus, as it does not contain an insertion in the NS2–3 gene. Downstream of position 7530 is found a sequence derived from a cp BVDV genome similar to that of BVDV Osloss or CP14. In the case of BVDV III-C this cp-BVDV-derived part of the genome starts already at a position corresponding to nucleotide 4535 of a regular genome. Thus, almost the entire NS2–3 gene is duplicated (position F′ to C_4 in Fig. 8). Within this second NS2–3 gene a ubiquitin-coding element is inserted. The insertion position corresponds again to residues 5152/5153 of a noncp BVDV genome. The ubiquitin insertion is very similar to the one found for BVDV Osloss. The 5′ and 3′ ends of the insertion correspond exactly to the ends of one ubiquitin monomer. Surprisingly, sequence analysis in our lab revealed that the ubiquitin-coding sequence of BVDV III-C has a length of 80 instead of 76 codons. This peculiarity is due to a duplication of codons 48–51 of the ubiquitin gene. In contrast to BVDV TGAC, 99.6% homology was found when the duplicated sequences in the III-C genome were compared with one another. Thus, the genomes of the noncp and cp viruses that probably recombined to generate BVDV III-C were derived from a common ancestor.

We intended to perform a more detailed analysis of the partial sequences of BVDV III-C, TGAC, and 190. However, it was not possible to assign the sequences deposited in the gene bank under accession numbers M87805–M87815 to the various virus isolates. Moreover, the evaluation of the sequence data was hampered by the fact that none of the respective sequences obtained from our version of the GeneBank contained a continuous open reading frame.

In all cases described above, the identified genome rearrangements are specific for the cp viruses. The noncp isolates of the respective pairs contain neither cellular insertions nor duplications.

e. Comparison of cp BVDV Strains with Ubiquitin-Coding Insertions. Up to now, nine cp BVDV isolates have been found to contain ubiquitin insertions varying in length from 228 nucleotides (exactly one ubiquitin monomer) to 498 residues. According to their genome organization these viruses can be divided into three groups. The first group has integrated the cellular sequence between formerly neighboring nucleotides (i.e., Osloss and CP14) and will be designated "Osloss type." The genome of the second group, termed "CP1 type," contains large duplications of variable length flanking the insertions (i.e., CP1, 190). The third group resembles the Osloss type, since the insertion is integrated between formerly neighboring nucleotides. However, the sequences flanking the insertions are duplicated as in the CP1 type. This third group comprises TGAC and IIIC and is named "TGAC type."

Despite prominent differences, some important features are common to all genomes with ubiquitin-coding insertions. All of the host-cell-derived sequences determined so far end with the last codon of a ubiquitin gene (Fig. 9). Thus, the differences in insert size can be regarded as variation of the 5′ position at which viral and cellular sequences have been fused. The situation is similar for the flanking viral sequences. In all cases, the nucleotides following the insertion correspond to the residues starting with position 5153 in a BVDV genome without insertion (B in Fig. 8), in the numbering of the sequence of BVDV SD-1 (Deng and Brook, 1992). However, the sequence preceding the cellular insertion in the viral genome can vary. It is conserved for the genomes of the Osloss type as well as isolates TGAC and IIIC (position 5152, A in Fig. 8), but it may differ in cases where duplication has occurred (C_1 and C_2 in Fig. 8). In conclusion, the arrangement of cellular and viral sequences seems to be fixed at the 3′ border but flexible at the 5′ border (Fig. 9).

One major difference between the three BVDV groups with ubiquitin-coding insertions is in the expression of NS3. It was initially believed that NS3 always represents a cleavage product of NS2–3 (Fig. 10). This

	5′ crossing over site	3′crossing over site
CP14:	...DLEHLGWVLR **KESTLHLVLR**...	...**LHLVLRLRGG** GPAVCKKITN...
Osloss:	...N******I*K **MQIFVKTLTG**...	...**********S *********E...
CP1:	...NV*AAKGYVH **QDKEGIPPDQ**...	...********** *********E...
190:	...T*VIKWLAFG **VQKEST*HPV**...	...********** *********E...
TGAC:	...RI******** **********...	...********** *********E...
III-C:	...********** **MQIFVKTLTG**...	...********** *********E...

FIG. 9. Comparison of the amino acid sequences deduced from the regions flanking the 5′ and 3′ crossing over sites in the genomes of various cpBVDV isolates with ubiquitin-coding insertions. For BVDV CP14 the complete sequence is given using the one-letter code, while for the other viruses only residues differing from this sequence are specified. Asterisks indicate residues identical with the CP14 sequence. Ubiquitin sequences are shown in bold.

conclusion was mainly based on the antigenic relationship between NS2–3 and NS3 (Purchio *et al.*, 1984; Donis and Dubovi, 1987a,b; Pocock *et al.*, 1987; Corapi *et al.*, 1988). Because of the lack of specific antibodies the second processing product could not be identified at that time. Antisera raised against bacterial fusion proteins or synthetic peptides led to the demonstration of both processing products of NS2–3, namely NS2 and NS3, for one BVDV strain (Collett *et al.*, 1988b, 1991). Protein analyses revealed that NS2–3 cleavage is also responsible for NS3 generation in cells infected with BVDV Osloss and CP14 (Meyers *et al.*, 1991; Tautz *et al.*, 1993). However, in equivalent studies with BVDV CP1, NS2 could not be demonstrated even though the respective anti-NS2 serum reacted with NS2–3 (Meyers *et al.*, 1991). Thus NS3, which is clearly detectable in cells infected with BVDV CP1, must be generated by a mechanism different from cleavage of NS2–3. In this context the genome organization of BVDV CP1 and other isolates of this type should be considered. Starting at the 5′ end, the respective genomes are colinear with a genome without insertion up to a region corresponding to position 7537 (BVDV CP1) to 7687 (BVDV 190). According to current estimations on genome organization of pestiviruses, these positions are all located downstream of the 3′ end of the NS2–3 gene (Collett *et al.*, 1991; Tautz *et al.*, unpublished). Following the ubiquitin-coding insertions the viral sequence starts again with codon 1590, which is located at position 5153–5155 of a noncp BVDV genome. Codon 1590 fits very well with the assumed 5′ end of the NS3 gene. Thus, the genome of these viruses contains a complete NS2–3 gene upstream of the host-cell-derived insertion, and an NS3 gene down-

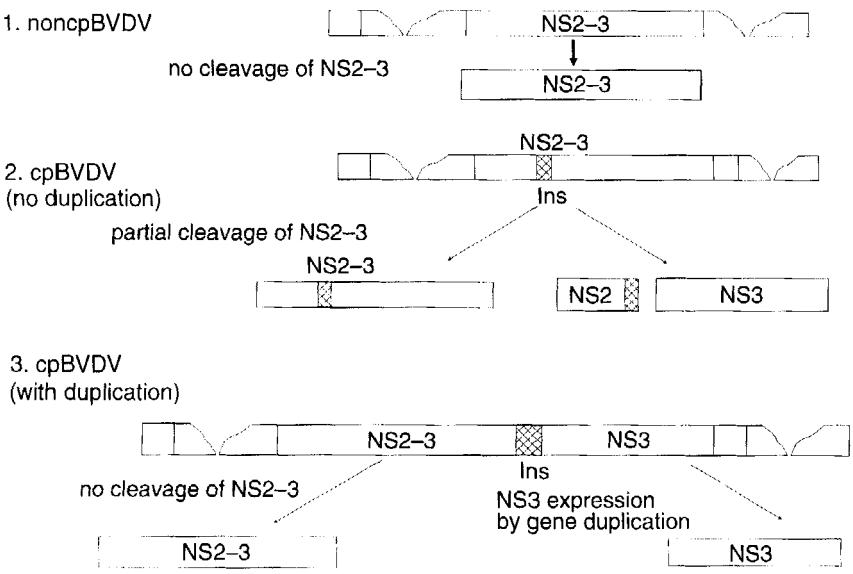


FIG. 10. Schematic drawing showing the generation of NS2–3 or NS2–3 and NS3 for noncp BVDV and cp BVDV, respectively. In the case of cp BVDV, the different processing schemes for viruses with and without duplication are indicated. Ins stands for inserted cellular sequence.

stream of the ubiquitin-coding sequence. The most likely interpretation of the protein data is that expression of NS2–3 occurs from the sequence preceding the insertion, while NS3 is translated from the duplicated region located downstream of the insertion (Fig. 10) (Meyers *et al.*, 1991).

BVDV III-C and TGAC again differ from the other viruses with ubiquitin insertions. As outlined above, the genomes of these viruses were presumably generated by nonhomologous recombination between a genome of the Osloss type and that of a BVD virus without insertion in the NS2–3 gene, most likely a noncp BVDV. The genomes of these two viruses contain two NS2–3 genes. With respect to the genomic 5′ end, the first NS2–3 gene codes for a complete NS2–3 protein (Fig. 10); the second version of the gene is 5′ terminally truncated and contains the ubiquitin-coding insertion. Even though data on BVDV III-C polyprotein processing are not available, the following model can be suggested: NS2–3 of BVDV III-C and TGAC is derived from the upstream version of the gene, while NS3 results from processing of the protein encoded by the downstream gene. The cleavage also gives rise

to an amino-terminally truncated NS2 protein, which for BVDV TGAC consists of only eight amino acids, while it has a length of 205 residues in the case of BVDV III-C. Because of the sequence rearrangement resulting from recombination, these truncated NS2-polypeptides are probably fused to proteins located upstream in the polyprotein.

With respect to expression of NS3, BVDV TGAC and III-C have to be regarded as intermediates between the Osloss group and the CP1 group since they generate NS3 by processing of NS2–3, which is encoded by a duplicated sequence.

f. The Possible Function of the Ubiquitin Insertions. It has been hypothesized that the insertion of cellular sequences into BVDV genomic RNA is responsible for the cytopathogenicity of the resulting recombinants. Along these lines the expression of the cp BVDV marker protein NS3 should be directly linked to the respective genome rearrangements. Discussion of the putative function of ubiquitin integrated into the polyproteins of the respective BVD viruses requires some background information on cellular ubiquitin and its genes.

Ubiquitin is a highly conserved protein consisting of 76 amino acids. The posttranslational attachment of ubiquitin to cellular proteins has been implicated in a variety of different functions (Rechsteiner, 1987; Hershko and Ciechanover, 1992). The best-characterized role of ubiquitin concerns protein catabolism, where attachment of ubiquitin tags proteins for selective breakdown (reviewed by Hershko and Ciechanover, 1992; Jentsch, 1992; Johnson *et al.*, 1992).

Two types of ubiquitin genes with different putative functions have been described for eukaryotic cells. Under normal growth conditions ubiquitin expression occurs from genes coding for a monomer of ubiquitin fused to carboxy-terminal extensions of 52–80 amino acids. These extensions represent ribosomal proteins (Finley *et al.*, 1989; Redman and Rechsteiner, 1989) (Fig. 11). Under stress conditions, enhanced expression of the other form of ubiquitin genes is observed. These genes code for polyubiquitin, consisting of multiple ubiquitin monomers arranged in a head-to-tail array (Finley *et al.*, 1987) (Fig. 11). The specific function of these genes seems to be rapid elevation of intracellular ubiquitin levels. Importantly, in both cases the primary product expressed from ubiquitin genes is not monomeric ubiquitin but a fusion protein of ubiquitin and a second component; in the case of polyubiquitin genes this second component is provided by additional ubiquitin monomers. The organization of ubiquitin genes implies that the monomeric form of this protein has to be generated by co- or posttranslational proteolytic processing (Fig. 11). The respective cleavage reaction(s) are mediated by a family of cellular proteases called ubiqui-

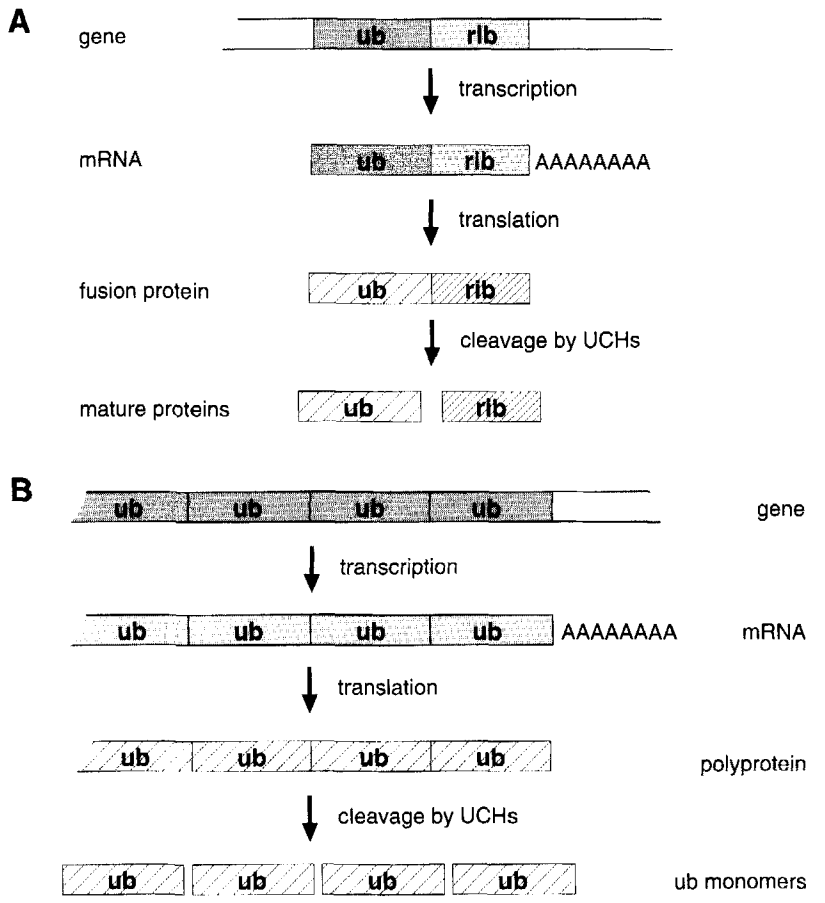


FIG. 11. Generation of monomeric ubiquitin by proteolytic processing of precursors derived from genes encoding fusions of ubiquitin (ub) with ribosomal proteins (rib) (A) or from polyubiquitin genes (B). UCHs, ubiquitin carboxy-terminal hydrolases.

tin carboxy-terminal hydrolases (UCHs) (Rose, 1988; Mayer and Wilkinson, 1989; Jonnalagadda *et al.*, 1989; Baker *et al.*, 1992). Cleavage occurs after glycine-76, the carboxy-terminal residue of ubiquitin. The requirements for cleavage of natural and recombinant ubiquitin fusion proteins have been analyzed *in vitro* and *in vivo* (Agell *et al.*, 1988; Butt *et al.*, 1988; Monia *et al.*, 1989; Tobias and Varshavsky, 1991; Baker et al., 1992). All of the tested substrates were cleaved. Exchanges of the amino acid following glycine-76 of ubiquitin in a recombinant fusion protein showed that only the presence of a proline residue at the P1′

position dramatically reduced cleavage efficiency (Butt *et al.*, 1988).

The cp phenotype of BVDV is correlated with expression of NS3. As outlined above, this protein is either generated by processing of NS2–3 (Osloss type) or expressed from a duplicated region of the genome (CP1 type). In both cases the expression of NS3 is dependent on generation of a protease cleavage site at its amino terminus. Within the BVDV polyprotein the ubiquitin could actually provide a processing signal recognized as substrate by the cellular UCHs. In consequence, cleavage would occur between residue 76 of ubiquitin and the following amino acid, which in all cases is a glycine. This model fits with the data on polyprotein processing of BVDV Osloss. Immunoprecipitation with antisera directed against the cellular sequence proved that the ubiquitin moiety is part of NS2 and not of NS3 (Meyers *et al.*, 1991).

Final evidence for the crucial role of ubiquitin as processing signal was obtained by *in vitro* translation assays based on cDNA constructs derived from BVDV CP14 (Tautz *et al.*, 1993). Different parts of the NS2–3 gene of this virus were cloned in frame in a plasmid providing an SP6 RNA polymerase promoter and a translation initiation codon. Translation of *in vitro* transcribed RNA containing the ubiquitin-coding sequences of BVDV CP14 in rabbit reticulocyte lysate did not result in full-length polypeptides. Instead, defined processing products were observed. The requirements for the respective cleavage reactions were determined by translation of constructs into which site-specific mutations had been introduced. The results of these experiments can be summarized as follows: (i) The ubiquitin-coding sequence of BVDV CP14 is necessary for the observed processing. (ii) In order to obtain cleavage the substrate has to contain at least one complete monomer of ubiquitin. (iii) The BVDV NS3 protease is not involved in the reaction, since deletion of the active center of the enzyme has no influence on processing. (iv) The active protease has properties typical for cellular ubiquitin C-terminal hydrolases, which are characterized by dependence on a substrate with a complete monomer of ubiquitin, cleavage after the C-terminal glycine-76 of ubiquitin, and reduction of cleavage of substrates with a proline following glycine-76 (Tautz *et al.*, 1993). In conclusion, the experimental data obtained with BVDV CP14 constructs argue for the following processing scheme: the presence of two complete monomers of ubiquitin preceded by 14 carboxy-terminal amino acids of a third monomer results in three UCH-cleavable bonds. The three cuts release not only NS3 but also the two complete ubiquitin proteins (Fig. 12). This processing scheme also offers an explanation for our failure to detect NS2 of CP14 with antisera directed against ubiquitin. The 14 amino acids of the truncated monomer repre-

sent the only cellular moiety of the fusion protein; this short stretch may not be recognized by the anti-ubiquitin antibodies.

The processing scheme determined for BVDV CP14 most likely applies for all BVD viruses with ubiquitin-coding insertions. In experiments similar to those described above for CP14, *in vitro* cleavage of translation products was also demonstrated for BVDV Osloss (Tautz *et al.*, 1993). This approach was particularly interesting since the ubiquitin-coding element of strain Osloss contains two mutations resulting in changes of the amino acid sequence. Moreover, one of these mutations concerns glycine-76, which is changed into serine. The *in vitro* data showed that cleavage of the mutated protein occurs less efficiently when compared to CP14 (Tautz *et al.*, 1993). Since exactly one complete monomer of ubiquitin is present the cellular sequence is not released but stays fused to NS2 (Fig. 12) (Meyers *et al.*, 1991).

Processing of the respective polyprotein region encoded by BVDV CP1 has also been analyzed. The CP1 genome encodes one amino-

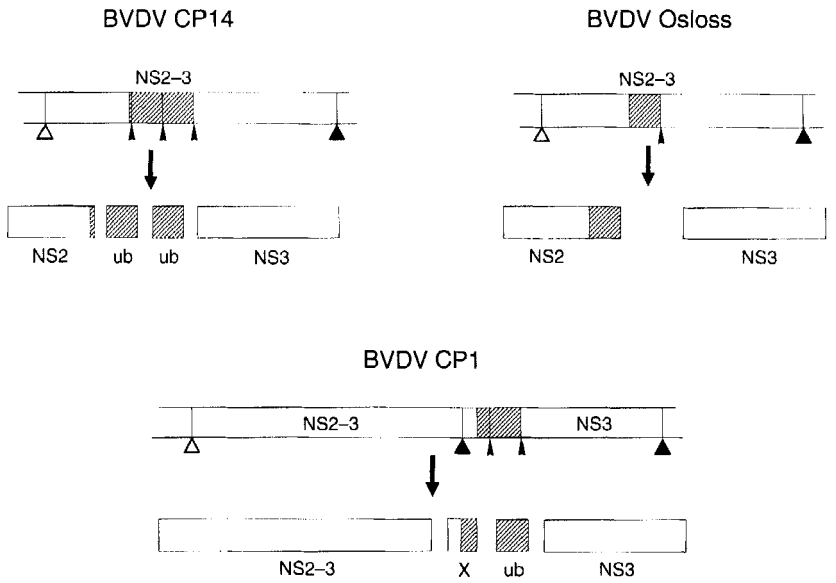


FIG. 12. Processing schemes for different BVDV polyproteins containing ubiquitin sequences. Assumed cleavage by cellular signalase is indicated by open triangles; filled triangles indicate cleavage by the NS3 protease itself. Arrowheads point at positions where processing by cellular ubiquitin–carboxy-terminal hydrolases occurs (see text). The polypeptide designated X represents a fusion protein composed of viral sequences (probably derived from NS4A) and a truncated ubiquitin protein.

terminally truncated and one complete monomer of ubiquitin. Since no amino acid exchanges with respect to animal ubiquitin were found, processing most likely leads to release of the complete monomer of ubiquitin, while the ubiquitin fragment should be part of a fusion protein together with a carboxy-terminally truncated NS4A protein (X in Fig. 12). In the case of BVDV 190, TGAC, and III-C, predictions concerning the processing products can hardly be made, as all three published sequences exhibit mutations within the encoded host-cell-derived polypeptides. However, since these three viruses supposedly express NS3, at least cleavage at the carboxy-terminal end of the complete ubiquitin monomer preceding NS3 should occur.

In summary, the integration of ubiquitin-coding sequences into BVDV genomes leads to an additional protease cleavage site in the viral polyprotein. The respective sequence is cut by a cellular protease with properties of known ubiquitin C-terminal hydrolases. The cleavage results in release of the N terminus of the cp BVDV marker protein NS3.

g. The Prevalence of cp BVD Viruses with Ubiquitin-Coding Insertions. Integration of cellular ubiquitin-coding sequences at a certain position of the NS2–3 gene of BVDV apparently represents a frequently used way to generate a cytopathogenic BVD virus. Among the 28 cp BVDV isolates analyzed so far, 9 contain ubiquitin-coding insertions. The reason for the prevalence of this specific type of genome modification is still a matter of speculation. Some interesting viewpoints may be considered. Ubiquitin provides a protease cleavage site which most likely can be cut in all cells of the respective host, since the ubiquitin pathway is regarded as general feature of all eukaryotic cells. Cleavage after the carboxy-terminal glycine of ubiquitin by UCHs represents an efficient reaction which is almost independent of the downstream sequence. Thus, integration of ubiquitin into a completely new context seems to have little if any influence on the processing reaction. Ubiquitin-coding sequences belong to a group of housekeeping genes with very high cellular mRNA levels; under stress conditions like heat shock and probably also viral infection the amount of these mRNAs is even raised above the standard level (Hershko and Ciechanover, 1992). Ubiquitin-coding sequences therefore represent abundantly available partners for recombination within a virus-infected cell.

5. A Second Player: cIns

As mentioned above, ubiquitin-coding sequences represent the most abundant host-cell-derived insertion in cp BVDV genomes. So far, only one other cellular sequence has been identified in BVDV RNA, which is called cIns. The cIns insertion was first detected in RNA from BVDV

NADL (Meyers *et al.*, 1990). It initially seemed that BVDV NADL was the only strain with the respective sequence. However, six additional cp BVDV strains containing cIns sequences have been identified by Northern hybridization and RT-PCR assays (Meyers *et al.*, unpublished; Tijssen, personal communication). Since these viruses have not been further characterized yet, discussion of cIns and its putative role in BVDV polyprotein processing is preliminary and mainly based on our knowledge about BVDV NADL.

The cIns insertion in the genome of BVDV NADL is at first glance comparable to the ubiquitin-coding sequences in the RNA of BVDV Osloss and CP14. In all three cases cellular sequences of similar sizes are integrated into the NS2–3 gene. However, a more thorough comparison elucidates clear differences. The cIns insertion is located 54 codons upstream of the position where ubiquitin-coding insertions are found (Fig. 4). As outlined above, all available data on viruses with ubiquitin-coding insertions indicate that the amino terminus of NS3 is located in the polyprotein directly downstream of ubiquitin. Cleavage of NS2–3 from BVDV NADL directly downstream of the cellular insert would result in an NS3 protein about 5 kDa larger than that of BVDV Osloss. However, according to gel electrophoretic analyses, NS3 proteins of different cp BVDV isolates, including BVDV NADL, have the same sizes. It is unlikely that a difference of 5 kDa would not be noticed on the respective gels (Greiser-Wilke *et al.*, 1992; Thiel *et al.*, unpublished observation). It therefore appears that the cleavage of NS2–3 of BVDV NADL does not occur in the immediate neighborhood of the cIns insertion but some 50 amino acids further downstream. Accordingly, processing of NS2–3 should occur by a mechanism completely different from that observed for BVDV strains with ubiquitin-coding insertions. Furthermore, preliminary data on the genome organization of other cp BVDV isolates with cIns insertions indicate that the inserted sequences are not necessarily located at the same position as found for BVDV NADL (Meyers *et al.*, unpublished observation). Future work will aim at further elucidation of the genome organization of these viruses and the mechanism of NS2–3 processing induced by cIns insertions.

One approach aimed at understanding the function of cIns insertions is to learn more about the cellular counterpart. So far, few data are available on the respective gene(s) and protein(s). Interestingly, three different bands homologous to the inserted sequence have been identified in Northern hybridization experiments with RNA from bovine (MDBK) cells (Meyers *et al.*, 1990). The significance of this finding is not known. Since complete cellular RNA was analyzed, the three

forms might represent splice intermediates or differentially spliced products derived from one gene. Alternatively, different genes sharing the cIns sequence could exist. One important result of the Northern hybridization experiments is that only small amounts of cIns mRNAs were detected, in particular when compared to ubiquitin mRNAs. If this finding also applies to cells in the animal usually infected by BVDV, recombination of viral RNA with cIns sequences would be expected to be unlikely because of low concentration of the cellular RNA. Nevertheless, 7 out of 28 cp BVDV isolates have so far been found to contain cIns sequences.

At present nothing is known about structure and function of the product(s) derived from cIns gene(s). The sequence inserted in the BVDV NADL genome was found to be highly homologous to a cDNA fragment derived from a bovine mRNA (Fig. 6) (Meyers *et al.*, 1990). According to this partial mRNA sequence, the cIns insertion represents an internal fragment of the mRNA. Interestingly, a GeneBank search led to identification of a human sequence which in a short region exhibits 98% homology with the bovine mRNA. This high degree of phylogenetic conservation argues in favor of an important role of the respective polypeptide. However, all approaches aiming at identification of cIns gene products with specific antisera have failed so far. Elucidation of the structure of cIns genes and further attempts toward identification and analysis of the respective gene products are part of future plans.

B. Viral Gene Autarky: cp BVDV without Cellular Insertions

1. The Giants: Double Duplication and Rearrangement

To learn more about the different ways leading to a cytopathogenic BVD virus, we specifically looked for viral genomes with cp-specific mutations differing from the already known ubiquitin-coding or cIns insertions. New cp BVDV isolates were screened by Northern hybridization with probes specific for the two known cellular inserts, and viruses which were double-negative in this assay were chosen for further analysis. Among these isolates several showed drastically enlarged genomes, while the RNAs of the corresponding noncp viruses had a size of 12.3 kb. Two of them, which are termed BVDV Pe515CP and BVDV CP6, were analyzed by partial cDNA cloning and sequencing (Meyers *et al.*, 1992). Both Pe515CP and CP6 belong to virus pairs isolated from animals that had died of MD (Brownlie *et al.*, 1984; Corapi *et al.*, 1988).

a. Genome Structure of Pe515CP and CP6. On Northern blots the genomic RNAs of BVDV Pe515CP and CP6 both appeared to be more

than 2 kb larger than those of the respective noncp viruses. This situation is reminiscent of viruses like CP1, 190, TGAC, and III-C, for which large duplications of viral sequences were identified. After cloning and sequencing this was also found to be the case for Pe515CP and CP6. The determined sequence of Pe515CP is colinear with the genome of noncp BVDV SD-1 until position 7456 (C_5 in Fig. 13). Downstream therefrom, an insertion of 462 nucleotides is present, which is followed by a region corresponding to the NS3 gene. The genome organization of CP6 is very similar to that of Pe515CP. However, the insertion is found far downstream of the position identified for Pe515CP and the other viruses with duplications. The CP6 genome is colinear with the SD-1 sequence up to position 8788 (C_6 in Fig. 13) which, according to our present knowledge, lies in the region of the 5′ end of the NS5 gene. The sequence inserted after position 8788 has a length of 507 nucleotides. Downstream of this inserted element again a sequence coding for NS3 follows (starting with B in Fig. 13). Analysis of the genome of Pe515NCP, the noncp counterpart to Pe515CP, revealed that the identified rearrangements were specific for the cp isolate (Meyers *et al.*, 1992).

Taken together, the genome organization of BVDV Pe515CP and CP6 is very similar to that of CP1 and 190, with an insertion of several hundred nucleotides flanked by large duplications of viral RNA. The major difference concerns the inserted sequence, which for Pe515CP and CP6 is not a host-cell-derived element but represents a BVDV sequence derived from the 5′ region of the genome. In the case of Pe515CP the insertion corresponds to nucleotides 428 to 889 (Fig. 13, positions D and E, respectively) while for CP6 the respective sequence can be aligned with nucleotides 383 to 889 of the BVDV SD-1 genome (Fig. 13, positions D′ and E, respectively). The inserted sequences code for the pestiviral protease N^{pro}. Interestingly, the CP6 insertion starts three nucleotides upstream of the translation initiation codon of the large open reading frame, while in the case of Pe515CP the inserted sequence begins with codon 15. Further analysis revealed that the inserted sequences indeed represent duplications. For each virus one version is located at the original position close to the genomic 5′ end, while the other is found in the new context preceding the NS3 gene. For reasons of simplification we denoted the duplicated N^{pro} sequences "SD," for small duplication, and the second duplication encompassing the NS3-coding sequence "LD," for large duplication (Fig. 13) (Meyers *et al.*, 1992).

A particularly important finding was that the sequence following the SD-encoded N^{pro} in the polyproteins of BVDV Pe515CP and CP6 starts again with glycine-1590, which also follows ubiquitin in the polypro-

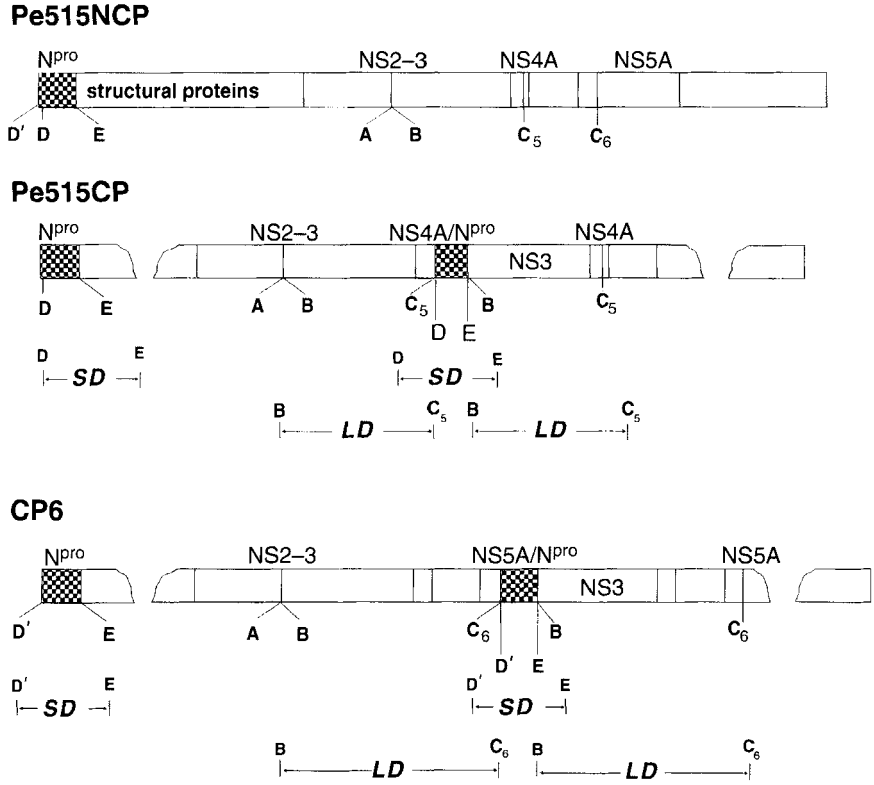


FIG. 13. Genome organization of BVDV Pe515CP and CP6. C_5 and C_6 represent the last residues of the large duplications (LD). D and D′ represent the 5′ ends and E the 3′ ends of the small duplications (SD). A and B designate the same positions as in Figs. 4, 5, and 7.

teins of Osloss, CP1, and the other viruses with ubiquitin sequences. Thus, with respect to their 3′ ends the different insertions are found at exactly the same genomic position (B in Fig. 13) (Meyers *et al.*, 1992).

b. Additional Viruses with Genomes of the Double Duplication Type. Our screening of cp BVDV isolates by Northern hybridization with pestivirus-specific probes resulted in identification of eight viruses with genome sizes of more than 14 kb, including CP1, Pe515CP, and CP6. Based on the data obtained by detailed analysis of isolates like CP1, TGAC, III-C, CP6, and Pe515CP, the genomes of all these isolates probably contain duplications of viral sequences. After hybridization with the respective probes, four of eight strains were found to contain ubiquitin sequences while none contained cIns-homologous RNA.

Thus, in addition to Pe515CP and CP6, two viruses with enlarged genomes were found which could have genome rearrangements without cellular insertions. Since genomes of the Pe515CP/CP6 type contain only viral sequences, an easy way for identification of additional members of this group is not available. As a first approach, we hybridized viral RNA against an oligonucleotide bridging the conserved 3′ recombination position composed of the 3′ end of the N^{pro} gene and the 5′ end of the NS3-coding sequence. Stable binding of such a probe should only occur with BVDV genomes of the respective type. In Northern hybridization experiments this probe specifically recognized BVDV CP11 (Corapi *et al.*, 1988), one of the two candidates mentioned above. Since the corresponding noncp BVDV isolate NCP11 was negative in this assay, CP11 can be regarded as member of the group with N^{pro} duplication (Meyers *et al.*, unpublished observation). BVDV CP4, the second putative candidate, was negative in the hybridization assay, and was therefore analyzed by cDNA cloning and sequencing. The genome organization of BVDV CP4 was found to be very similar to that of Pe515CP and CP6. Again a duplicated N^{pro} gene (SD) was found inserted between large duplications of viral sequences (LD). In the case of CP4, the SD has a length of 489 nucleotides, starting with codon 6 (position 401) and ending again with nucleotide 889. Once more, the LD starts with codon 1590 (position 5153). For CP4, the end of LD was located at position 8491 with respect to BVDV SD-1 (Fig. 14). The failure to recognize BVDV CP4 with the oligonucleotide probe is probably due to a lower degree of sequence homology (Meyers *et al.*, unpublished data).

Among all analyzed pestiviruses BVDV CP6 has by far the largest genome. Adding up the LD with a length of 3636 nucleotides and the SD, which comprises 507 nucleotides, the RNA of BVDV CP6 is 4143 nucleotides longer that that of a noncp BVD virus. The second largest BVDV RNA was also generated through double duplication and rearrangement: the genome of BVDV CP4 contains 3828 additional residues. It is interesting that genome enlargements of about one-third do not influence the viability or growth characteristics of these viruses in an obvious way.

c. Protein Processing Scheme for Viruses with Duplicated N^{pro} Gene. A more detailed analysis of the sequences determined for Pe515CP, CP6, and CP4 reveals some highly interesting similarities among them and viruses with ubiquitin-coding insertions. As already mentioned above, the 5′ end of SD varies for the three viruses with duplication of the N^{pro} sequences (D, D′, and D″ in Figs. 13 and 14). In contrast, the SD insertions end with the same sequence, which corresponds to codon 168 of the large pestiviral ORF (E in Figs. 13 and 14). Experiments

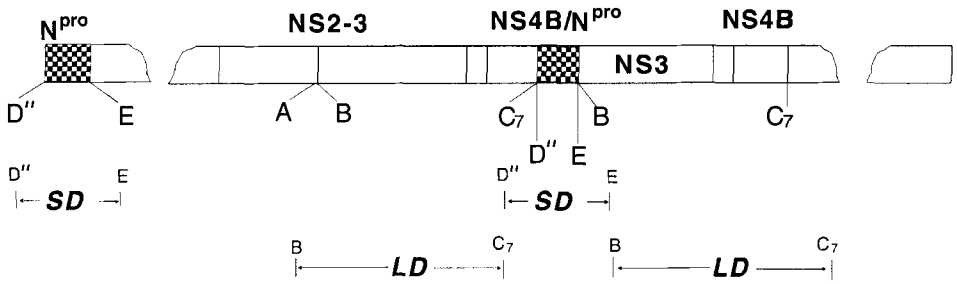


FIG. 14. Genome organization of BVDV CP4. The designation of genomic positions is as in Fig. 13 (A, B, and E) or equivalent to Fig. 13 (C_7 and D'').

with CSFV have revealed that this codon represents the 3′ end of the N^{pro} gene (Stark *et al.*, 1993). N^{pro} was found to cleave at its own carboxy terminus, thereby generating the amino terminus of the following capsid protein C (Wiskerchen *et al.*, 1991; Stark *et al.*, 1993). The amino acids flanking the cleavage site are highly conserved among pestiviruses (see Section II,C). The NS3 sequence following the SD-encoded N^{pro} differs from that in the original context; nevertheless, it is tempting to speculate that N^{pro} could be active even as an internal part of a polyprotein and cleave at the altered site. In radioimmunoprecipitation experiments with an antiserum specific for N^{pro}, a protein of 28 kDa was identified in BVDV Pe515CP-infected cells which was not present in cells infected with the noncytopathogenic virus Pe515NCP. The same polypeptide was recognized by antisera directed against NS4A and again this reaction was cp-specific. Thus, the 28-kDa protein represents a fusion protein composed of NS4A and N^{pro} sequences (Meyers *et al.*, 1992). According to estimations about the location of cleavage sites in the pestiviral polyprotein, the molecular weight of 28 kDa fits well with the assumption that cleavage occurs at the carboxy terminus of the SD-encoded N^{pro}. After *in vitro* translation of pestiviral RNA, efficient cleavage between N^{pro} and C was detected (Stark *et al.*, 1993). There is evidence that variation of the amino acid at the P1′ position of the N^{pro} cleavage site can be tolerated (Stark *et al.*, 1993). It therefore seems possible that in the new context N^{pro} is still active. Indeed, cleavage of polypeptides derived from cDNA constructs corresponding to the SD/LD region of the Pe515CP genome has been observed *in vitro* (Meyers *et al.*, 1992). The proposed processing scheme for BVDV Pe515CP therefore includes autoproteolytic cleavage at the carboxy terminus of the duplicated N^{pro} encoded by SD. This reaction

generates the carboxy terminus of the 28 kDa polypeptide and the amino terminus of the downstream protein (Fig. 15). In analogy to CP1 and other viruses with duplications of NS2–3-coding sequences it is likely that upstream sequences code for NS2–3 while NS3 is derived from the LD downstream of SD. NS2–3 and NS3 of BVDV CP6 could be generated in a similar way, even though a hypothetical fusion protein composed of N^{pro} and a very small portion of NS5A could not be demonstrated (Meyers *et al.*, 1992).

The processing schemes outlined above for cp BVDV with ubiquitin insertions and for the isolates of the double duplication type suggest that NS3 encoded by all these viruses has the same amino terminus, which at the genome level corresponds to position B in the schematic presentations (Figs. 4, 7, 8, 13, and 14).

2. The Dwarfs: Deletion Leads to cp BVDV

For two cp BVD viruses, namely CP9 and CP13, two pestivirus-specific bands were observed during hybridization screening. One of these bands had a size of about 12.3 kb and comigrated with the genomic RNA of the corresponding noncp BVD virus. The other RNA, however, was only about 8 kb in size (Tautz *et al.*, 1994; Kupfermann *et al.*, submitted). This finding was surprising, as pestivirus gene expression does not rely on subgenomic mRNAs. Based on data obtained for other RNA viruses, the most plausible explanation for the subgenomic RNAs is that they represent genomes of defective interfering particles (DI) present in the virus preparation. DIs are commonly

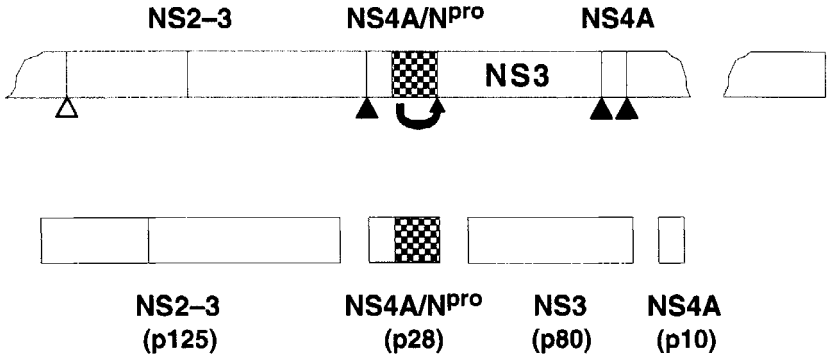


FIG. 15. Processing of the BVDV Pe515CP polyprotein in the region containing the duplicated N^{pro} (SD). Open triangle, cleavage assumed to be effected by cellular signalase; filled triangle, cleavage by NS2–3/NS3; curved arrow, cleavage by N^{pro}.

found for a variety of viruses (reviewed in Barrett and Dimmock, 1986; Schlesinger, 1988; Holland, 1991; Roux *et al.,* 1991) but have not been described for pestiviruses. One reason might be that virus yields and consequently also multiplicities of infection (m.o.i.) are usually low in the pestivirus system. The generation of defective genomes is, however, predominantly observed when viruses are serially passaged with high m.o.i.

Using different approaches, which were all based on the fact that DIs are dependent on helper viruses and therefore can be eliminated by simple dilution experiments, it was shown that the subgenomic RNAs found after infection with BVDV CP9 and CP13 indeed represent the genomes of BVDV DIs (Tautz *et al.,* 1994; Kupfermann *et al.,* in preparation). A much more interesting result of these experiments was that removal of the DIs always led to loss of cytopathogenicity. Autonomously replicating BVDV cloned from the CP9 and CP13 stocks by end-point dilution turned out to be noncytopathogenic. Transfection of cells with RNA from BVDV CP9 or CP13 did not result in recovery of cytopathogenic virus. Only when target cells already infected with a noncp BVDV were used for transfection experiments was a cytopathic effect readily observed. Similar results were obtained after plaque purification: infection of cells with material obtained from isolated plaques rescued cytopathogenicity only after superinfection with a noncp helper virus. In conclusion, the BVDV isolates CP9 and CP13 are each composed of a cytopathogenic DI and a noncytopathogenic helper virus (Tautz *et al.,* 1994; Kupfermann *et al.,* in preparation).

a. The Simple Way: One-Step Deletion Generates the CP9 DI. After determination of the genome organization of the CP9 DI by cDNA cloning and sequencing, a deletion of 4263 nucleotides was detected spanning the complete structural protein-coding region and the NS2 gene. Interestingly, this deletion starts exactly downstream of the N^{pro} gene and ends just before the putative 5′ end of the NS3-coding sequence, leading to a direct fusion of N^{pro} and NS3 in the encoded polyprotein (Fig. 16, E and B, respectively) (Tautz *et al.,* 1994). This arrangement of the two viral proteases is in principle the same as that observed in the polyproteins of BVDV Pe515CP, CP4, and CP6, in which N^{pro} and NS3 encoded by the SD and downstream LD region are also fused (Fig. 16). In contrast to CP9, the duplicated N^{pro} of the other three viruses is altered at the amino terminus. Nevertheless, as shown for Pe515CP, the duplicated N^{pro} present in the hypothetical fusion protein probably cleaves at its own carboxy terminus and thereby generates the amino terminus of NS3. For BVDV CP9 DI the same mechanism is responsible for the generation of NS3, as shown by experiments using *in vitro* translation and transient expression via vaccinia virus.

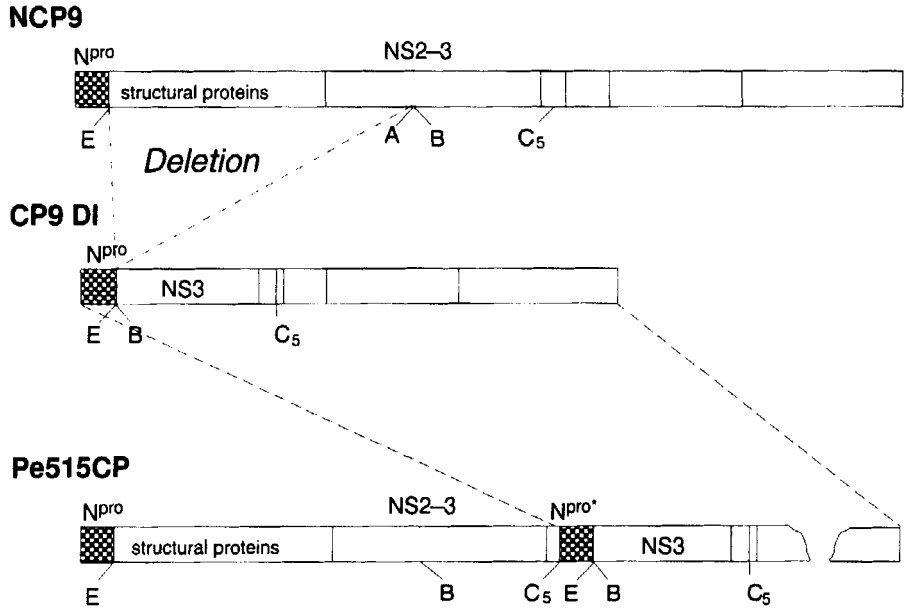


FIG. 16. Genome organization of the BVDV CP9 DI in comparison with that of BVDV NCP9 and BVDV Pe515CP. The designation of genomic positions is the same as in Fig. 13.

b. Scrambled Sequences: Two-Step Deletion Leads to the CP13 DI. The genome of the BVDV CP13 DI was found to lack 4713 nucleotides with respect to NCP13, the noncytopathogenic virus from the same animal. In contrast to the CP9 DI, this difference is not due to one large deletion, but results from two separate deletions. Starting with the 5′ end of the long ORF, 13 codons of the N^{pro} gene are present, followed by 10 codons derived from the genomic region coding for glycoprotein E1 (Fig. 17) (Kupfermann *et al.*, submitted). Downstream of this gene fragment, NS3-coding sequences are found. Interestingly, the NS3-coding region of the DI genome starts 5 codons downstream of codon 1590, which follows ubiquitin- or N^{pro}-coding sequences in the genomes of BVDV isolates with the respective insertions or deletion (Fig. 17, B′ and B, respectively). Under the assumption that codon 1590 represents the 5′ end of the NS3 gene, the CP13 DI encodes an amino-terminally truncated NS3 protein. Protein analyses have shown that the 23 amino acids derived from the fragments of the N^{pro} and E1 genes which precede the NS3-coding region in the CP13 DI genome are not removed from the NS3 protein. Thus, NS3 expressed by this DI is truncated and

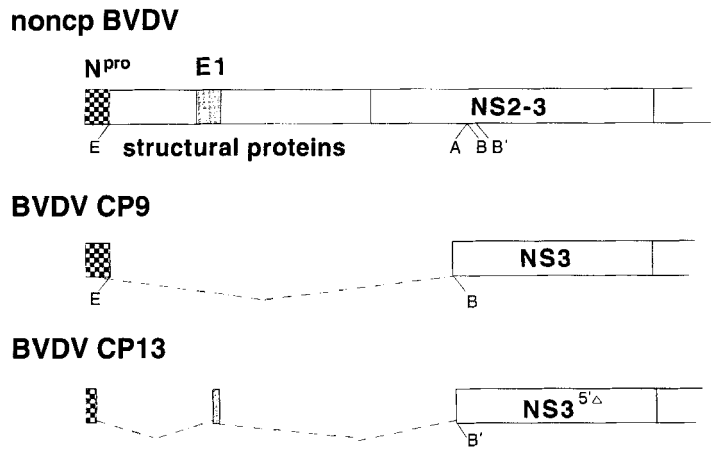


FIG. 17. Genome structure of the BVDV CP13 DI in comparison to that of noncp BVDV and the BVDV CP9 DI. E1, glycoprotein 1 (gp25); B and E mark the same genomic positions as in the preceding figures. B' marks the first nucleotide of the amino-terminally truncated NS3 gene of the CP13 DI designated $NS3^{5'\Delta}$.

contains additional amino acids at its amino terminus. It is therefore designated $NS3^{5'\Delta}$ in the figures. Transient expression studies showed that in spite of these changes the DI-encoded NS3 is able to cleave two of its substrate sites in the polyprotein (Kupfermann *et al.*, submitted).

c. Lose One Third and Kill Your Host: The Role of DI Particles in the Induction of MD. It has long been suggested that DIs play a role in the pathogenesis of virus diseases. Ameliorating effects of DIs on disease symptoms and attenuation of viruses by DIs have been shown for influenza virus (Morgan and Dimmock, 1992), vesicular stomatitis virus (Fultz *et al.*, 1982a,b), Semliki Forest virus (Dimmock and Kennedy, 1978), and some other virus systems (reviewed in Barrett and Dimmock, 1986; Huang, 1988; Huang and Baltimore, 1970; Roux *et al.*, 1991). However, to our knowledge a DI causing a disease of humans or animals has not yet been described. BVDV CP9 and CP13 have been isolated from animals suffering from mucosal disease (MD). Since the cytopathogenic virus present in such animals is regarded as causative for the disease and the DIs represent the only cytopathogenic agent identified in BVDV isolates CP9 and CP13, it can be concluded that in these cases the DIs are responsible for induction of the disease. Comparison of corresponding regions of the DI genomes and the noncp helper viruses showed almost 100% sequence identity, which strongly suggests that the DI is derived from the respective noncp virus by recom-

bination (Tautz *et al.*, 1994; Kupfermann *et al.*, submitted). Thus, development of MD is most likely initiated by the generation of a DI derived from the noncp virus already present within the persistently infected animal.

One interesting question about the role of cytopathogenic DIs with regard to induction of MD concerns horizontal transmission. It has been suggested by experimental infection of persistently infected animals that a high degree of serological similarity between the persisting noncp virus and the cp isolate used for superinfection is a prerequisite for induction of MD (Brownlie *et al.*, 1984). This phenomenon is probably connected with the fact that the persistently infected animals have acquired immunotolerance to the persisting virus. As these cattle are competent to clear superinfection with serologically different BVD viruses (Coria and McClurkin, 1978; Bolin *et al.*, 1985; Moennig *et al.*, 1990; Brownlie, 1991), it can be hypothesized that the immune response prevents extensive replication of the cp virus and thereby induction of MD. Accordingly, horizontal spread of newly generated cp BVDV strains can lead to MD only in a limited subset of animals already persistently infected with a serologically fitting noncp BVDV. Importantly, the two cp BVDV DIs lack structural protein genes and their genomes are packaged into virions whose structural proteins are provided by the respective helper virus. Changing from one helper virus to another would theoretically provide a means to avoid the immune response in any persistently infected animal, at least with regard to structural proteins. In principle, cp BVDV DIs should therefore easily spread among persistently infected animals and cause disease. Future investigations are necessary to elucidate whether BVDV DIs can be held responsible for larger outbreaks of MD.

3. The Minimalist: 27 Nucleotides Do the Job

Screening of virus pairs by Northern hybridization led to identification of one cytopathogenic virus which did not fit into any of the previously identified groups. The respective virus, called CP7, has a genome that perfectly comigrates with the RNA of the corresponding noncp isolate NCP7 (Corapi *et al.*, 1988; Tautz *et al.*, submitted). Probes specific for ubiquitin- or cIns-insertions did not recognize the CP7 genome. We therefore started cloning and sequencing to identify the genetic basis for cytopathogenicity of this virus. Analysis of the genomic region flanking the frequent insertion position between nucleotides 5152 and 5153 did not result in identification of insertions, deletions, or striking mutations. However, after sequencing the complete NS2 gene, a small insertion of 27 nucleotides was identified,

which was located between nucleotides 4353 and 4354 (Tautz *et al.*, submitted). To identify the origin of the CP7 insert a data library search for homologous sequences was performed. A cellular sequence with more than random homology was not found. The highest degree of similarity was obtained with a sequence of BVDV Osloss corresponding to residues 4066 to 4092 of the genome. A more detailed analysis of the CP7 sequence revealed that the insertion actually is 100% identical with a sequence located about 300 nucleotides upstream. The 27 duplicated nucleotides are inserted in a different reading frame, and therefore encode nine amino acids which are not found elsewere in the polyprotein (Fig. 18) (Tautz *et al.*, submitted).

Analysis of the NS2 gene sequence from NCP7 has shown that the 27-nucleotide insertion was specific for the cp virus of pair #7. It was difficult to imagine that such a small change is responsible for cytopathogenicity. Elaborate studies were conducted to analyze the significance of the CP7 insertion for expression of NS3. Using a vaccinia-virus-based transient expression system, cDNA constructs covering the NS2–3 gene of CP7 were expressed (Tautz *et al.*, submitted). In these

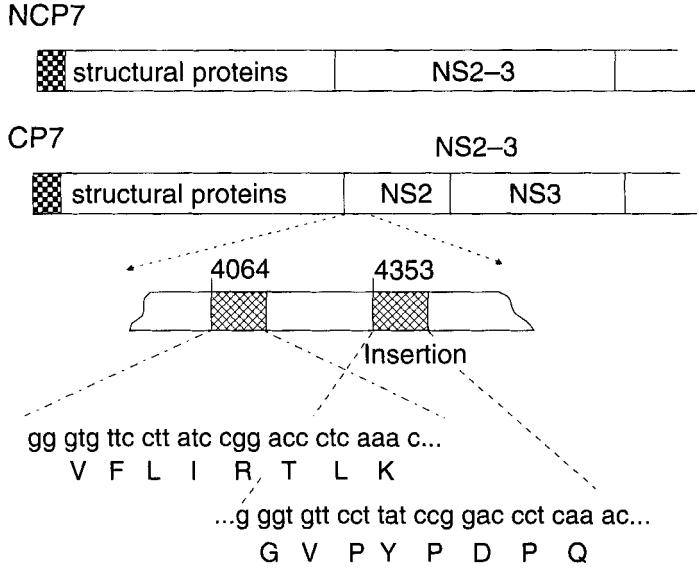


FIG. 18. Genome organization of BVDV NCP7 and CP7, the members of pair #7. The numbers correspond to the genomic positions (based on BVDV SD-1) of the nucleotides preceding the sequence duplicated in the CP7 genome and the position at which the duplicated sequence is found as an insertion. The duplicated sequence and the deduced amino acid sequences are shown.

assays cleavage of NS2–3 resulting in NS2 and NS3 was observed. However, after exchange of part of the NS2-coding region with the corresponding fragment of NCP7, processing of NS2–3 could not be detected. Further analyses showed that (i) the insertion of 27 nucleotides is necessary and sufficient to induce the NS2–3 cleavage, and (ii) the NS3 protease activity can be knocked out without any effect on the processing of NS2–3 (Tautz *et al.*, submitted). Future experiments will address the influence of the amino acid sequence and the position of the insertion. A very interesting question concerns the mechanism of the observed cleavage. The available data suggest that a cellular protease processes NS2–3. If so, one wonders about the difference between the two putative substrates, namely NS2–3 of both viruses from this pair.

C. The Outsiders: cp BVDV without Genome Rearrangement?

Shortly after the first reports on the presence of cellular insertions in the genomes of cp BVDV isolates and the presentation of the hypothesis that recombination leads to cytopathogenic viruses, contradictions were put forward. PCR analyses of cp BVDV strains Singer and Lamspringe did not result in identification of recombination-induced changes in the genomic region surrounding the putative cleavage site within NS2–3 (de Moerlooze *et al.*, 1990). However, in the respective experiments only a minor part of the NS2–3-coding region was analyzed. It is therefore possible that these viruses contain in their NS2–3 gene small insertions similar to the one of CP7. We have also analyzed several cp BVDV strains which are commonly used. For all these strains the genomic RNAs were found to be about 12.3 kb in length, and they did not hybridize to ubiquitin- or cIns-probes. Accordingly, the genomes of these viruses contain neither these cellular sequences nor larger duplications or deletions. Cloning and sequencing revealed that the complete NS2–3 gene was apparently free of recombination-induced changes (Meyers *et al.*, unpublished data). It is therefore very likely that these strains belong to an additional group of cp BVDV for which the genetic basis of cytopathogenicity has yet to be determined. Interestingly, expression of NS3 can be detected for all these viruses. Elucidation of the mechanism responsible for generation of NS3 by these isolates will be an interesting task for future work. It has to be mentioned that the presence of duplicated sequences or insertions is not restricted to cp BVDV. For noncp BVDV2-890, a type 4 pestivirus, a duplication of 228 nucelotides has been identified in the NS2 gene (Ridpath and Bolin, 1995). In addition, the noncp BVDV isolate New York was found to contain a 12 nucleotide insertion in the NS2 gene (de Moerlooze *et al.*, 1990).

D. Development of Mucosal Disease: More Questions than Answers

It was pointed out above that mere BVDV infection is not sufficient to induce mucosal disease. One prerequisite is that transplacental infection of a fetus with a noncp BVD virus occurs during the first 125 days of gestation; this can lead to acquired immunotolerance accompanied by life-long persistent infection. In a second step, a cytopathogenic variant of this virus has to be generated. The data reviewed in the preceding sections show that this event usually requires recombination of sequences from different origins, leading to a variety of genome rearrangements. It is generally assumed that generation of the cp virus precedes the outbreak of MD by a not yet determined time lag. There are still many open questions concerning the processes responsible for generation of cp BVDV and the molecular basis of cytopathogenicity. Unfortunately, our knowledge about pathogenesis of MD is very limited. It has been shown for persistently infected calves that the noncp virus is present in several tissues, including the gastrointestinal tract and the central nervous system. Virus can be isolated from peripheral blood mononuclear leukocytes (PBLs) and also from serum (Bielefeldt Ohmann *et al.*, 1987; Bielefeldt Ohmann, 1988a,b; Liebler *et al.*, 1991). However, it seems that only a limited number of cells are productively infected (Bielefeldt Ohmann, 1988a). The reason for this apparent restriction of noncp virus replication is not yet known, but it is attractive to speculate about a mechanism regulating virus replication. Once a cp virus has been generated the situation changes. The number of infected cells is increased in animals with experimental MD in comparison to persistently viremic calves, and most of these cells harbor cp virus (Liebler *et al.*, 1991). The nature of the advantage of cp viruses with regard to virus production is not clear. There are reports on differences in tissue tropism between cp and noncp BVDV (Brownlie *et al.*, 1988; Clarke *et al.*, 1987; Greiser-Wilke *et al.*, 1993). However, since there is no evidence for different receptors for the two biotypes it is likely that these findings reflect relative virus yields rather than absolute differences in the range of target cells. The data on the molecular characteristics of cp BVDV available now do not explain any replication advantage of cp viruses. It would rather be expected that viruses with huge sequence duplications would have reduced replication rates because of the increased amount of RNA to be copied. From this point of view, it is very interesting that the respective viruses apparently show no tendency to lose the additional sequences but seem to be genetically stable both in the infected animal and in tissue culture. Accordingly, the assumed advantage of cp BVDV seems to compensate for the disadvantage of a much larger genome. Experiments aiming at

comparison of growth characteristics of cp and noncp BVDV belonging to virus pairs are currently conducted and will hopefully broaden our basis for discussion of this interesting issue.

E. New Teams Enter the Game: Cytopathogenic BDV and CSFV

In comparison to the BVDV system, cytopathogenicity of BDV and CSFV has been studied with much less intensity. This is probably due to a lower prevalance of cp isolates from the field and the fact that a disease similar to MD is not known for pigs. For sheep, however, an MD-like syndrome has been described (see Section II,B,3; Nettleton *et al.*, 1992). For a long time almost nothing was known about the molecular biology of cp BDV and cp CSFV isolates. However, recent analyses have provided answers to the question about the probable molecular basis of cytopathogenicity of several BD and CSF viruses.

1. Cytopathogenic BDV: Again a Cellular Gift

Recently, two cp BDV isolates, namely BDV Cumnock and BDV Moredun, were analyzed at the molecular level. Both viruses were isolated from infected animals together with accompanying noncp viruses (Vantsis *et al.*, 1976; Becher *et al.*, 1996). Cp and noncp viruses derived from one animal were found to be antigenically closely related. Thus, in analogy to the BVDV system, cp and noncp Moredun viruses as well as cp and noncp Cumnock viruses are regarded as virus pairs. Genome analyses by RT-PCR revealed 99% nucleotide sequence homology for the members of each pair, indicating that in these cases also the cp viruses represent mutants of the noncp viruses (Becher *et al.*, 1996). Again insertions were identified in the NS2–3 genes of the cp viruses. The respective sequences are missing in the genomes of the noncp isolates. The insertion in the RNA of cp BDV Cumnock has a length of 387 nucleotides, while that of cp BDV Moredun consists of 333 residues (Becher *et al.*, 1996). Interestingly, the two insertions are highly homologous to the cellular sequence of 270 nucleotides called cIns, which is inserted into the BVDV NADL genome. Thus, the RNAs of the two cp BDV isolates analyzed so far also contain cellular sequences and therefore represent products of recombination between viral and cellular RNA. Both BDV insertions encompass the complete cellular element of BVDV NADL together with additional 5′ and 3′ sequences. The three cIns insertions all have different 5′ and 3′ ends (Fig. 19). Analyses of the deduced amino acid sequences revealed only one or three exchanges between the putative bovine protein and the cellular elements in the polyproteins of BDV Cumnock or Moredun, respectively. Since

the BDV insertions most likely are derived from ovine mRNA(s) this finding again shows the high degree of phylogenetic conservation of the respective cellular protein.

An interesting difference between the BVDV NADL and the two BDV insertions concerns their localization within the genome. The two BDV RNAs have integrated the cellular sequences at a position which in comparison to the BVDV SD-1 ORF is located downstream of codon 1532 (Fig. 19). Thus, the insertions of the ovine isolates are located five codons upstream of the position where the NADL cIns is found. Compared to ubiquitin-coding elements in the genomes of BVDV isolates, for which both the 3′ end of the inserted sequence and the 3′ integration position are strictly conserved, the cIns insertions allow a higher degree of flexibility with respect to both the inserted sequence and the genomic location.

2. Cytopathogenic CSFV: No Protease, Please

Cytopathogenic CSFV isolates have only rarely been described (Gillespie *et al.*, 1960; van Bekkum and Barteling, 1970; de Castro, 1973; Laude, 1978). Many publications actually state that CSFV is generally considered to be noncytopathogenic. Molecular characterization of three

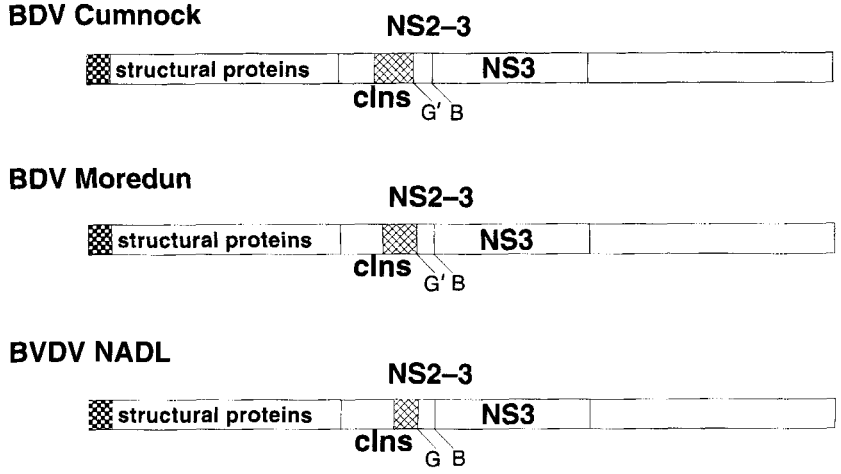


FIG. 19. Genome organization of the cytopathogenic BDV isolates Cumnock CP and Moredun CP in comparison with that of BVDV NADL. B designates the same genomic position as in the preceding figures; G and G′, residue following the cIns insertions in the genomes of BVDV NADL and the cp BDV isolates, respectively.

independent cp CSFV isolates termed CSFV Alfort/M, CSFV ATCC, and CSFV Steiermark resulted in some interesting findings. For all three isolates, two virus-specific RNAs were identified within infected cells. One of these RNAs had the size of genomic CSFV RNA (12.3 kb), while the other was only about 7.5 kb (Meyers and Thiel, 1995). The presence of a subgenomic RNA is reminiscent of results obtained for BVDV isolates CP9 and CP13, which contain cytopathogenic DIs. With approaches analogous to those used for characterization of these BVD viruses, it was shown that the cp CSFV isolates are also composed of cytopathogenic DIs and noncytopathogenic helper viruses.

To determine the genome organization of the cp CSFV DIs, cDNA cloning and sequencing was performed. Surprisingly, all three DIs contain exactly the same deletion of 4764 nucleotides, which starts with the first nucleotide following the translation initiation codon and ends with the codon preceding the probable 5′ end of the NS3 gene (Fig. 20) (Meyers and Thiel, 1995). Thus, the 3′ end of the deletion is exactly the same as that observed for the BVDV CP9 DI. In contrast to the latter, which generates the amino terminus of NS3 through the activity of the preceding protease N^{pro}, the N^{pro} gene is also deleted (Fig. 20). Thus, the translation initiation codon is fused with the NS3 coding

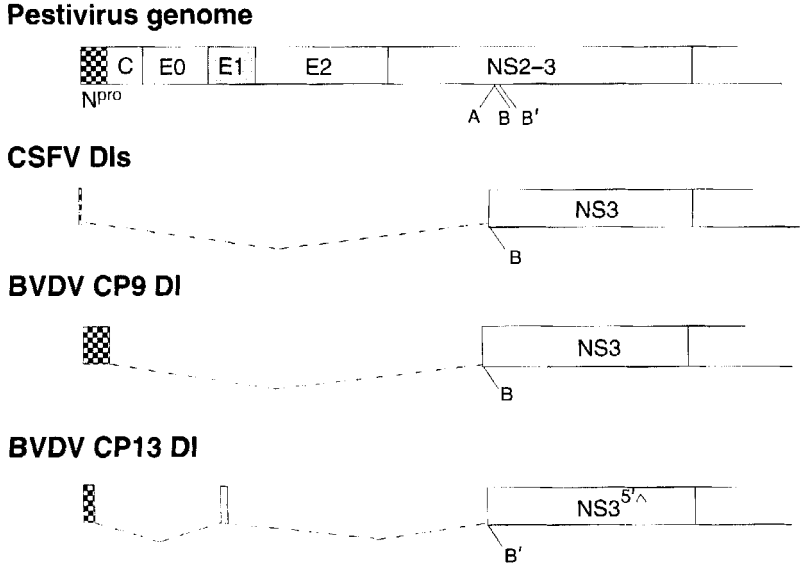


FIG. 20. Genome structure of the CSFV DIs in comparison with the BVDV DIs CP9 and CP13. See also legend to Fig. 17. C, E0, E1, and E2 indicate structural proteins.

sequence, and the first protein expressed from the genome of the CSFV DIs is NS3 with an extra methionine at its amino terminus (Meyers and Thiel, 1995). In contrast to all other cp pestiviruses analyzed so far, the specific kind of genome rearrangement found for the cp CSFV DIs eliminates the requirement of a protease for generation of the amino terminus of NS3. The only other exception is the BVDV CP13 DI, which, however, expresses an NS3 protein with altered amino terminus (designated $NS3^{5'\Delta}$ in the figures).

While the cytopathogenic strain Alfort/M was generated during tissue culture passages of the noncp virus CSFV Alfort, CSFV ATCC and CSFV Steiermark were most likely obtained directly from infected animals. Thus, cp pestiviruses can probably arise by recombination in infected pigs. Since persistent CSFV infections of pigs are apparently limited to several weeks or months, the frequency of such reactions per infected animal might be lower than in cattle. In addition, a marker for generation of a cp virus comparable to outbreak of MD in cattle is not known for pigs. It remains to be determined whether the virulence of the cp CSFV viruses is different from that of the corresponding noncp viruses. Generation of cp CSFV during an outbreak of classical swine fever might actually only be recognized by chance. Elaborate animal experiments are required to elucidate a possible role of cp CSFV in the pathogenesis of classical swine fever.

3. *It's Not All the Same: One Interesting Difference between cp BVDV and Other cp Pestiviruses*

The analyses of cp BDV and cp CSFV isolates resulted in findings very similar to those obtained before in the BVDV system. The cytopathogenic viruses were generated by recombination, which is also true for cp BVD viruses, at least for the majority. Both types of recombination identified in the BDV and CSFV systems, namely integration of cIns-coding cellular sequences for BDV and generation of cp DIs by specific deletions in the case of CSFV, were known from cp BVDV. Regarding the CSFV DIs it is noteworthy that the 3′ end of the deletion is exactly the same as that of the BVDV CP9 DI, and that this position corresponds to the position at which integration of ubiquitin- or N^{pro}-coding sequences has been detected.

In the case of cp BVDV one obvious result of the recombination-induced changes is the expression of NS3, which is not observed in cells infected with noncp BVDV. In contrast, BDV and CSFV can express NS3 in addition to NS2–3, whether or not these viruses are cytopathogenic (Thiel *et al.*, 1991; Becher *et al.*, 1994, 1996). Thus, at least in the case of BDV and CSFV, the mere presence of NS3 within an infected cell seems not to be sufficient for the induction of CPE. However, even

though reliable quantitative data are missing, protein analyses indicate that the amount of NS3 is dramatically higher in cells infected with cp CSFV or cp BDV than in those replicating noncp CSFV or noncp BDV (Meyers and Thiel, 1995; Becher *et al.*, 1994, 1996). It is possible that a certain concentration of NS3 is required for induction of CPE which is not reached after infection with noncp BDV or noncp CSFV. Alternatively, NS3 expressed by the noncp viruses could be structurally different from NS3 of cp isolates. As outlined already, the 3′ recombination position often determines the amino terminus of the encoded NS3. One interpretation of the conservation of the 3′ recombination position identified for different cp pestiviruses is that a particular amino terminus of NS3 is critical for cytopathogenicity. To verify this hypothesis, knowledge about the amino-terminal residues of NS3 from different origins would be advantageous, including (i) viruses with cIns insertions, (ii) CP7, and (iii) noncp BDV and noncp CSFV.

F. Start with One, Leave for a Second, Return to the First: A Mechanism Resembling a Well-Known Love Story

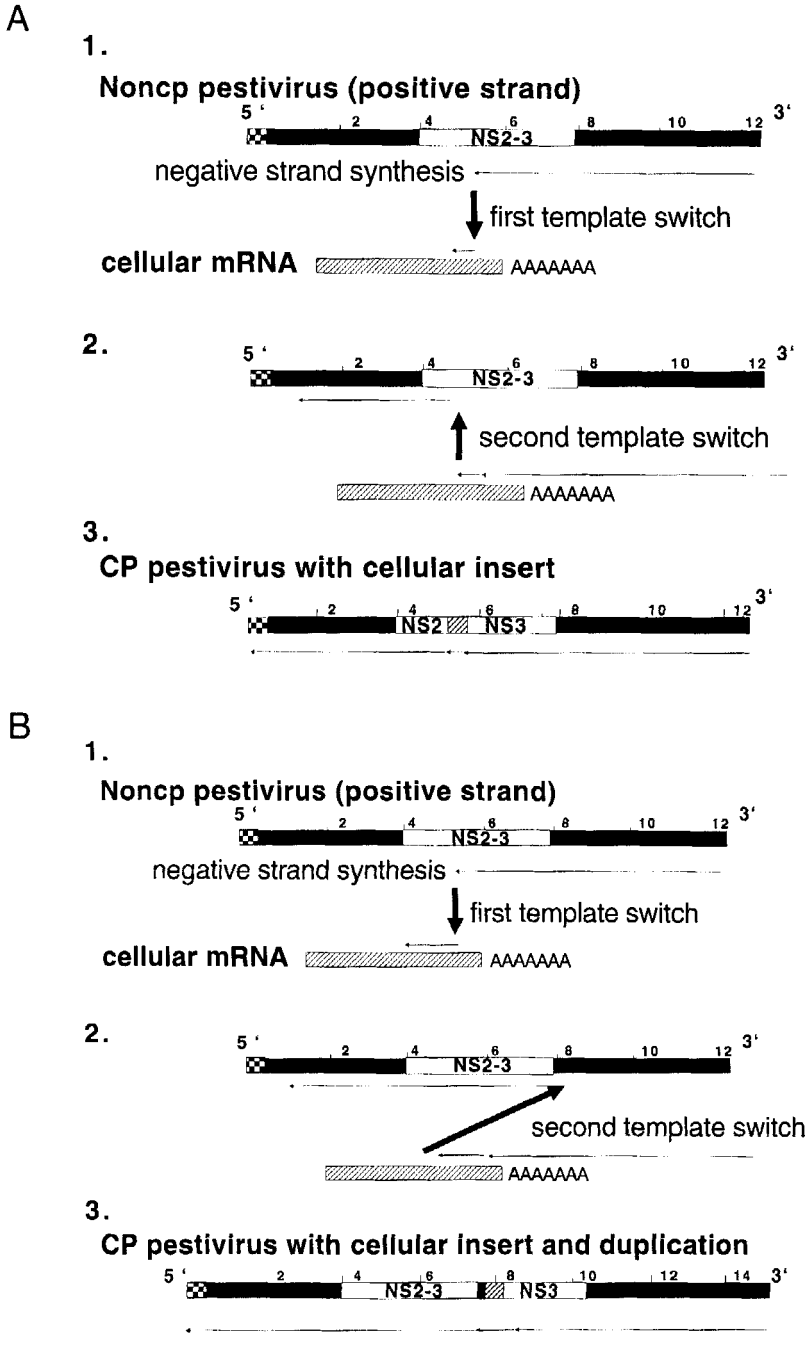
Pestiviruses are considered to replicate in the cytoplasm of the host cell, and neither reverse transcription nor detection of viral antigen or nucleic acid in nuclei has ever been reported for these viruses. Thus, the integration of cellular sequences or rearrangement of viral sequences responsible for generation of the genomes of cp pestiviruses should occur at the RNA level in the cytoplasm of infected cells. RNA recombination has been reported for different viruses (for review see King *et al.*, 1987; Jarvis and Kirkegaard, 1991). Both homologous reactions between corresponding regions of two molecules of viral RNA (Lai *et al.*, 1985; Saunders *et al.*, 1985; Bujarsky and Kaesberg, 1986; Kirkegaard and Baltimore, 1986; Makino *et al.*, 1986; Romanova *et al.*, 1986; Tolskaya *et al.*, 1987; Jarvis and Kirkegaard, 1992) and nonhomologous recombinations between viral RNAs and other partners or two noncorresponding regions of viral RNA have been reported (Lazzarini *et al.*, 1981; Jennings *et al.*, 1983; Monroe and Schlesinger, 1983; Lai *et al.*, 1985; Khatchikian *et al.*, 1989; Cascone *et al.*, 1990, 1993; Bujarsky and Dzianott, 1991; Zhang *et al.*, 1991; Carpenter and Simon, 1994; Charini *et al.*, 1994). A so-called "copy choice" mechanism has been proposed as the enzymatic basis for both types of reactions (Lazzarini *et al.*, 1981; Lai *et al.*, 1985; Kirkegaard and Baltimore, 1986; Makino *et al.*, 1986; King *et al.*, 1987). The "copy-choice" model proposes that recombination results from a template switch of the viral RNA polymerase during genome replication or mRNA transcription. Based on

the copy choice mechanism most RNA recombinations can be explained by a single template switch. However, double crossover has also been observed (Bujarsky and Dzianott, 1991).

It has been found that RNA recombination can occur during synthesis of the negative as well as the positive strand of the genome, but in several cases recombination during negative-strand synthesis is predominant. According to the results of experiments with poliovirus, the bias toward recombination during negative-strand synthesis is due to the much higher concentration of positive strands, which represent the acceptor templates (Jarvis and Kirkegaard, 1992). The same explanation probably holds true for other virus systems. In the case of pestiviral genomes with host-cell-derived insertions, the recombination must have happened during negative-strand synthesis because the cellular sequences are found in the viral genomes in coding orientation, and the corresponding cellular mRNAs are present only as positive strands.

Considering the recombination mechanism, the genomes of cp pestiviruses analyzed so far have all been generated by nonhomologous recombination. The different recombinants can be divided into three groups. The first group includes most viruses with cellular insertions and all genomes of the double duplication type (*i.e.,* Pe515CP, CP4, CP6). These genome structures can be explained by two template switches: the polymerase first copies part of the genome, then leaves its template, reinitiates polymerization at a second template, which can be of cellular or viral origin, and, after having copied several hundred nucleotides thereof, returns to a viral genome template (Fig. 21). The position to which the polymerase jumps back during the second switch determines whether a genome with or without duplication is generated. A jump back to the nucleotide located directly upstream of the first crossing-over position leads to a genome of the Osloss or NADL type (Fig. 21A). Returning to a position located in the positive strand downstream of the first crossing-over site results in duplication of part of the genome (Fig. 21B). For the latter a high degree of flexibility is observed. The range of duplications identified so far extends from 2304 to 3636 nucleotides (BVDV Pe515CP and CP6, respectively). This flexibility may be one important reason for the predominance of viruses with duplications.

The second group consists of BVDV CP9 and the CSFV DIs. These four defective genomes have been generated by a single template switch, which can easily be imagined as an intramolecular reaction (Fig. 22). The generation of such simple recombinants should happen with higher frequencies than the two-step processes responsible for establishment of viruses belonging to the first group. At first sight, it ap-


A
1.
Noncp pestivirus (positive strand)
5'
2
4
6
NS2-3
8
10
12
3'
negative strand synthesis
first template switch
cellular mRNA
AAAAAAA
2.
second template switch
AAAAAAA
3.
CP pestivirus with cellular insert
NS2
NS3
B
1.
Noncp pestivirus (positive strand)
negative strand synthesis
first template switch
cellular mRNA
AAAAAAA
2.
second template switch
AAAAAAA
3.
CP pestivirus with cellular insert and duplication
14

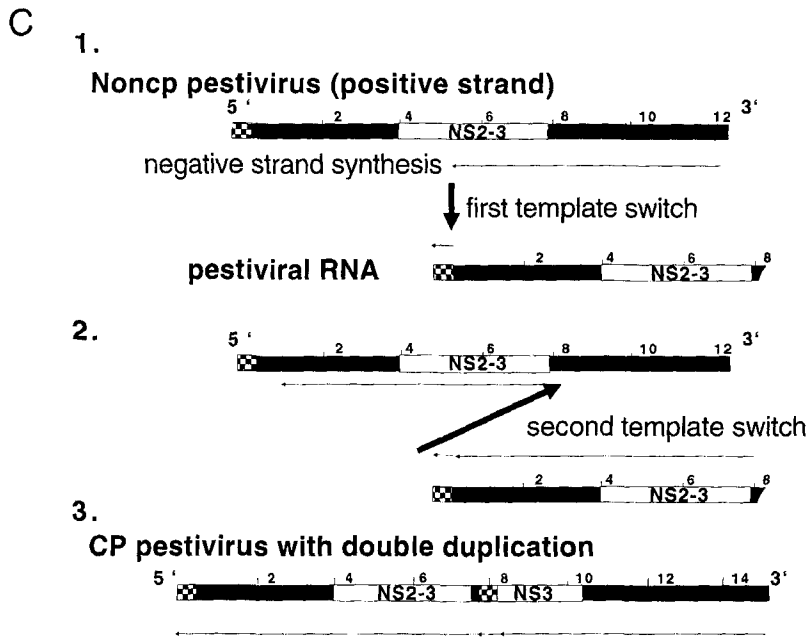


FIG. 21. Schematic presentation of the two-step copy choice recombination leading to pestivirus recombinants with cellular insert (A), with cellular insert and duplication (B), or with double duplication (C). Thick arrows indicate template switches of the viral polymerase; thin arrows represent newly synthesized RNA.

pears to be surprising that among 33 analyzed cp pestivirus isolates only 4 belong to this type. However, for viruses derived from diseased animals, disadvantages of DIs within the host could compensate for the possibly higher frequency of these recombinations. One very obvious disadvantage of DIs compared with replication-competent nondefective cp viruses is their strict helper-virus dependence, which hampers spread of the cp virus under low titer conditions. It also has to be kept in mind that DIs could easily get lost during virus isolation and plaque purification. BVDV CP13 DI does not belong to the second group. Since the genome of this DI contains two deletions, two template switches are necessary to generate it. Therefore, the CP13 DI belongs to the first group.

So far the third group consists of only two defined members. The genomes of BVDV isolates TGAC and III-C have been generated in a three-step process (Fig. 23). The first two template switches are equivalent to those described above for the first group in cases when no duplication is introduced (Osloss type). Then, in a third step the poly-

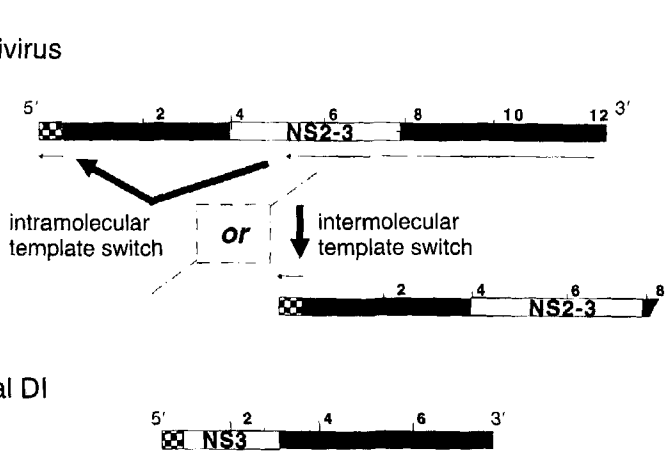


FIG. 22. Model for a recombination with one template switch leading to a DI with one deletion. The two possible mechanisms, namely intramolecular template switch (left part of 1) and intermolecular template switch (right part of 1), are indicated. In principle, intramolecular reactions are possible for all recombinations involving only viral RNA.

merase switches again, leaving its template in the NS2-coding region to start synthesis again downstream of NS3. At first glance such a recombination seems highly unlikely, since it apparently requires three consecutive template switches (Fig. 23A). However, it can be assumed that all three polymerase jumps have not occurred in one single reaction. The genomes of these viruses could have been generated in two separate recombinations. A first reaction could lead to a cp virus corresponding to those of the first group. After replication this RNA could then undergo a second recombination with one template switch from one position of the viral RNA to a downstream position (Fig. 23B). This second recombination would be very similar to that leading to formation of the DI genomes belonging to the second group. Thus, the mechanism responsible for generation of these genomes probably represents just the first group's process followed by the second group's recombination. Since the product of the first process could be amplified by replication, the probability of such an event should be considerably higher than for a consecutive three-step process.

1. The Site Specificity of Recombination

In many cases of homologous recombination, the distribution of crossing-over sites was found to be more or less random; however, preference for certain sites has also been observed (Kirkegaard and

Baltimore, 1986; Tolskaya *et al.*, 1987; Keck *et al.*, 1987; Bujarsky *et al.*, 1994; Lai, 1992). A combination of random recombination and subsequent selection or of recombination at preferred sites combined with selection has been considered responsible for site specificity (Banner and Lai, 1991; Bujarsky *et al.*, 1994; or Tolskaya *et al.*, 1987, respectively). Determination of crossover sites in nonhomologous recombination has been assigned to base-pairing between the nascent RNA strand and the recombination partner (Lazzarini *et al.*, 1981; Lai *et al.*, 1985; Makino *et al.*, 1986; Romanova *et al.*, 1986; King *et al.*, 1987; Bujarsky and Dzianott, 1991), but in some other cases specific sequence elements have been identified adjacent to crossing-over sites (Zhang *et al.*, 1991; Cascone *et al.*, 1993; Carpenter and Simon, 1994).

All cp BVDV strains with ubiquitin-coding insertions and genomes of the double duplication type as well as the BVDV CP9 DI and the CSFV DIs show conservation of the viral sequence at the 3′ crossing-over site (Fig. 24). It is an interesting question whether this site specificity is due to the recombination mechanism or is based on selection. Sequence analyses have been conducted to identify putative signals responsible for this conservation. To analyze one of the putative cellular recombination partners, two bovine polyubiquitin mRNAs were partially cloned and sequenced (Meyers *et al.*, 1991). Sequence homology between viral and cellular RNA in the crossing-over region could not be identified. This was also true for the viral sequences in recombinations of the double duplication type. In the case of one of the polyubiquitin mRNAs, called rpubl, complementarity was observed between the sequence downstream of the last ubiquitin monomer and the region upstream of the crossing-over site in the viral RNA (Meyers *et al.*, 1991). Thus, a heteroduplex could be formed between the two recombining RNAs directly at the 3′ recombination site. However, the differences in codon usage found in rpubl and the ubiquitin-coding sequences present in the viral genomes indicate that the integrated cellular sequences most likely are not derived from rpubl. Sequence complementarity could not be detected for any other recombination of pestiviruses. Since no striking secondary structures or sequence elements like direct or inverted repeats were found at the crossing-over site either, the high incidence of recombinations at this position cannot easily be explained on the basis of nucleotide sequence. It is interesting that the residues downstream of the recombination site are homologous to the reverse of the genomic 3′ end, which is 5′–A/G–G–C–C–C–3′, sometimes followed by one or two more C residues (Fig. 24) (Brock *et al.*, 1992; Moormann *et al.*, 1990; Meyers *et al.*, 1996). Thus, the viral polymerase leaves its template after having synthesized a sequence

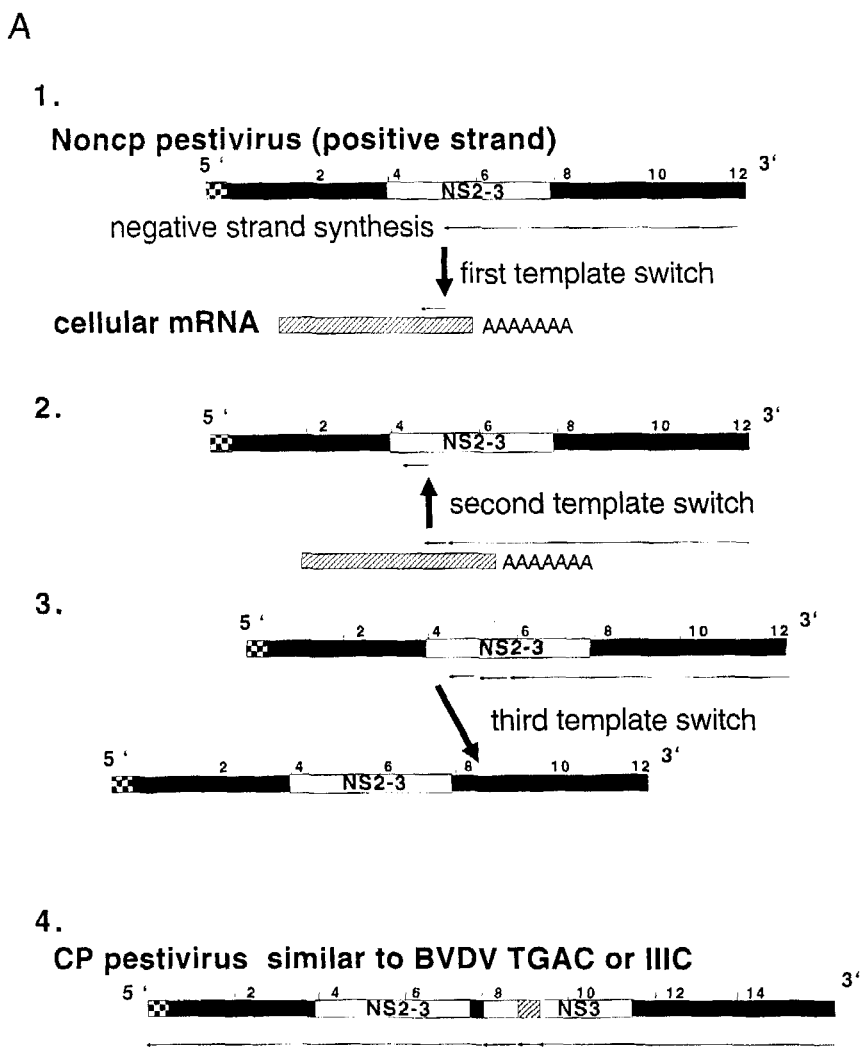
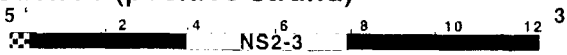


FIG. 23. The two possibilities for generation of a three-step recombinant via template switches. A shows the model for a recombination with three consecutive template switches; in B the possibility of two separate recombination events with a replication competent intermediate is illustrated.

B

1.

Noncp pestivirus (positive strand)

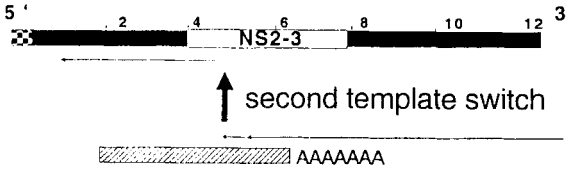


negative strand synthesis

first template switch

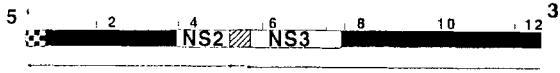
cellular mRNA AAAAAAA

2.



3.

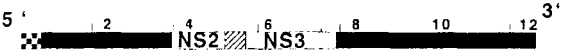
CP pestivirus with cellular insert



4. Replication

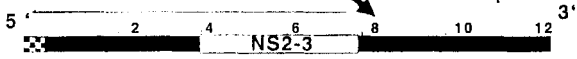
5.

CP pestivirus with cellular insert



negative strand synthesis

third template switch



pestiviral RNA

6.

CP pestivirus similar to BVDV TGAC or IIIC

```
Sequences at the 3' recombination site (G residue
downstream of blank corresponds to position 5153
in a genome without insertion):

BVDV Osloss:          ...gaggggtagt GGGCCTGCCGT...
BVDV CP1:             ...+++++++gg+ ++++++++++...
BVDV 190:             ...+++++++gg+ ++++++++++...
BVDV Pe515CP:         ...c+t+a+ct+c ++++++++++...
BVDV CP6:             ...t+caa+ct+c ++++++++++...
BVDV DI9:             ...t+t+a+ct+c ++++++++++...

CSFV Alfort DI:       ...at++cac+tg ++A++A++T++...
CSFV ATCC DI:         ...at++cac+tg ++A++T++A++...
CSFV Steiermark DI:   ...at++cac+tg ++A++C++T++...

-------------------------------------------------

Consensus:            ...---------- GGGCCTGCCGT...
                                      A  A  T
                                         C  A

-------------------------------------------------

Reverse complement
of genomic 3' ends:
                      BVDV:     (GG)GGGCAGTT....
                      CSFV:     (GG)GGGCCGTT....

-------------------------------------------------

Sequences at 3' recombination site
(genomic locations different from position 5153):

BVDV NADL:            ...ttcacggatg CCTTTCAGGCA...
BVDV CP7:             ...accctcaaag CTGGGGGGGTT...
BVDV CP13 DI:         ...tccctggtga AGAAAATCACC...
```

FIG. 24. Comparison of the sequences at the 3′ crossing-over site of different pestivirus recombinants and the genomic 3′ end. The sequence downstream of the 3′ recombination position is capitalized, while the preceding residues are given in lower case. Note that position 5153 is equivalent to position B in the preceding schematic drawings showing genome structures of the different viruses.

very similar to the end of a positive-strand viral RNA. It might turn out that learning more about the RNA replication of pestiviruses will also help in better understanding their recombination.

It has to be kept in mind that all pestivirus recombinants analyzed so far exhibit a cytopathogenic phenotype. The cytopathogenicity of these viruses is in all cases correlated with expression of NS3 or at least enhanced production of this protein. The ability to express NS3 is linked to the recombination-induced change at the genome level. Therefore, in addition to putative sequence elements guiding crossing over, functional pressure has to be considered as a reason for site-specific recombination in pestiviruses. This functional pressure could

be based on the requirement of a given amino terminus of NS3 in connection with the cp phenotype. The proposed role of NS3 in induction of cytopathogenicity most likely is based on a defined function of the protein. Maintenance of this function may allow only limited flexibility of the protein, including its boundaries. A combination of certain sequence elements, somehow suited as crossing-over sites, and functional pressure on the recombination product could result in a high number of site-specific recombinants, as observed for pestiviruses.

2. The Role of NS3 for Cytopathogenicity of Pestiviruses

The molecular basis of cytopathogenicity still awaits experimental analysis. The data presented above show that the cp phenotype is usually correlated with genome rearrangements resulting from recombination. For CSFV the correlation between genome rearrangement and cytopathogenicity was recently demonstrated by generation of a cp DI from cloned sequences (Meyers *et al.,* 1996). It is noteworthy that all the different changes identified at the genome level directly influence the NS2–3 gene. As a result of these rearrangements, expression of NS3 can be observed in the BVDV system (Fig. 25). For CSFV and BDV, the effect of the genomic changes is more difficult to interpret. It is very clear, however, that in these cases also the expression of NS3 changes dramatically. The fact that a variety of different mutations is correlated with two phenotypic changes, namely specific alteration in the viral polypeptide pattern and cytopathogenicity of the respective viruses, strongly argues for a causal connection between recombination, NS3 expression, and cytopathogenicity. It seems likely that NS3 is responsible for the cytopathogenic effect of the various cp pestiviruses. Future experiments will be conducted to support this hypothesis and to analyze whether NS3 is sufficient to provoke a cytopathic effect and what kind of mechanism leads to death of the infected host cells.

IV. Large-Scale Jumps Instead of Tiny Steps: Considerations about Viral Evolution

Acquisition of new properties from the host cell by recombination has long been proposed as an important force in the evolution of RNA viruses (Steinhauer and Holland, 1987; Strauss and Strauss, 1988). While point mutations in general allow only slow evolutionary changes, recombination represents a powerful means to quickly gain new properties. As outlined above, recombination is observed in a variety of virus systems. Nonhomologous recombinations with cellular sequences

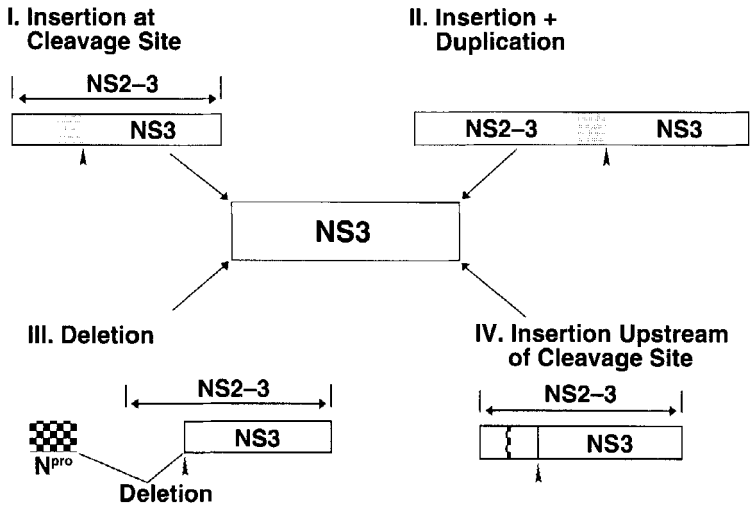


FIG. 25. Different changes at the genome level all lead to (enhanced) expression of NS3. Arrowheads point to the position of the cleavage site at the amino terminus of NS3.

have also been reported for other viruses (Monroe and Schlesinger, 1983; Tsiang *et al.*, 1985; Munishkin *et al.*, 1988; Khatchikian *et al.*, 1989; Charini *et al.*, 1994). For influenza virus and poliovirus, integration of short stretches of cellular 28S ribosomal RNA has been found, leading in both cases to phenotypic changes. The influenza virus recombinant contains 54 cellular nucleotides inserted into the hemagglutinin gene. As a consequence of the recombination the hemagglutinin is cleaved by a cellular protease. Because of this cleavage the phenotype of the respective virus variant changes from apathogenic to highly pathogenic for chicken (Khatchikian *et al.*, 1989). In the case of poliovirus, a lethal mutant with an artificial 3-nucleotide insertion in the protease 3C-gene regained viability by a 15-nucleotide insertion at the position of the original mutation. This insertion was found to be derived from the 28S ribosomal RNA (Charini *et al.*, 1994). In both cases the recombination with cellular sequences resulted in a phenotypic change which was based on restoration of a protease cleavage site. In some aspects these findings are similar to the observations in the pestivirus system. However, the respective poliovirus and influenza virus recombinants arose during tissue culture passages. In contrast, almost all pestivirus recombinants were generated within infected animals. Moreover, pestiviruses represent the first "classical" positive-stranded RNA viruses for which integration of host-cellular protein-

coding sequences has been demonstrated. The large number of natural recombinants available in the pestivirus system provides the opportunity to study RNA virus evolution in action. A whole set of different recombinations, including integration of cellular sequences, duplication and rearrangement of viral sequences, and introduction of deletions, are employed to produce a certain phenotype. There is certainly a lot more to learn about the molecular mechanisms involved in RNA virus evolution by further studies of this fascinating system.

ADDENDUM

The designations of the pestiviral proteins used in this article are mostly as recently proposed by the Flaviviridae study group of the International Committee on Taxonomy of Viruses; only EO will be replaced by E^{rns} (for RNase soluble).

ACKNOWLEDGMENTS

We thank Norbert Tautz and Robert Stark for their help with the figures as well as many colleagues for sharing with us unpublished information.

REFERENCES

Agell, N., Bond, U., and Schlesinger, M. J. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3693–3697.

Ahmed, R., Salmi, A., Butler, L. D., Chiller, J. M., and Oldstone, M. B. A. (1984). *J. Exp. Med.* **60**, 521–540.

Baker, J. A. (1946). *Proc. Soc. Exp. Biol. Med.* **63**, 183.

Baker, J. C. (1987). *J. Am. Vet. Med. Assoc.* **190**, 1449–1458.

Baker, R. T., Tobias, J. W., and Varshavsky, A. (1992). *J. Biol. Chem.* **267**, 23364–23375.

Banner, L. R., and Lai, M. M. C. (1991). Virology **185**, 441–445.

Barlow, R. M. (1972). *J. Comp. Pathol.* **82**, 151–157.

Barrett, A. D. T., and Dimmock, N. J. (1986). *Curr. Top. Microbiol.* **128**, 55–84.

Becher, P., Shannon, A. D., Tautz, N., and Thiel, H.-J. (1994). *Virology* **198**, 542–551.

Becher, P., König, M., Paton, D. J., and Thiel, H.-J. (1995). *Virology* **209**, 200–206.

Becher, P., Meyers, G., Shannon, A. D., and Thiel, H.-J. (1996). *J. Virol.* **70**, 2992–2998.

Bielefeldt Ohmann, H. (1988a). *Acta Vet. Scand.* **29**, 77–84.

Bielefeldt Ohmann, H. (1988b). *Vet. Pathol.* **25**, 304–309.

Bielefeldt Ohmann, H., and Bloch, B. (1981). *Arch. Virol.* **71**, 57–74.

Bielefeldt Ohmann, H., Ronsholt, L., and Bloch, B. (1987). *J. Gen. Virol.* **68**, 1971–1982.

Bolin, S. R., McClurkin, A. W., Cutlip, R. C., and Coria, M. F. (1985). *Am. J. Vet. Res.* **46**, 573–576.

Bolin, S., Moennig, V., Kelso Gourley, N. E., and Ridpath, J. (1988). *Arch. Virol.* **99**, 117–123.

Brock, K. V., Deng, R., and Riblet, S. M. (1992). *J. Virol. Method.* **38**, 39–46.

Brownlie, J. (1991). *Arch. Virol.* **Suppl. 3**, 79–96.

Brownlie, J., Clarke, M. C., and Howard, C. J. (1984). *Vet. Rec.* **114**, 535–536.
Brownlie, J., Clarke, M. C., and Howard, C. J. (1988). 15th World Buiatrics Contress, Spain. pp. 899–904.
Bujarski, J. J., and Dzianott, A. M. (1991). *J. Virol.* **65**, 4153–4159.
Bujarski, J. J., Nagy, P. D., and Flasinski, S. (1994). *Adv. Virus Res.* **43**, 275–302.
Bujarsky, J. J., and Kaesberg, P. (1986). *Nature* **321**, 528–531.
Butt, T. R., Khan, M. I., Marsh, J., Ecker, D. J., and Crooke, S. T. (1988). *J. Biol. Chem.* **262**, 16364–16371.
Carpenter, C. D., and Simon, A. E. (1994). *Virology* **201**, 419–423.
Cascone, P. J., Carpenter, C. D., Li, X. H., and Simon, A. E. (1990). *EMBO J.* **9**, 1709–1715.
Cascone, P. J., Haydar, T., and Simon, A. E. (1993). *Science* **260**, 801–805.
Cay, B., Chappuis, G., Coulibaly, C., Dinter, Z., Edwards, S., Greiser-Wilke, I., Gunn, M., Have, P., Hess, G., Junti, N., Liess, B., Mateo, A., McHugh, P., Moennig, V., Nettleton, P., and Wensvoort, G. (1989). *Vet. Microbiol.* **20**, 123–129.
Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990). *Annu. Rev. Microbiol.* **44**, 649–688.
Charini, W. C., Todd, S., Gutman, G. A., and Semler, B. L. (1994). *J. Virol.* **68**, 6547–6552.
Clarke, M. C., Brownlie, J., and Howard, C. J. (1987). *In* "Pestvirus Infections of Ruminants" (J. W. Harkness, ed.), pp. 3–12. CEC Seminar, Brussels.
Collett, M. S., Larson, R., Gold, C., Strinck, D., Anderson, D. K., and Purchio, A. F. (1988a). *Virology* **195**, 191–199.
Collett, M. S., Larson, R., Belzer, S. K., and Retzel, E. (1988b). *Virology* **165**, 200–208.
Collett, M. S., Wiskerchen, M. A., Welniak, E., and Belzer, S. K. (1991). *Arch. Virol.* **Suppl 3**, 19–27.
Corapi, W. V., Donis, R. O., and Dubovi E. J. (1988). *J. Virol.* **62**, 2823–2827.
Corapi, W. V., French, T. W., and Dubovi, E. J. (1989). *J. Virol.* **63**, 3934–3943.
Corapi, W. V., Donis, R. O., and Dubovi, E. J. (1990). *J. Am. Vet. Med. Assoc.* **51**, 1388–1394.
Coria, M. F., and McClurkin, A. W. (1978). *J. Am. Vet. Med. Assoc.* **172**, 449–451.
Darbyshire, J. A. (1960). *Vet. Rec.* **72**, 331.
De Castro, M. P. (1973). *In Vitro* **9**, 8–16.
De Moerlooze, L., Desport, M., Renard, A., Lecomte, C., Brownlie, J., and Martial, J. A. (1990). *Virology* **177**, 812–815.
De Moerlooze, L., Lecomte, C., Brown-Skimmer, S., Schmetz, D., Guiot, C., Vandenburg, D., Allaer, D., Rossius, M., Chappuis, G., Dina, D., Renard, A., and Martial, J. A. (1993). *J. Gen. Virol.* **74**, 1433–1438.
Deng, R., and Brock, K.V. (1992). *Virology* **191**, 867–879.
Dimmock, N. J., and Kennedy, S. I. T. (1978). *J. Gen. Virol.* **39**, 231–242.
Dinter, Z. (1963). *Zbl. Bakt. Mikrobiol., Hyg. A.* **188**, 475–486.
Donis, R. O., and Dubovi, E. J. (1987a). *Am. J. Vet. Res.* **48**, 1549–1554.
Donis, R. O., and Dubovi, E. J. (1987b). *J. Gen. Virol.* **68**, 1597–1605.
Donis, R. O, and Dubovi, E. J. (1987c). *Virology* **158**, 168–173.
Donis, R. O., Corapi, W., and Dubovi, E. J. (1988). *J. Gen. Virol.* **69**, 77–86.
Dunne, H. W. (1973). *Adv. Vet. Sci. Comp. Med.* **17**, 315–359.
Edwards, S., Sands, J. J., and Harkness, J. W. (1989). *Arch. Virol.* **102**, 197–206.
Elbers, K., Tautz, N., Becher, P., Stoll, D., Rümenapf, T., and Thiel, H.-J. (1996). *J. Virol.* **70**, 4131–4135.
Finley, D., Özkaynak, E., and Varshavsky, A. (1987). *Cell* **48**, 1035–1046.
Finley, D., Bartel, B., and Varshavsky, A. (1989). *Nature* **338**, 394–401.

Fultz, P. N., Shadduck, J. A., Kang, C.-Y., and Streilein, J. W. (1982a). *Infect. Immun.* **37**, 679–686.

Fultz, P. N., Shadduck, J. A., Kang, C.-Y., and Streilein, J. W. (1982b). *J. Gen. Virol.* **63**, 493–497.

Gillespie, J. H., Sheffy, B. E., and Baker, J. A. (1960). *Proc. Soc. Exp. Biol. Med.* **105**, 679–681.

Gorbalenya, A. E., Koonin, E. V., and Lai, M. M.-C. (1991). *FEBS Lett.* **288**, 201–205.

Gray, E. W., and Nettleton, P. F. (1987). *J. Gen. Virol.* **68**, 2339–2346.

Greiser-Wilke, I., Moennig, V., Coulibaly, C. O. Z., Leder, L., Dahle, J., and Liess, B. (1990). *Arch. Virol.* **111**, 213–225.

Greiser-Wilke, I., Liess, B., Schepers, J., Stahl-Hennig, C., and Moennig, V. (1991). *Arch. Virol.* Suppl **3**, 55–65.

Greiser-Wilke, I., Dittmar, K. E., Liess, B., and Moennig, V. (1992). *J. Gen. Virol.* **73**, 47–52.

Greiser-Wilke, I., Haas, L., Dittmar, K., Liess, B., and Moennig, V. (1993). *Virology* **193**, 977–980.

Harasawa, R., and Tomiyama, T. (1994). *J. Clin. Microbiol.* **32**, 1604–1605.

Hershko, A., and Ciechanover, A. (1992). *Annu. Rev. Biochem.* **61**, 761–807.

Hofmann, M. A., Brechtbühl, K., and Stäuber, N. (1994). *Arch. Virol.* **139**, 217–229.

Holland, J. J. (1991). *In* "Fundamental Virology" (B. N. Fields and D. M. Knipe, eds.), 2nd Ed., pp. 151–165. Raven Press, New York.

Horzinek, M. C., Reczko, E., and Petzoldt, K. (1967). *Arch. Ges. Virusforsch.* **21**, 475–478.

Houghton, M. (1996). *In* "Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), 3rd Ed., pp. 1035–1058. Raven Press, New York.

Howard, C. J., Brownlie, J., and Clarke, M. C. (1987). *Vet. Microbiol.* **13**, 361–369.

Huang, A. S. (1988). *In* "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, eds.), vol. III, pp. 195–208. CRC Press, Boca Raton, FL.

Huang, A. S., and Baltimore, D. (1970). *Nature* **226**, 325–327.

Hulst, M. M., Himes, G., Newbigin, E., and Moormann, R. J. M. (1994). *Virology* **200**, 558–565.

Jarvis, T. C., and Kirkegaard, K. (1991). *Trends Genet.* **7**, 186–191.

Jarvis, T. C., and Kirkegaard, K. (1992). *EMBO J.* **11**, 3135–3145.

Jennings, P. A., Finch, J. T., Winter, G., and Robertson, J. S. (1983). *Cell* **34**, 619–627.

Jentsch, S. (1992). *Annu. Rev. Genet.* **26**, 179–207.

Johnson, E. S., Bartel, B., and Varshafsky, A. (1992). *EMBO J.* **11**, 497–505.

Jonnalagadda, S., Butt, T. R., Monia, B. P., Mirabelli, C. K., Gotlib, L., Ecker, D. J., and Crooke, S. T. (1989). *J. Biol. Chem.* **264**, 10637–10642.

Keck, J. G., Stohlmann, S. A., Soe, S., Makino, S., and Lai, M. M. C. (1987). *Virology* **156**, 331–334.

Khatchikian, D., Orlich, M., and Rott, R. (1989). *Nature* **340**, 156–157.

King, A. M. Q., Ortlepp, S. A., Newman, J. W. I., and McCahon, D. (1987). *In* "The Molecular Biology of the Positive Stranded RNA Viruses" (D. J. Rowlands, M. A. Mayo, and B. W. J. Mahy, eds.), pp. 129–152. Academic Press, London.

Kirkegaard, K., and Baltimore, D. (1986). *Cell* **47**, 433–443.

Koprowski, H., James, T. R., and Cox, H. R. (1946). *Proc. Soc. Exp. Biol. Med.* **63**, 178–183.

Kosmidou, A., Ahl, R., Thiel, H.-J., and Weiland, E. (1995). *Vet. Microbiol.* **47**, 111–118.

Kupfermann, H., Thiel, H.-J., and Meyers, G., submitted.

Lai, M. M. C. (1992). *Microbiol. Rev.* **56**, 61–79.

Lai, M. M. C., Baic, R. S., Makino, S., Keck, J. G., Egbert, J., Leibowitz, J. L., and Stohlmann, S. A. (1985). *J. Virol.* **56**, 449–456.

Laude, H. (1977). *Arch. Virol.* **54**, 41–51.
Laude, H. (1978). *Ann. Microbiol. (Inst. Pasteur)* **129A**, 553–561.
Lazzarini, R. A., Keene, J. D., and Schubert, M. (1981). *Cell* **26**, 145–154.
Liebler, E. M., Waschbüsch, J., Pohlenz, J. F., Moennig, V., and Liess, B. (1991). *Arch. Virol.* **3**, 109–124.
Liess, B., Frey, H. R., Kittsteiner, H., Baumann, F., and Neumann, W. (1974). *Dtsch. Tierärztl. Wschr.* **81**, 481–487.
Makino, S., Keck, J. G., Stohlmann, S. A., and Lai, M. M. C. (1986). *J. Virol.* **57**, 729–737.
Mayer, A. N., and Wilkinson, K. D. (1989). *Biochemistry* **28**, 166–172.
McClurkin, A. W, Bolin, S. R., and Coria, M. F. (1985). *J. Am. Vet. Med. Assoc.* **186**, 568–569.
McKercher, D. G., Saito, J. K., Crenshaw, G. L., and Bushnell, R. B. (1968). *J. Am. Vet. Med. Assoc.* **152**, 1621–1624.
Meyers, G., and Thiel, H.-J. (1995). *J. Virol.* **69**, 3683–3689.
Meyers, G., Rümenapf, T., and Thiel, H.-J. (1989a). *Virology* **171**, 555–567.
Meyers, G., Rümenapf, T., and Thiel, H.-J. (1989b). *Nature* **341**, 491.
Meyers, G., Rümenapf, T., and Thiel, H.-J. (1990). *In* "New Aspects of Positive Strand RNA Viruses" (M. A. Brinton and F. X. Heinz, eds.), pp. 25–29. American Society for Microbiology, Washington, DC.
Meyers, G., Tautz, N., Dubovi, E. J., and Thiel, H. J. (1991). *Virology* **180**, 602–616.
Meyers, G., Tautz, N., Stark, R., Brownlie, J., Dubovi, E. J., Collett, M. S., and Thiel, H.-J. (1992). *Virology* **191**, 368–386.
Meyers, G., Thiel, H.-J., and Rümenapf, T. (1996). *J. Virol.* **70**, 1588–1595.
Moennig, V., Mateo, A., Greiser-Wilke, I., Bolin, S. R., Kelso, N. E., and Liess, B. (1988). Abstracts of the 89th Annual Meeting of the American Society for Microbiology 326.
Moennig, V., and Plagemann, G. W. (1992). The pestiviruses. *Adv. Virus Res.* **41**, 53–91.
Moennig, V., Frey, H.-R., Liebler, E., Pohlenz, P., and Liess, B. (1990). *Vet. Rec.* **127**, 200–203.
Moennig, V., Greiser-Wilke, I., Frey, H. R., Haas, L., Liebler, E., Pohlenz, J., and Liess, B. (1993). *J. Vet. Med.* **40**, 371–377.
Monath, T. P., and Heinz, F. X. *In* "Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), 3rd Ed., pp. 961–1034. Raven Press, New York.
Monia, B. P., Ecker, D. J., Jonnalagadda, S., Marsh, J., Gotlib, L., Butt, T. R., and Crooke, S. T. (1989). *J. Biol. Chem.* **264**, 4093–4103.
Monroe, S. S., and Schlesinger, S. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3279–3283.
Moormann, R. J. M., Warmerdam, P. A. M., Meer, B. V. D., Schaaper, W. M. M., Wensvoort, G., and Hulst, M. M. (1990). *Virology* **177**, 812–815.
Morgan, D. J., and Dimmock, N. J. (1992). *J. Virol.* **66**, 1188–1192.
Munishkin, A. V., Veronin, L. A., and Chetverin, A. B. (1988). *Nature* **333**, 473–475.
Nettleton, P. F., Gilmour, J. S., Herring, J. A., and Sinclair, J. A. (1992). *Comp. Immun. Microbiol. Infect. Dis.* **15**, 179–188.
Özkaynak, E., Finley, D., Solomon, M. S., and Varshavaky, A. (1987). *EMBO J.* **6**, 1429–1439.
Paton, D. J., Sands, J. J., and Roehe, P. M. (1991). *Arch. Virol.* Suppl **3**, 47–54.
Paton, D. J., Lowings, J. P., and Barrett, A. D. T. (1992). *Virology* **190**, 763–772.
Pellerin, C., Van Den Hurk, J., Lecomte, J., and Tijssen, P. (1994). *Virology* **203**, 260–268.
Peters, W., Greiser-Wilke, I., Moennig, V., and Liess, B. (1986). *Vet. Microbiol.* **12**, 195–200.
Pocock, D. H, Howard, C. J., Clarke, M. C., and Brownlie, J. (1987). *Arch. Virol.* **94**, 43–53.

Poole, T. L., Wang, C., Popp, A., Potgieter, L. N. D., Siddiqui, A., and Collett, M. S. (1995). *Virology* **206**, 750–754.

Purchio, A. F., Larson, R., and Collett, M. S. (1984). *J. Virol.* **50**, 666–669.

Qi, F., Ridpath, J. F., Lewis, R., Bolin, S. R., and Berry, E. S. (1992). *Virology* **189**, 285–292.

Rebhun, W. C., French, T. W., Perdrizet, J. A., Dubovi, E. J., Dill, S. G., and Karcher, L. F. (1989). *J. Vet. Intern. Med.* **3**, 42–46.

Rechsteiner, M. (1987). *Annu. Rev. Cell Biol.* **3**, 1–30.

Redmann, K. L., and Rechsteiner, M. (1989). *Nature* **383**, 438–440.

Renard, A., Dino, D., and Martial, J. (1987). Vaccines and diagnostics derived from bovine diarrhea virus. European Patent Application number 86870095. 6. Publication number 0208672, 14 January 1987.

Rice, C. M.. (1996). In "Fields Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), 3rd Ed., pp. 931–959. Lippincott–Raven, Philadelphia.

Ridpath, J. F., Bolin, S. R., and Dubovi, E. J. (1994). *Virology* **205**, 66–74.

Ridpath, J. F., and Bolin, S. R. (1995). *Virology* **212**, 39–46.

Romanova, L., Blinov, V. M., Tolskaya, E. A., Viktorova, E. G., Kolesnikova, M. S., Guseva, E. A., and Agol, V. I. (1986). *Virology* **155**, 202–213.

Rose, I. A. (1988). *In* "Ubiquitin" (M. Rechsteiner, ed.), pp. 111–114. Plenum, New York.

Roux, L., Simon, A. E., and Holland, J. J. (1991). *Adv. Virus Res.* **40**, 181–211.

Rümenapf, T., Meyers, G., Stark, R., and Thiel, H.-J. (1991). *Arch. Virol.* **3**, 7–18.

Rümenapf, T., Unger, G., Strauss, J. H., and Thiel, H.-J. (1993). *J. Virol.* **67**, 3288–3294.

Saunders, K., King, A. M. Q., McCahon, D., Neumann, J. W. I., Slade, W. R., and Forss, S. (1985). *J. Virol.* **56**, 921–929.

Schlesinger, S. (1988). *In* "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, eds.), Vol. 2. CRC Press, Boca Raton, FL.

Schneider, R., Unger, G., Stark, R., Schneider-Scherzer, E., and Thiel, H.-J. (1993). *Science* **261**, 1169–1171.

Shaw, I. G., Winkler, C. E., and Terlecki, S. (1967). *Vet. Rec.* **81**, 115–116.

Stark, R., Meyers, G., Rümenapf, T., and Thiel, H.-J. (1993). *J. Virol.* **67**, 7088–7095.

Steinhauer, D. A., and Holland, J. J. (1987). *Annu. Rev. Microbiol.* **42**, 657–683.

Strauss, J. H., and Strauss, E. G. (1988). *Annu. Rev. Microbiol.* **42**, 657–683.

Tamura, J. K., Warrener, P., and Collett, M. S. (1993). *Virology* **193**, 1–10.

Tautz, N., Meyers, G., and Thiel, H.-J. (1993). *Virology* **197**, 74–85.

Tautz, N., Thiel, H.-J., Dubovi, E. J., and Meyers, G. (1994). *J. Virol.* **68**, 3289–3297.

Tautz, N., Meyers, G., Stark, R., Dubovi, E. J., and Thiel, H.-J. Submitted.

Thiel, H.-J., Stark, R., Weiland, E., Rümenapf, T., and Meyers, G. (1991). *J. Virol.* **65**, 4705–4712.

Thiel, H.-J., Plagemann, G. W., and Moennig, V. (1996). *In* "Fields Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, eds.), 3rd Ed., pp. 1059–1073. Lippincott–Raven, Philadelphia.

Tobias, J. W., and Varshavsky, A. (1991). *Biol. Chem.* **266**, 12021–12028.

Tolskaya, E. A., Romanova, L. I., Blinov, V. M., Viktorova, E. G., Sinyakov, A. N., Kolesnikova, M. S., and Agol, V. I. (1987). *Virology* **161**, 54–61.

Tsiang, M., Monroe, S., and Schlesinger, S. (1985). *J. Virol.* **54**, 38–44.

van Bekkum, J. G., and Barteling, S. J. (1970). *Arch. Gesamte Virusforsch.* **32**, 185–200.

van Oirschot, J. T. (1992). *In* "Diseases of Swine" (Lerman *et al.*, eds.), 7th Ed., pp. 274–285. Iowa State University Press, Ames, IA.

van Rijn, P. A., van Gennip, H. G. P., de Meijer, E. J., and Moorman, R. J. M. (1993). *J. Gen. Virol.* **74**, 2053–2060.

van Rijn, P. A., Miedema, K. W., Wensvoort, G., Gennip, H. G. P., and Moormann, R. J. M. (1994). *J. Virol.* **68**, 3934–3942.

van Zijl, M., Wensvoort, G., de Kluyver, E., Hulst, M., van der Gulden, H. Gielkens, A., Berns, A., and Moormann, R. (1991). *J. Virol.* **65**, 2761–2765.

Vantsis, J. T., Barlow, R. M., Fraser, J., Rennie, J. C., and Mould, D. L. (1976). *J. Comp. Pathol.* **86**, 111–120.

Vantsis, J. T., Linklater, K. A., Rennie, J. C., and Barlow, R. M. (1979). *J. Comp. Pathol.* **83**, 331–339.

Warrener, P., and Collett, M. (1995). *J. Virol.* **69**, 1720–1726.

Weiland, E., Stark, R., Haas, B., Rümenapf, T., Meyers, G., Thiel, H.-J. (1990). *J. Virol.* **64**, 3563–3569.

Weiland, E., Ahl, R., Stark, R., Weiland, F., and Thiel, H.-J. (1992). *J. Virol.* **66**, 3677–3682.

Wengler, G., Bradley, D. W., Collett, M. S., Heinz, F. X., Schlesinger, R. W., and Strauss, J. H. (1995). *In* "Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, eds.), pp. 415–427, Springer-Verlag, New York.

Wensvoort, G. (1989). *J. Gen. Virol.* **21**, 9–20.

Wensvoort, G., and Terpstra, C. (1985). *Tijdschr. Diergeneesk.* **110**, 263–269.

Wensvoort, G., Terpstra, C., de Kluijver, E. P., Kragten, C., and Warnaar, J. C. (1989). *Vet. Microbiol.* **21**, 9–20.

Westaway, E. G., Brinton, M. A., Gaidamovich, S. Y. A., Horzinek, M. C., Igarashi, A., Kääriäinen, L., Lvov, D. K., Porterfield, J. S., Russel, P. K., and Trent, D. W. (1985). *Intervirology* **24**, 125–139.

Wilhelmsen, C. L., Bolin, S. R., Ridpath, J. F., Cheville, N. F., and Kluge, J. P. (1991). *Am. J. Vet. Res.* **52**, 269–275.

Windisch, J. M., Schneider, R., Stark, R., Weiland, E., Meyers, G., and Thiel, H.-J. (1996). *J. Virol.* **70**, 352–358.

Wiskerchen, M. A., and Collett, M. S. (1991). *Virology* **184**, 341–350.

Wiskerchen, M. A., Belzer, S. K., and Collett, M. S. (1991). *J. Virol.* **65**, 4508–4514.

Zhang, C., Cascone, P. J., and Simon, A. E. (1991). *Virology* **814**, 791–794.

ADVANCES IN VIRUS RESEARCH, VOL. 47

PRINCIPLES OF MOLECULAR ORGANIZATION, EXPRESSION, AND EVOLUTION OF CLOSTEROVIRUSES: OVER THE BARRIERS

A. A. Agranovsky

Belozersky Institute of Physico-Chemical Biology
and the Department of Virology
Moscow State University
119899 Moscow, Russia

I. Introduction

The closterovirus group (having its name from κλωστερ, Greek for "thread") combines several positive-strand RNA viruses with very flexuous filamentous particles, of which beet yellows virus (BYV) is the type virus (Bar-Joseph and Hull, 1974; Bar-Joseph and Murant, 1982). Closteroviruses are distinct from other RNA viruses of plants in some important phenomenological aspects. They have genomes of up to 20 kilobases (kb), a value comparable only to those of the animal coronaviruses and toroviruses, which have the largest RNA genomes of all positive-strand RNA viruses. The existence of such genomes having a coding capacity several times that of an average RNA virus genome (e.g., TMV) raises questions as to the trend whereby the long genomes

have evolved and the possible novel functions they have acquired. The dramatic increase in the closterovirus genome coding capacity may be linked to the distinct ecological niche they occupy. Thus, closteroviruses are the only elongated plant viruses known so far to cause phloem-limited infections in plants and to persist in their insect vectors for many hours, in contrast to only minutes. Closteroviruses also attract interest from an applied standpoint, since the diseases they cause in crops like citrus trees and sugar beet are listed among the most economically important plant viral diseases (Duffus, 1973; Falk and Duffus, 1988). Understanding the molecular mechanisms that underlie the distinct biological patterns of closteroviruses may help in developing measures against crop losses.

Details of closterovirus cytopathology, ecology, relationships with vectors, and disease control may be found in other comprehensive reviews (Bar-Joseph *et al.*, 1979; Duffus, 1973; Falk and Duffus, 1988; Lesemann, 1988; Lister and Bar-Joseph, 1981; Murant *et al.*, 1988; Tollin and Wilson, 1988). The molecular organization, evolution, and taxonomy of closteroviruses have been reviewed in considerable depth (Coffin and Coutts, 1993; Dolja *et al.*, 1994). However, the picture changes rapidly, and thus there is a need for updating. In this review, I will focus on the molecular organization, evolution, and expression of closterovirus genomes, as well as on their unique particle structure.

II. Draft of Closterovirus Taxonomy

Until recently, the taxonomic status of closteroviruses has been ill-defined, and the group was a quite heterogeneous collection, with the main distinction being between the so-called typical and atypical closteroviruses (Bar-Joseph *et al.*, 1979; Bar-Joseph and Murant, 1982; Francki *et al.*, 1991). Typical closteroviruses have particles of 1000–2000 nm in length, induce characteristic BYV-type membranaceous vesicles in infected tissues, and are transmitted by insect vectors (Table I). Atypical (or "clostero-like") viruses have shorter particles of about 750 nm, do not induce BYV-type vesicles, and have no known vector. Sequencing of the typical and atypical closteroviruses and comparison of the encoded protein sequences confirmed the disparity between them (Agranovsky *et al.*, 1991a,b, 1994a; Dolja *et al.*, 1991, 1994; German *et al.*, 1990; Koonin, 1991; Yoshikawa *et al.*, 1992). Former closteroviruses such as apple chlorotic leafspot virus, apple stem grooving virus and citrus tatter leaf virus have already been reclassified as trichoviruses and capilloviruses, respectively (Francki *et al.*, 1991;

Mayo and Martelli, 1993; Zaccomer *et al.*, 1995), and others will possibly follow. Hence, the "atypical closteroviruses" will not be considered in this review (for an updated list, see Coffin and Coutts, 1993).

A stimulating draft of the phylogenetic taxonomy of positive-strand RNA viruses has been proposed on the strength of sequence comparisons for proteins involved in genome replication and expression, particularly RNA-dependent RNA polymerases (Koonin, 1991; Koonin and Dolja, 1993). With respect to closteroviruses, it is proposed that a family *Closteroviridae* be established which, with other groups of plant viruses (tobamoviruses, tobraviruses, hordeiviruses, bromoviruses, and wheat soil-borne furovirus), has been included in a putative order called *Tobamovirales* (Dolja *et al.*, 1994, Koonin and Dolja, 1993). As recently approved by the International Committee on Taxonomy of Viruses, the family *Closteroviridae* is divided into two genera comprising monopartite (genus *Closterovirus*; type species, BYV) and bipartite closteroviruses (unnamed genus; type species, lettuce infectious yellows virus, LIYV) (Dolja *et al.*, 1994; J. P. Martelli, personal communication). This latest revision of closterovirus taxonomy is used here (Table I). More divisions at the generic level have been proposed in the family, based on differences in genome size and the number of ORFs in monopartite closteroviruses (Dolja *et al.*, 1994; Karasev *et al.*, 1995). Sequencing of more closterovirus genomes should help to define whether these differences may serve a rationale for further subdivisions.

III. Biological Patterns and Cytopathic Effects

Closteroviruses are distributed worldwide, and some of them cause devastating crop losses. The most typical symptoms in herbaceous species are yellowing, veinal necrosis, and leaf-rolling. In woody species, disease symptoms are described as seedling yellows, stem-pitting, and die-back (Bar-Joseph *et al.*, 1979; Milne, 1988). Natural host ranges reportedly vary from narrow to moderate; in artificial inoculation, however, at least one closterovirus, BYV, infects over 100 species in 15 families (Duffus, 1973). Closteroviruses are not seed-borne and are not readily transmitted mechanically, with insects being the only principal vectors. Transmission of BYV and citrus tristeza virus (CTV) by aphids takes place in a semipersistent mode, with acquisition and inoculation feeding times of 15 min to 1 hr, and retention of the virus in the insect for up to 3 days (Bar-Joseph *et al.*, 1979; Murant *et al.*, 1988). A similar transmission mode has been demonstrated for the whitefly-borne closteroviruses LIYV (Duffus *et al.*, 1986), sweet potato sunken vein virus

TABLE I

SOME PROPERTIES OF CURRENTLY RECOGNIZED AND TENTATIVE *CLOSTEROVIRIDAE* FAMILY MEMBERS[a]

Virus (abbreviation)	Particle length (nm)	Genome size (kb)	CP mol. wt (kDa)[b]	Vector[c]	BYV-type vesicles	References
Genus *Closterovirus*						
Alligator weed stunting (AWSV)[d]	1700	NK[e]	NK	NK	Yes	Hill and Zetler (1973)
Beet pseudo-yellows (BPYV)[d,f]	1500–1800	NK	28	W	Yes	Liu and Duffus (1990)
Beet yellows (BYV)	1250–1450	15.5	22, 24	A	Yes	Bar-Joseph and Hull (1974); Agranovsky *et al.* (1994a, 1995)
Beet yellow stunt (BYSV)[d]	1400	NK	NK	A	Yes	Duffus (1972)
Burdock yellows (BuYV)	1600–1750	NK	NK	A	Yes	Nakano and Inoye (1980)
Carnation necrotic flock (CNFV)	1250–1450	NK	23.5	A	Yes	Bar-Joseph *et al.* (1976); Inouye (1974)
Carrot yellow leaf (CYLV)	1600	NK	NK	A	Yes	Yamashita *et al.* (1976)
Citrus tristeza (CTV)[d]	2000	19.3	25	A	Yes	Bar-Joseph and Lee (1989); Karasev *et al.* (1995)
Dendrobium vein necrosis (DVNV)[d]	1865	NK	NK	NK	No	Lesemann (1977)
Diodia yellow vein (DYVV)[d]	NK	NK	NK	W	Yes	Larsen *et al.* (1991)
Festuca necrosis (FNV)[d]	1725	NK	NK	NK	Yes	Schmidt *et al.* (1963)
Grapevine leafroll-associated 1 (GLRaV-1)[d]	1200–2000	NK	39	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)
Grapevine leafroll-associated 2 (GLRaV-2)[d,g]	1400-2200	NK	26	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)
Grapevine leafroll-associated 3 (GLRaV-3)[d]	1400-2200	NK	43	M	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990); Belli *et al.* (1994)
Grapevine leafroll-associated 4 (GLRaV-4)[d]	1400-2200	NK	36	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)

Heracleum virus 6 (HV6)[d]	1400	NK	NK	A	NK	Bem and Murant (1979)
Little cherry (LCV)[d]	1670	~17	46	M	Yes	Ragetti *et al.* (1982); Raine *et al.* (1986); K. Eastwell, personal communication; W. Jelkmann, personal communication
Pineapple mealybug wilt-associated (PMWaV)[d]	1200	NK	24	M	NK	Gunasinghe and German (1989)
Sugarcane mild mosaic (SCMMV)[d]	1500–1600	NK	NK	M	Yes	Lockhart *et al.* (1992)
Wheat yellow leaf (WYLV)	1600–1850	NK	NK	A	Yes	Inouye (1976)
Unnamed genus (bipartite closteroviruses)						
Lettuce infectious yellows (LIYV)	1800–2000[h]	7+8.1	28	W	Yes	Duffus *et al.* (1986); Klaasen *et al.* (1995)
Sweet potato sunken vein (SPSVV)[d]	950	~7.6+8.4	29	W	Yes	Cohen et al. (1992); Winter *et al.* (1992); U. Hoyer, personal communication

[a] Some viruses currently recognized as closteroviruses, i.e., lettuce chlorosis virus (LeCV), tomato infectious chlorosis virus (TICV), cucurbit yellow stunting disorder virus (CYSDV), and grapevine leafroll-associated viruses 5 and 6 (GLRaV-5 and -6), are not included because of the lack of descriptive data; on the other hand, LCV, which is not yet considered a family member by ICTV, is included here since its features appear to be typical of the *Closteroviridae*.

[b] The exact molecular weight values are indicated wherever sequence data were available.

[c] Abbreviations used for vectors are A, aphids; M, mealybugs; W, whiteflies.

[d] Tentative species.

[e] NK, not known.

[f] BPYV has several synonymic designations, i.e., cucumber yellows, muskmelon yellows, or melon yellows virus (J. P. Martelli, personal communication); the question of whether BPYV and cucumber chlorotic spot virus (CCSV; Section V,B) are identical is uncertain and thus awaits comparative tests (L. P. Woudt, personal communication).

[g] As indicated by recent analysis (Boscia *et al.*, 1995), GLRV-2 is identical to the virus earlier described as "grapevine corky bark-associated virus" (GCBaV; Namba *et al.*, 1991).

[h] LIYV has particles of 950 nm modal length according to other measurements (D.-E. Lesemann, personal communication).

(SPSVV; Cohen *et al.*, 1992), and beet pseudo-yellows virus (BPYV; Duffus, 1973). A semipersistent transmission pattern implies a stricter virus–vector specificity as compared with, for example, the non-persistent aphid transmission of potyviruses (Falk and Duffus, 1988). Indeed, among seven aphid species that can transmit CTV, *Toxoptera citridicis* is the most efficient vector; likewise, *Myzus persicae* is the best among 24 species transmitting BYV (Duffus, 1973; Bar-Joseph *et al.*, 1979; Lister and Bar-Joseph, 1981). Interactions of the whitefly-borne closteroviruses with their vectors seem to be even more specific; thus, LIYV and SPSVV can only be transmitted by *Bemicia tabaci* (Cohen *et al.*, 1992; Duffus *et al.*, 1986), whereas BPYV can only be transmitted by *Trialeurodes vaporariorum* (Duffus, 1973). Likewise, the mealybug-borne grapevine leafroll-associated virus 3 (GLRaV-3) and the little cherry virus (LCV) are specifically transmitted by *Pulvinaria vitis* and *Phenacoccus aceris*, respectively (Belli *et al.*, 1994; Raine *et al.*, 1986).

The time needed for aphids to successfully inoculate closteroviruses and the type of the disease symptoms may both reflect the phloem-limited nature of closteroviral infections (Bar-Joseph *et al.*, 1979; Milne, 1988). Data obtained with the help of electronic monitoring indicate that, on BYV transmission, the aphid stylet reaches the phloem in 9–38 min (Limburg *et al.*, 1994).

In the cells, closteroviruses give rise to massive particle aggregates often organized as banded inclusions. The most characteristic type of intracellular inclusions are numerous vesicles surrounded by a membrane, possibly of mitochondrial origin (Essau and Hoefert, 1971; Lesemann, 1988; Kim *et al.*, 1989). These structures, referred to as BYV-type vesicles, are considered an important taxonomic feature of the family, and are suitable for closterovirus diagnosis (Table I). The BYV-type vesicles presumably contain double-stranded RNA (Coffin and Coutts, 1993); it would be interesting to determine whether these structures have anything to do with closterovirus replication.

IV. Particle Structure: Closteroviruses Are Rattlesnakes, Not Just Threads

Closterovirus particles are 950–2000 nm in length, depending on the virus (Table I), and about 12 nm in diameter (Tollin and Wilson, 1988). Virions consist of a single RNA molecule coated by capsid protein (CP) subunits of 22–28 kDa; in some grapevine leafroll-associated viruses, however, the CPs are about 40 kDa (Table I). The particles of BYV and

CTV have a primary helix pitch of 3.5–3.7 nm, with 8.5 and 10 subunits of CP per helix turn, respectively (Bar-Joseph *et al.,* 1972; Chevallier *et al.,* 1983). In the BYV particle, the structure repeats in two helical turns (Chevallier *et al.,* 1983), in contrast to the heracleum virus 6 structure with a five-turns repeat (Tollin *et al.,* 1992).

The fact that closteroviruses possess the most flexible particles among elongated RNA viruses has been a key feature for their recognition as a distinct group (Brandes and Bercks, 1965). This apparent flexibility is reflected in a lower ratio of closterovirus RNA mass to modal particle length (2831–3230/nm) as compared with that for potex-, poty-, and carlaviruses (4038–4112/nm), or tobamoviruses (6666/nm) (Bar-Joseph and Hull, 1974; Bar-Joseph *et al.,* 1979). A loosely wound helix of BYV and CTV particles may account for their sensitivity to ribonuclease (Bar-Joseph and Hull, 1974).

It has been found that the BYV genome encodes a 24-kDa protein (24K) strikingly similar to the 22-kDa CP (22K), and that both these proteins have counterparts encoded in the CTV genome (Boyko *et al.,* 1992). It has been suggested that the genes for CP homologues arose by gene duplication that probably occurred in a common closterovirus ancestor; it is noteworthy that, despite significant divergence, the CP homologues of BYV and CTV have retained the profile of conserved amino acid residues that are believed to ensure the characteristic fold of the filamentous plant virus CPs (Boyko *et al.,* 1992, Dolja *et al.,* 1991). This discovery posed a question of the presence of a second structural protein in the virions. Initially, there were doubts as to whether this is theoretically possible, and researchers in this field (the author being no exception) have long been under the spell of the "single CP" paradigm of elongated plant virus structure (Boyko *et al.,* 1992; Dolja *et al.,* 1994). Then, quite unexpectedly, immunoelectron microscopy (IEM) revealed two serologically distinct segments in the BYV particles: a "tail" selectively labeled with antibodies to the 24-kDa protein (Fig. 1A), and the main segment labeled with antibodies to purified virus and to the recombinant 22-kDa protein (Fig. 1B; D.-E. Lesemann and author, unpublished observations). After decoration with antibodies to the 24-kDa protein, the BYV particles exhibited two peaks of modal lengths corresponding to tailed 1370-nm particles and to 1293-nm tailless particles; the lengths of the anti-24K serum-decorated tails peaked at 75 nm (Agranovsky *et al.,* 1995). The good balance between these lengths illustrates the propensity of the tails to break off and the relative stability of the 22K-coated part against further degradation. In line with this, purified BYV preparations contained a lower portion of the tailed particles as compared with crude

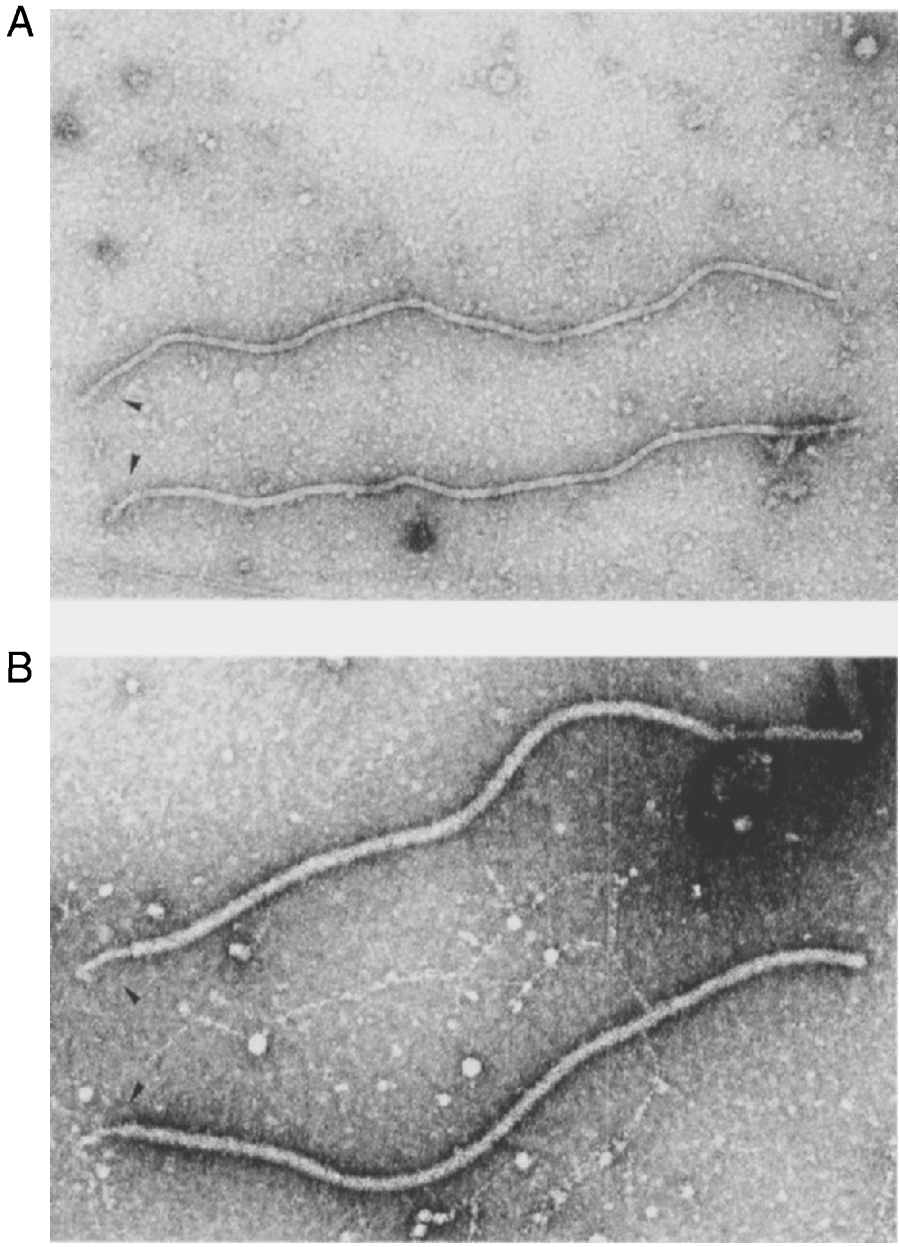


FIG. 1. Immunoelectron micrographs of BYV particles in crude sap of infected *Tetragonia expansa* plants (×95,000 magnification). (A) Decoration with mouse polyclonal antiserum against the N-terminal peptide of the BYV 24K protein. (B) Decoration with rabbit antiserum against purified BYV particles. Grids with the adsorbed virus particles were incubated with the antisera for 15 min (A) and 30 min (B), followed by negative staining with uranyl acetate. Arrows tag the distinct vinon tail. (Courtesy of D.-E. Lesemann.)

sap extracts, and the tails were no longer observed in the particles in the infected plant sap after overnight incubation at room temperature (D.-E. Lesemann, unpublished observations). The discrepancy between the BYV particle lengths determined for purified virus preparations and for leaf-dips (1250 versus 1370–1450 nm; reviewed in Bar-Joseph *et al.*, 1979) may also be at least partially attributed to preservation or loss of the tails.

Strange as it may seem, the morphological polarity of the BYV particles was first noticed much earlier. With the aid of electron microscopy of the methylamine tungstate-stained BYV particles, Hills and Gay (1976) observed an 83-nm terminal region with a helix pitch of 4.05 nm, which was clearly distinct from the main part of the 1390-nm virion having a pitch of 3.45 nm. These observations, which long remained unexplained, are consistent with the IEM data discussed above.

Thus, BYV particles, unlike those of other elongated plant viruses, possess a distinct tail built of multiple subunits of a minor CP, and hence have a "rattlesnake" rather than uniform structure (Agranovsky *et al.*, 1995). Comparison of the lengths of the 24K- and 22K-encapsidated segments (75 vs. 1293 nm) gives a rough estimate of one 24-kDa molecule per 17 molecules of 22K, which is consistent with the relative proportions of the respective subgenomic mRNAs in BYV-infected tissues (Dolja *et al.*, 1990). Moreover, the putative subgenomic promoters for the 22K and 24K mRNAs are similar (see Section VI,C), thus suggesting concerted expression of both structural proteins in viral infection. The involvement of 24K in virion formation is in line with previous computer predictions that the closterovirus CP duplicates have a spatial fold conserved in the monophyletic family of filamentous virus CPs (Boyko *et al.*, 1992; Dolja *et al.*, 1991). Specific decoration of the BYV particles with antibodies against the N-terminal peptide of 24K strongly indicates that the N terminus of the minor CP is exposed on the virion surface, as is the case with CPs of filamentous potex- and potyviruses (Agranovsky *et al.*, 1995; Koenig and Torrance, 1986; Shukla *et al.*, 1988).

It is possible that other closteroviruses have a similar virion structure. There is remarkable CP size heterogeneity in purified CTV preparations, albeit at least partly due to posttranslational modification of the (major) CP (Sekiya *et al.*, 1991; Lee *et al.*, 1988). Intriguingly, LIYV preparations purified in Cs_2SO_4–sucrose gradients contained minor amounts of an approximately 55-kDa protein along with the 28-kDa CP (Klaassen *et al.*, 1994). The putative CP duplicate encoded in the RNA-2 of this virus has a deduced molecular weight of 52 kDa (Fig. 2; Klaassen *et al.*, 1995).

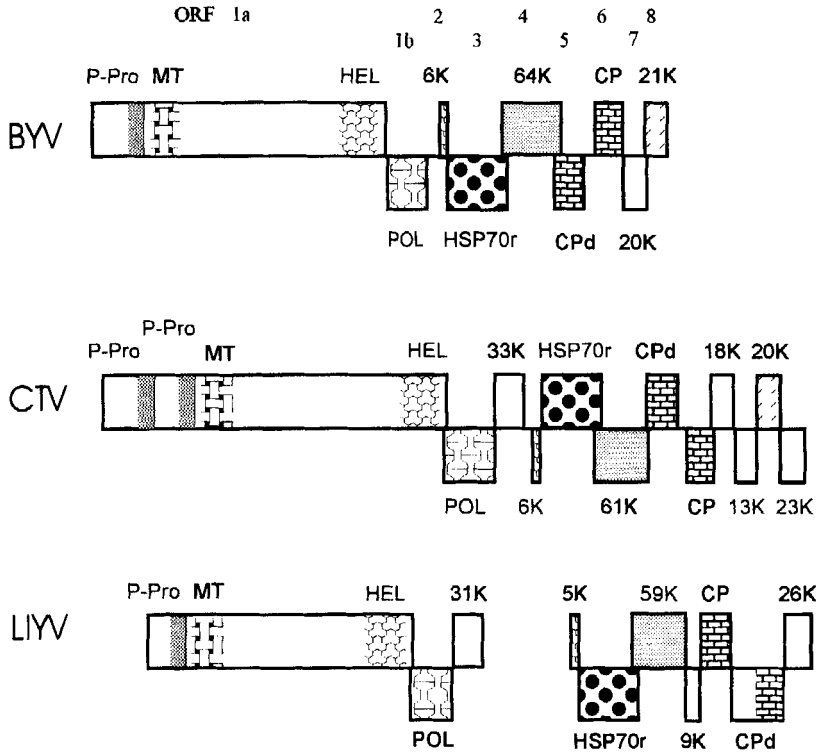


FIG. 2. Comparison of the genome maps of the closteroviruses BYV, CTV, and LIYV. ORFs are shown as boxes, with related domains indicated by the same fill-pattern. P-Pro, papain-like proteinase; MT, methyltransferase; HEL, helicase; POL, RNA-dependent RNA polymerase; HSP70r, HSP70-related protein; CP and CPd, capsid protein and its diverged duplicate.

How is the "rattlesnake" particle assembled? The principal mechanism might follow the classical scheme of TMV self-assembly, which starts at an internal origin of assembly (OAs) in the genomic RNA, and continues by adding CP disks or smaller aggregates in both the 5′ and 3′ directions (reviewed in Dobrov and Atabekov, 1989; Butler, 1984; Mathews, 1991; Lomonossoff and Wilson, 1985). We have assumed the existence of a nucleation region inherent in the BYV RNA that might discriminate between the capsid proteins or their disks, whereupon they proceed to encapsidate the RNA (Agranovsky *et al.,* 1995). In order to identify the BYV particle's end made up of 24K, we have run the following experiment: purified particles were sonicated and treated with antibodies to 24K, and the antibody–virion fragment complexes

were isolated on Protein A-Sepharose. The RNA extracted from the virion fragments thus selected was 5′-end labeled with [γ-^{32}P]ATP and used as a probe to develop Southern blots with cDNA clones representing the 5′-terminal, middle, and 3′-terminal regions of the BYV genome. Compared with the RNA probe prepared from nonfractionated virion fragments, the antibody-selected probe apparently hybridized more strongly to the 5′-terminal clones, thus suggesting that this was the 5′-terminal region of the BYV RNA associated with 24K, and that the putative nucleation signal might be mapped to this region (unpublished data). This is consistent with the fact that none of the 3′-coterminal subgenomic RNAs of BYV produced in infected plants are found in purified virus preparations (Dolja *et al.*, 1990). Closterovirus particle formation may be assisted by virus-encoded nonstructural protein(s); in some filamentous DNA phages carrying a few copies of minor CPs at their ends, assembly is chaperoned by phage-encoded proteins that are not part of the mature particles (reviewed in Russel, 1993). BYV encodes at least one nonstructural protein likely to be instrumental in protein–protein interactions, a 65-kDa homologue of the HSP70 cell heat-shock proteins (see Section V,A).

V. Genome Organization and Functions of Gene Products in Closteroviruses

A. Aphid-Transmissible Monopartite Closteroviruses

1. Beet Yellows Virus

The genome of the Ukrainian strain of BYV (BYV-U) consists of 15,480 nucleotides (nt), is 5′-capped and contains no 3′-poly(A) (Karasev *et al.*, 1989; Agranovsky *et al.*, 1991b, 1994a). Computer translation reveals nine ORFs in the sequence, flanked by 5′- and 3′-untranslated regions of 107 and 141 nt, respectively (Fig. 2). For the 3′ region, two potentially stable hairpins were predicted, and it has been speculated that these may serve as a recognition signal for viral replicase. BYV RNA cannot be aminoacylated or adenylylated *in vitro* and thus apparently has no 3′-tRNA-like structure (Agranovsky *et al.*, 1991a).

The 5′-proximal ORF1a codes for the 295-kDa product which encompasses the domains with methyltransferase and RNA helicase sequence motifs (MT and HEL; Fig. 2) that are conserved in the large subsets of positive-strand RNA viruses (Gorbalenya and Koonin, 1993; Rozanov *et al.*, 1992). The MT domain is believed to be involved in the capping of viral mRNAs. This activity has been experimentally demon-

strated for the alphavirus nsP1 protein (Mi and Stollar, 1991) and TMV 126-kDa protein (Dunigan and Zaitlin, 1990), and suggested for the closely related domains in the Sindbis-like supergroup virus replicases (Ahlquist *et al.*, 1985; Rozanov *et al.*, 1992). Likewise, the HEL domain, whose strand-separating activity was shown experimentally for the potyvirus CI protein (Lain *et al.*, 1990) and the pestivirus NS3 protein (Warrener and Collett, 1995), is implicated in unwinding of RNA duplexes on replication of many other virus groups, based on clear conservation of its sequence in the established and putative helicases (Gorbalenya *et al.*, 1988; Gorbalenya and Koonin, 1989, 1993; Hodgman, 1988). The ORF1b overlaps the last 40 triplets of ORF1a and codes for a product of approximately 53 kDa (Fig. 2) containing the domains of RNA-dependent RNA polymerase (POL) (Kamer and Argos, 1984; Koonin, 1991). Putative BYV replicase, which is likely to be expressed as an ORF1a/1b 348-kDa fusion protein (see below), should have a size about twice that of the other related viral replicases. This difference is due to two unique regions in the putative ORF1a/1b fusion product: a 600-residue N-terminal overhang containing a domain of cystein papain-like proteinase (P-Pro), and a 700-residue central insert harboring a 100-residue stretch that may be related to retrovirus aspartyl proteases (Agranovsky *et al.*, 1994a). The BYV P-Pro was found to be moderately similar to the C-terminal P-Pro domain in helper component proteases (HC-Pro) of potyviruses (Carrington *et al.*, 1989). HC-Pro proteases are multifunctional proteins required for potyvirus transmission by aphids (Atreya *et al.*, 1992; Pirone, 1991) and long-distance spread in plants (Cronin *et al.*, 1995). By analogy, similar function(s) may be proposed for the BYV leader proteinase.

The next downstream ORFs (2, 3, and 4) are arranged as an overlapping block and encode 6.4-, 65-, and 64-kDa products, respectively (Fig. 2). The BYV 6.4-kDa protein (6.4K) shows marginal similarity to the small hydrophobic proteins encoded in the "triple gene block" (TGB) of potex- and carlaviruses (Agranovsky *et al.*, 1991b; Morozov *et al.*, 1989). However, only a part of the residues conserved in the BYV 6.4K and the TGB-encoded proteins may be found in the approximately 6-kDa proteins encoded in CTV, beet yellow stunt virus (BYSV), and LIYV (Karasev *et al.*, 1994b; Klaassen *et al.*, 1995). Though their common origin is thus questionable, the small hydrophobic proteins of closteroviruses and potexviruses may be functionally equivalent. The TGB-encoded proteins of potexviruses bind to membranes *in vitro* (Morozov *et al.*, 1990) and mediate the cell-to-cell transport of the viral infection *in vivo* (Beck *et al.*, 1991). The BYV 6.4K synthesized in

rabbit reticulocyte lysates also showed affinity to cell membranes, and its involvement in the virus infection transport has been suggested (reviewed in Dolja *et al.,* 1994).

The 65-kDa protein (65K) is strikingly similar to the HSP70 family of cell heat-shock proteins (Agranovsky *et al.,* 1991a). HSP70s are ubiquitous molecular chaperones which assist proper folding, oligomerization, and transmembrane transport of other proteins (reviewed in Gething and Sambrook, 1992). Structurally, HSP70s consist of two parts, the N-terminal ATPase domain and the C-terminal peptide-binding domain (reviewed in Craig *et al.,* 1993). The BYV 65K protein contains an N-terminal domain whose sequence and tentative spatial fold are very similar to the HSP70 ATPase, and a unique C-terminal domain that cannot be folded into the $\beta_4\alpha$ (HLA-like) structure typical of the HSP70 peptide-binding domains (Agranovsky *et al.,* 1991a; Rippmann *et al.,* 1991; F. Rippmann, personal communication). Hence, the structure of the putative protein-binding domain of 65K suggests a function different from that of classical chaperones. Karasev *et al.* (1992) first reported that BYV 65K expressed in a cell-free transcription–translation system coprecipitates with purified bovine brain microtubules. The binding of 65K was abolished by pretreatment of microtubule preparations with subtilisin, thus suggesting its specificity. Very recently, bacterially expressed 65K and its fragments have been produced in our laboratory (Nikiphorova *et al.,* 1995). Using a polyclonal antiserum to the C-terminal 13-kDa fragment of 65K, the protein was detected in BYV-infected *Tetragonia expansa* plants (Agranovsky *et al.,* manuscript in preparation). *In vitro* assays showed that the purified BYV 65K, like the cell HSP70s, has magnesium-dependent ATPase activity associated with its N-terminal 40-kDa fragment. However, 65K, unlike its cell homologues, was found to be unable to bind to immobilized denatured protein, and its ATPase activity was not stimulated *in vitro* by sequence-nonspecific peptides (A. Agranovsky, S. Nikiphorova, O. Denisenko, and A. Folimonov, unpublished data). Although these data establish some biochemical characters of the BYV 65K pertinent to its function, the possible involvement of 65K in the cell-to-cell movement of the closterovirus infection (Agranovsky *et al.,* 1991a), which may involve specific interactions with the cell cytoskeleton and translocation machinery (Karasev *et al.,* 1992), awaits experimental support.

Internal segments in the BYV 64-kDa protein and in the equivalent CTV 61-kDa protein reportedly show similarity to a domain in the HSP90 heat-shock proteins (Koonin *et al.,* 1991; Pappu *et al.,* 1994).

However, the related approximately 60-kDa proteins of LIYV and SPSVV fail to display this similarity (Klaassen *et al.,* 1995; author, unpublished observation).

ORF5 and ORF6 code for the 24K and 22K capsid proteins of BYV, respectively (Agranovsky *et al.,* 1991b, 1995). The bacterially expressed BYV 24K and 22K share some common epitopes. Upon tissue fractionation, both proteins bulk in the soluble fraction of the BYV-infected cells, but are also found in the cell wall and membrane fractions (Agranovsky *et al.,* 1994b). The structure of the BYV virions built of two CPs is reminiscent of some other plant RNA viruses, thus implying functional analogy. First, one cannot but recall the rod-shaped furoviruses and spherical luteoviruses harboring a few copies of CP extended by readthrough of a leaky terminator codon in the CP gene (Bahner *et al.,* 1990; Cheng *et al.,* 1994; Filichkin *et al.,* 1994; Richards and Tamada, 1992). Notably, such an aberrant protein has recently been mapped to one end of furovirus particles (Haeberle *et al.,* 1994). The readthrough CP species are held to ensure the persistence of furoviruses and luteoviruses in their respective vectors: fungal zoospores and aphids. Likewise, the semipersistent mode of BYV transmission may be due to the ability of the assembled 24K to cling tightly to cell membranes lining the aphid's alimentary tract. Another (and not necessarily alternative) possibility may be that the 24K tail directs the closterovirus particle to a host (phloem) cell receptor. Conceivably, the fact that p24 is involved in formation of mature virions does not discredit the earlier suggestion that it might participate in the formation of nonvirion ribonucleoproteins adapted for the cell-to-cell transport of genomic RNA (Boyko *et al.,* 1992; Dolja *et al.,* 1994).

ORFs 7 and 8 encode 20- and 21-kDa products, respectively. The latter is related to a 20-kDa protein encoded in the CTV genome (Pappu *et al.,* 1994). Apart from this, these products have shown no significant similarities to any proteins in the current database. Recently we produced a polyclonal antiserum against the BYV 21-kDa protein purified from bacteria; using this antiserum to develop Western blots, the 21-kDa protein was detected in soluble and membrane fractions of BYV-infected plants (R. Zinovkin and author, unpublished data).

The German and British strains of BYV (BYV-G and BYV-B) have been partially sequenced (Agranovsky *et al.,* 1994a; Brunstedt *et al.,* 1991), allowing their comparison with BYV-U. BYV-U and BYV-G showed 88.5% identity of the nucleotide sequences and the same organization of ORFs 2 to 8 within the 3′-terminal 6-kb region. The majority of nucleotide substitutions in the BYV-G sequence are in the third

positions of codons; even when the substitutions change the coding, only about half of the resulting amino acid changes are nonconservative. Nevertheless, the data compiled in Table II indicate some differences in the extent of conservation of individual protein sequences in the two BYV strains. Proteins 65K, 64K, 24K, and 22K are the best conserved among the strains, whereas the low-molecular-weight proteins (6.4K, 20K, and 21K) are apparently more variable (Table II). There is also a remarkable nucleotide sequence conservation of the intergenic and 3′-untranslated regions among the two strains, suggesting the functional importance of these regions (Table II). The partial sequence of BYV-B shows the same disposition of ORFs 4 to 7 (Brunstedt *et al.*, 1991). The West European strains are apparently closer to each other than to the Ukrainian strain; in particular, the 22-kDa capsid protein sequences of the BYV-G and BYV-B are identical (Table II).

TABLE II

COMPARISON OF THE NUCLEOTIDE AND AMINO ACID SEQUENCES AMONG THE THREE BYV STRAINS

Gene product	ORF2 64K	ORF3 65K	ORF4 64K	ORF5 24K	ORF6 22K	ORF7 20K	ORF8 21K
Strain							
U/G[a]	11[b]	4	8	2	5	18	12
U/B	—	—	7	2	5	—	—
G/B	—	—	2	2	0	—	—
Nontranslated regions			I		II		3′ end
Strain							
U/G			14		14		4
U/B			—		13		—
G/B			—		1		—

[a] U, G, and B stand for the Ukrainian, German, and British strains of BYV, respectively. The available nucleotide sequences of the BYV-G and BYV-B align with nt 9375–15353 and 11684–14407, respectively, in the complete BYV-U sequence.

[b] Percent of amino acid (for polypeptide products) or nucleotide (for noncoding regions) substitutions revealed on pairwise comparisons of strains. Dashes indicate positions where no sequence for the British strain was available. Nontranslated region 1 is between ORFs 1b and 2, and nontranslated region II is between ORFs 5 and 6.

2. Citrus Tristeza Virus

Among plant RNA viruses, CTV has the largest undivided genome (19,296 nt for the Florida T36 isolate), exceeding that of BYV by about 4 kb (Karasev *et al.*, 1994b, 1995; Pappu *et al.*, 1994). The overall genome structure of CTV is similar to that of BYV, comprising the P-Pro, MT, HEL, and POL domains; the small hydrophobic protein; the HSP70 homologue; the 61-kDa protein related to BYV 64K; the 27-kDa CP homologue (27K); the 25-kDa CP; and the 20-kDa protein homologous to the BYV 21-kDa protein (Fig. 2). On the other hand, the CTV genome encodes some proteins or polyprotein domains that are not conserved in BYV. Interestingly, the P-Pro domain in the ORF1a product of CTV is duplicated (Fig. 2; Karasev *et al.*, 1995). Pairwise comparisons of the putative CTV leader proteins of predicted molecular weights 54 and 55 kDa and the BYV 66-kDa protein revealed no sequence similarity among the three proteins apart from the C-terminal 150-residue part encompassing the P-Pro domain. Among the products encoded by the 3′-proximal genes of CTV, the 20-kDa protein (20K) is related to the BYV 21-kDa protein (21K), whereas the 33-, 13-, 18-, and 23-kDa proteins have no homologues in other sequenced closterovirus genomes (Karasev *et al.*, 1995; Pappu *et al.*, 1994). The 23-kDa protein contains a sequence motif enriched in cysteine and basic residues, which is conserved in putative nucleic acid binding proteins encoded in the 3′-proximal genes in carlaviruses and allied viruses. Therefore, this putative protein has been implicated in RNA binding and regulation of CTV gene expression (Dolja *et al.*, 1994).

Severely pathogenic CTV isolates share a common epitope on their particles not found on the particles of mild isolates, thus suggesting the CTV pathogenicity may have some of its determinants associated with the 25-kDa CP (Pappu *et al.*, 1993). The CTV 27K and 20K proteins were detected in infected plants with polyclonal antibodies against the recombinant proteins (Febres *et al.*, 1994; Pappu *et al.*, 1994). The bulk of 27K was found in the cell wall fraction of infected citrus leaves, although the protein was also detectable in the soluble and membrane fractions (Febres *et al.*, 1994). Thus, the 27K association with subcellular fractions differs somewhat from that reported for the homologous BYV 24K which, like BYV 22K, is predominantly found in the soluble fraction (Agranovsky *et al.*, 1994b). Recent experiments with yeast two-hybrid system have shown that the CTV 20K (a homologue of the BYV 21K) is capable of homologous interactions, thus suggesting that this protein might function as a di- or multimer (S. Gowda, personal communication).

Sequencing of a 3′-terminal 2.5-kb portion of another CTV isolate, Israeli VT or "seedling yellows" isolate, revealed four ORFs encoding 18-, 13-, 20-, and 23.5-kDa proteins that showed close relatedness to the respective products of CTV-T36 (Mawassi *et al.*, 1995a).

3. Carnation Necrotic Fleck and Beet Yellow Stunt Viruses

Comparisons of partial sequences of BYSV and carnation necrotic fleck virus (CNFV) indicate their relatedness to CTV and BYV (Dolja *et al.*, 1994; Karasev *et al.*, 1994a,b; Klaassen *et al.*, 1995). Both the BYSV and CNFV genomes contain a conserved array of ORFs coding for POL, a small hydrophobic protein, and an HSP70 homologue. In addition, the BYSV genome bears an approximately 30-kDa protein gene inserted between the POL and small hydrophobic protein genes. This configuration is similar to that in the respective part of the CTV genome (Fig. 2). The fact that BYV and CNFV induce very similar patterns of dsRNAs in the infected plant cells (Dodds and Bar-Joseph, 1983) is indicative of the overall similarity of their gene layouts.

B. Whitefly-Transmissible Mono- and Bipartite Closteroviruses

The genome of lettuce infectious yellows virus (LIYV) is divided among RNA-1 and RNA-2 components of 8.1 and 7.2 kb, respectively (Klaassen *et al.*, 1995). Interestingly, LIYV RNA-1 and RNA-2 show no similarity between their respective terminal untranslated regions (which may be expected to contain putative recognition signals for the replicase), with the exception of the 5′-terminal pentanucleotide, which is identical in both genomic components (Klaassen *et al.*, 1995). LIYV RNA-1 encompasses the overlapping ORFs 1a and 1b coding for the putative P-Pro and the replicative domains, and the 3′-terminal ORF for a 31-kDa protein (Fig. 2). Very recently, a full-length cDNA copy of the LIYV genome was produced, and it was found that the RNA-1 T7 transcript is necessary and sufficient to support the replication in protoplasts (B. W. Falk, personal communication). This is the first experimental evidence for the assignment of closterovirus replicative functions to the domains conserved in ORF 1a/1b. LIYV RNA-2 contains genes for the small hydrophobic protein, the HSP70 homologue, the 59-kDa protein distantly related to the BYV 64-kDa and CTV 61-kDa products, the 9-kDa protein, the 28-kDa CP, the 52-kDa protein whose C-terminal domain is homologous to the CP, and the 26-kDa protein (Fig. 2; Klaassen *et al.*, 1995). SPSVV RNA-1 and RNA-2 have sizes comparable to those of their counterparts in the genome of LIYV and show a similar organization of ORFs, with the biggest difference being

that SPSVV RNA-2 encodes a putative CP duplicate of 79-kDa (U. Hoyer, E. Maiss, W. Jelkmann, and J. Vetten, unpublished data).

The monopartite genome of another whitefly-transmissible closterovirus, cucumber chlorotic spot virus (CCSV), has a size of approximately 15.5 kb (Woudt *et al.*, 1993a,b). The sequence of its coding part shows 5′-terminal overlapping ORFs encoding the domains of the leader P-Pro, MT, HEL, and POL; 3′-proximal ORFs code for the small hydrophobic protein, the HSP70 homologue, the approximately 60-kDa protein, the 9-kDa protein, the 28-kDa CP, the 74-kDa protein containing the C-terminal domain homologous to the CP, and the 23-kDa protein (L. P. Woudt, personal communication). Thus, the undivided genome of CCSV shows an overall arrangement of genes unexpectedly similar to that of the bipartite genomes of LIYV and SPSVV. In line with this, comparisons of the encoded proteins showed close relatedness among CCSV, LIYV, and SPSVV, suggesting that the three whitefly-transmissible closteroviruses constitute a distinct evolutionary lineage (Fig. 3). The LIYV, SPSVV, and CCSV ORF 1a and 1b products (including the N-terminal leader proteins) can be confidently aligned with high statistical scores over almost the entire protein length, whereas their similarity to the equivalent products of BYV and CTV is essentially confined to the core replicative domains. The same is true for the encoded HSP70 homologues, the approximately 60-kDa proteins, and the 28K-kDa CPs. The LIYV 52-kDa, SPSVV 79-kDa, and CCSV 74-kDa capsid protein duplicates show closest relatedness within the approximately 200-residue C-terminal segments, including the CP-like core domains. The gene for a putative 9-kDa protein located upstream of the CP gene is unique for the whitefly-borne closteroviruses; comparison of the 9-kDa sequences encoded in LIYV, SPSVV, and CCSV showed moderate conservation. The presence of a gene for an approximately 30-kDa product located downstream of the POL gene is common for CTV, BYSV, and the bipartite closteroviruses, although the relatedness of the encoded products is not apparent (Dolja *et al.*, 1994; Karasev *et al.*, 1994b; Klaassen *et al.*, 1995). The CCSV genome does not encode a product of similar size and location.

C. Other Closteroviruses

The 3′-terminal 8.3-kb sequence of the mealybug-transmissible LCV has been recently determined (R. Keim-Konrad and W. Jelkmann, manuscript in preparation). In the 5′ to 3′ direction, the sequence encompasses the conserved ORFs for the HSP70 homologue, the ~60K protein, the 46-kDa (putative) CP, the 76-kDa CP duplicate, and

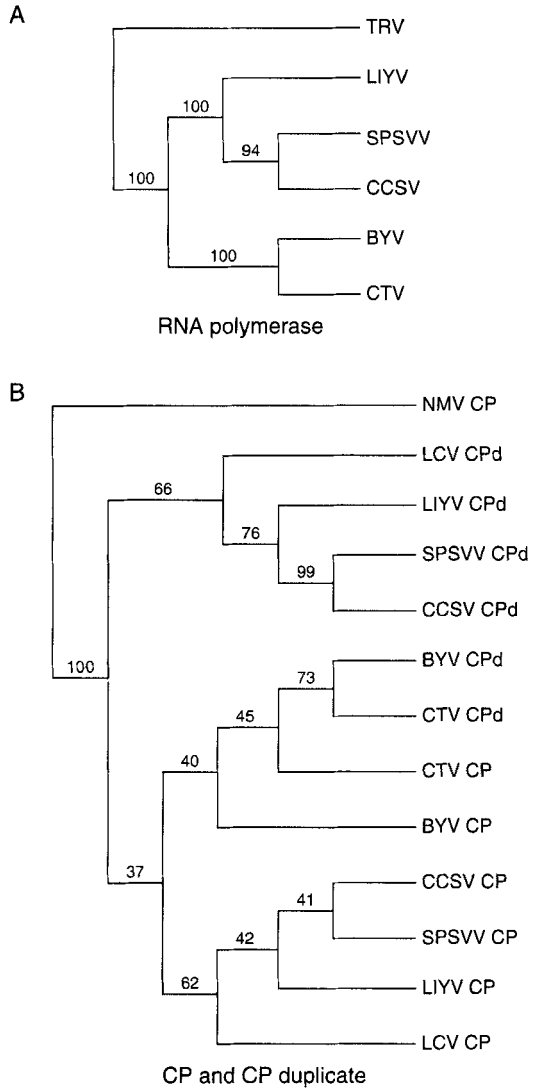


FIG. 3. Tentative phylogenetic reconstructions derived from aligned sequences of RNA-dependent RNA polymerase (A) and capsid protein core (B) domains of closteroviruses. Multiple alignments were produced by the program OPTAL (Gorbalenya *et al.*, 1989). Trees were constructed using the program PROTPARS of the PHYLIP package (Felsenstein, 1989) from the alignments of 294 and 98 amino acid residues of polymerases and CPs, respectively, excluding the positions containing gaps. The number above each node shows the percentage of bootstrap replicates in which a given node was recovered. Branch lengths are arbitrary. The related sequences of tobacco rattle tobravirus (TRV) polymerase and narcissus mosaic potexvirus (NMV) CP were used as the outgroups in the respective trees. Protein sequences were extracted from database, except for CCSV, SPSVV, and LCV (personal communications from L. P. Woudt, U. Hoyer, and W. Jelkmann, respectively).

unique ORFs for 21- and 27-kDa proteins (W. Jelkmann, personal communication). The genome of GLRaV-3 reportedly contains the HEL and POL domains, a gene for the HSP70 homologue, and the 43-kDa CP gene (Ling *et al.,* 1994), but their arrangement has not been yet described. Future sequencing and comparison efforts are expected to elucidate relationships among mealybug-transmissible closteroviruses.

D. General Outlook of Genome Structure and Gene Functions

In closterovirus genome organization, variation of a common theme is evident. The invariant elements include the P-Pro, MT, HEL, and POL domains; the small hydrophobic protein; the HSP70 homologue, the ~60K protein; and the CP and its duplicate. Interestingly, the 65K, 64K and CP genes that are conserved at the species level are apparently least divergent in the BYV strains. It is plausible that these genomic elements provide for the characteristic biological patterns common to all the members of this family. Wherever differences in the conserved genes occur, they intriguingly parallel modifications of the biological features. For example, the genes for the (major) CP and its duplicate are much more divergent and are transposed in the genomes of whitefly-transmissible CCSV, LIYV, and SPSVV as compared to aphid-transmissible BYV and CTV. In the genome of a mealybug-transmissible LCV, the CP and CP-duplicate ORFs are arranged like those in the whitefly-transmissible species; however, their amino acid sequences are divergent from those of the other family members (Fig. 3), and the LCV CP is notably large. Could the disparity among the closterovirus CPs and CP duplicates be due to their involvement in and specific adaptation to transmission by different types of vector?

The obvious similarities in gene arrangement and the encoded protein sequences between the whitefly-borne monopartite and bipartite closteroviruses clearly indicate that the lineages represented by CCSV, LIYV, and SPSVV on the one hand, and by BYV and CTV on the other, had diverged at an evolutionary stage preceding the splitting of the closterovirus genomes. This has important implications in the taxonomy of *Closteroviridae*; clearly, the second genus (Table I) should include not only the bipartite closteroviruses but also CCSV (and perhaps other whitefly-transmitted closteroviruses). Broadly, sequencing of more closterovirus genomes is expected to shed some light on whether the adaptation to vectors has been a key factor of their molecular evolution.

In closterovirus genomes, the 3′-proximal ORFs vary in number and, as a rule, encode nonconserved protein sequences. The 3′-most ORFs

have perhaps diverged most rapidly among the BYV strains (Table II). This might indicate that their products are involved in functions connected with fast environmental response, such as modulation of symptom expression or adaptation of the virus to changes in the host or vector populations. A number of BYV and CTV isolates have been reported, differing in the severity of symptoms they cause in host plants (reviewed in Moseley and Hull, 1990; Pappu *et al.*, 1993; Rogov *et al.*, 1993).

Closterovirus genomes appear to have modular organization (reviewed in Dolja *et al.*, 1994). In the beginning, the existence of three modules in the BYV genome was envisaged, represented by overlapping gene blocks separated by two uridine-rich spacers (Agranovsky *et al.*, 1991b). It has been suggested that these modules have evolved as distinct entities and that they encode proteins expressed very early (replicase), early (putative transport proteins), and late in the infection (CP and two proteins of unknown function). Revisions of this scheme have been proposed (Dolja *et al.*, 1994; Karasev *et al.*, 1995); they agree in placing the genes for the CP and its duplicate into the 3′-terminal module, which seems reasonable. Further, it may be speculated that some closterovirus genes (the approximately 30-kDa ORF and the unique 3′-terminal ORFS) evolved independently of the conserved modules.

VI. Expression Strategy

A. Papain-like Proteinase

Computer-assisted predictions and *in vitro* experiments have demonstrated that the 295-kDa product encoded in the 5′-most gene of BYV is in fact a polyprotein (Agranovsky *et al.*, 1994a). The P-Pro domain located in its N-terminal portion mediates autoproteolysis at the Gly–Gly bond to release a 588-residue (66-kDa) leader protein and a C-terminal 229-kDa protein with MT and HEL domains. The catalytic Cys and His residues in the P-Pro active center and its cleavage site (inferred from alignment with potyviral P-Pro domains) have been confirmed experimentally using point mutagenesis and *in vitro* translation. It has been found that deletion of 245 residues from the N terminus of the leader protein does not impair but rather stimulates the cleavage, and that the His and Cys residues, which are not conserved in the related thiol proteinases, have different effects on autoproteolysis of the BYV polyprotein (Agranovsky *et al.*, 1994a). This is consistent with the results obtained for a related P-Pro of the chestnut blight hypovirulence-associated dsRNA virus (Choi *et al.*, 1991).

In the CTV ORF1a product, the P-Pro domain is duplicated (Karasev *et al.*, 1995). Based on alignment of the CTV and BYV P-Pro sequences, the cleavage sites in the CTV polyprotein have been predicted at the Gly–Gly doublets at positions 484–485 and 976–977 from the N terminus. Hence, processing of the 349-kDa ORF1a product of CTV would yield two leader proteins of 54 and 55 kDa, and a 240-kDa (C-terminal) protein with MT and HEL domains (Karasev *et al.*, 1995). Putative leader P-Pro domains may also be revealed on computer analysis of the ORF1a products of LIYV (Klaassen *et al.*, 1995), SPSVV (U. Hoyer, personal communication), and CCSV (L. P. Woudt, personal communication). The tentative P-Pro cleavage sites in the ORF1a products of LIYV, SPSVV, and CCSV deviate from the consensus drawn for the BYV, CTV, and potyvirus proteinases, being VG/A, LG/V, and VG/V, respectively. If autocatalysis indeed occurs at these sites, the respective leader proteins should have 412, 496, and 402 residues, respectively.

B. Ribosomal Frameshifting

In the genomes of BYV (Agranovsky *et al.*, 1994a), CTV (Karasev *et al.*, 1995), CCSV (ten Dam, 1995), and bipartite closteroviruses (Klaassen *et al.*, 1995; U. Hoyer, E. Maiss, W. Jelkmann, and J. Vetten, unpublished data), the HEL and POL domains are split between the products of overlapping 5′-proximal ORFs found in 0/+1 configuration, thus indicating that the polymerase may be expressed via +1 ribosomal frameshifting. Although many viral RNA polymerases are expressed as frameshift fusions resulting from translation of overlapping genes, the 0/+1 configuration of the closterovirus replication-associated ORFs is quite unusual [to my knowledge, the only other example is the dsRNA virus of *Leishmania* (Stuart *et al.*, 1992)]. In all other cases, including retroviruses (Jacks and Varmus, 1985; Jacks *et al.*, 1988), dsRNA-containing viruses (Dinman *et al.*, 1991), and the diverse groups of positive-strand RNA viruses of animals and plants (Brierly *et al.*, 1987; Godeny *et al.*, 1993; Jiang *et al.*, 1993; Mäkinen *et al.*, 1995; Miller *et al.*, 1988; Xiong and Lommel, 1989), an upstream ORF and the downstream (POL) ORF are found in 0/–1 configuration. Hence, the tentative frameshifting mechanism in closteroviruses deserves a special comment.

The canonical mechanism of leftward (or –1) frameshifting postulates a one-step-back movement ("simultaneous slippage") of two tRNAs bound to a "shifty" mRNA sequence X XXY YYZ, to decode it as XXX YYY (Jacks *et al.*, 1988). The reading-frame switching is stimulated by a pseudoknotted secondary structure (Brault and Miller, 1992; Prüfer *et al.*, 1992; ten Dam *et al.*, 1990); this effect is probably con-

nected with the ability of such a structure to impede the progress of ribosomes along the template (Tu *et al.,* 1992). Rightward (or +1) frameshifting has been described for the yeast *Ty* retrotransposons and the *E. coli* release factor gene (Clare *et al.,* 1988; Craigen *et al.,* 1985; reviewed in Farabaugh, 1993). In retrotransposons, the frameshifting is enhanced by a rare "hungry" codon adjacent to the shifty codon (Farabaugh *et al.,* 1993), whereas in the bacterial gene stimulation is provided by the in-frame stop codon and a downstream Shine-Daigarno-like sequence transiently interacting with ribosomal 16S RNA (Weiss *et al.,* 1988). In short, as follows from comparisons of different frameshifting mechanisms, the reading-frame switching requires some signal(s) to slow down the translating ribosome, thus increasing the chances of the out-of-frame triplet recognition.

The BYV ORF1a ends in a GGGUUUA sequence resembling the "shifty" heptamers of the retroviral type. This resemblance, which we could not but mention in an earlier work (Agranovsky *et al.,* 1994a), is probably fortuitous, as such a heptamer is not conserved in the other closterovirus genomes (Fig. 4; Karasev *et al.,* 1995; Klaassen *et al.,* 1995; ten Dam, 1995). Notwithstanding, we did not suggest the "slippery" consensus to provide for the +1 frameshifting in the BYV svstem; rather, our explanation was based on the "U33 grapple" pairing model (Weiss, 1984). Specifically, offset pairing was postulated between U-7998 in the ORF1a UAG stop codon and the nucleotide located leftward to the anticodon of $tRNA^{Val}$ to mediate transition of a subset of translating ribosomes into ORF1b (Agranovsky *et al.,* 1994a). In accord with this, the (G/C)UU U** consensus (where ** designates the last two bases in stop codons) is seen at the 3′ termini of ORF1a in BYV, LIYV, SPSVV (Fig. 4), and CCSV (ten Dam, 1995). Further, it may be speculated that putative secondary structure elements at the BYV ORF1a stop codon (Agranovsky *et al.,* 1994a) serve to stall the ribosome, thus promoting the frameshifting. At least partially, this RNA fold is conserved in the respective genome regions of CTV and CCSV (Fig. 5; ten Dam *et al.,* 1995), but not in the LIYV genome (Klaassen *et al.,* 1995).

Alternative frameshifting models have been proposed for CTV and BYV (Karasev *et al.,* 1995) and for LIYV (Klaassen *et al.,* 1995). Superposition of the nucleotide and protein sequences in the CTV and BYV HEL/POL gene overlaps reveals a remarkable amino acid conservation profile, suggesting that the frameshifting in the CTV gene occurs after the GUU valine codon, which is not the penultimate triplet there (Fig. 4; Karasev *et al.,* 1995). By analogy with the yeast retrotransposon system (Farabaugh *et al.,* 1993), the putative +1 frameshifting signal is postulated to be simple, not to include any secondary

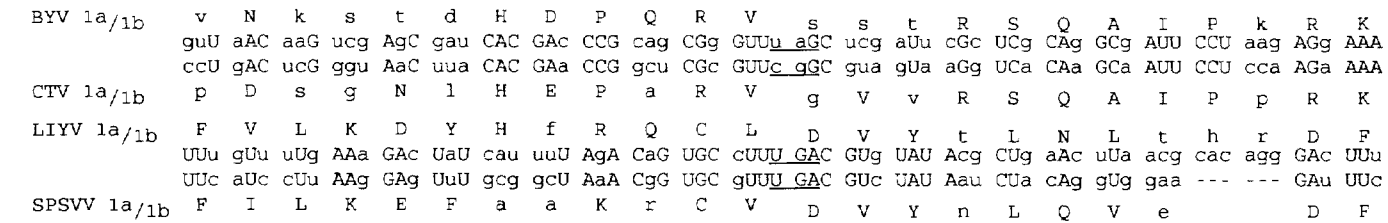


FIG. 4. Alignment of the nucleotide and amino acid sequences in the region of the putative ribosomal frameshift in the genomes of BYV, CTV, LIYV, and SPSVV. Residues conserved in two or more sequences are capitalized. Stop codons in the BYV, LIYV and SPSVV ORF1a and the respective Arg codon in the CTV ORF 1a are underlined. The ORF1a of CCSV ends in the CGG CGA GUUUGA sequence (L. P. Woudt, personal comrnunication) that is similar to those in LIYV and SPSVV.

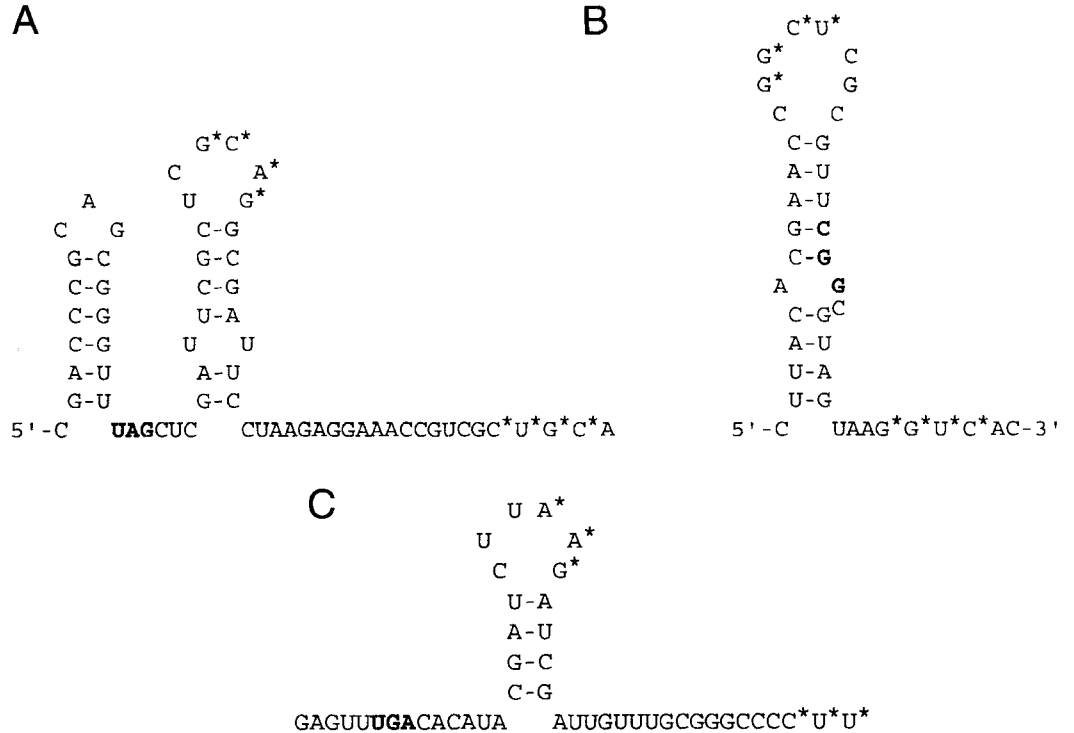


FIG. 5. Proposed structural fold for the regions of the putative ribosomal frameshift in the genomes of BYV (A), CTV (B), and CCSV (C). The UAG stop codon in the BYV ORF1a and the respective CGG arginine codon in the genome of CTV are shown in bold. Nucleotides potentially involved in pseudoknot formation are marked by asterisks. CCSV fold has been proposed by ten Dam (1995).

structure, and to use codons that may cause ribosome pausing, namely the rare CGG (arginine) codon in the CTV ORF1a or UAG stop codon in the BYV RNA (Dolja *et al.*, 1994; Karasev *et al.*, 1995). For LIYV, Klaassen *et al.* (1995) proposed that the +1 frameshift occurs by slippage of $tRNA^{Lys}$ on the AAAG string located eight triplets upstream of the stop codon in the ORF1a.

Experimental evidence has been obtained for ribosomal frameshifting on expression of closterovirus ORFs 1a and 1b, and we may not have to wait long for elucidation of the frameshift mechanism. An expression cDNA clone was produced, which contained, in a heterologous context, a 113-nt CCSV-specific fragment encompassing the potential GUU UGA "shifty stop" sequence and a tentative pseudoknot downstream (ten Dam, 1995). Translation of the SP6 transcript of this clone in wheat germ and rabbit reticulocyte-cell-free systems yielded ^{35}S-methionine-labeled products consistent with the expression of CCSV ORF1b via ribosomal frameshifting with an efficiency of about 2%. It is worth mentioning that the minimal "shifty stop" sequence of CCSV was found to be incapable of frameshifting *in vitro*, suggesting that a more elaborate signal must be involved (ten Dam, 1995). Likewise, we have produced an expression clone containing the BYV-specific insert encompassing the ORF1a/1b overlap. Translation of the T7 transcript of this clone in rabbit reticuloctye lysate resulted in ribosomal frameshifting with an efficiency of less than 1% (Agranovsky, Zelenina, and Morozov, unpublished data).

C. Subgenomization of 3′-Proximal Genes

Closterovirus genes located 3′-ward of the POL gene are likely to be expressed via formation of 3′-coterminal subgenomic (sg) RNA species. Plants infected with CTV, BYV, CNFV, SPSVV, LIYV, BPYV, LCV, and other clostero-like viruses contain a variety of dsRNA species, of which some may correspond to the subgenomic size messengers (Coffin and Coutts, 1992; Dodds and Bar-Joseph, 1983; Gunasinghe and German, 1989; Hu *et al.*, 1990; Larsen *et al.*, 1991; Namba *et al.*, 1991; Winter *et al.*, 1992; K. Eastwell, personal communication).

In BYV-infected plants, six species of double- and single-stranded RNAs have been identified by Northern blot hybridization, corresponding to the full-sized genomic RNA and to sgRNAs for the 65, 64, 24, 22, and 21K ORFs. The identity of the BYV sgRNAs for the 24, 22, and 21K ORFs is supported by *in vitro* translation of the respective dsRNA species denatured with methyl mercuric hydroxide, which yielded proteins compatible in size with those deduced for the ORF 5, 6, and 8

products; the product of the most abundant 1.6-kb dsRNA (corresponding to the sgRNA for 22-kDa CP) was immunoprecipitable with an antiserum to BYV particles (Dolja *et al.*, 1990). None of the BYV sgRNAs was found to be encapsidated (Dolja *et al.*, 1990).

The dsRNA patterns produced in CTV-infected plants vary greatly from strain to strain (Dodds *et al.*, 1987; Guerri *et al.*, 1991; Moreno *et al.*, 1990). For at least one particular CTV isolate, T36, this pattern was found to be stable when the virus had been propagated in different citrus hosts (Hilf *et al.*, 1995). Comprehensive Northern blot analysis of single- and double-stranded RNAs from CTV-T36-infected plants demonstrated the presence of nine 3′-coterminal sgRNA species representing the 33, 65, 61, 27, 25, 18, 20, and 23K ORFs (Hilf *et al.*, 1995). The 3.2-kb (CP) sgRNA of CTV-T36 (Hilf *et al.*, 1995), as well as the 3.2-, 1.6-, and 0.9-kb sgRNAs coding, respectively, for the CP, 20-kDa, and 23-kDa proteins of CTV-VT (Mawassi *et al.*, 1995a), were found to be encapsidated. In addition, the CTV-VT encapsidates a 2.4-kb RNA that possesses properties of a defective RNA; as revealed by sequencing, this species is composed of 1.1-kb and 1.3-kb regions derived from the 5′ and 3′ termini of the CTV genome (Mawassi *et al.*, 1995b). Conceivably, the presence of this defective RNA in virions implies that the putative origin of assembly is in the outskirts of the CTV genome (Mawassi *et al.*, 1995b).

The 5′ termini of the BYV sgRNAs for the major and minor CPs (22 and 24K) were mapped by primer extension to the adenosine residues found 52 and 105 nt upstream of the respective initiating codons (Agranovsky *et al.*, 1994b). The sequence at the starts of both sgRNAs of BYV is conserved (CCAUUUYA; Y for pyrimidine) and may thus represent a core element of the subgenomic promoter. Interestingly, this element resembles the sequences at the 5′ ends of the CP sgRNAs of tobamoviruses and *Bromoviridae* family members. Bearing in mind that the BYV repilcase is most closely related to those of the tobamolike viruses, it is tempting to speculate on parallel conservation, in the process of evolution, of template-binding domains in viral replicases and the signals they recognize in viral RNAs (Agranovsky *et al.*, 1994b). Recently, we mapped the 5′ end of the BYV sgRNA for the 64-kDa protein at the adenosine residue located 141 nt upstream of the ORF4 initiating codon (M. Vitushkina and author, unpublished data). The sequence at the respective start site, ACAUAAUU, significantly deviates from the consensus derived for the 22 and 24K sgRNAs. This, together with the fact that no sequence elements conforming to the CCAUUUYA consensus can be seen in the BYV genome sequence upstream of the AUG codons in ORFs 2, 3, 4, 7, and 8, suggests that the

BYV replicase may interact with different types of subgenomic promoters, thus providing for transcriptional regulation of closterovirus gene expression.

D. Peculiarities of Closterovirus Genome Expression

For genome expression, BYV and possibly other closteroviruses combine autoproteolysis by a papain-like proteinase, ribosomal frameshifting, and sgRNA formation, thus resembling the animal viruses belonging to the corona-like superfamily rather than any other known plant virus group (Agranovsky *et al.*, 1994a). The situation with BYV and CTV, whose genomes contain single and double P-Pro domains, respectively, further parallels that in coronaviruses and arteriviruses, some of which show similar P-Pro duplication (Godeny *et al.*, 1993; Karasev *et al.*, 1995; Lee *et al.*, 1991: Snijder and Horzinek, 1993). As closteroviruses and corona-like viruses represent evolutionarily disparate lineages (Koonin, 1991; Koonin and Dolja, 1993), it seems plausible that similar expression strategies in these groups have evolved independently to confer an advantage in expression of large RNA genomes.

Expression of the 5′-proximal genes in closterovirus genomes should produce some proteins in unequal amounts. Thus, translation of the BYV genomic RNA should yield the major 295-kDa protein and a fusion 348-kDa protein processed into the 66-, 229-, and 282-kDa proteins. In CTV, translation should result in 349- and 401-kDa polyproteins further processed into the 53- and 54-kDa cleaved leaders, and 240- and 290-kDa proteins (Karasev *et al.*, 1995). The synthesis of closterovirus ORF1a/1b fusion proteins (containing the complete array of replication-associated domains) is perhaps down-regulated (see Section VI,B), as is the case with other virus systems employing translational frameshifting (Brault and Miller, 1992, and references therein).

In many positive-strand RNA virus genomes, one can discern a trend to regulate the expression of POL and other replication-associated domains (1) by using a leaky nonsense codon or a frameshift signal to isolate the sequence coding for POL from the upstream coding sequence, and (2) by expressing the POL, MT, and HEL (or MT+HEL) domains as distinct products resulting either from polyprotein processing or from translation of individual genomic RNAs. Splitting of the viral replicase into distinct components, whose expression may be regulated separately, is likely to provide the required flexibility in performing different enzymatic functions in RNA replication, namely, unwinding of duplexes, asymmetric synthesis of (+) and (–) strands, synthesis of subgenomic RNAs, and RNA capping (Agranovsky *et al.*,

1994a). Unlike closteroviruses, all other plant viruses that utilize frameshift for POL expression have small genomes and encode neither MT nor HEL (Koonin and Dolja, 1993), whereas in corona-like virus genomes both the POL and HEL domains (in this order) are located 3′-ward of the frameshift site (Snijder and Horzinek, 1993). Thus, closteroviruses are the only viruses known so far in which the frameshift occurs between the sequences coding for HEL and POL.

VII. Evolution of Closterovirus Genomes

A. RNA Genome and RNA Replicase: Coordinated Evolution?

Comparisons of the MT, HEL, and POL sequences reveal close similarity of closteroviruses to tobamo-, tobra-, furo-, hordei-, idaeo-, bromo-, and ilarviruses, which comprise a compact "tobamo" lineage (Koonin and Dolja, 1993) within the Sindbis-like supergroup of positive-strand RNA viruses (Goldbach *et al.,* 1991). Apart from the conserved replicative core that has been vertically inherited from an ancestor shared with tobamo-like viruses, closterovirus genomes show elements of most probably horizontal acquisition (Fig. 6). This concerns the 65-kDa protein evidently homologous to the HSP70 family of cell chaperones, the CP and its duplicate, and (less likely) the leader P-Pro related to the potyvirus HC-Pro. The capture of foreign genes and intragenomic sequence duplication might be driven by the same mechanism, i.e., copy-choice RNA recombination (Kirkegaard and Baltimore, 1986; Wang and Walker, 1993). On evolutionary divergence of closteroviruses, some of these elements underwent further shuffling. Thus the sequence coding for the leader P-Pro was duplicated in the CTV genome, and the gene for the CP homologue was extended and moved downstream of the (major) CP gene in the genomes of mono- and bipartite whitefly-borne closteroviruses (or vice versa). The N- and C-terminal domains of the CP homologue of LIYV both showed similarity to the CP, so it cannot be ruled out that a triplication of the CP gene has occurred in the LIYV genome (Klaassen *et al.,* 1995). Consistent with this hypothesis, the sizes of the CP duplicates of SPSVV (79 kDa) and CCSV (74 kDa), as well as that of LIYV (52 kDa), are rough multiples of their CP sizes (~28 kDa), and the repeated segments of marginal similarity to the CP core may also be found in the N-terminal parts of the 79- and 74-kDa proteins (author, unpublished observation). Thus, expansion of the closterovirus genomes may be partly attributed to insertions and tandem duplications at both ends of the conserved replicative core.

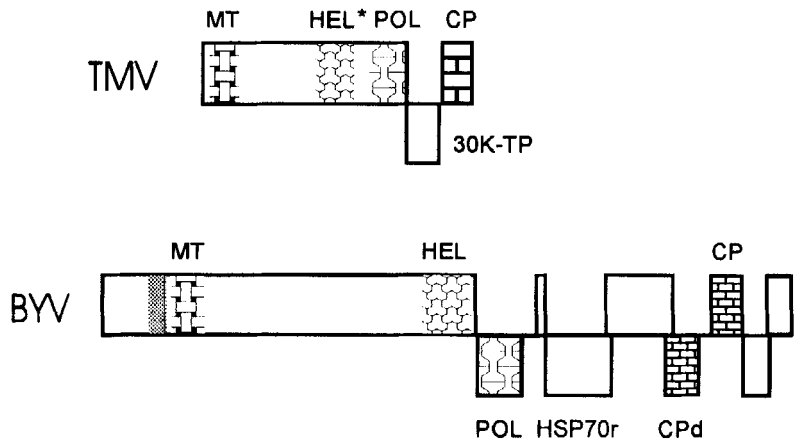


FIG. 6. Schematic comparison of the genomes of a closterovirus (BYV) and a tobamovirus (TMV). Related domains are shown by identical fill patterns and are marked as in Fig. 2. Asterisk marks the leaky termination codon in the TMV replicase gene; 30K, transport or movement protein gene of TMV; CP, capsid protein.

Another large insertion in the closterovirus genomes lies within the replicative core, between the MT and HEL domains (Fig. 6). The sequence of this region is significantly diverged in the BYV, CTV, and LIYV replicases (Dolja *et al.*, 1994; Karasev *et al.*, 1995; Klaassen *et al.*, 1995). Interestingly, among closteroviruses and related plant viruses the size of the MT–HEL span grows almost linearly with the increase of the genome size; in the case of viruses with divided genomes, it is related rather to the size of the largest genomic component (Table III). Generally, a rule "the larger the genome, the larger the replicase" inferred from these comparisons may also be applicable to arteri-, toro-, and coronaviruses of animals (den Boon *et al.*, 1991; Godeny *et al.*, 1993; Snijder and Horzinek, 1993). This relationship is not trivial, as the overall increase in genome size in these cases is not solely due to insertions in the replicase gene(s), but also to the appearance of new coding sequences flanking the replicative core. It seems quite likely that, in the process of evolution, expansion of closterovirus genomes was attended by an increase in the size of their replicases (Agranovsky, *et al.*, 1994a). At least one apparent obstacle in maintaining large RNA genomes must be accumulation of mutations in the progeny strands due to low fidelity of viral RNA polymerases (Holland *et al.*, 1982; Steinhauer and Holland, 1987). It would be interesting to see if, for example, the inserted domains or the leader proteins serve as "spell-

TABLE III

RELATIONSHIPS BETWEEN THE GENOME LENGTH AND REPLICASE SIZE IN THE TOBAMO-LIKE LINEAGE OF PLANT VIRUSES

Virus (group)	Genome size (kb)	Replicase size (aa)	MT–HEL distance (aa)[a]	HEL–POL distance (aa)[b]
BMV (bromo)	3.2+2.9+2.1	960+820[c]	430	
BSMV (hordei)	3.8+3.3+3.2	1140+770[c]	540	
RBDV (idaeo)	5.5+2.2	1690	420	680
TMV (tobamo)	6.4	1620	560	640
TRV (tobra)	6.8+3.4	1710	610	640
SBWMV (furo)	7.1+3.6	1830	700	650
LIYV (clostero)	8.1+7.2	2070[d]	1000	660
BYV (clostero)	15.5	2510[d]	1400	660
CTV (clostero)	19.3	2610[d]	1470	660

[a] The MT–HEL distance is measured between the C-terminal part of methyltransferase motif IV (Rozanov *et al.*, 1992) and the GKS/T signature in helicase motif I (Gorbalenya and Koonin, 1993).

[b] The HEL–POL distance is measured between the GKS/T and the GDD polymerase signature (Kamer and Argos, 1984).

[c] For BMV and BSMV, in which MT–HEL and POL domains are assigned to two individual proteins, sizes are given for both the putative methyltransferase-helicase and polymerase.

[d] The sizes of closterovirus replicases are given assuming translational frameshifting for ORFs 1a/1b expression and the cleavage of the N-terminal leader protein.

checkers" on strand copying. In DNA-dependent DNA polymerases, the 3′–5′ exonuclease activity assigned to a distinct protein domain is crucial for high replication fidelity (reviewed in Kunkel, 1988). Also, these domains might mediate homologous recombination between the virus RNA molecules to get rid of incorrigible errors, thus maintaining viable progeny. In coronavirus replication, recombination is believed to be a key mechanism to combat high-frequency errors (Jarvis and Kirkegaard, 1991; Lai, 1990). Finally, there is a possibility that the insert between MT and HEL contains a set of distinct domains to recognize the replication signals on an RNA template. The fact that up to six and nine sgRNA species may be synthesized on BYV and CTV infection, respectively, as compared with only two sgRNAs in the case of TMV, may again be a corollary to the increased complexity of the closterovirus replicases. Naturally, the possibilities mentioned above do not exclude one another.

In sharp contrast to the MT–HEL span, the distance between HEL and POL is essentially the same, about 650 residues, in all the tobamo-like virus replicases, despite the POL expression mode (Table III). Conservation of this arrangement may reflect constraints imposed on the replicase architectonics that must ensure concerted action of the strand-separating helicase "wedge" and the copying polymerase unit. The fact that the HEL and POL domains are found in two distinct gene products of bromo- and hordeiviruses does not contradict this rule as it would seem; at least for brome mosaic bromovirus, it has been demonstrated that the helicase-like and polymerase-like proteins form a complex in which the HEL and POL domains are juxtaposed in a fashion very similar to that in the TMV replicase (Kao and Ahlquist, 1992; Kao *et al.*, 1992).

B. Packaging Constraints and RNA Genome Evolution

Apart from replication constraints, evolution of closteroviruses toward increasing the genome size would have had to overcome packaging constraints. Comparisons of particle and genome structure of closteroviruses with those of other plant RNA viruses reveal some tendencies that may help one imagine how this could happen. Mono- and multipartite RNA viruses can be subdivided into those having "compressed" and "stretched" genomes, and this may be related to the virion type. Some spherical viruses, namely luteoviruses, tombusviruses, and tymoviruses, have compressed monopartite genomes in which ORFs extensively overlap to form "doubledecker" gene arrangements (Miller *et al.*, 1988; Morch *et al.*, 1988; Rochon and Tremaine, 1989). Such economical use of the coding sequence perhaps reflects a compromise between the necessity to widen the repertoire of viral genes and the limited size of an RNA molecule that would fit a spherical particle (Bransom *et al.*, 1995). However, the maintenance of overlapping genes has an apparent drawback, as this precludes each gene from being optimally adapted (Keese and Gibbs, 1992). The genome splitting seen in many RNA virus groups perhaps allows lifting of packaging constraints and minimizing the use of overlapping ORFs (or decompressing the preexisting gene overlaps). Thus, in spherical comoviruses, nepoviruses, dianthoviruses, and *Bromoviridae* members, the genomes are divided among separately encapsidated RNA components, each containing nonoverlapping gene(s) (Fig. 7).

An elongated helical capsid is less restrictive for the size of enveloped RNA. As proposed for the corona-like viruses, transition from a spherical (arterivirus-type) to helical (coronavirus-type) nucleocapsid

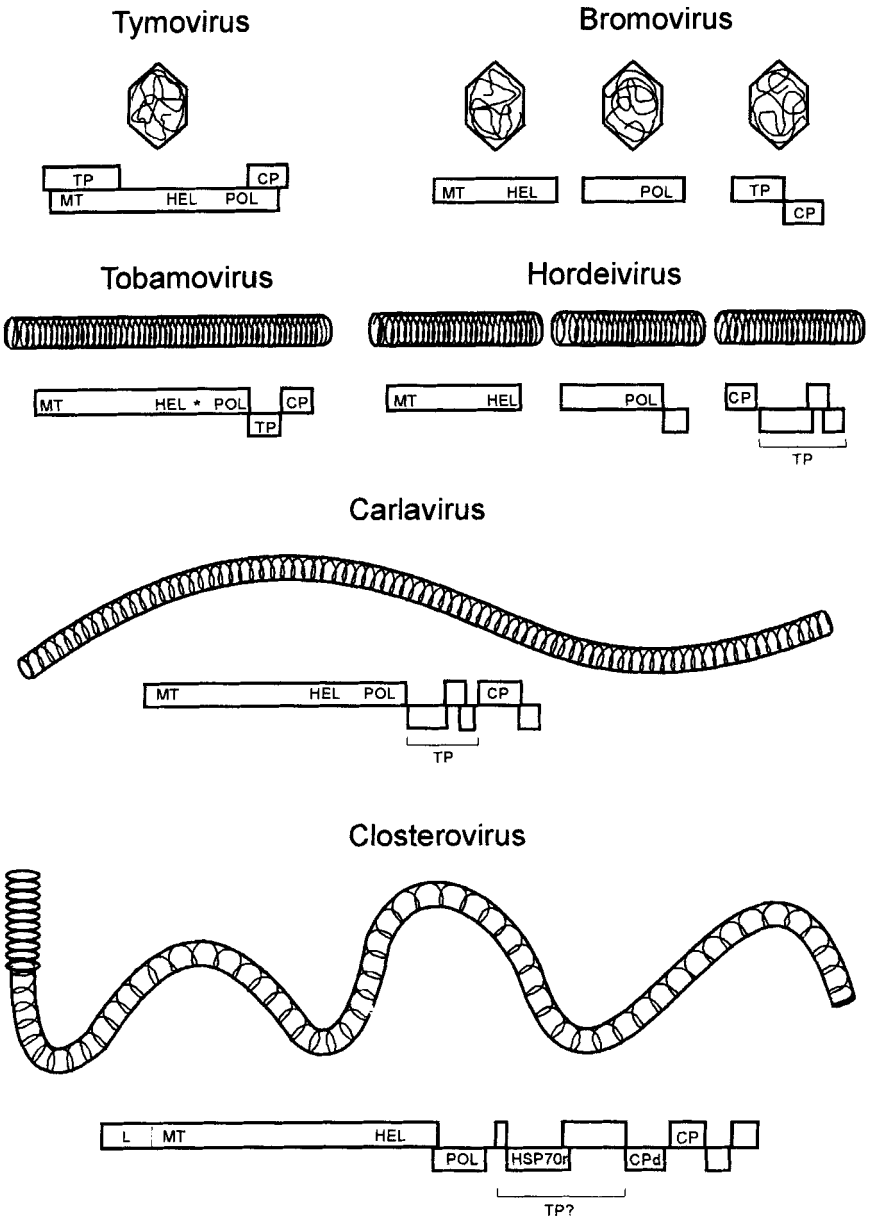


FIG. 7. Relationships between capsid type and genome organization in closteroviruses and other Sindbis-like supergroup viruses. Designations of the conserved protein domains are as in Fig. 2; other designations are TP, transport or movement protein(s); L, leader protein, *, leaky stop codon. Depicted roughly to scale.

exempted a progenitor of the toro- and coronaviruses from packaging constraints, thus allowing a nearly twofold genome expansion (Godeny *et al.*, 1993). Characteristically, most of the plant RNA viruses with elongated particles (tobamo-, tobra-, furo-, poty-, carla-, and closteroviruses) have stretched genomes, with modest overlapping of genes if any (Fig. 7). However, a subset of these viruses having rigid rodlike particles do not encapsidate RNA molecules of more than approximately 7 kb, possibly because of restrictions imposed on the particle length by sterical hindrances in the cell (Dolja *et al.*, 1994) and/or their mechanical fragility. This could have forced the splitting of the genomes of tobra-, hordei-, and furoviruses. In contrast to tobamoviruses having a monopartite 6.4-kb genome with only three genes, these viruses have genomes of 9–10 kb that encompass four to seven genes (Fig. 7).

It has been reasonably hypothesized that acquisition of flexible and superflexible helical capsids by the ancestors of carlaviruses, potyviruses, and closteroviruses allowed their genomes to grow to 10 kb and 20 kb, respectively (Dolja *et al.*, 1994). Again, these viruses possess stretched monopartite genomes (Fig. 7). As for the possible relationship between genome division and capsid type, the existence of bipartite filamentous viruses allied with the last two groups (bymoviruses and bicomponent closteroviruses) suggests the involvement of evolutionary factors other than packaging constraints that might have driven the genome splitting. In closteroviruses, the capsid evolution was crowned by employing the second CP. This conferred on their particles a structural complexity unprecedented among simple elongated viruses, which may be expected to require unusual assembly mechanisms.

VIII. Conclusions

The borrowing from Boris Pasternak's book of verse ("Over the Barriers," 1914–1916) in the title emphasizes that closteroviruses evolved by surmounting the restraints imposed on the genome and particle structure of positive-strand RNA viruses. Closteroviruses have large RNA genomes whose size and coding potential may only be compared to those of the corona-like viruses. Nevertheless, despite similar expression strategies and genome layouts developed in these two groups, closteroviruses cannot be considered as "plant coronaviruses" of a kind, since these similarities do not extend to amino acid sequences. Rather, they reflect independent adaptation to handling large RNA genomes in the two evolutionarily distant lineages. Colinearity and conservation of the main replicative domains clearly suggest the common ancestry of

closteroviruses and other plant tobamo-like viruses. However, closteroviruses have followed a distinct evolutionary pathway that has led to dramatic expansion of their genomes. Along with this, their evolution would have had to solve problems connected with replication and packaging of large RNA molecules; it is plausibe that this has been achieved by increasing the size (and functional complexity) of RNA replicase and by using a superflexible capsid made up of two CPs. Expansion of the closterovirus genomes has partially resulted from RNA recombination. It is possible that the horizontally acquired elements brought in novel enzymatic activities and structural elements advantageous for closterovirus adaptation to a distinct ecological niche, distinguished by the phloem-limited nature of infection and the semipersistent mode of insect transmission. In this respect, the most intriguing products are the HSP70-related protein, having the properties of a microtubule-binding ATPase, and the capsid protein duplicate involved in particle formation. There are many more closterovirus gene products whose functions remain enigmatic, since they have neither sequence-related counterparts in a current database nor known functional motifs. Hence, we may have more surprises. At present, we are making only the first steps in perceiving how closterovirus infection proceeds at the molecular level, despite some progress that has made it possible at least to address these questions. Further studies of the functions encoded by the large RNA genomes of closteroviruses are expected to provide a better understanding of the molecular mechanisms of their interactions with the genomes of their hosts and vectors.

ACKNOWLEDGMENTS

I am especially grateful to U. Hoyer, E. Maiss, and H. J. Vetten; to W. Jelkmann; and to L. P. Woudt and S&G Seeds B.V. for communicating their sequence data prior to publication. I wish to thank V. P. Boyko, K. Eastwell, B. W. Falk, S. Gowda, and J. P. Martelli for providing their unpublished data and preprints, D.-E. Lesemann for electron micrographs, and F. Rippmann for communicating his protein modelbuilds. Thanks are due to J. G. Atabekov, Yu. L. Dorokhov, A. V. Galkin, A. E. Gorbalenya, E. V. Koonin, and S. Yu. Morozov for fruitful discussions and critical reading of the manuscript.

REFERENCES

Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Koonin, E. V., and Dolja, V. V. (1991a). *J. Mol. Biol.* **217**, 603–610.

Agranovsky, A. A., Boyko, Y. P., Karasev, A. V., Lunina, N. A., Koonin, E. V., and Dolja, V. V. (1991b). *J. Gen. Virol.* **72**, 15–23.

Agranovsky, A. A., Koonin, E. V., Boyko. V. P., Maiss, E., Frötschl, R., Lunina, N. A., and Atabekov, J. G. (1994a). *Virology* **198**, 311–324.

Agranovsky, A. A., Koenig, R., Maiss, E., Boyko, V. P., Casper, R., and Atabekov, J. G. (1994b). *J. Gen. Virology* **75**, 1431–1439.

Agranovsky, A. A., Lesemann, D.-E., Maiss, E., Hull, R., and Atabekov, J. G. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2470–2473.

Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. M., Haseloff, J., Zimmern, D. (1985). *J. Virol.* **53**, 536–542.

Atreya, C. D., Atreya, P. L., Thornbury, D. W., and Pirone, T. P. (1992). *Virology* **191**, 106–111.

Bahner, I., Lamb, J., Mayo, M. A., and Hay, R. T. (1990). *J. Gen. Virol.* **71**, 2251–2256.

Bar-Joseph, M., and Hull, R. (1974). *Virology* **62**, 552–562.

Bar-Joseph, M., and Lee, R. F. (1989). CMI/AAB Descript. Plant Viruses. No. 353.

Bar-Joseph, M., and Murant, A. F. (1982). CMI/AAB Descript. Plant Viruses. No. 260.

Bar-Joseph, M., Loebenstein, G., and Cohen, J. (1972). *Virology* **50**, 821–828.

Bar-Joseph, M., Inouye, T., and Sutton, J. (1976). *Plant Dis. Rep.* **60**, 851.

Bar-Joseph, M., Garnsey, S. M., and Gonsalves, D. (1979). *Adv. Virus Res.* **25**, 93–168.

Beck, D. L., Guilford, P. J., Voot, D. M., Andersen, M. T., and Forster, R. L. S. (1991). *Virology* **183**, 695–702.

Belli, G., Fortusini, A., Casati, P., Belli, L., Bianco, P. A., and Pratti, S. (1994). *Riv. Pat. Veg. S.V.* **4**, 105–108.

Bem, F., and Murant, A. F. (1979). *Ann. Appl. Biol.* **92**, 237–242.

Boscia, D., Greif, C., Gugerli, P., Martelli, G. P., Walter, B., and Gonsalves, D. (1995). *VITIS* **34**, in press.

Boyko, V. P., Karasev, A. V., Agranovsky, A. A., Koonin, E. V., and Dolja, V. V. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9156–9160.

Brandes, J., and Bercks, R. (1965). *Adv. Virus Res.* **11**, 1–12.

Bransom, K. L., Weiland, J. J., Tsai, C.-H., and Dreher, T. W. (1995). *Virology* **206**, 403–412.

Brault, V., and Miller, W. A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2262–2266.

Brierly, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Block, V. C., Brown, T. D. K., and Inglis, S. C. (1987). *EMBO J.* **6**, 3779–3785.

Brunstedt, J., Moseley, J., and Hull, R. (1991). *Virus Genes* **5**, 267–272.

Butler, P. J. G. (1984). *J. Gen. Virol.* **65**, 253–279.

Carrington, J. C., Cary, S. M., Parks, T. D., and Dougherty, W. G. (1989). *EMBO J.* **8**, 365–370.

Cheng, S.-L., Domier, L. L., and D'Arcy, C. (1994). *Virology* **202**, 1003–1006.

Chevallier, D., Engle, A., Wurtz, M., and Charles, P. (1983). *J. Gen. Virol.* **64**, 2289–2293.

Choi, G. H., Pawlyk, D. M., and Nuss, D. L. (1991). *Virology* **183**, 747–752.

Clare, J. J., Belcourt, M., and Farabaugh, P. J. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6816–6820.

Coffin, R. S., and Coutts, R. H. A. (1992). *Intervirology* **33**, 197–203.

Coffin, R. S., and Coutts, R. H. A. (1993). *J. Gen. Virol.* **74**, 1475–1483.

Cohen, J., Franck, A., Yetten, H. J., Lesemann, D.-E., and Loebenstein, G. (1992). *Ann. Appl. Biol.* **121**, 257–268.

Craig, E. A., Gambill, B. D., and Nelson, R. J. (1993). *Microbiol. Rev.* **82**, 3616–3620.

Craigen, W. J., Cook, R. G., Tate, W. P., and Caskey, C. T. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3616–3620.

Cronin, S., Verchot, J., Haldeman-Cahill, R. Schaad, M. G., and Carrington, J. C. (1995). *Plant Cell* **7**, 549–559.

den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C., and Spaan, W. J. M. (1991). *J. Virol.* **65**, 2910–2920.

Dinman, J. D., Icho, T., and Wickner, R. B. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 174–178.
Dobrov, E. V., and Atabekov, J. O. (1989). *In* "Plant Viruses," Vol. I, "Structure and Replication" (C. L. Mandahar, ed.), pp. 174–205. CRC Press, Boca Raton, FL.
Dodds, J. A., and Bar-Joseph, M. (1983). *Phytopathology* **73**, 419–423.
Dodds, J. A., Jarupat, T., Lee, J. G., and Roistacher, C. N. (1987). *Phytopathology* **77**, 442–447.
Dolja, V. V., Karasev, A. V., and Agranovsky, A. A. (1990). *In* "New Aspects of Positive-Strand RNA Viruses" (R. Rueckert and M. Brinton, eds.), pp. 31–35. ASM Publications, Washington, DC.
Dolja, V. V., Boyko, V. P., Agranovsky, A. A., and Koonin, E. V. (1991). *Virology* **184**, 79–86.
Dolja, V. V., Karasev, A. V., and Koonin, E. V. (1994). *Annu. Rev. Phytopathol.* **32**, 261–285.
Domingo, E., Sabo, T., Taniguchi, T., and Weissmann, C. (1978). *Cell* **13**, 735–744.
Duffus, J. E. (1972). *Phytopathology* **62**, 3161–3165.
Duffus, J. E. (1973). *Adv. Virus Res.* **18**, 347–386.
Duffus, J. E., Larsen, R. C., and Liu, H. Y. (1986). *Phytopathology* **76**, 97–100.
Dunigan, D. D., and Zaitlin, M. (1990). *J. Biol. Chem.* **265**, 7779–7786.
Essau, K., and Hoefert, L. L. (1971). *Protoplasma* **72**, 255.
Falk, B. W., and Duffus, J. E. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 275–296. Plenum Press, New York.
Farabaugh, P. J. (1993). *Cell* **74**, 591–596.
Farabaugh, P. J., Zhao, H., and Vimaladithan, A. (1993). *Cell* **74**, 93–103.
Febres, V. J., Pappu, H. R., Anderson, E. J., Pappu, S. S., Lee, R. F., and Niblett, C. L. (1994). *Virology* **201**, 178–181.
Felsenstein, J. (1989). "PHYLIP 3.2 Manual." Herbarium, University of California, Berkeley.
Filichkin, S. A., Lister, R. M., McGrath, P. F., and Young, M. J. (1994). *Virology* **205**, 290–299.
Francki, R. I. B., Faquet, C. M., Knudson, D. L., and Brown, F., eds. (1991). *Arch. Virol.* **Suppl. 2**, 1.
German, S. Candresse, T., Lanneau, M., Huet, J. C., Pernollet, J. C., and Dunez, J. (1990). *Virology* **179**, 104–112.
Gething, M. J., and Sambrook, J. (1992). *Nature* **355**, 33–45.
Gibbs, A. (1987). *J. Cell Sci.* (Suppl.) **7**, 319–337.
Godeny, E. K., Chen, L., Kumar, S., Methven, S. L., Koonin, E. V., and Brinton, M. A. (1993). *Virology* **194**, 585–596.
Goldbach, R., Le Gall, O., and Wellink, J. (1991). *Semin. Virol.* **2**, 19–25.
Gorbalenya, A. E. (1992). *Semin. Virol.* **3**, 359–371.
Gorbalenya, A. E., and Koonin, E. V. (1989). *Nucleic Acids Res.* **17**, 8413–8440.
Gorbalenya, A. E., and Koonin, E. V. (1993). *Curr. Opin. Struct. Biol.* **3**, 419–429.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). *Nature* **333**, 22.
Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P., and Koonin, E. V. (1989). *J. Mol. Evol.* **28**, 256–268.
Guerri, J., Moreno, P., Munoz, N., and Martinez, M. E. (1991). *Plant Pathol.* **40**, 38–44.
Gunasinghe, U. B., and German, T. L. (1989). *Phytopathology* **79**, 1337–1341.
Haeberle, A.-M., Stussi-Garaud, C., Schmitt, C., Garaud, J.-C., Richards, K. E., Guilley, H., and Jonard, G. (1994). *Arch. Virol.* **134**, 195–203.
Hilf, M. E., Karasev, A. V., Pappu, H. R., Gumpf, D. J., Niblett, C. L., and Garnsey, S. M. (1995). *Virology* **208**, 576–582.
Hill, H. R., and Zetler, F. W. (1973). *Phytopathology* **63**, 443.

Hills, G. J., and Gay, M. R. (1976). *John Innes Inst. Annu. Report* **67**, 107–109.
Hodgman, T. C. (1988). *Nature* **333**, 22–23; 578 (Erratum).
Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and Van de Pol, S. (1982). *Science* **215**, 1577–1585.
Hu, J. S., Gonsalves, D., and Tellis, O. (1990). *J. Phytopath. Z.* **128**, 1–14.
Hull, R. (1992). *Semin. Virol.* **3**, 373–382.
Inouye, T. (1974). CMI/AAB Descript. Plant Viruses. No. 136.
Inouye, T. (1976). CMI/AAB Descript. Plant Viruses. No. 157.
Jacks, T., and Varmus, H. E. (1985). *Science* **230**, 1237–1242.
Jacks, T., Madhani, H. D., Masiarz, F. R., and Varmus, H. E. (1988). *Cell* **55**, 447–458.
Jarvis, T. C., and Kirkegaard, K. (1991). *Trends Genet.* **7**, 186–191.
Jiang, B. M., Monroe, S. S., Koonin, E. V., Stine, S. E., and Glass, R. I. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10539–10543.
Kamer, G., and Argos, P. (1984). *Nucleic Acids Res.* **12**, 7269–7283.
Kao, C. C., and Ahlquist, P. (1992). *J. Virol.* **66**, 7293–7302.
Kao, C. C., Quadt, R., Hershberger, R. P., and Ahlquist, P. (1992). *J. Virol.* **66**, 7293–7302.
Karasev, A. V., Agranovsky, A. A., Rogov, V. V., Miroshnichenko, N. A., Dolja, V. V., and Atabekov, J. G. (1989). *J. Gen. Virol.* **70**, 241–245.
Karasev, A. V., Kashina, A. S., Gelfand, V. I., and Dolja, V. V. (1992). *FEBS Lett.* **304**, 12–14.
Karasev, A. V., Nikolaeva, O. N., Gumpf, D. J., Garnsey, S. M., and Dawson, W. O. (1994a). *Phytopathology* **84**, 1156.
Karasev, A. V., Nikolaeva, O. N., Koonin, E. V., Gumpf, D. J., and Garnsey, S. M. (1994b). *J. Gen. Virol.* **75**, 1415–1422.
Karasev, A. V., Boyko, V. P., Gowda, S., Nikolaeva, O. N., Hilf, M. E., Koonin, E. V., Niblett, C .L., Cline, K. C., Gumpf, R. F., Gamsey, S. M., Lewandowski, D. J., and Dawson, W. O. (1995). *Virology* **208**, 511–520.
Keese, P. K., and Gibbs, A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9489–9493.
Kim, K. S., Gonsalves, D., Teliz, D., and Lee, K. W. (1989). *Phytopathology* **79**, 357–360.
Kirkegaard, K., and Baltimore, D. (1986). *Cell* **47**, 433–443.
Klaassen, V. A., Boeshore, M. L., Dolja, V. V., and Falk, B. W. (1994). *J. Gen. Virol.* **75**, 1525–1533.
Klaassen, V. A., Boeshore, M. L., Koonin, E. V., Tian, T., and Falk, B. W. (1995). *Virology* **208**, 99–101.
Koenig, R., and Torrance, L. (1986). *J. Gen. Virol.* **67**, 2145–2151.
Koonin, E. V. (1991). *J. Gen. Virol.* **72**, 2197–2206.
Koonin, E. V., and Dolja, V. V. (1993). *Crit. Rev. Biochem. Mol. Biol.* **28**, 375–430.
Koonin, E. V., Mushegian, A. R., Ryabov, E. V., and Dolja, V. V. (1991). *J. Gen. Virol.* **72**, 2895–2903.
Kunkel, T. A. (1988). *Cell* **53**, 837–840.
Lai, M. M. C. (1990). *Annu. Rev. Microbiol.* **44**, 303–333.
Lain, S., Riechmann, J. L., and Garcia, J. A. (1990). *Nucleic Acids Res.* **18**, 7003–7006.
Larsen, R. S., Kim, K. S., and Scott, H. A. (1991). *Phytopathology* **81**, 227–232.
Lee, H.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). *Virology* **180**, 567–582.
Lee, H.-J., Calvert, L. A., Nagel, J., Hubbard, J. D. (1988). *Phytopathotogy* **78**, 1221–1226.
Lesemann, D.-E. (1977). *Phytopathol. Z.* **89**, 330–334.
Lesemann, D.-E. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 179–236. Plenum Press, New York.
Limburg, D. D., Mauk, P. A., and Godfrey, L. D. (1994). *Phytopathology* **84**, 1066.

Ling, K.-S., Drong, R. F., Slightom, J. L., and Gonsalves, D. (1994). *Phytopathology* **84**, 1372.

Lister, R. M., and Bar-Joseph, M. (1981). *In* "Handbook of Plant Virus Infections and Comparative Diagnosis" (E. Kurstak, ed.), pp. 809–844. Elsevier/North-Holland, Amsterdam.

Liu, H. Y., and Duffus, J. E. (1990). *Phytopathology* **80**, 866–869.

Lockhart, B. E. L., Autrey, L. J. C., and Comstock, J. C. (1992). *Phytopathology* **82**, 691–695.

Lomonossoff, G. P., and Wilson, T. M. A. (1985). *In* "Molecular Plant Pathology" (J. W. Davies, ed.), Vol. 1, pp. 43–83. CRC Press, Boca Raton, FL.

Mäkinen, K., Næss, V., Tamm, T., Truve, E., Aaspõlu, A., and Saarma, M. (1995). *Virology* **207**, 566–571.

Mathews, R. E. F. (1991). "Plant Virology" (3rd Ed.). Academic Press, New York.

Mawassi, M., Gafny, R., Gagliardi, D., and Bar-Joseph, M. (1995a). *J. Gen. Virol.* **76**, 651–659.

Mawassi, M., Karasev, A. V. , Mietkiewska, E., Gafny, R., Lee, R. F., Dawson, W. O., and Bar-Joseph, M. (1995b). *Virology* **208**, 383–387.

Mayo, M. A., and Martelli, G. P. (1993). *Arch. Virol.* **133**, 496–498.

Mi, S., and Stollar, V. (1991). *Virology* **178**, 429–434.

Miller, W. A., Waterhouse, P. M., and Gerlach, W. L. (1988). *Nucleic Acids Res.* **16**, 6097–6111.

Milne, R. G. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4. Plenum Press, New York.

Morch, M.-D., Boyer, J.-C., and Haenni, A.-L. (1988). *Nucleic Acids Res.* **16**, 6157–6173.

Moreno, P., Guero, J., and Munoz, N. (1990). *Phytopathology* **80**, 477–482.

Morozov, S. Yu., Dolja, V. V., and Atabekov, J. G. (1989). *J. Mol. Evol.* **297**, 52–62.

Morozov, S. Yu., Miroshnichenko, N. A., Zelenina, D. A., Fedorkin, O, N., Soloviev, A. G., Lukasheva, L. I., and Atabekov, J. G. (1990). *Biochemie* **72**, 677–684.

Moseley, J., and Hull, R. (1990). *Ann. Appl. Biol.* **118**, 605–613.

Murant, A .F., Raccah, B., and Pirone, T. P. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 237–274. Plenum Press, New York.

Nakano, M., and Inouye, T. (1980). *Ann. Phytopath. Soc. Jpn.* **46**, 7–10.

Namba, S., Boscia, D., Azzam, O., Maixner, M., Hu, J. S., Golino, D., and Gonsalves, D. (1991). *Phytopathology* **81**, 964–970.

Nikiphorova, S. Yu., Agranovsky, A. A., and Atabekov, J. C. (1995). *Dokladi Rossiiskoi Akademii Nauk* **340**, 416–418 (in Russian).

Pappu, H. R., Pappu, S. S., Manjunath, K. L., Lee, R. F., and Niblett, C. L. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3641–3644.

Pappu, H. R., Karasev, A. V., Anderson, E. J., Pappu, S. S., Hilf, M. E, Febres, M. E., Eckloff, R. M. G., McCaffery, M., Boyko, V., Gowda, S., Dolja, V. V., Koonin, E. V., Gumpf, D. J., Cline, K. C., Garnsey, S. M., Dawson, W. O., Lee, R. F., and Niblett, C. L. (1994). *Virology* **199**, 35–46.

Pirone, T. P. (1991). *Semin. Virol.* **2**, 81–87.

Prüfer, D, Tacke, E., Schimtz, J., Kull, B., Kaufmann, A., and Rohde, W. (1992). *EMBO J.* **11**, 1111–1117.

Ragetti, H. W. J., Elder, M., and Schroeder, B. K. (1982). *Can. J. Bot.* **60**, 1235–1248.

Raine, J., McMullen, R. D., and Forbes, A. R. (1986). *Can. J. Plant Pathol.* **8**, 6–11.

Richards, K. E., and Tamada, T. (1992). *Annu. Rev. Phytopathol.* **30**, 291–313.

Rippmann, F., Taylor, W. R., Rothbard, J. B., and Green, M. N. (1991). *EMBO J.* **10**, 1053–1059.

Rochon, D. M., and Tremaine, J. H. (1989). *Virology* **169**, 251–259.
Rogov, V. V., Karasev, A. A., and Agranovsky, A. A. (1993). *Phytopathol. Z.* **137**, 79–88.
Rozanov, M. N., Koonin, E. V., and Gorbalenya, A. E. (1992). *J. Gen. Virol.* **73**, 2129–2134.
Russel, M. (1993). *J. Mol. Biol.* **231**, 689–697.
Sekiya, M. E., Lawrence, S. D., McCaffery, M., and Cline, K. (1991). *J. Gen. Virol.* **72**, 1013–1020.
Schmidt, H. B., Richter, J., Hertsch, W., and Klinkowski, M. (1963). *Phytopathol. Z.* **47**, 66–74.
Shukla, D. D., Strike, P. M., Tracy, S. L., Gough, K. H., and Ward, C. W. (1988). *J. Gen. Virol.* **69**, 1497–1508.
Snijder, E. J., and Horzinek, M. C. (1993). *J. Gen. Virol.* **74**, 2305–2316.
Steinhauer, D. A., and Holland, J. J. (1987). *Annu. Rev. Microbiol.* **41**, 409–433.
Strauss, E. G., and Strauss, J. H. (1986). *In* "The Togaviridae and Flaviviridae" (E.G. Strauss and J. H. Strauss, eds.), pp. 35–90. Plenum Press, New York.
Stuart, K. D., Weeks, R., Gilbride, L., and Myler, P. J. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8596–8600.
ten Dam, E. B. (1995). Ph.D. Thesis, Leiden University, pp. 93–106.
ten Dam, E. B., Pleij, C. W., and Bosch, L. (1990). *Virus Genes* **4**, 121–136.
Tollin, P., and Wilson, H. R. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 51–83. Plenum Press, New York.
Tollin, P., Wilson, H. R., Roberts, I. M., and Murant, A. F. (1992). *J. Gen. Virol.* **75**, 3045–3048.
Tu, C., Tzeng, T.-H., and Bruenn, J. A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8636–8640.
Wang, Y., and Walker, P. J. (1993). *Virology* **195**, 719–731.
Warrener, P., and Collett, M. S. (1995). *J. Virol.* **69**, 1720–1726.
Weiss, R. B. (1984). *Proc. Natl. Accad. Sci. U.S.A.* **81**, 5797–5801.
Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., and Gesteland, R. F. (1988). *EMBO J.* **7**, 1503–1507.
Winter, S., Purac, A., Legget, F., Frison, E. A., Rossel, H. W., and Hamilton, R. I. (1992). *Phytopathology* **82**, 869–875.
Woudt, L. P., de Rover, A. P., and Van Grinsven, M. Q. J. M. (1993a). "Proc. Int. Congr Plant Pathol., 6th, Montreal," p. 313.
Woudt, L. P., de Rover, A. P., de Haan, P. T., and Van Grinsven, M. Q. J. M. (1993b). "Int. Congr. Virol, 9th, Glasgow," p. 326.
Xiong, Z., and Lommel, S. A. (1989). *Virology* **171**, 543–554.
Yamashita, S., Ohki, S. T., Doi, Y., and Yora, K. (1976). *Ann. Phytopathol. Soc. Jpn.* **42**, 382–390.
Yoshikawa, N., Sasaki, E., Kato, M., and Takahashi, T. (1992). *Virology* **191**, 98–105.
Zaccomer, B., Haenni, A.-L., and Macaya, G. (1995). *J. Gen. Virol.* **76**, 231–247.
Zimmern, D. (1988). *In* "RNA Genetics" (J. J. Holland, E. R. Domingo, and P. Ahlquist, eds.), pp. 211–240. CRC Press, Boca Raton, FL.
Zimmerman, D., Bass, P., Legin, R., and Walter, B. (1990). *J. Phytopathol. Z.* **130**, 205–218.

ADVANCES IN VIRUS RESEARCH, VOL. 47

COMPARISON OF THE REPLICATION OF POSITIVE-STRANDED RNA VIRUSES OF PLANTS AND ANIMALS

Kenneth W. Buck

Department of Biology
Imperial College of Science, Technology and Medicine
London SW7 2BB, England

I. Introduction

Many important animal, bacterial, fungal, and plant viruses have genomes of positive-stranded (messenger-sense) RNA (Murphy *et al.*, 1995). Over the last 12 years it has become clear that some viruses in these different types of hosts have related genes, and sometimes similar arrangements of genes and modes of gene expression, leading to suggestions about evolutionary pathways (Haseloff *et al.*, 1984; Kamer and Argos, 1984; Ahlquist *et al.*, 1985; Goldbach, 1986; Zimmern, 1988; Strauss and Strauss, 1988; Habili and Symons, 1989; Bruenn, 1991; Goldbach *et al.*, 1991; Koonin, 1991a; Dolja and Carrington, 1992; Koonin and Dolja, 1993; Dolja *et al.*, 1994; Goldbach and de Haan, 1994; Ward, 1994). Goldbach and de Haan (1994) outlined four possible evolu-

tionary pathways to account for the observed relationships between viruses of different host types: convergent evolution, transduction (introgression) of (conserved) host genes, common ancestry, interviral recombination. Nucleotide and amino acid sequence similarities provide evidence for common ancestries of genes and, in cases where gene orders and modes of gene expression are similar (e.g., cowpea mosaic virus and poliovirus; Goldbach, 1986), possibly of viruses. However, there is also good evidence that recombination between RNAs of different viruses has played an important role in virus evolution (Lai, 1992, 1995; Simon and Bujarski, 1994). Since mixed infections of positive-stranded RNA viruses and viruses with different genome types, such as single-stranded (ss) DNA, double-stranded (ds) DNA, negative-stranded RNA, or dsRNA, are also possible, acquisition of genes from such disparate viruses is also possible. For example, some animal coronaviruses (positive-stranded RNA) and influenza virus C (negative-stranded RNA) have related hemagglutinin-esterase proteins (Cavanagh and MacNaughton, 1994). Recent studies have also indicated that some positive-stranded RNA viruses have acquired genes from the host, e.g., the HSP70 heat-shock chaperone protein gene analogues found in viruses of the plant closterovirus genus (Dolja *et al.,* 1994; Klaasen *et al.,* 1995); other examples are reviewed by Lai (1995). Recombination between positive-stranded viral RNA and transgenic plant transcripts has been demonstrated (Greene and Allison, 1994). The concept has therefore developed of modular evolution of positive-stranded RNA viruses, with the possibility of genes or groups of genes with related functions, e.g., replication genes or structural genes, having been acquired from different sources.

It is also noteworthy that RNA polymerases that replicate RNA apparently lack proofreading activities, leading to error rates of 10^{-3} to 10^{-5} (Domingo and Holland, 1994). This has led to the concept of the viral RNA "quasi" species (Eigen and Biebriecher, 1988), consisting of a population of closely related sequences, the composition of the population being influenced by positive and negative selection pressures. Such high mutation rates have probably limited the size of (unsegmented) RNA genomes, the largest being the animal coronaviruses (up to ~30 kb) and the plant closteroviruses (up to ~20 kb) and most being <10 kb, although some form of editing remains a possibility for the largest RNA genomes.

Positive-stranded RNA viruses have also evolved (or acquired) genes with functions specific for their hosts. A large number of positive-stranded RNA viruses (over 600 species) are adapted to replication and spread in plant hosts (Goldbach *et al.,* 1994). One important adaptation

is the requirement of plant viruses to spread from cell to cell through the plasmodesmatal connections. Plant viruses encode specialized movement proteins which modify the plasmodesmata in a variety of ways to allow the (active) passage of viral nucleic acids, nucleoproteins, or particles (Waigmann and Zambryski, 1994; Lucas and Gilbertson, 1994). Since there is evidence that plant proteins and/or mRNAs move through plasmodesmata during the course of normal plant development (Lucas, 1995), it is possible that plant viruses have become adapted to enable them to utilize an endogenous transport system for their own movement through the plant. Animal and bacterial viruses also encode specialized proteins, not needed by plant viruses, for attachment and entry into cells and for release from cells. Examples include virus attachment proteins which bind to receptors on the surface of cells, and proteins which direct budding (exocytosis) of enveloped viruses from the cell surface or which induce cell lysis. Because of the high mutation rate of RNA viruses (Domingo and Holland, 1994) and the similarities of present-day viruses, Goldbach and de Haan (1994) considered it unlikely that common ancestors of plant and animal viruses predated the evolutionary separation of plants and animals. Insect vectors of both plant and animal positive-stranded RNA viruses are known, in some of which the viruses are able to replicate, and these were considered as the likely bridge between plant and animal viruses. Additionally, the ability of a virus to replicate in two different types of host [vertebrate and invertebrate, e.g., alphaviruses (Strauss and Strauss, 1994); plant and invertebrate, e.g., marafiviruses (Gamez and Leon, 1988)] is likely to be a significant driving force in evolution. Although the insect vector hypothesis is persuasive, the possibility of evolutionary stasis has also to be considered and the additional hypothesis that some positive-stranded RNA viruses, which predated the separation of plants and animals, have provided ancestral gene clusters for current plant and animal viruses cannot be discounted.

The replication of the genome can be considered to be the most fundamental aspect of the biology of positive-stranded RNA viruses. This can be envisaged to take place in two main stages: (i) synthesis of a complementary (negative-stranded) RNA using the genomic positive-stranded RNA as a template; (ii) synthesis of progeny positive-stranded RNA using the negative-stranded RNA as a template. At least one virus-encoded protein, an RNA-dependent RNA polymerase (RdRp), is required to catalyze RNA synthesis directed by the RNA template and using the four ribonucleoside triphosphates as substrates. Another activity is required to unwind the duplex formed between the template and newly synthesized strands. The poliovirus RdRp ($3D^{pol}$ protein) has

been shown to have unwinding activity (Cho *et al.*, 1993). However, it is not clear whether the role of this activity in poliovirus RNA replication is to remove secondary structure in the template RNA, thereby increasing processivity of the enzyme, or to unwind duplexes formed as a result of RNA synthesis. Many positive-stranded RNA viruses also encode helicases (or putative helicases), which may function in duplex unwinding during replication. Other proteins probably have specialized roles in the initiation of RNA synthesis. By analogy with the phage Qβ RdRp holoenzyme, which contains four host (bacterial) protein subunits as well as the virus-encoded polymerase subunit (Blumenthal and Carmichael, 1979; van Duin, 1988), it is likely that RdRp holoenzymes of positive-stranded RNA viruses of eukaryotic hosts will also consist of both virus-encoded and host-encoded protein subunits. Hence the replication proteins of animal and plant positive-stranded RNA viruses, which are related in terms of amino acid sequences or gene order, may nevertheless have adapted to their particular host in terms of their interactions with host proteins. One aim of this review is to identify similarities and differences in the replication of positive-stranded RNA viruses of plants and animals. As the organization and expression of replication proteins differ considerably between different groups of viruses within a particular host, a further aim will be to compare replication strategies of different virus groups. Some positive-stranded RNA plant viruses also produce subgenomic RNAs which are used as mRNAs for translation of open reading frames (ORFs) which are internal in the virus genome. The different strategies which have evolved for this purpose will be compared. The replication of satellite RNAs, which by definition generally contain substantial nucleotide sequence distinct from that of the helper virus genome (reviewed by Roossinck *et al.*, 1992; Kaper, 1995) will only be discussed when the results are relevant to replication of the genomic RNA. The subject of replicase-mediated resistance, which has been recently reviewed (Carr and Zaitlin, 1993; Baulcombe, 1994; Lomonossoff, 1995; Mueller *et al.*, 1995), will not be considered in this article. RNA replication is important in RNA recombination in generating hybrid RNA molecules and defective-interfering (DI) RNAs. This topic is the subject of a number of excellent recent reviews (Lai, 1992; Bujarski *et al.*, 1994; Simon and Bujarski, 1994) and will be covered here only insofar as data on recombination impact on replication mechanisms. The retroviruses encapsidate single-stranded, messenger-sense RNA, but replicate via DNA intermediates; the replication of these viruses will not be considered here, except for brief comparisons of reverse transcriptases with other types of RNA and DNA polymerases (see Section II,A).

II. Virus-Encoded Replication Proteins

A. RNA-Dependent RNA Polymerases

An early debate in the study of plant positive-stranded RNA replication was whether RNA synthesis was catalyzed by host-encoded or virus-encoded RNA polymerases. Infection of plants with a range of positive-stranded RNA viruses from different families or genera leads to increase in activity and amounts of plant RNA polymerases which are able to initiate RNA synthesis utilizing an RNA template (Fraenkel-Conrat, 1986; Schiebel *et al.,* 1993a,b). Similar activities have been reported from animal cells (Volloch, 1986; Volloch *et al.,* 1987). Crude preparations of viral RNA polymerases isolated from infected plants are often contaminated with the host RNA polymerase. The purified enzymes from several plants consist of single polypeptide chains with molecular masses in the range 100–140 kDa, they can utilize ssRNA and ssDNA as templates with no marked sequence specificity, and synthesize relatively short, heterogeneous RNA products (<500 nt). The isolation of replication complexes, containing virus-encoded RdRp but free from the host RdRp, able to synthesize full-length viral RNA, e.g., for cowpea mosaic virus (van der Meer *et al.,* 1984) and cucumber mosaic virus (Hayes and Buck, 1990), argues against a role of the host RdRp in RNA replication, at least for the viruses studied. The biological role of the host RdRp is unknown, although it has been suggested that it may play a role in posttranscriptional suppression of gene expression (Dougherty and Parks, 1995).

It is now clear that all replication-competent positive-stranded RNA viruses which have been sequenced encode an RdRp. For many viruses, putative RdRps have been identified from conserved amino acid sequence motifs in polypeptides whose sequences have been deduced from ORFs in the nucleotide sequences of the viral RNA and the requirement for the ORFs containing these motifs in virus replication has been established by mutation. RNA polymerase activity has been demonstrated biochemically for the specific proteins encoded by these ORFs for only a few viruses, e.g., $Q\beta$ replicase subunit II (Landers *et al.,* 1974), poliovirus 3D protein (van Dyke and Flanegan, 1980; Rothstein *et al.,* 1988; Neufeld *et al.,* 1991), hepatitis C virus NS5B protein (Behrens *et al.,* 1996).

Kamer and Argos (1984) identified several similar motifs between the known poliovirus RdRp ($3D^{pol}$ protein) and putative RdRps of several other positive-stranded RNA viruses of animals and plants. The most conserved of these consisted of a central Gly–Asp–Asp (GDD)

triplet flanked by pentapeptides consisting of mainly hydrophobic amino acids, suggesting a β-hairpin structure composed of two hydrogen-bonded antiparallel β-strands connected to a short exposed loop containing the GDD amino acids. Subsequent analyses have extended the range of viruses and identified further conserved motifs (Poch *et al.*, 1989; Habili and Symons, 1989; Bruenn, 1991; Koonin, 1991a; Koonin and Dolja, 1993; Dolja *et al.*, 1994). Koonin (1991a) and Koonin and Dolja (1993) identified eight such motifs, three of which (IV, V, and VI) showed unequivocal conservation, allowing the signature DX_3[FYWLCA]$X_{0-1}DX_n$[STM]GX_3TX_3[NE]X_n[GS]DD to be proposed as an identifier of RdRps of positive-stranded RNA viruses and some related dsRNA viruses (X indicates an unspecified amino acid residue; alternative amino acids at particular sites are shown in square brackets). The regions of this signature separated by X_n correspond to parts of motifs IV, V, and VI.

The importance of some of these motifs has been confirmed by *in vitro* mutagenesis. Mills *et al.* (1988) showed that many linker insertions in a central region of the phage Qβ replicase gene, which included motifs I to VIII, were lethal. Furthermore, mutation of the G residue of the GDD box in motif VI to M, P, S, or V reduced the phage Qβ replication *in vivo* to <1% of wild-type; mutation of the G to A abolished replication (Inokuchi and Hirashima, 1987). Similar mutations in the GDD box of the poliovirus 3D protein considerably reduced its RNA polymerase activity in an *in vitro* assay, although the quantitative effects of the different mutations were different from those in Qβ; mutation of the G to A or S gave 5 to 20% of wild-type activity, while mutation of the G to C, M, P, or V abolished enzyme activity (Jablonski *et al.*, 1991). Furthermore, it was shown that mutation of the Y residue, which flanks the GDD box (YGDD) in the 3D protein of poliovirus and equivalent proteins of some related viruses (Koonin, 1991a), to F had no effect on *in vitro* polymerase activity or virus viability, whereas mutations of Y to S, I, or H considerably reduced or abolished *in vitro* polymerase activity and were lethal. Interestingly, a Y to M mutation had no effect on *in vitro* (primed) polymerase activity, but reduced virus infectivity, giving rise *in vivo* to a second compensatory mutation upstream of the conserved polymerase motifs; this suggested possible interaction of an upstream region of the polymerase with the YGDD sequence of the conserved polymerase motif VI (Jablonski and Morrow, 1993). Sankar and Porter (1992) mutated seven amino acid residues in the encephalomyocarditis 3D protein, which showed a high degree of conservation in motifs IV to VI (D235→E, motif IV; D240→E, motif IV; G294→A, motif V; T298→S, motif V; G332→E, motif VI; D333→E, motif VI; D334→E,

motif VI). All of these mutations abolished, or reduced to a very low level, the *in vitro* RNA polymerase activity of this protein. Mutations to chemically similar residues in regions of these motifs which are conserved only in their hydrophobicity had either no effect on, or a less pronounced reduction in, enzyme activity. Mutations of the GDD box of the potato virus X 166-kDa protein to GED, ADD, or GAD abolished infectivity for plants and reduced RNA replication in protoplasts to undetectable levels (Longstaff *et al.,* 1993). Similarly, a G→R substitution in the GDD box of turnip yellow mosaic virus 66-kDa protein abolished RNA replication (Weiland and Dreher, 1993). Mutations in the central polymerase-like domain of the brome mosaic virus 2a protein, encompassing motif IV and flanking sequences, also gave rise to mutants in which replication became temperature sensitive or was abolished (Kroner *et al.,* 1989; Traynor *et al.,* 1991). A G→E mutation in motif II in the polymerase-like domain of the Sindbis virus nsP4 protein rendered the virus temperature-sensitive for RNA replication (Hahn *et al.,* 1989a).

Argos (1988), Poch *et al.* (1989), Delarue *et al.* (1990), and Heringa and Argos (1994) extended the amino acid sequence comparisons to include RNA-dependent DNA polymerases (reverse transcriptases), DNA-dependent RNA polymerases, and DNA-dependent DNA polymerases, as well as RdRps of negative-stranded and dsRNA viruses. It was found that counterparts to the Koonin (1991a) RdRp sequence motifs IV and VI are present in many of these other types of polymerases (Table I), leading to the suggestion of similar types of protein folds (Delarue *et al.,* 1990; Heringa and Argos, 1994). The three-dimensional structures of the Klenow fragment of *Escherichia coli* DNA polymerase I (Ollis *et al.,* 1985), phage T7 RNA polymerase (Sousa *et al.,* 1993), and human immunodeficiency virus (HIV) type 1 reverse transcriptase (Koehlstaedt *et al.,* 1992) indicate that all these polymerases contain a "hand" structure with a cleft formed between "fingers," "palm," and "thumb" subdomains. Further analysis of the structure of the HIV reverse transcriptase complexed with a template and primer (Jacoba-Molina *et al.,* 1993) indicated that the template and primer are located in the cleft. Sequence motifs equivalent to IV (β-strand) and VI (β-strand-loop-β-strand) were located close together on the floor of the cleft within the palm subdomain, with the three conserved D residues in these motifs (Table I), lying close to the 3′-OH of the primer and the NTP binding site in the polymerase catalytically active site; it was suggested that the function of the three D residues may be to bind the Mg^{2+} ions needed for polymerase activity. Recent studies of the poliovirus RdRp (3D protein) indicate that it also has a "hand" structure (Schultz *et al.,* 1995). Richards *et al.* (1992) located the NTP binding site of the polio-

virus 3D protein to a region spanning motifs II to III, the C-terminal part of the region being 12 residues from the conserved D residue in motif IV. Hayes *et al.* (1994a) and Bates *et al.* (1995) showed that antibodies raised to peptides corresponding to motif VI in the RdRp proteins of cucumber mosaic virus (2a protein) and red clover necrotic mosaic virus (88-kDa protein) inhibited initiation of RNA synthesis by

TABLE I

CONSERVATION OF ASPARTIC ACID (D) RESIDUES IN AMINO ACID SEQUENCE MOTIFS IN RNA-DEPENDENT RNA POLYMERASES AND OTHER TYPES OF NUCLEIC ACID POLYMERASES

Polymerase[a]	Genome[b]	Lineage[c]	Virus or host[d]	Motif IV[e]	Number[f]	Motif VI[e]
RdRp	RNA (+)	1. Picorna	PV	F **D** YTGY.DASLS	80	MIAYG **D.D** VIAS
			CPMV	C **D** YSSF.DGLLS	95	LVTYG **D.D** NLIS
		1. Poty	TEV	A **D** GSQF.DSSLT	85	YYVNG **D.D** LLIA
			BYMV	G **D** GSRF.DSSID	89	FVCNG **D.D** NKFA
		1. Sobemo	SBMV	A **D** ISGF.DWSVQ	78	CIAMG **D.D** SVEG
			PLRV	T **D** CSGF.DWSVA	79	AMAMG **D.D** ALEA
		1. Arteri	IBV	W **D** YPKC.DRAMP	127	LMILS **D.D** GVVC
			EAV	T **D** LESC.DRSTP	100	VYIYS **D.D** VVL.
		1. Astro	HAstV	F **D** WTRY.DGTIP	92	TVVYG **D.D** RLST
		2. Phage	phage Qβ	V **D** LSAASDS.IS	66	VTVYG **D.D** IILP
			phage MS2	I **D** LSSASDS.IS	67	IGIYG **D.D** IICP
		2. Flavi	YFV	D **D** TAGW.DTRIT	117	MAVSG **D.D** CVVR
			TBEV	D **D** TAGW.DTKVT	114	MLVSG **D.D** CVVR
		2. Pesti	HCV	Y **D** TRCF.DSTVT	86	MLVCG **D.D** LVVI
			BVDV	F **D** TKAW.DTQVT	88	IHVCG **D.D** GFLI
		2. Carmo	TBSV	L **D** ASRF.DQHCS	82	LANCG **D.D** CVLI
			RCNMV	L **D** ASRF.DQHCS	83	LANNG **D.D** CVLV
		3. Tymo	TYMV	N **D** YTAF.DQSQH	71	IMVSG **D.D** SLID
			PVX	N **D** YTAF.DQSQD	72	IMVSG **D.D** SLID
		3. Tobamo	TMV	L **D** ISKY.DKSQN	78	GAFCG **D.D** SLLY
			BMV	A **D** LSKF.DKSQG	78	AIFSG **D.D** SLI.
		3. Rubi	SINV	T **D** IASF.DKSQD	80	AAFIG **D.D** NIIH
			BNYVV	I **D** AAAC.DSGQC	76	MAMKG **D.D** GFK.
RdRp	RNA (–)		FLUA	G **D** NTKW.NENQN	124	GLQSS **D.D** FALI
			SENV	T **D** LKKY.CNLWR	94	AMVQG **D.N** QAIA
RdRp	RNA (ds)		BTV	I **D** FGYG.EGRVA	110	EQYVG **D.D** TLFY
			IBDV	I **D** LEKG.EANCT	103	IERSI **D.D** IRGK

TABLE I *(continued)*

Polymerase[a]	Genome[b]	Lineage[c]	Virus or host[d]	Motif IV[e]	Number[f]	Motif VI[e]
RdDp	RNA (+)		MMTV	I **D** LQDC.FFNI	61	IVHYM **D.D** ILLA
			HIV	C **D** VGDA.YSFV	61	IYQYM **D.D** LYVG
RdDp	DNA		HBV	L **D** VSAA.FYHL	109	AFSYM **D.D** VVLG
			CaMV	F **D** CKSG.FWQV	49	CCVYV **D.D** ILVF
DdDp	DNA		EBV	F **D** FASL.YPSI	158	RIIYG **DTD** SIFV
	DNA		Humα	L **D** FNSL.YPSI	129	EVIYG **DTD** SIMI
	DNA		phage T7	I **D** ASGL.ELRC	168	MAWVH **D.E** IQVG
	DNA		*E. coli* I	A **D** YSQI.ELRI	169	INQVH **D.E** ELVF

[a] RdRp, RNA-dependent RNA polymerase; RdDp, RNA-dependent DNA polymerase (reverse transcriptase); DdRp, DNA-dependent RNA polymerase; DdDp, DNA-dependent DNA polymerase.

[b] RNA (+), positive-stranded RNA; RNA (ds), double-stranded RNA; RNA (–), negative-stranded RNA.

[c] The numbers and names (given for the positive-stranded RNA viruses only) refer to the RdRp supergroups and lineages from Koonin (1991) and Koonin and Dolja (1993). Two examples of each group are given.

[d] Abbreviations: PV, poliovirus; CPMV, cowpea mosaic virus; TEV, tobacco etch virus; BYMV, barley yellow mosaic virus; SBMV, southern bean mosaic virus; PLRV, potato leafroll virus; IBV, avian infectious bronchitis virus; EAV, equine arteritis virus; HAstV, human astrovirus; YFV, yellow fever virus; BVDV, bovine diarrhea virus; TBSV, tomato bushy stunt virus; RCNMV, red clover necrotic mosaic virus; TYMV, turnip yellow mosaic virus; PVX, potato virus X; TMV, tobacco mosaic virus; BMV, brome mosaic virus; SINV, Sindbis virus; BNYVV, beet necrotic yellow vein virus; FLUA, influenza A virus; SENV, Sendai virus; BTV, Blue tongue virus; IBDV, infectious bursal disease virus; MMTV, mouse mammary tumor virus; HIV, human immunodeficiency virus; HBV, hepatitis B virus; CaMV, cauliflower mosaic virus; EBV, Epstein-Barr virus; Humα, human DNA polymerase α; *E. coli* I, *E. coli* DNA polymerase I.

[e] Sequence data and motifs are from Koonin (1991a), Koonin and Dolja (1993), Poch *et al.* (1989), Delarue *et al.* (1990), Heringa and Argos (1994). Sequence motifs IV and VI of Koonin (1991a), and Koonin and Dolja (1993) are counterparts of motifs A and C of Poch *et al.* (1989), Delarue *et al.* (1990), and Heringa and Argos (1994). The human astrovirus sequences were from Lewis *et al.* (1994). Aspartic acid (D) residues that are invariant or nearly invariant are shown in bold.

[f] Numbers of amino acid residues between motifs IV and VI.

template-dependent RNA polymerase complexes isolated from plants infected with the respective viruses (shown to contain the 2a and 88-kDa proteins respectively), suggesting that the inhibitory effect of the antibodies may be to interfere with the binding of the template at the catalytically active site of the enzyme. The observation that the antibodies did not inhibit the activity of polymerase complexes con-

taining bound RNA template is consistent with this hypothesis.

On the basis of sequence similarities extending over 300 amino acids, Koonin (1991a) and Koonin and Dolja (1993) classified the RdRps of positive-stranded RNA viruses into three supergroups (1, 2, and 3), with a number of different lineages within each supergroup (Table II). All three supergroups contain RdRps of viruses of animals and plants, and supergroup 2 contains additionally RdRps of the bacterial positive-stranded RNA virus *Leviviridae* family. Although the supergroups and lineages are undoubtably important in establishing likely evolutionary relationships between viral RdRps, it is not yet clear to what extent the sequence differences which define the supergroups reflect any significant differences in the biological properties of the RdRps or their roles in RNA replication. Evolution of RNA replication systems is probably best considered in terms of coevolution of the RNA polymerase proteins with other proteins in the replication complex and with cis-acting sequences required for RNA replication. Isolated polymerase subunits [e.g., Qβ RNA polymerase subunit 2 (Landers *et al.*, 1974); poliovirus 3D protein (van Dyke and Flanegan, 1980; Cho *et al.*, 1993)] are able to catalyze RNA synthesis only on primed templates and synthesis is not template-specific. Initiation of positive-stranded and negative-stranded RNA synthesis on unprimed templates requires interaction of the core polymerase with additional proteins [host proteins in the case of Qβ (Blumenthal and Carmichael, 1979; van Duin, 1988), both host- and virus-encoded proteins in the case of poliovirus (Barton *et al.*, 1995; Xiang *et al.*, 1995a,b; McBride *et al.*, 1996)], some of which will also interact specifically with the viral RNA (which may be terminal or internal structures), conferring template specificity on the replication system. Some replication proteins are clearly multifunctional and may need to interact with different viral or host proteins at various stages of the virus replication cycle. Charged-to-alanine mutagenesis of the poliovirus RdRp (3D protein) yielded multiple temperature-sensitive mutants defective in RNA synthesis, many of which mapped in the N-terminal third of the protein well away from the conserved polymerase motifs (Diamond and Kirkegaard, 1994). Since clustered charged-to-alanine mutagenesis is designed to target residues on the surface of folded proteins, these mutants may be indicative of sites of interaction of the 3D polymerase with other viral or host proteins. Another ts mutant with a M→T mutation in the C-terminal portion of the $3D^{pol}$ protein was shown to be defective in the initiation of RNA replication, but not in the elongation of nascent chains (Barton *et al.*, 1996). This could also be indicative of a site of interaction with another viral or host protein, or the RNA template. Different modes of expression of the

genes encoding replication proteins, which control the stoichiometry of their synthesis, and the synthesis of subgenomic RNAs by some viruses which requires additional recognitions and controls are clearly also important in considering the evolution of replication systems. Brief details of some additional properties, relevant to replication, of the viruses in the three supergroups are included in Table II, together with virus supergroup designations based on these and other properties of the viruses (Goldbach and de Haan, 1994). The classification of hepatitis E virus in the *Caliciviridae* family (Murphy *et al.,* 1995) may need to be revised because the virus shares many properties with viruses in the alpha-like supergroup.

B. Helicases

Most positive-stranded RNA viruses contain ORFs with the potential to encode proteins with amino acid sequence motifs characteristic of well-defined RNA helicases and in many cases such ORFs have been shown to be essential for RNA replication, e.g., linker insertions in the helicase-like domain of the brome mosaic virus 1a protein were either lethal or rendered the virus temperature-sensitive for replication of all classes of RNA (positive-strand, negative-strand, subgenomic) (Kroner *et al.,* 1990), and some monoclonal antibodies which mapped to the helicase-like domain of the cucumber mosaic virus (CMV) 1a protein partially inhibited a purified CMV replicase complex, shown to contain the 1a protein (Hayes and Buck, 1990; Hayes *et al.,* 1994a). Helicases could function to unwind duplexes formed during RNA replication to allow strands to act as templates for further replication. They could also have an important function in removing secondary structure from RNA templates, e.g., to aid in initiation of negative-strand synthesis on templates with extensive 3′ secondary structure, such as those with tRNA-like 3′ termini, and to increase the processivity of RNA polymerases through regions of internal secondary structure. In DNA replication, duplex unwinding is generally carried out by helicases and removal of secondary structure from the resultant single-stranded templates by single-stranded DNA-binding proteins (Kornberg and Baker, 1992). However, cellular RNA helicases often have the function of removing secondary structure from ssRNA, e.g., the eukaryotic translational initiation factor, eIF-4A (Merrick, 1992). On the basis of conserved motifs, cellular and viral RNA and DNA helicases (and putative helicases) have been classified into a number of superfamilies (Gorbalenya *et al.,* 1988; Hodgeman, 1988; Habili and Symons, 1989; Gorbalenya and Koonin, 1989; Gorbalenya *et al.,* 1989, 1990; Lain *et al.,* 1989; Koonin,

TABLE II

PROPERTIES OF POSITIVE-STRANDED RNA VIRUSES

RdRp Supergroup[a]	RdRp Lineage[a]	Virus family or genus[b]	Virus examples	Host[c]	No. of genome segments	Mode of expression of RdRp[d]	5′ End of RNA	3′ End of RNA[e]	Production of subgenomic RNA	Helicase superfamily[a]	Helicase lineage[a]	Virus supergroup[f]	Reference[g]
1	Picorna	*Picornaviridae*	Polio, human rhino, hepatitis A, foot and mouth disease, encephalomyocarditis	A	1	PP	VPg	poly(A)	No	3	Picorna	Picorna-like	1
1	Picorna	*Sequiviridae*	Parsnip yellow fleck	P	1	PP	VPg?	S		3	?	Picorna-like	2, 3
			Rice tungro spherical	P	1	PP	VPg?	poly(A)		3	Como		3, 4
1	Picorna	*Comoviridae*	Cowpea mosaic, tobacco ringspot	P	2	PP	VPg	poly(A)	No	3	Como	Picorna-like	5, 6
	Picorna	*Caliciviridae*	Feline calici, rabbit hemorrhagic disease	A	1	PP	VPg	poly(A)	Yes	3	Calici	Picorna-like	7
1	Poty	*Potyviridae*	Potato Y, tobacco etch, plum pox	P	1	PP	VPg	poly(A)	No	2	Poty	Picorna-like	8
			Barley yellow mosaic	P	2	PP	VPg	poly(A)	No	2	Poty		8
1	Sobemo	*Sobemovirus*	Southern bean mosaic	P	1	PP	VPg	S	Yes			Sobemo-like	9
			Cocksfoot mottle	P	1	FS, PP?	VPg	S	Yes				10
1	Sobemo	*Luteovirus* subgroup II	Potato leafroll, beet western yellows, barley yellow dwarf (RGV, RMV, RPV)	P	1	FS, PP?	VPg	S	Yes			Sobemo-like	11
1	Sobemo	*Enamovirus*	Pea enation mosaic RNA 1	P		FS, PP?	VPg?	S	Yes				12, 13
1	Sobemo	*Barnaviridae*	Mushroom bacilliform	F	1	FS?	?	S	Probably				14

1	Sobemo	*Nodaviridae*	Black beetle, flock-house, Nodamura	A	2	D	Cap	S	Yes				15
1	Arteri	*Coronaviridae*	Avian infectious bronchitis, mouse hepatitis, Berne	A	1	FS, PP	Cap	poly(A)	Yes	1	Arteri	Corona-like	16, 17, 54
1	Arteri	*Arterivirus*	Equine arteritis, simian hemor-rhagic fever	A	1	FS, PP	Cap	poly(A)	Yes	1	Arteri	Corona-like	18, 19
1	Astro	*Astroviridae*	Human astro	A	1	FS, PP?	?	poly(A)	Yes				20, 21
2	Phage	*Leviviridae*	Phage Qβ, R17, MS2	B	1	D	ppp	S	No				22
2	Flavi	*Flaviviridae*	Yellow fever, tick-borne encephalitis	A	1	PP	Cap	S	No	2	Flavi	Flavi-like	23
2	Pesti	*Flaviviridae*	Bovine viral diar-rhea, hepatitis C	A	1	PP	?	S	No	2	Flavi	Flavi-like	24, 25
2	Carmo	*Tombusviridae*	Tomato bushy stunt, cucumber necrosis, cymbidium ring-spot, carnation mottle, turnip crinkle	P	1	RT	Cap	S	Yes			Carmo-like	26
2	Carmo	*Machlomovirus*	Maize chlorotic mottle	P	1	RT	Cap	S	Yes			Carmo-like	26–28
2	Carmo	*Necrovirus*	Tobacco necrosis	P	1	RT	pp	S	Yes			Carmo-like	26, 29
2	Carmo	*Dianthovirus*	Red clover necrotic mosaic	P	2	FS	Cap	S	Yes			Carmo-like	26, 30
2	Carmo	*Luteovirus* subgroup I	Barley yellow dwarf (MAV, PAV, SGV)	P	1	FS	?	S	Yes			Carmo-like	11, 26
2	Carmo	*Enamovirus*	Pea enation mosaic RNA 2	P		FS	?	S	Probably				12, 13
2	Carmo	*Umbravirus*	Carrot mottle , groundnut rosette	P	1	?	?	S	Probably				31
2	Carmo	Unclassified	Beet western yellows ST9-associated RNA	P		FS	?	S	Probably				11, 32

(continued)

TABLE II *(continued)*

RdRp Supergroup[a]	RdRp Lineage[a]	Virus family or genus[b]	Virus examples	Host[c]	No. of genome segments	Mode of expression of RdRp[d]	5′ End of RNA	3′ End of RNA[e]	Production of subgenomic RNA	Helicase superfamily[a]	Helicase lineage[a]	Virus supergroup[f]	Reference[g]
3	Tymo	*Capillovirus*	Apple stem grooving	P	1	D	?	poly(A)	Probably	1	Tymo	Alpha-like	33-35
3	Tymo	*Carlavirus*	Carnation latent, potato M, blueberry scorch	P	1	PP?	Cap?	poly(A)	Yes	1	Tymo	Alpha-like	36, 55
3	Tymo	*Trichovirus*	Apple chlorotic leafspot, potato T	P	1	D	Cap?	poly(A)	Probably	1	Tymo	Alpha-like	34, 35
3	Tymo	*Tymovirus*	Turnip yellow mosaic, eggplant mosaic	P	1	PP	Cap	tRNA-like	Yes	1	Tymo	Alpha-like	37, 38
3	Tymo	*Potexvirus*	Potato X, cymbidium mosaic, foxtail mosaic	P	1	D	Cap	poly(A)	Yes	1	Tymo	Alpha-like	39
3	Tobamo	*Tobamovirus*	Tobacco mosaic, pepper mild mottle	P	1	RT	Cap	tRNA-like	Yes	1	Tobamo	Alpha-like	40
3	Tobamo	*Tobravirus*	Tobacco rattle, pea early browning, pepper ringspot	P	2	RT	Cap	tRNA-like	Yes	1	Tobamo	Alpha-like	41
3	Tobamo	*Hordeivirus*	Barley stripe mosaic	P	3	D	Cap	tRNA-like	Yes	1	Tobamo	Alpha-like	42
3	Tobamo	*Furovirus*	Soil-borne wheat mosaic	P	2	RT	Cap	tRNA-like	Probably	1	Tobamo	Alpha-like	43
3	Tobamo	*Idaeovirus*	Raspberry bushy dwarf	P	2	D	Cap?	S	Yes	1	Tobamo	Alpha-like	44, 45
3	Tobamo	*Bromoviridae*	Brome mosaic, cucumber mosaic	P	3	D	Cap	tRNA-like	Yes	1	Tobamo	Alpha-like	46, 47
			Alfalfa mosaic, tobacco streak	P	3	D	Cap	S	Yes	1			48

3	Tobamo	*Closterovirus*	Beet yellows, citrus tristeza	P	1	FS, PP	Cap	S	Yes	1	Tobamo	Alpha-like	33, 34
3	Tobamo	*Closterovirus* (Tentative)	Lettuce infectious yellows	P	2	FS, PP	Cap?	S	Yes	1	Tobamo	Alpha-like	49
3	Rubi	*Togaviridae*	Semliki Forest, Sindbis, rubella	A	1	PP	Cap	poly(A)	Yes	1	Rubi	Alpha-like	50, 51
3	Rubi	*Caliciviridae*	Hepatitis E	A	1	PP	Cap	poly(A)	Yes	1	Rubi	Alpha-like?	52
3	Rubi	*Furovirus* (Tentative)	Beet necrotic yellow vein virus	P	4	D	Cap	poly(A)	Yes	1	Rubi	Alpha-like	53

[a] RNA-dependent RNA polymerase supergroups and lineages, and helicase superfamilies are from Koonin and Dolja (1993).

[b] Families and genera are from Murphy *et al.* (1995). A genus is only given when not assigned to a family.

[c] A, animal; B,bacterium; F, fungus; P, plant.

[d] PP, polyprotein processing; RT, readthrough; FS, frameshift; D, direct translation.

[e] S, structure other than tRNA-like or poly(A).

[f] Virus supergroups as defined by Goldbach and de Haan (1994).

[g] Data are from Murphy *et al.* (1995) and the following. 1. Wimmer *et al.* (1993). 2. Turnbull-Ross *et al.* (1993). 3. Reavy *et al.* (1993). 4. Shen *et al.* (1993). 5. Peters *et al.* (1995). 6. Mayo and Fritsch (1994). 7. Lambden and Clarke (1995). 8. Riechman, J. L. *et al.* (1992). 9. Othman and Hull (1995). 10. Makinen *et al.* (1995). 11. Miller *et al.* (1995). 12. Demler *et al.* (1993). 13. Demler *et al.* (1994). 14. Revill *et al.* (1994). 15. Ball, L. A. (1995). *J. Virol.* **69**, 720–727. 16. Lai (1990). 17. Snijder and Horzinek (1993). 18. Plagemann and Moenning (1992). 19. Snijder *et al.* (1995). 20. Jiang *et al.* (1993). 21. Willcocks *et al.* (1994). 22. Van Duin (1988). 23. Chambers *et al.* (1990). 24. Collett (1992). 25. Matsuura and Miyamura (1993). 26. Russo *et al.* (1994). 27. Nutter *et al.* (1989). 28. Lommel *et al.* (1991). 29. Meulewater *et al.* (1992). *J. Virol.* **66**, 6419–6428. 30. Giesman-Cookmeyer *et al.* (1995). 31. Gibbs (1995). 32. Passmore *et al.* (1993). 33. Coffin and Coutts (1993). 34. Dolja *et al.* (1994). 35. Candresse (1993). 36. Foster (1992). 37. Kadare *et al.* (1992). 38. Kadare *et al.* (1995). 39. Solovyev *et al.* (1994). 40. Dawson and Lehto (1990). 41. Zerfass and Beier (1992). 42.Donald and Jackson, 1994. 43. Shirako and Wilson (1993). 44. Natsuaki *et al.* (1991). 45. Ziegler *et al.* (1992). 46. Ahlquist (1992). 47. Palukaitis *et al.* (1992). 48. Reusken *et al.* (1995). 49. Klaasen *et al.* (1995). 50. Strauss and Strauss (1994). 51. Frey (1994). 52. Purdy *et al.* (1993). 53. Richards and Tamada (1992). 54. Cavanagh and MacNaughton (1994). 55. Hillman and Lawrence (1995).

1991b; Bork and Koonin, 1993; Koonin and Dolja, 1993), with several lineages or subfamilies associated with each superfamily. The superfamily and lineage designations for the positive-stranded virus RNA helicases and putative helicases (Koonin and Dolja, 1993) are given in Table II. Superfamilies 1 and 2, which have been shown to be distantly related, were each characterized by seven motifs (which could all be aligned between the two superfamilies), whereas superfamily 3 has just three motifs. Two of the motifs, versions of which are present in all three superfamilies (designated I and II for superfamilies 1 and 2, and A and B for superfamily 3), are variants of the ATP-binding motifs, first described by Walker *et al.* (1982): A, GXXXXGK[TS] and B, ΦΦΦΦD where X is an unspecified amino acid residue, Φ is a hydrophobic residue, and the residues in square brackets are alternates. Variants of motif A are present in a vast class of both ATP-binding proteins and GTP-binding proteins; crystallographic structural studies on several of these indicate that motif A (I) is in the NTP-binding site located in a phosphate-binding loop (P-loop) between a β-strand and an α-helix. The B (II) motif may provide D residues required for chelation of Mg^{2+} ions which in turn bind to the terminal phosphates and promote NTP hydrolysis (Saraste *et al.*, 1990; Schulz, 1992). The consensus sequences for the motifs for the three viral superfamilies (Koonin and Dolja, 1993) are shown in Table III.

The (putative) helicases of the positive-stranded RNA viruses in superfamily 2 constitute a subset of proteins containing variants of the sequence DEAD (DEAD box proteins), present in motif II as DEXH, some of which have been demonstrated to have RNA-dependent ATPase and RNA helicase activity, e.g., eIF-4A (Rozen *et al.*, 1990; Jaramillo *et al.*, 1990), the human nuclear protein p68 (Hirling *et al.*, 1989), vaccinia virus nucleoside triphosphohydrolase (I8R protein) (Schuman, 1992, 1993; Bayliss and Smith, 1996), and human RNA helicase A (Lee and Hurwitz, 1993). These helicases function in a 3′ to 5′ direction, requiring a duplex RNA with a 3′ overhang, except for eIF-4A which has bidirectional activity (but requires also eIF-4B for helicase activity). Mutagenesis of eIF-4A has shown that the AXXXXGKT motif (see motif I, Table III) is required for ATP binding, whereas the DEAD sequence (see motif II, Table III) functions in ATP hydrolysis and coupling of ATP hydrolysis to helicase activity. Two other motifs, SAT (present in motif IV, Table III) and **HRIGRXXR** (the residues shown in bold are present in motif VI, Table III) were shown to be needed for eIF-4A RNA helicase activity, the latter sequence functioning specifically in RNA binding (Pause and Sonenberg, 1992; Pause *et al.*, 1993). The counterpart of the latter motif in the vaccinia virus DEXH box RNA

TABLE III

CONSENSUS SEQUENCE OF CONSERVED MOTIFS IN SUPERFAMILIES OF POSITIVE-STRANDED RNA VIRUS HELICASE-LIKE PROTEINS

Super-family[a]		N[c]		N		N		N
	Motif I[b]		Motif IA		Motif II		Motif III	
1	XXX&X**G**XP**G**X**GKT**XX&XX A A **S**	1–9	XXXX&XXXXXXXXXX&	12–52	X&&&**DE**&X **D**A	8–40	&&&&GDXXQ AACC	9–30
2	&&XXXX**G**S**GKT**XXX&P A **S**	6–14	XRX&UUXPTRXUXXE& K A SK A N	43–51	&&&&**DE**XH	16–24	X&XUTATPP S	39–52
	Motif A				Motif B		Motif C	
3	XEP&X&&&X**G**XX**G**X**GKS**XXX D C **T**	26–32			XQX&&U&**DD** E AC **E**	30–36	KGXX@XSX&U&XSTNX TS	

	Motif IV	N	Motif V	N	Motif VI
1	XXXXXXXRX	46–86	XXXX**T**&XXXQ**G**XT&XXVX&&X SA S AC	6–14	XXXXX&VAUT**R**XXX G S
2	X&XUTATPP S	31–40	&U&X**T**D&XEX**G**UX&XXXXXUU N A	20–28	TXXXXXQRXG**R**UGR

[a] Superfamilies and data from Koonin and Dolja (1993).

[b] X, unspecified residue; U, bulky aliphatic residue (I, L, V, M); @, aromatic residue (F, Y, W); &, bulky hydrophobic residue (aliphatic or aromatic). Alternate residues are shown below the sequence. Residues conserved in motifs between more than one family are shown in bold. Residues underlined in superfamily 2 are conserved in corresponding motifs in DEAD box proteins (Fuller-Pace, 1994).

[c] N, Number of amino acid residues between the motifs (range).

helicase was required for ATP hyrolysis and RNA unwinding, but not for RNA binding, suggesting that the contribution of conserved helicase motifs to overall protein function may be context-dependent (Gross and Schuman, 1996).

RNA-stimulated ATPase, RNA binding, and RNA helicase activity have also been shown for a number of positive-stranded RNA virus proteins in the DEXH subset of helicase superfamily 2, the CI protein of plum pox virus (PPV) (Table II, Poty helicase lineage) (Lain *et al.*, 1990, 1991; Fernandez *et al.*, 1995), the tamarillo mosaic virus (TaMV) CI protein (poty helicase lineage) (Eagles *et al.*, 1994); the bovine viral diarrhea virus (BVDV) p80 (NS3) protein, and equivalent proteins of yellow fever virus and hepatitis C virus (all flavi helicase lineage) (Suzich *et al.*, 1993; Tamura *et al.*, 1993; Warrener *et al.*, 1993; Warrener and Collett, 1995; unpublished results quoted in Warrener and Collett, 1995). The helicase activities of the PPV CI protein (isolated from infected plants) and the BVDV NS3 protein (expressed in insect cells) were shown to require a dsRNA substrate with a 3′ single-stranded overhang and to function in a 3′ to 5′ direction; dsRNA with a 5′ overhang or blunt-ended dsRNA was not a substrate for the helicase. Helicase activity of the TaMV CI protein (isolated from infected plants) also functioned in a 3′ to 5′ direction (other substrates were not tested). The helicases of all three viruses could be stimulated by all four NTPs, with a slight preference in the case of PPV protein for purine NTPs. Binding of the PPV and TaMV CI proteins to RNA did not require an NTP; deletion mutagenesis of the PPV CI protein (expressed in *E. coli* as a maltose-binding fusion) located the RNA binding site to a region containing motif VI. The RNA-dependent ATPases of West Nile and yellow fever flaviviruses (expressed in *E. coli*) were located to a C-terminal fragment of the NS3 protein containing the seven conserved helicase motifs (Table III) (Warrener *et al.*, 1993; Wengler and Wengler, 1993). It is noteworthy that the basal level of ATPase activity of the viral helicases in the absence of RNA was much higher than that of the cellular RNA helicases in superfamily 2.

In the helicase superfamily 3, RNA (and DNA) helicase activity has been demonstrated for the simian virus 40 (SV40) large T antigen (Scheffner *et al.*, 1989). Like the viral RNA helicases of superfamily 2, the SV40 RNA helicase activity functioned in a 3′ to 5′ direction; furthermore, mutations of DE to AA in motif B abolished the ATPase activity of this protein (Weiner and Bradley, 1991). No helicase activity has yet been demonstrated for a positive-stranded RNA virus protein in this superfamily, but studies have indicated the importance of the putative helicases of cowpea mosaic virus (58-kDa protein or region of

B polyprotein precursor; como helicase lineage) and poliovirus (2C protein; picorna helicase lineage) in virus replication. Mutation of the conserved K500 residue to T in the motif A sequence GKSRTGK500S in the 58-kDa domain reduced cowpea mosaic virus B RNA replication in cowpea protoplasts to an undetectable level (Peters *et al.*, 1994). A mutation of the conserved D545 to P in the motif B sequence MDD545 was also lethal. Both mutations were shown to act at the protein, rather than RNA, level. However, whereas the K→T mutation in motif A caused a small reduction in ATP binding and an altered distibution of viral proteins, which failed to aggregate into the large cytopathic structures observed in protoplasts infected with wild-type B-RNA, no such effects were observed with the D→P mutation in motif B.

The poliovirus 2C protein and its NTP binding site have been shown to be essential for RNA replication. The 2C protein is present in membranous replication complexes, isolated either from infected cells (Bienz *et al.*, 1990) or from a combined *in vitro* translation–replication system (Barton *et al.*, 1995). Poliovirus RNA replication is inhibited by guanidine hydrochloride at the level of initiation of RNA synthesis (Caliguiri and Tamm, 1973; Barton *et al.*, 1995) and guanidine-resistant (gr) and guanidine-dependent (gd) mutants map to the 2C protein (Wimmer *et al.*, 1993; Tolskaya *et al.*, 1994), implying a role for the 2C protein in the initiation of RNA synthesis. Mutations in the conserved amino acids of the NTP-binding motifs A (129GSPGTGKS136) and B (176DD177), e.g., G129I, K135Q, K135R, S136T, S136A, D176L, and D177L, abolished or greatly reduced RNA synthesis (Mirzayan and Wimmer, 1992; Teterina *et al.*, 1992). The 2C protein, expressed in *E. coli* as a maltose-binding protein fusion (Rodriguez and Carrasco, 1993) or in insect cells not as a fusion (Mirzayan and Wimmer, 1994a), has been shown to have ATPase activity. A K135Q mutation in motif A abolished ATPase activity of the insect-expressed protein. Taken together with the previous observation that this mutation in the virus greatly reduces RNA replication, this result implies that ATP hydrolysis, dependent on this site, is required for RNA replication. The defect in the mutant was probably in ATP binding since the wild-type, but not the mutant, 2C protein bound to ATP-agarose. The guanidine-resistant and guanidine-sensitive mutants mapped to a region overlapping the conserved helicase motifs, although none of the conserved amino acids was altered (Wimmer *et al.*, 1993; Tolskaya *et al.*, 1994). The 2C ATPase activity was not sensitive to concentrations of guanidine which inhibit poliovirus RNA replication *in vivo* (Mirzayan and Wimmer, 1994a) and it was suggested that the region targeted by guanidine, while not being directly involved in ATP binding or hydrolysis, may nevertheless

be important in a coupled downstream event, such as RNA binding and duplex unwinding (Tolskaya *et al.,* 1994), as described above for some of the superfamily 2 helicases. However, the 2C ATPase activity was not stimulated (Mirzayan and Wimmer, 1994a), or modestly stimulated (twofold) (Rodriguez and Carrasco, 1993), by addition of RNA, although the 2C protein contains a consensus dsRNA-binding sequence (Paul *et al.,* 1994b). Hence, if the 2C protein is a helicase, it differs from the superfamily 2 helicases described above, although the possibility that the (putative) 2C helicase activity requires interaction with another viral or cellular protein for interaction with RNA and unwinding activity cannot be discounted. It is also noteworthy that the poliovirus $3D^{pol}$ protein has an unwinding (strand-displacement activity) (Cho *et al.,* 1993), as described in Section I. Although none of these experiments rules out the hypothesis that the 2C protein is a helicase with a role in the initiation of RNA replication, the possibility arises that ATP hydrolysis might be coupled to some other function of the 2C protein, such as virion uncoating (Li and Baltimore, 1988, 1990), encapsidation of viral RNA, or membrane trafficking (both of which are likely to be important in RNA replication), as suggested by Mirzayan and Wimmer (1994a).

RNA helicase activity has not yet been demonstrated for any of the proteins in superfamily 1. Duplex unwinding in a 3′→5′ direction has been demonstrated for a purified RdRp complex, isolated from plants infected with alfalfa mosaic virus (de Graaf *et al.,* 1995a). Such RdRp preparations contain the helicase-like P1 protein (Quadt *et al.,* 1991), although it was not demonstrated whether the unwinding was due to the P1 protein, the P2 polymerase, or some other protein component of the complex. Several members of helicase superfamily 2 have DNA helicase activity, e.g., the *E. coli* Rep protein, which has DNA-dependent ATPase activity, translocates in the 3′→5′ direction, and can unwind RNA/DNA hybrids (Kornberg and Baker, 1992).

The nsP2 protein of Semliki Forest virus is a multifunctional protein which has several essential functions in RNA synthesis, as well as a C-terminal papain-like protease domain (Strauss and Strauss, 1994). Its N-terminal half contains sequence motifs characteristic of the helicase superfamily 1 (rubi lineage, Table II) and a mutant with a C→Y mutation at residue 304 between motifs IV and V was temperature-sensitive for subgenomic RNA synthesis (Hahn *et al.,* 1989b). The nsP2 protein when expressed in *E. coli* as a fusion protein with a short N-terminal addition containing a histidine tag was found to have ATP and GTP binding activity, ATPase and GTPase activity further stimulated by the presence of ssRNA, and RNA binding capacity. Similar activities were found with a truncated nsP2 protein containing the

N-terminal 470 amino acid residues, which contained the helicase-like motifs but lacked the C-terminal protease domain. A mutant nsP2 protein in which the conserved lysine residue of motif I was replaced by asparagine exhibited no ATPase or GTPase activity. By analogy with the superfamily 2 and 3 NTPases discussed above, it is likely that this residue functions in NTP binding. It is likely that the NTPase activity of the nsP2 protein is coupled in some way to its function in RNA synthesis, but attempts to demonstrate helicase activity in the *E. coli*–expressed fusion protein were unsuccessful. As discussed for the poliovirus 2C protein, it is possible that interaction with another viral or host protein is needed for helicase activity.

A segment of the rubella virus nonstructural protein containing the helicase-like domain, when expressed as part of a fusion protein in *E. coli*, was shown to have RNA-stimulated NTPase activity (Gros and Wengler, 1996). Although helicase or NTPase activity was not demonstrated for the turnip yellow mosaic virus 140-kDa protein, a K→S substitution in the motif I sequence GFAGCGKT of its helicase-like domain resulted in complete loss of detectable RNA replication (Weiland and Dreher, 1993).

It is noteworthy that all the supergroup 3 RNA polymerases were found in viruses with superfamily 1 helicase-like proteins (Table II). There was also complete correspondence between the polymerase and helicase lineages of these viruses. This probably indicates coevolution and a long-standing association of the polymerase-like and helicase-like genes of these viruses, some of which have animal hosts and others plant hosts, and which constitute the alpha-like (Sindbis-like) virus supergroup (Goldbach and de Haan, 1994). Similarly, the supergroup 1 RNA polymerases of the picorna lineage were found in animal and plant viruses with superfamily 3 helicase-like proteins (Table II), again indicating likely coevolution of these genes. However, plant viruses with supergroup 1 RNA polymerases of the poty lineage and animal viruses with supergroup 2 polymerases of the flavi and pesti lineages both had helicase-like proteins of superfamily 2 (Table II) (Lain *et al.*, 1989). This could indicate creation of new combinations of polymerase and helicase genes in the evolution of these viruses. The coronaviruses, toroviruses, and arteriviruses, which are somewhat divergent from other positive-stranded RNA viruses, contain a unique combination of supergroup 1 polymerases and superfamily 1 helicase-like proteins. As helicase activity has as yet only been demonstrated for positive-stranded RNA viruses of the helicase superfamily 2 and the role of such activities in the replication of the viruses concerned has yet to be confirmed, it is premature to speculate as to what extent the sequences

which define the different helicase superfamilies are reflected in any significant differences in their activities or roles in RNA replication.

It is noteworthy that several families of positive-strand RNA viruses appear to lack the characteristic NTP-binding motifs A (I) and B (II) (Table III). These include some viruses with supergroup I polymerases of the sobemo lineage [the *Sobemovirus* and *Luteovirus* subgroup II genera of plant viruses and the *Nodaviridae* family of animal (mainly insect) viruses] and the animal astroviruses, and some viruses with supergroup II polymerases (the *Tombusviridae* family and *Machlomovirus, Necrovirus, Dianthovirus,* and *Luteovirus* subgroup I genera of plant viruses, and the *Leviviridae* family of ssRNA bacteriophages). All of these viruses, except the astroviruses, have genomes smaller than 6 kb. Typical purine NTP-binding sites could also not be detected in negative-stranded RNA viruses (*Orthomyxoviridae, Paramyxoviridae, Rhabdoviridae*), some dsRNA viruses (*Birnaviridae*), and the retrovirus/pararetrovirus group (*Retroviridae, Hepadnaviridae, Caulimovirus*) (Gorbalenya and Koonin, 1989). There are several possible explanations for this. (i) It is possible that NTP-binding motifs are present, but have diverged too much to be easily recognized by primary sequence comparisons. Such variants in cellular NTP-binding proteins are known (Traut, 1994), e.g., the counterparts of motifs A (I) and B (II) in the ATP binding site of actin are **DNGSGLCKA** and GIVL**DSGDGV**, where the residues in bold have been shown from crystallographic structural analysis to make contact with the ligand (Kabsch *et al.*, 1990). Possible variants of motif A (I) (soybean dwarf luteovirus, SDV; southern bean mosaic sobemovirus, SBMV), motif B (II) (SDV), motif IV (maize chlorotic mottle machlomovirus, MCMV; carnation mottle carmovirus, CMoV; cucumber necrosis tombusvirus, CNV; barley yellow dwarf luteovirus, BYDV-PAV; SDV; beet western yellows luteovirus, BWYV; potato leafroll luteovirus, PLRV), and motif VI (MCMV; CMoV; CNV; BYDV-PAV; SDV; BWYV; PLRV; SBMV) have been recognized (Habili and Symons, 1989), but their significance remains to be determined. (ii) The virus polymerase may have unwinding (strand-displacement) activity, as shown for poliovirus RdRp (Cho *et al.*, 1993). (iii) Unwinding might be accomplished by a helix-destabilizing protein which utilizes the energy of stoichiometric binding to single-stranded nucleic acid to drive melting of a duplex in the absence of NTP hydrolysis. The poliovirus $3D^{pol}$ protein has been shown to bind cooperatively to single-stranded RNA under certain conditions (Pata *et al.*, 1995). (iv) The virus might co-opt a cellular helicase to aid in duplex unwinding.

Some viruses with superfamily 1 putative helicases, such as hordeiviruses, potexviruses, and beet necrotic yellow vein virus (BNYVV),

encode a second protein with helicase-like motifs, which is not required for RNA replication. Such proteins have been considered to arise from the replicative helicase-like protein by a gene duplication event, followed by diversification of function (Koonin and Dolja, 1993). The additional helicase-like proteins of hordeiviruses, potexviruses, and BNYVV have been shown to be involved in virus cell-to-cell movement in infected plants (Petty *et al.*, 1990; Beck *et al.*, 1991; Gilmer *et al.*, 1992a).

1. Models for RNA Replication: The Role of Helicases

Different basic models have been proposed for the replication of positive-stranded RNA viruses, which involve intermediates with different structures and which have implications for the involvement of helicases in the replication process. Three models are shown in Fig. 1. In model 1, the RdRp recognizes a promoter at the 3′ end of the positive-strand RNA template (a) and starts to synthesize a complementary negative strand. The nascent negative strand only remains base-paired to the positive strand in the region where the polymerase binds to the template and is actively synthesizing RNA. The 5′ tail of the nascent negative strand is not base-paired to the template; hence most of the replicative intermediate (b) is in a single-stranded form. Continuation of the reaction leads to the formation of a free negative strand product (c), and releases the positive-strand template. The polymerase then recognizes a promoter at the 3′ end of the negative strand and, using the negative strand as a template, starts to synthesize a progeny positive strand, giving a second type of replicative intermediate (d). As before, the nascent strand is only base-paired to the template in the region of the active site of the polymerase where RNA synthesis is taking place, so that this replicative intermediate is also mainly single-stranded. Before the synthesis of the first progeny positive strand has been completed, initiation of synthesis of further positive strands takes place, giving a replicative intermediate consisting of a full-length negative-stranded template, to which are attached several nascent positive strands which again are largely in a single-stranded form (e). The process continues to synthesize and release multiple copies of the progeny positive-stranded RNA (f).

The first stage of model 2 is essentially the same as that of model 1, except that the negative strand formed remains base-paired with the positive-stranded template, giving a replicative intermediate (b), consisting of partially double-stranded and partially single-stranded structure. The reaction continues to give a fully double-stranded RNA (replicative form) (c). In this model, no free negative strand is synthesized. The polymerase then recognizes a promoter at the end of the

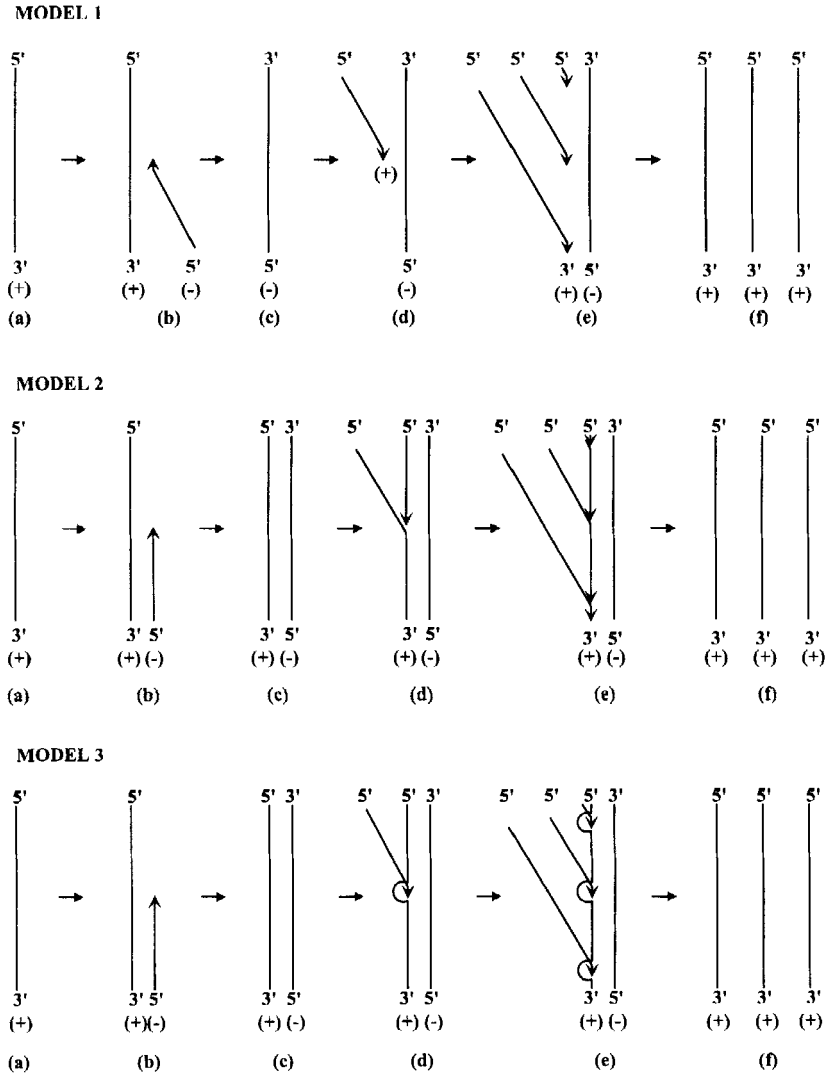


FIG. 1. Models for RNA replication. See text for detail.

replicative form dsRNA containing the 3′ end of the negative strand and the 5′ end of the positive strand. Synthesis of progeny positive-stranded RNA commences using the negative strand as a template by a strand-displacement mechanism, giving rise to replicative intermediates consisting of double-stranded RNA with one (d), or following reinitiations several (e), single-stranded 5′ tails of the full-length posi-

tive strands. The first full-length positive strand to be released from the replicative intermediate will be the original template strand (a); continued reaction will then result in the synthesis and release of multiple progeny positive strands (f). The formation of the double-stranded replicative form RNA in model 3 is exactly the same as in model 2. However, synthesis of progeny positive-stranded RNA using the negative strand of the dsRNA only displaces the positive strand of the dsRNA transiently in the region where RNA synthesis is taking place. The replicative intermediates formed (d, e) consist of double-stranded RNA with one or several single-stranded tails, but unlike the replicative intermediates in model 2 (d, e) in which the single-stranded tails are the displaced 5′ tails of full-length positive strands, these single-stranded tails belong to the nascent, incomplete progeny positive strands. The synthesis of progeny positive strands from a dsRNA replicative form RNA in model 2 is analogous to the semiconservative transcription of dsRNA by strand displacement characteristic of dsRNA viruses of the *Birnaviridae, Cystoviridae,* and *Partitiviridae* families (Buck, 1979; van Etten *et al.,* 1980; Dobos and Roberts, 1983), whereas that in model 3 is analogous to conservative transcription of dsRNA characteristic of dsRNA viruses of the *Reoviridae* and *Totiviridae* families (Shatkin and Kozak, 1983; Fujimura and Wickner, 1989).

Model 1 has been shown to operate in the replication of Qβ RNA. Qβ RNA replicase holoenzyme catalyzes the complete replication of Qβ RNA; it can utilize either the positive or negative strands as templates and produces free progeny positive and negative strands. Replicative intermediates are largely single-stranded and the replicase cannot utilize Qβ dsRNA as a template (Blumenthal and Carmichael, 1979). There is evidence that the ability of the progeny nascent strands to form stable secondary structures is important in the production of single-stranded RNA progeny. Templates with little secondary structure had a tendency to form extended RNA–RNA duplexes during replication, resulting in reduced synthesis of new RNA strands (Priano *et al.,* 1987; Axelrod *et al.,* 1991). It appears that immediately after the complementary base pairs are formed during Qβ RNA replication, they are rapidly unwound, presumably by one of the proteins in the replication complex (which does not include a recognizable helicase), although the mechanism is unknown.

The nature of the replicative intermediates in the replication of the eukaryotic positive-stranded RNA viruses is less clear. Double-stranded RNAs (usually called replicative form or RF dsRNAs) are usually found in nucleic preparations from cells infected with positive-strand RNA viruses or in the extracted products of *in vitro* RNA synthesis reactions with crude or partially purified polymerase preparations from infected

plants or animal cells (de Graaf and Jaspars, 1994), irrespective of the polymerase or helicase type. Double-stranded RNA could arise *in vivo* as dead-end products formed by annealing of positive and negative strands; even in the case of Qβ, for which model 1 (Fig. 1) is well established, dsRNA accumulates as the concentrations of viral RNAs in the cell increase (Priano *et al.*, 1987). Furthermore, annealing of positive and negative strands can also occur on extraction of nucleic acids from cells by deproteinization. Similar considerations apply to the different types of replicative intermediates. The "closed" intermediates (Fig. 1, model 2, d, e) could be formed by collapse of the "open" intermediates (Fig. 1, model 1, d, e). dsRNAs extracted from *in vitro* RNA synthesis reactions could be formed in the same way, or could result from inadequacies in the *in vitro* systems. Many *in vitro* template-dependent systems synthesize only negative-stranded RNA, indicating a lack of one or more essential components (or the presence of an inhibitor).

Isolated poliovirus replicative intermediates consist of a full-length negative-strand template to which are attached six to eight nascent positive strands, but conflicting evidence has been presented that such intermediates have either an "open," mainly single-stranded structure (Oberg and Philipson, 1971; Richards *et al.*, 1984; Bienz *et al.*, 1994) or a "closed" structure (Meyer *et al.*, 1978; Nilsen *et al.*, 1981; Troxler *et al.*, 1992). Using electron microscopy and other techniques, Garnier *et al.* (1980) showed that the turnip yellow mosaic virus replicative intermediates formed *in vivo* were mainly in the single-stranded form and considered that the dsRNA isolated from infected leaves was mainly an isolation artifact, although some dsRNA may be formed *in vivo* late in infection. Using an *in vitro* system able to catalyze the complete replication of cucumber mosaic virus RNA, Hayes and Buck (1990) detected free positive and negative strands in the ratio of 7:1, as well as some dsRNA, suggesting that replication occurred according to model 1 and showing that some free negative strands could survive the phenol extraction procedure in the presence of an excess of positive strands. DsRNA was not a template for the replicase complex (Hayes and Buck, 1993), but this could merely indicate that the replicase lacked an essential component required to initiate synthesis on a dsRNA template. Evidence was obtained that flavivirus RNA positive-strand RNA molecules are synthesized from negative-strand templates by a semiconservative mechanism, as in model 2 (Fig. 1) (Chu and Westaway, 1985; Cleaves *et al.*, 1981). Wu and Kaesberg (1991), using a flockhouse virus template-dependent *in vitro* RNA replication system, showed that labelled nucleotides incorporated into dsRNA during a short pulse could be chased into ssRNA using excess unlabelled nucleotides, sug-

gesting "closed" replication intermediates (Fig. 1, model 2). Although free negative-stranded RNA has generally not been isolated from eukaryotic cells infected with positive-stranded RNA viruses, replication of flockhouse virus RNA 2 could be initiated from a negative-stranded template in cells in which replication proteins were provided in trans (Ball, 1994). Furthermore, an isolated cucumber mosaic virus RdRp complex was capable of replicating a satellite RNA when provided with the negative-strand satellite RNA (Wu and Kaper, 1994), an isolated brome mosaic virus RdRp complex could use negative-strand templates to produce subgenomic RNA (Dreher and Hall, 1988), and an isolated alfalfa mosaic virus RdRp could use a negative-strand template to produce full-length and subgenomic RNA (de Graaf *et al.,* 1995a,b). However, attempts to use negative strands for the replication of alphaviruses were not successful (reviewed by Strauss and Strauss, 1994). Finally, some evidence for the formation of double-stranded replication intermediates comes from the observation that transgenic plants expressing a yeast double-stranded RNA-specific ribonuclease showed resistance (albeit incomplete) to tomato mosaic virus, cucumber mosaic virus, and potato virus Y (Watanabe *et al.,* 1995). In model 1, it would be anticipated that the short double-stranded regions where the nascent transcript is bound to the template would be protected by the replicase complex and hence not accessible to the dsRNA ribonuclease.

Model 3 can be eliminated in some cases, e.g., labelled nucleotides incorporated into tobacco mosaic virus RF by an isolated polymerase preparation containing endogenous template were predominantly in the positive strand (Young and Zaitlin, 1986), which is not consistent with model 3. Generally, the available evidence does not allow the "open" and "closed" models (models 1 and 2) to be distinguished unequivocally for any eukaryotic positive-stranded RNA virus and further work is needed in this area. It is possible that different mechanisms operate for different viruses or groups of viruses. This could be related to different mechanisms for the initiation of positive-stranded RNA synthesis and to the requirement for some viruses to synthesize subgenomic RNAs. For example, viruses that initiate with a VPg may employ a different mechanism than those that are capped. Model 2 requires that the double-stranded replicative form RNA is unwound at the end containing the 3′ terminus of the negative strand before synthesis can begin. Most RNA helicases, including the cellular and viral helicases discussed above, and most DNA helicases (Kornberg and Baker, 1992) bind to single-stranded regions adjacent to the duplex region to be unwound, and do not act on completely double-stranded structures. Unwinding of a completely double-stranded RF could therefore involve another protein. Conversion of the *E. coli* replication origin to an open

structure requires the sequence-specific DNA-binding protein, the dnaA protein (Bramhill and Kornberg, 1988). Further unwinding of the duplexes as DNA synthesis proceeds at the replication forks utilizes helicases. In the case of positive-stranded RNA replication, a 3′→5′ helicase bound to the negative template strand of the unwound origin (promoter) could perform the subsequent duplex unwinding. The plum pox and bovine diarrhea virus helicases (Lain *et al.*, 1990; Warrener and Collett, 1995) have the required strand-displacement activity, as does the poliovirus $3D^{pol}$ protein (Cho *et al.*, 1993) and an alfalfa mosaic virus RdRp complex (de Graaf *et al.*, 1995a). Model 2 requires recognition of a double-stranded structure for initiation of positive-strand synthesis (and for subgenomic RNA synthesis for those viruses that utilize subgenomic promoters), whereas model 1 requires recognition of a single-stranded structure. Since secondary structure in the single-stranded RNA appears to be important for recognition of at least some genomic and subgenomic promoters (see Section V,D), model 1 would appear to be favored in this aspect. Model 1 requires an explanation, not required for model 2, as to how the base pairs formed at RNA synthesis are almost immediately unwound by the replicase complex. This could probably be accomplished by a helix-destabilizing protein, or a second molecule of the helicase, suitably positioned in the replicase complex. In this respect, it is noteworthy that purified replication complexes of cucumber mosaic virus (Hayes and Buck, 1990) and brome mosaic virus (Quadt *et al.*, 1993) appear to contain more of the respective 1a protein (which contains the helicase-like domain) than the 2a protein (which contains the polymerase-like domain), and brome mosaic virus RNA virus replication was found to be more sensitive to reductions in the expression of the 1a protein than to reductions in the expression of the 2a protein (Kroner *et al.*, 1990; Traynor and Ahlquist, 1990; Dinant *et al.*, 1993). It should be noted, however, that the 1a protein also has capping functions (see Section II,C).

The different replication models also have some implications for RNA recombination. There is good evidence that, at least in some cases, RNA recombination leading to the formation of viable recombinants or DI RNAs requires RNA replication (reviewed by Lai, 1992; Bujarski *et al.*, 1994; Simon and Bujarski, 1994). The favored mechanism, called "copy choice," requires the polymerase, together with the nascent strand, to switch from one template (the donor) to another template (the recipient), either after first dissociating from the donor template (nonprocessive model) or in a processive fashion without first dissociating. Recombination can occur during negative-strand synthesis, as for poliovirus (Kirkegaard and Baltimore, 1986), brome mosaic

virus (Bujarski *et al.*, 1994), and flockhouse virus (Li and Ball, 1993), or during positive-strand synthesis, as for turnip crinkle virus (Carpenter *et al.*, 1995). In replication model 2 (Fig. 1) in the negative-strand-synthesizing replicative intermediate (b) and the positive-strand-synthesizing replicative intermediate (d, e), the nascent strand remains completely base-paired to its template. This would appear to make it more difficult, in the nonprocessive model, for the polymerase and nascent strand to dissociate from the template or, in the processive model, could create a structure in which the newly synthesized recombinant strand is base-paired to both donor and recipient templates. The problem would be overcome if a second RdRp complex were advancing along the template close behind, carrying out displacement synthesis. If a second polymerase complex "caught up" with a paused complex, the resulting change of structure could facilitate detachment or strand switching, depending on the type of recombination event. Some processive models of recombination require the recipient template strand to be base-paired to the donor template strand in front of the advancing replicase. This would appear to present a problem in replication model 2 for recombination occurring during positive-strand synthesis, unless the recipient template were able to specifically invade the unwound RNA where RNA synthesis is taking place. This problem would not occur with any of the models of RNA replication for recombination taking place during negative-strand synthesis.

In recombination models in which the replicase pauses at a duplex region in the template to allow the helicase to unwind it (or to switch templates), it might be expected that mutations in the helicase protein could affect the site of recombination. Evidence that this occurs has been obtained using mutants of brome mosaic virus with amino acid insertions in the helicase-like domain of the 1a protein (Nagy *et al.*, 1995). It has been suggested that the apparent great propensity of viruses such as the tombus-, carmo-, and luteoviruses to form recombinants and DI RNAs could be due to their missing recognizable NTP-binding and helicase domains (Simon and Bujarski, 1994; Gibbs, 1995). The apparent lack of a virus-encoded helicase could render the RdRps of these viruses less processive and hence more prone to strand switching.

C. Capping and Methylation Enzymes

The presence of a 5′ cap structure is characteristic of most eukaryotic mRNAs, and mRNAs of DNA viruses, retro- and pararetroviruses, negative-strand RNA viruses, some double-stranded RNA viruses, and some positive-stranded RNA viruses. The simplest cap structure (cap 0)

consists of 7MeGpppN, where N is usually A or G. The cap 1 and cap 2 structures are 7MeGpppNmpN′ and 7MeGpppNmpN′m, where m represents methylation of the 2′-OH group of the ribose moiety of the nucleoside. Capping of the cellular mRNAs and those of most DNA viruses, retroviruses, and pararetroviruses occurs in the nucleus using cellular capping enzymes; capping of influenza virus mRNAs also occurs in the nucleus and involves endonucleolytic cleavage of capped oligonucleotides from cellular mRNAs for use as primers in transcription of the negative-stranded RNA genome. Capping of mRNAs of some other negative-stranded RNA viruses, such as the rhabdoviruses, and of dsRNA viruses and positive-stranded RNA viruses that possess a cap, and of some DNA viruses, such as the pox viruses, occurs in the cytoplasm, and utilizes virus-encoded enzymes (Mizumoto and Kaziro, 1987; Murphy *et al.,* 1995). The positive-stranded RNA viruses that possess capped genomic RNAs include those in the alpha-like, corona-like, flavi-like, and carmo-like virus supergroups (Table II). The capping reaction of cellular (Mizumoto and Kaziro, 1987), vaccinia virus (Shuman and Moss, 1990), and reovirus (Furiuchi *et al.,* 1976) mRNAs involves the following reactions: (i) pppNpN′..... → ppNpN′..... + p_i (hydrolysis of the 5′-terminal phosphate of the nascent RNA transcript by an RNA triphosphatase); (ii) guanylyltransferase + GTP → guanylyltransferase-GMP + pp_i (interaction of GTP with mRNA:guanylyltransferase to form a covalently bound guanylyltransferase-GMP intermediate); (iii) guanylyltransferase-GMP + ppNpN′..... → GpppNpN′..... + guanylyltransferase (donation of GMP from the guanylyltransferase-GMP intermediate to form a 5′-5′ triphosphate linkage); (iv) GpppNpN′..... + AdoMet → 7MeGpppNpN′..... + AdoHcy (methylation of position 7 of the terminal G cap using a specific methyltransferase and *S*-adenosylmethionine [AdoMet] as the methyl donor to generate the cap 0 structure and *S*-adenosylhomocysteine [AdoCys]). For mRNAs that contain a cap 1 or cap 2 structure, further methylation of the 2′-OH groups of the ribose moieties of N and N′ is carried out using specific methyltransferases. It has been shown that for reovirus mRNAs, which have a cap 1 structure, capping is coupled to, and occurs at an early stage of, mRNA synthesis (Furiuchi *et al.,* 1976).

Using mutants resistant to methionine starvation (which reduces cellular AdoMet levels) and mycophenolic acid (which reduces cellular GTP levels), evidence was obtained that capping of Semliki Forest virus (SFV) genomic and subgenomic RNAs is linked to RNA synthesis and that these activities reside in the nsP1 protein (Scheidel *et al.,* 1989; Mi *et al.,* 1989; Mi and Stollar, 1990; Scheidel and Stollar, 1991).

The methyltransferase activity of the nsP1 protein of SFV has been expressed in *E. coli* (Mi and Stollar, 1991) and the nsP1 protein of the closely related Sindbis virus in *E. coli* and insect cells (Laakkonen *et al.,* 1994). The Sindbis virus enzymes, as well as an enzyme preparation from Sindbis virus-infected cells, were able to catalyze the transfer of a methyl group from AdoMet to GTP, dGTP and GpppG, but not to 7MeGTP, GpppA, or *in vitro* transcribed RNAs with GpppA or GpppG caps. Subsequently it was shown that the nsP1 protein formed a covalent complex with 7MeGMP, but not with GTP (Ahola and Kääriänen, 1995). It therefore appears that the nsP1 protein has both capping and methylation activities, but that, unlike the cellular and viral capping reactions discussed above, the GTP is first converted to 7MeGTP by the methylase and then the 7MeGTP reacts with the guanylyltransferase to form a guanylyltransferase-7MeGMP complex, which in turn transfers the 7MeGMP to the ppNpN'..... 5' end of the nascent RNA.

Mapping of nsP1 mutants indicated that the methyltransferase domain was located close to its N terminus (Mi *et al.,* 1989; Mi and Stollar, 1990). Sequence comparisons with other positive-stranded RNA viruses has led to the identification of a methyltransferase domain near the N terminus of replication proteins of all members of the alpha-like virus supergroup (Rozanov *et al.,* 1992). Four conserved motifs were identified, with motifs I, II, and IV having an invariant H residue, a DXXR signature and an invariant Y residue respectively. Additional conserved motifs enabled two groups of viruses with the methyltransferase-like domain to be distinguished, the "altovirus" group (*Togaviridae* and *Bromoviridae* families, *Tobamovirus, Tobravirus, Hordeivirus, Furovirus* genera) and the "typovirus" group (*Tymovirus, Potexvirus, Carlavirus, Trichovirus* genera). Mutations which increased the affinity of the SFV nsP1 for AdoMet mapped close to motif II (Mi *et al.,* 1989; Mi and Stollar, 1990), suggesting that this motif may be part of the AdoMet binding site. Furthermore, mutation of any of the conserved H, D, R, and Y residues to A in Sindbis virus abolished virus infectivity and the methyltransferase activity of the *E. coli*–expressed nsP1 protein (Wang *et al.,* 1996).

A guanylyltransferase-like enzyme has been isolated from plants infected with tobacco mosaic virus and identified as the 126-kDa replication protein, which has an N-terminal methyltransferase-like domain and a C-terminal helicase-like domain (Dunigan and Zaitlin, 1990). This transferase formed a covalent complex with GMP in the absence of AdoMet and therefore appears to be different from the Sindbis virus nsP1 protein which only forms a complex with 7MeGMP.

This is surprising since both viruses belong to the alpha-like supergroup and may represent some divergence in capping mechanism between the animal alphaviruses and the plant tobamoviruses.

The motifs which characterize the methyltransferase-like domain of the alpha-like virus supergroup could not be found in any of the other virus superfamilies whose members have a 5′-capped RNA (Rozanov *et al.*, 1992). However, the flavivirus NS5 protein and reovirus λ2 protein have been shown to have methyltransferase-like domains related to a class of cellular methyltransferases (Koonin, 1993) and a flavivirus NS3 protein has been shown to have an RNA triphosphatase activity (Wengler and Wengler, 1993). Sequences near the C terminus of this protein had some sequence similarity to a vaccinia virus RNA triphosphatase. Hence it is likely that the flavivirus group also encode-capping enzymes. The other groups of capped RNA viruses probably also encode such enzymes, which may be recognized by different, as yet undetermined, motifs.

Transgenic tobacco plants expressing antisense RNA for AdoCys hydrolase, which controls the cellular AdoCys/AdoMet ratio, were shown to be resistant to infection by tobacco mosaic mosaic virus, potato virus X, and cucumber mosaic virus (all with 5′-capped RNAs), but much less resistant to infection by potato virus Y (5′ VPg) (Masuta *et al.*, 1995). Resistance was thought to be due partially to undermethylation of the viral cap structure, although induction of host resistance as a result of excess cytokinin levels may also have played a role.

D. Genome-Linked Virus Proteins

Genome-linked virus proteins (VPgs) are virus-encoded proteins that are covalently linked by a phosphodiester linkage to the 5′-terminal nucleotide of the virus genomic RNA. They are found in viruses of the picorna-like and sobemo-like virus supergroups (Table II). Removal of the VPg of some viruses, such as nepoviruses (Hellen and Cooper, 1987) and caliciviruses (Burroughs and Brown, 1978), by protease treatment destroys the infectivity of the RNA. However, this effect is probably due to increased sensitivity of the RNA to exonuclease attack and does not represent a specific requirement for VPg for infectivity. Capped transcripts produced *in vitro* from full-length clones of caliciviruses were infectious (Sosnovtsev and Green, 1995); infectious *in vitro* transcripts have also been produced from comoviruses (Vos *et al.*, 1988), picornaviruses (van der Werf *et al.*, 1986; Sarnow, 1989; Lee *et al.*, 1993), and potyviruses (Dolja *et al.*, 1992) showing that VPg is not needed for infectivity.

The poliovirus and cowpea mosaic virus VPgs consist of 22 and 28 amino acid residues, repectively, and are linked to the 5′ end of the virus RNAs by tyrosine and serine phosphodiester linkages, respectively (Wimmer *et al.,* 1993; Jaegle *et al.,* 1987). The VPg of tobacco etch potyvirus is larger, with a mass of 21 kDa (Riechmann *et al.,* 1992). The poliovirus Vg is removed from the genomic RNA by a host enzyme soon after infection and the 5′ ends of viral RNA isolated from polysomes mainly consist of pUpU..... (Wimmer *et al.,* 1993). Mutational and other analyses have shown that the poliovirus VPg (and/or its precursor 3AB) are essential for RNA replication (Reuer *et al.,* 1990; Xiang *et al.,* 1995a). Viral VPgs are probably involved in the initiation of both negative-strand and positive-strand synthesis; VPg is found covalently linked to the 5′ ends of newly synthesized negative and positive strands of poliovirus (Wimmer *et al.,* 1993) and cowpea mosaic virus (Lomonossoff *et al.,* 1985). VPg of poliovirus has been shown to be present in infected cells both as free VPg and as VPg-pUpU (Crawford and Baltimore, 1983). Furthermore, VPg-pU can be synthesized in membrane-bound replication complexes isolated from infected cells and elongated to longer VPg-RNA molecules (Takeda *et al.,* 1986; Toyoda *et al.,* 1987). Such a molecule (or a precursor 3AB-pU (see Section II,E,4) could act as a primer for RNA synthesis, although other models have been suggested (see Section III).

E. Expression of, and Interactions between, Replication Proteins

1. Alpha-like Virus Supergroup

The polymerase, helicase, and methyltransferase functions of viruses in the alphavirus-like virus supergroup can be expressed in several different ways. For some viruses, the methyltransferase-like, helicase-like, and polymerase-like domains are within the same protein (in that order), which is translated from the 5′ ORF. An example is the 166-kDa protein encoded by ORF 1 of the genomic RNA of potato virus X (Huisman *et al.,* 1988; Longstaff *et al.,* 1993). The other downstream genes of the virus, which are required for cell-to-cell movement and encapsidation, are translated from subgenomic RNAs.

For some viruses, which have divided genomes, the methyltransferase and helicase-like domains and the polymerase-like domain are in two separate proteins, translated from two separate RNAs. Viruses in the *Bromoviridae* family, such as brome mosaic virus (BMV) (Ahlquist, 1992), cucumber mosaic virus, and alfalfa mosaic virus, fit into this category. BMV has three genomic RNAs, of which only RNAs 1 and 2

are essential for RNA replication. The BMV 1a protein, encoded by RNA 1, contains the methyltranferase-like domain near its N terminus and the helicase-like domain near its C terminus, and the 2a protein, encoded by RNA 2, contains the polymerase-like domain. RNA 3 encodes a cell-to-cell movement protein and the coat protein, the latter being translated from a subgenomic RNA (RNA 4).

For some other viruses, the methyltransferase-like and helicase-like domains are present on a protein which is translated from the 5′ ORF and the polymerase domain, which is located downstream, is translated by readthrough (suppression) of the stop codon to give a fusion protein, e.g., the 126-kDa protein of tobacco mosaic virus (TMV) contains the methyltransferase-like and helicase-like domains, and the 183-kDa protein, translated by readthrough of an amber stop codon (UAG), additionally contains the polymerase-like domain (Dawson and Lehto, 1990). As a result, the methyltransferase-like and helicase-like domains are present in excess over the polymerase-like domain, being present both in the more abundant 126-kDa protein and in the 183-kDa protein. Mutational analysis has shown that both the 126-kDa protein and the 183-kDa protein are required for efficient RNA replication (Ishikawa *et al.*, 1986, 1991b). This could imply, as suggested for BMV and cucumber mosaic virus (CMV) (see Section II,B,1), that two or more copies of the helicase domain are required for different unwinding functions of RNA replication, or that the helicase functions in the 126-kDa protein and the methyltransferase functions in the 183-kDa protein (or vice versa). The other two downstream ORFs, which encode the cell-to-cell movement protein and the capsid protein, are not required for RNA replication; they are translated from subgenomic RNAs.

The BMV 1a and 2a proteins copurify with an RdRp complex isolated from infected plants (Quadt and Jaspars, 1990; Quadt *et al.*, 1993). Furthermore, the 1a and 2a proteins, produced by translation *in vitro* of RNA 1 and RNA2, form a complex; a complex was also formed when purified 2a protein expressed in insect cells was mixed with purified helicase-like domain expressed in *E. coli*, showing that no other proteins were needed to mediate the 1a–2a interaction. Certain mutations in the 1a protein which blocked virus RNA replication or made replication temperature-sensitive also prevented or made temperature-sensitive, respectively, the interaction between the 1a and 2a proteins, suggesting that this interaction is required for BMV RNA replication (Kao *et al.*, 1992). The region of interaction has been mapped to the N-terminal 115 amino acids of the 2a protein and the helicase-like domain of more than 50 kDa of the 1a protein (Kao and Ahlquist, 1992; O'Reilly *et al.*, 1995). Mutational analysis has shown that the N-

terminal region of the 2a protein, which lies upstream of the conserved polymerase-like domain, is essential for virus RNA replication (Traynor *et al.*, 1991). Requirement of the large size of the helicase-like domain of the 1a protein for the interaction suggested involvement of a higher-order structure and it was shown that this domain, and analogous domains of three other viruses in the *Bromoviridae* family [alfalfa mosaic virus, cowpea chlorotic mottle virus (CCMV), CMV], folded into protease-resistant structures; mutations which rendered the region protease-sensitive were defective in RNA replication (O'Reilly *et al.*, 1995). The BMV 1a protein is homologous to the TMV 126-kDa protein and the BMV 1a–2a protein complex is homologous in structure to the TMV 183-kDa protein. The homologous proteins may serve similar functions in the replication complexes of the two viruses. Using different combinations of the 1a and 2a proteins of BMV and CMV expressed transiently in plant protoplasts to support replication of BMV or CCMV RNA 3 templates, Dinant *et al.* (1993) found that the combination of CCMV 1a and BMV 2a did not support detectable synthesis of negative-strand, positive-strand, or subgenomic RNA, whereas the combination of BMV 1a and CCMV 2a was preferentially defective in positive-strand and subgenomic RNA accumulation, and negative-strand RNA was only slightly affected. The latter combination suggested that partial incompatibility of 1a–2a can distinguish some aspect of negative-stranded and positive-stranded RNA synthesis (including synthesis of subgenomic RNA), possibly the capping function of the 1a protein, since only positive-stranded RNAs are capped. Isolation of replication complexes from yeast expressing the 1a and 2a proteins has shown that the replication complexes, which probably contain host and viral proteins (see Section III), can only form in the presence of an RNA template (Quadt *et al.*, 1995).

A different strategy is used by animal viruses of the *Togaviridae* family (Strauss and Strauss, 1994; Frey, 1994), which are the only enveloped viruses in the alpha-like virus supergroup. The viruses of the *Alphavirus* genus are the best studied. The nonstructural proteins are encoded at the 5′ end of the genome and are translated as a polyprotein which is cleaved by virus-encoded proteinase activity. The structural proteins (the capsid and envelope proteins) are encoded at the 3′ end of the genome and are translated from a subgenomic RNA (26S RNA) as a polyprotein, which is cleaved by virus and host proteinases. Only the nonstructural proteins are required for RNA replication. Translation of Sindbis virus RNA gives a polyprotein P123, containing the sequences of nsP1, nsP2, and nsP3. Readthrough of an opal termination codon (UGA), which is followed by an essential (C), results in

the production of P1234. The readthrough rate in translation *in vitro* was found to be 10–20% at 30°C and <5% at 40°C. Since nsP1 is associated with methyltransferase and guanylyltransferase (see Section II,C), nsP2 has a helicase-like domain (see Section II,B) and nsP4 has a polymerase-like domain (see Section II,A), the relationship between P123 and P1234 is similar to that of the TMV 126-kDa and 183-kDa proteins. However, unlike the TMV proteins, there is a domain for a papain-like cysteine proteinase near the C terminus of nsP2. The specificity of nsP2 is modulated by the surrounding sequences, so that the pattern of cleavage of the polyproteins changes through the replication cycle (reviewed by Strauss and Strauss, 1994). In P1234, the proteinase can cleave in cis between nsP3 and nsP4 to give P123 and nsP4. The cleavage site is six amino acid residues downstream of the opal terminal codon, but as the C-terminal region of nsP3 is variable and deletions in it are well tolerated (Lastarza *et al.*, 1994b), P123 formed by primary translation and P123 formed by cleavage of P1234 are considered equivalent. P123 cannot be cleaved in cis, but once sufficient concentrations have built up, it is cleaved in trans to produce nsP1 and P23. The P23 proteinase can cleave the P2/P3 site in polyproteins (including itself) very efficiently, and nsP2 and nsP3 are produced. The P3/P4 site can only be cleaved by nsP2 polyproteins containing nsP3 sequences. Late in infection it is likely that that the P2/P3 site is cleaved (in trans) in the nascent polypeptide; hence the P3/P4 site is not cleaved and the main products are nsP1, nsP2, and P34, but the proteins produced late in infection are not thought to be involved in replication.

Mutations which prevented the cleavage to form nsP4 were lethal (reviewed by Strauss and Strauss, 1994), and the nsP4 protein is believed to be active in elongation of negative and positive RNA strands. However, mutations which inactivated the nsP2 proteinase and resulted in uncleaved P123 led to increased accumulation of negative-stranded RNA and decreased accumulation of positive-stranded genomic and subgenomic RNAs. Another mutation which greatly increased the rate of cleavage of P123 was lethal. These results, together with further mutational analysis, led to the proposal that P123 + nsP4 function in negative-stranded RNA synthesis (Shirako and Strauss, 1994; Lemm *et al.*, 1994). However, it is likely that nsP1 + P23 + nsP4 can also function in negative-stranded RNA synthesis, since temperature-sensitive mutations with lesions in nsP1 and nsP3 complement, whereas those in nsP2 and nsP3 do not rescue defects in negative-stranded RNA synthesis (Sawicki and Sawicki, 1994; Lastarza *et al.*, 1994a; Wang *et al.*, 1994). Complete cleavage of P123 to form nsP1, nsP2, and nsP3 not only results in inactivation of negative-strand RNA synthesis, but also

switches on positive-stranded genomic and subgenomic RNA synthesis (Shirako and Strauss, 1994; Lemm *et al.,* 1994).

Although polyprotein processing is common in the expression of viral genes (Dougherty and Semler, 1993), it is relatively uncommon in the alpha-like virus superfamily. In the case of the alphaviruses, and probably also rubella virus (Frey, 1994), it may have evolved as a method of inducing conformational changes in replication complexes to switch from negative-strand to positive-strand synthesis. After cleavage, all four proteins remain in association because they are found in isolated replication complexes (Barton *et al.,* 1991), although their relative juxtapositions may alter. All the nonstructural alphavirus proteins are multifunctional and have functions in both negative-strand and positive-strand synthesis, which have been distinguished by mutational analysis (reviewed in Strauss and Strauss, 1994). The nsP1 protein functions as a guanylyltransferase and methyltransferase in the capping of positive-stranded RNA (see Section II,C), it is required for negative-strand synthesis and in polyproteins it modulates the proteinase activity of the nsP2 protein by inhibiting cleavage of the site between nsP2 and nsP3. The nsP2 protein has an N-terminal helicase domain (see Section II,B) and a C-terminal proteinase domain; it is specifically required for the synthesis of subgenomic RNA and plays a role in the conversion of negative-strand-synthesizing to positive-strand-synthesizing enzymes (Dé *et al.,* 1996). The nsP3 protein has separate functions in negative-stranded RNA and subgenomic RNA synthesis (Lastarza *et al.,* 1994a; Wang *et al.,* 1994), some of which may be host-specific (Lastarza *et al.,* 1994b). Its presence in polyproteins modulates the activity of the nsP2 proteinase (see above). The nsP4 protein is likely to be the RNA polymerase that functions in negative-stranded, positive-stranded, and subgenomic RNA synthesis. Presumably all these proteins must also contain domains that keep them associated after proteolytic cleavage. The N-terminal region of the nsP4 protein is thought to be important for this purpose. Another possibility is that complete cleavage of the polyprotein could enable an additional copy of an individual subunit, such as the nsP2 protein which contains the helicase-like domain, to be recruited for a specific function in the synthesis of positive-stranded RNA, which was not needed for negative-stranded RNA synthesis.

Turnip yellow mosaic virus is an example of a plant alpha-like virus in which cleavage of a polyprotein is essential for virus RNA replication. Its genome contains three ORFs. ORF 1 encodes a 206-kDa protein which contains methyltransferase-like, helicase-like, and polymerase-like domains. ORF 2 encodes a cell-to-cell movement protein and ORF 3

encodes the coat protein, which is translated from a subgenomic RNA; neither of these proteins is required for RNA replication. The 206-kDa primary *in vitro* translation product is cleaved in cis by a papain-like cysteine proteinase, located between the methyltransferase-like and helicase-like domains, to give a protein of 66 kDa from the C-terminal end (which contains the polymerase-like domain) and a protein of 140 kDa from the N-terminal end, which contains the other three domains (Bransom *et al.*, 1991; Bransom and Dreher, 1994; Kadaré *et al.*, 1995). The proteinase was active and cleaved the same site when expressed in *E. coli* (Kadaré *et al.*, 1995) Although this cleavage is needed for virus RNA replication (Weiland and Dreher, 1993), its precise function has not been determined. Purified preparations of a RdRp complex isolated from infected plants were reported to contain a virus-encoded protein of 115 kDa (Mouches *et al.*, 1984; Candresse *et al.*, 1986), but the relationship of this protein to the 140-kDa and 66-kDa proteins is not clear.

The genome of beet yellows closterovirus contains a long ORF (1a) which could encode a protein of 295 kDa, which has a papain-like cysteine proteinase domain, a methyltransferase-like domain, and a helicase-like domain. The next ORF (1b), which has the capacity to encode a 53-kDa protein with a polymerase-like domain, overlaps the end of ORF 1a and is probably translated by a +1 frameshift (Agranovsky *et al.*, 1994). The leader proteinase is cleaved autocatalytically after translation *in vitro*. There is also a potential aspartic proteinase between the methyltranferase-like and helicase domains, but the significance of this for proteolytic processing or RNA replication is not known.

It is noteworthy that the order of the homologous methyltransferase-like, helicase-like, and polymerase-like domains is conserved in all the viruses in the alpha-like supergroup. It may be speculated that this could reflect some basic common structural feature in the organization of the replicase complex, particularly as the three domains can apparently function in one protein in some viruses, e.g., potato virus X. However, a single replication protein is not sufficient for tobacco mosaic virus replication. A tobacco mosaic virus mutant, engineered to remove the stop codon at the end of the ORF for the 126-kDa protein so that only the 183-kDa protein was produced, replicated poorly and the stop codon was reintroduced by reversion (Ishikawa *et al.*, 1986). The stop codon near the end of nsP3 in Sindbis virus, however, is not essential. Some alphaviruses, such as Semliki Forest virus, do not have the stop codon, so that the only primary translation product is P1234. It has been suggested that the extra amount of P123 produced when the stop codon is present provides extra proteinase to accelerate the switch from negative-strand to positive-strand synthesis (Strauss and Strauss,

1994). The phenotype of Sindbis virus mutants in which the stop codon had been replaced by a sense codon (Li and Rice, 1989) were consistent with this hypothesis. It is noteworthy that there is no apparent counterpart of the alphavirus nsP3 protein in the plant alpha-like viruses. This protein has functions distinct from its effect on nsP2 proteinase specificity, perhaps in protein folding or protein–protein or protein–RNA interactions, which must be circumvented or achieved in different ways in the plant virus replication complexes. Switches in functions to achieve synthesis of negative-stranded, positive-stranded, or subgenomic RNA may perhaps be achieved by a variety of methods, including proteolytic cleavage, protein–protein interactions, protein–RNA interactions, modification of proteins, for instance by phosphorylation, or changes in the stoichiometry of the different components in replication complexes. Further studies will be required to determine whether or not the considerable diversity in expression of replicase functions within the alpha-like supergroup merely reflects variations on a common theme.

A subset of the plant alpha-like viruses, carla-, furo-, hordei-, and tobraviruses, contain 3′-proximal genes for small cysteine-rich proteins, some of which have been shown to have RNA-binding activity and display some sequence similarity to other nucleic-acid-binding proteins (Koonin *et al.*, 1991). One of the functions of these proteins is to regulate the synthesis of the capsid protein, which in these viruses is encoded by 5′-proximal genes. Although these proteins are not absolute requirements for RNA replication, they have been shown to affect replication in some instances. Beet necrotic yellow vein virus has four RNA components. RNA 1 can replicate alone in protoplasts and contains a long ORF containing methyltransferase-like, helicase-like, and polymerase-like functions (Richards and Tamada, 1992). RNA 2 has several ORFs, including the coat protein gene at the 5′ end and a gene for a small cysteine-rich protein (P14) at the 3′ end. Null mutations in P14 not only reduced coat protein accumulation, but also had a specific inhibitory effect on RNA 2 accumulation (Hehn *et al.*, 1995). The defect in RNA 2 accumulation could not be complemented in trans, which is surprising since it is expressed as a single protein via a subgenomic RNA, and the defect in coat protein accumulation could be complemented in trans; it was suggested that translation of P14 and replication of RNA 2 were tightly coupled.

2. *Carmo-like Virus Supergroup*

Viruses in the *Tombusviridae* family (*Carmovirus* and *Tombusvirus* genera) and the *Dianthovirus, Luteovirus* (subgroup I), and *Necrovirus* genera have a 5′-proximal ORF which encodes a protein of 22–33 kDa

(Russo *et al.*, 1994). Readthrough of a stop codon at the end of this ORF (carmoviruses, necroviruses, and tombusviruses), or a ribosomal frameshift prior to the stop codon (dianthoviruses and subgroup I luteoviruses), gives rise to a fusion protein (82–99 kDa). The C-terminal region of the fusion protein contains a polymerase-like domain. Mutants of cymbidium ringspot virus, red clover necrotic mosaic virus, tomato bushy stunt virus, and turnip crinkle virus which produce only the smaller protein, or only the fusion protein, failed to replicate (Hacker *et al.*, 1992; Wei *et al.*, 1992; Dalmay *et al.*, 1993a; Gieseman-Cookmeyer *et al.*, 1995; White *et al.*, 1995), indicating that both proteins are required for RNA replication. For cymbidium ringspot virus and barley yellow dwarf virus-PAV, it was shown that the 5′-encoded protein and the readthrough or frameshift protein were the only virus-encoded proteins needed for RNA replication (Kollar and Burgyan, 1994; Mohan *et al.*, 1995). It was also shown that when these two proteins were expressed in transgenic plants from chromosomal insertions of the genes, replication of DI RNAs was supported. Furthermore, nonviable mutants of turnip crinkle virus with mutations in either the gene for the 28-kDa protein or the 88-kDa readthrough protein could complement each other (White *et al.*, 1995). These two experiments show that the two replication proteins act in trans. A purified RdRp complex isolated from plants infected with red clover necrotic mosaic virus was shown to contain both the 27-kDa protein and the 88-kDa frameshift fusion protein (Bates *et al.*, 1995). The function of the smaller replication protein, which is produced in a large excess over the fusion protein, is not known. It could be a helix-destabilizing protein which substitutes for the apparent absence of a helicase-like domain in these viruses. Evidence also suggests that it is involved in membrane localization of the replicase complex (see Section IV). RNAs of viruses in the carmo-like supergroup are capped and presumable capping enzymes are encoded by one or both of the essential ORFs, although no sequence similarity to capping enzymes of the alpha-like viruses has been detected. The other proteins encoded by these viruses, which are not needed for replication, are translated from subgenomic RNAs or additionally in the case of the dianthoviruses, from an additional RNA segment.

Carmo-like viruses have not been found in animals. The polymerase-like domains of carmo-like viruses fall into supergroup 2, like those of the animal flaviviruses, pestiviruses and hepatitis C virus. The *Flaviviridae* are clearly very different from the carmo-like viruses. They produce several structural and nonstructural proteins by proteinase cleavage of a polyprotein. They possess helicase-like domains of superfamily 2, with similarity to those of the potyviruses (see Section II,B),

but are clearly also different from the picorna-like viruses, e.g., they possess a 5′ cap. Very little is known about proteins involved in *Flaviviridae* replication, beyond that discussed in Section II,A–C, the requirement for certain proteolytic cleavages (e.g., Chambers *et al.*, 1993; Nestorowicz *et al.*, 1994), and the inhibition of *in vitro* conversion of RF to RI by an isolated dengue virus RNA polymerase fraction using antibodies to the NS3 (helicase) and NS5 (polymerase) regions (Bartholomeusz and Wright, 1993). However, replication of viruses in the Flavi-like supergroup is likely to have significant differences from that of the other virus supergroups.

3. Sobemo-like Virus Supergroup

These include the sobemoviruses and the subgroup II luteoviruses. Viruses of these two genera are similar at the 5′ end of the genome, which contains three overlapping ORFs, 1, 2, and 3, but differ at the 3′ ends. The 3′ ends of subgroup II luteoviruses more closely resemble those of subgroup I luteoviruses and it has been suggested that the subgroup II luteoviruses were derived by recombination between a sobemovirus and a subgroup I luteovirus (Miller *et al.*, 1995). For beet western yellows luteovirus, it has been shown that only ORFs 2 and 3 are required for RNA replication (Reutenauer *et al.*, 1993). For both viruses in both genera, ORF 3, which contains a polymerase-like domain, is translated by ribosomal frameshifting from near the end of ORF 2. ORF 2 contains motifs suggestive of a VPg and characteristic of a serine proteinase (Miller *et al.*, 1995; Makinen *et al.*, 1995). Genes downsteam of ORF 3 are translated from subgenomic RNAs. The sobemo-like virus supergroup contains no animal viruses, although the gene order VPg-proteinase-polymerase is similar to that in the picorna-like supergroup and the human astroviruses have a similar arrangement of a proteinase motif fused to a polymerase by a probable frameshift. Viruses in the *Nodaviridae* family have a polymerase-like domain which has been classified in the supergroup 1 sobemo-like lineage (Table II). However these viruses have a capped 5′ terminus and a simpler bipartite genome structure. Only one protein (the polymerase), which is expressed as a 5′-proximal ORF of RNA 1, is required for RNA replication (Ball, 1995). Their replication is likely to be significantly different from that of the sobemoviruses.

4. Picorna-like Virus Supergroup

Viruses in this family have VPg at their 5′ ends and poly(A) at their 3′ ends, their gene products are expressed by proteinase cleavage of a polyprotein translated from the genomic RNA, and a gene order of

helicase-VPg-proteinase-polymerase is conserved. The positions of the structural proteins differ, being translated at the 5′ end (*Picornaviridae*) or 3′ end (*Potyviridae*) of a polyprotein encoding the replication genes, on a separate RNA (*Comoviridae*), or from a subgenomic RNA (*Caliciviridae*).

By far the best-studied viruses in this supergroup are the picornaviruses (Wimmer *et al.*, 1993; Mirzayan and Wimmer, 1994b). The genome map of poliovirus may be divided into three regions, P1, P2, and P3. P1 encodes the virus capsid proteins and P2 and P3 encode the nonstructural proteins. The order of the genome is (N terminus to C terminus) 1A–1B–1C–1D–2A–2B–2C–3A–3B(=VPg)–3C–3D. Translation starts at an initiation site several hundred nucleotides from the 5′ end and continues until a stop codon is reached in the 3′-terminal region. Theoretically, a polyprotein of 2209 amino acids could be produced, but this is partially cleaved while still nascent by proteinases present within the polyprotein, which probably act in cis. The first cleavage is mediated by 2A, a chymotrypsin-like cysteine proteinase, which cleaves a YG linkage between the P1 and P2 regions. Once the 3C region, which encodes another chymotrypsin-like proteinase specific for QG linkages, has been translated, further cleavages can occur. Many intermediates are formed, some of which have functions distinct from the completely cleaved products. Most subsequent cleavages are mediated by $2A^{pro}$, $3CD^{pro}$, and, to a lesser extent, $3C^{pro}$. Many of the partly or completely cleaved products have roles in RNA replication and interact at different stages of replication. The product $2A^{pro}$ is a multifunctional protein which has functions in replication which are distinct from its functions as a proteinase (cleavage of poliovirus polyprotein, inhibition of host protein synthesis) and as an enhancer of poliovirus translation (Yu *et al.*, 1995; Lu *et al.*, 1995). Protein 2B (or its precursor 2BC) may be required in cis during assembly of the replication complex, because RNA-negative 2B mutants are noncomplementable (Johnson and Sarnow, 1991; van Kuppeveld *et al.*, 1995). The NTPase and putative helicase function of the 2C protein was discussed in Section II,B. The 3AB protein is the precursor of VPg (3B). 3AB has also been shown to stimulate the activity of the polymerase ($3D^{pol}$) in catalyzing RNA synthesis on various primed templates by up to 100-fold, possibly because of its ability to bind both the template-primer and $3D^{pol}$ (Lama *et al.*, 1994; Paul *et al.*, 1994a; Plotch and Palant, 1995). 3AB also forms a complex with $3CD^{pro}$ (which has no polymerase activity), resulting in accelerated autoprocessing of $3CD^{pro}$ to form $3C^{pro}$ and the active polymerase $3D^{pol}$ (Molla *et al.*, 1994). The 3AB–CD^{pro} complex was found to form a complex with a cloverleaf structure at the

5′ terminus of poliovirus RNA and formation of the complex was essential for RNA replication (Harris *et al.*, 1994; Xiang *et al.*, 1995a,b). 3AB or $3CD^{pro}$ alone did not bind to the cloverleaf. The 3AB–CD^{pro} complex, or 3AB alone, formed a complex with the 3′-terminal sequence of poliovirus. These observations led Harris *et al.* (1994) to propose a model for poliovirus RNA replication. Initiation of negative-strand synthesis using VPg-pU as a primer occurs following formation of a 3AB–$3CD^{pro}$ complex at the 3′ end of poliovirus RNA, cleavage of $3CD^{pro}$, and formation of $[3AB–D]^{superpol}$. After negative-strand synthesis to form a double-stranded duplex, the end of the duplex is unwound, possibly by the putative 2C helicase, allowing formation of the 3AB–$3CD^{pro}$-cloverleaf complex and initiation of positive-strand synthesis by $[3AB–3D]^{superpol}$ on the free 3′ end of the negative-stranded RNA. The multiplicity of functions of precursors and fully cleaved proteins, combined with the formation of functional complexes and their membrane locations (see Section IV), may lead to explanations of why some mutations, for instance in 3A or $3D^{pol}$, are complementable and others are not. Evidence has been obtained that poliovirus translation and replication are coupled, probably in the early stages of replication (Novak and Kirkegaard, 1994); such coupling could be required for the assembly of new replication complexes. Complete replication of poliovirus has been achieved *in vitro* using a combined translation and replication system (Molla *et al.*, 1991).

The plant comoviruses encode their replication proteins on one RNA segment (RNA 1 or B), which can replicate in protoplasts alone, and their capsid and cell-to-cell movement proteins on a second segment (RNA 2 or M), which requires RNA B for its replication (Eggen and van Kammen, 1988). The gene order of RNA B is (N terminus to C terminus) 32K–58K–VPg–24K–87K. Both RNAs are translated as polyproteins, which are cleaved in cis and trans by a 24K proteinase which resembles the poliovirus $3C^{pro}$. There appears to be no direct equivalent to the poliovirus $2A^{pro}$, but there is a 32K protein at the N terminus of the B polyprotein that acts as a regulator for processing of the B polyprotein and as a cofactor in cleavages in the M polyprotein. The 58K protein contains the helicase-like domain. The 87K protein contains a polymerase-like domain like that in poliovirus $3D^{pol}$, but the active form, detected in isolated replication complexes, appears to be the 110K precursor (24K–87K) (Dorssers *et al.*, 1984). This represents a significant difference from poliovirus, because $3CD^{pro}$ has no polymerase activity. VPg is probably involved in initiation of RNA synthesis (see Section II,D), but its mode of processing may be different from that in poliovirus. In the latest model, it is suggested that initiation may involve cleavage of the 112K polymerase precursor (VPg–24K–87K)

(Peters *et al.*, 1995), although in a crude RNA polymerase preparation, the 60K precursor (58K–VPg) was the most prominent protein detected by anti-VPg serum (Eggen *et al.*, 1988).

Various insertional and frameshift mutants of RNA B, one of which did not contain any of the B-encoded genes, could not be replicated in trans by wild-type RNA B, suggesting that translation and replication may be coupled (van Bokhoven *et al.*, 1993). M RNA must be replicated in trans, but mutational analysis of RNA M indicated that the N-terminal region of the M-encoded 58K protein must be translated to allow replication of RNA M to proceed in the presence of B replication proteins, again indicating a link between translation and replication (van Bokhoven *et al.*, 1993).

The gene order of a typical potyvirus, tobacco etch virus (TEV), is (N terminus to C terminus) P1–HCPro–P3–CI–6K–N1a–N1b–coat protein. The polyprotein translation product is cleaved by three proteinases (Dougherty and Semler, 1993). The P1 proteinase autocatalytically cleaves itself from the rest of the polyprotein. The HC-proteinase (HC-Pro) also autocatalytically cleaves itself at its C terminus in a cotranslational event. The remainder of the cleavages are carried out by the NIa proteinase, which has similarities with the picornavirus and comovirus proteinases. Like the picornavirus proteins, potyvirus proteins are multifunctional and several function in RNA replication (Riechmann *et al.*, 1992). The NIb protein contains the RNA polymerase-like domain. Mutants in the NIb region could be complemented by wild-type NIb protein expressed in transgenic plants with efficiencies varying from 1% to 100% (Li and Carrington, 1995). VPg constitutes the N-terminal region of the NIa proteinase and both NIa and the cleaved VPg have been found linked to the 5′ end of the viral RNA. A large proportion of the NIa and NIb proteins is found in the nucleus. A predominantly nuclear location was also found for the alphavirus nsP2 protein, but this is irrelevant for replication, since a proportion of the protein is localized to the cytoplasmic membranes where replication tales place (see Section IV). Localization of a proportion of the NIa and NIb proteins to the membranes where TEV RNA replication takes place may be mediated by the adjacent 6K protein, which has been shown to be essential for RNA replication (Restrepo-Hartwig and Carrington, 1994). 6K–VPg and 6K–NIa polyproteins were detected in extracts from infected plants. The 6K protein was shown to have a membrane location and may function in a similar way to the poliovirus 3A protein as a membrane anchor (see Section IV). The NIa and NIb proteins have been shown to interact in yeast cells (Hong *et al.*, 1995). The CI helicase was described in Section II,B. The CI, NIb, and NIb–

NIa proteins have been detected in polymerase preparations from infected plants (Martin *et al.,* 1995). HC-Pro also has an essential role in RNA replication, which requires its cleavage activity (Kasschau and Carrington, 1995). The P1 protein has RNA-binding activities (Soumounou and Laliberte, 1994) and has a stimulatory effect on RNA replication (Verchot and Carrington, 1995).

5. Corona-like Virus Supergoup

The coronavirus, torovirus, and arterivirus replication proteins are encoded in gene 1, which consists of two overlapping ORFs, 1a and 1b. ORF 1a is expressed as a polyprotein by direct translation of the viral RNA, while ORF 1b is expressed by ribosomal frameshifting to give a 1a–1b fusion protein. The latter ranges in size from 345 kDa (arteriviruses) up to 800 kDa (coronaviruses) (Snijder and Horzinek, 1993). Polymerase-like and helicase-like domains are found in that order in the frameshift part of the polyprotein. Clearly, this order is different from that in the other virus supergroups and the location of the helicase-like domain in the frameshift portion is also unusual. The gene 1 polyprotein is broken down into functional products by virus-encoded proteinases, some of which are papain-like, another resembles the poliovirus 3C proteinase, and another has similarities to both types (Kim *et al.,* 1995; Snijder *et al.,* 1995). It has been shown for mouse hepatitis virus that RNA replication is inhibited by a cysteine proteinase inhibitor (Kim *et al.,* 1995). Genes downstream of gene 1 are expressed via subgenomic RNAs which are synthesized by a mechanism different from that of other positive-stranded RNA viruses (see Section V,D,3).

F. Effect of Capsid Proteins on RNA Replication

Capsid proteins (CPs) play an important regulatory role in the replication cycle of positive-stranded RNA phages (Witherell *et al.,* 1991). However, it has been found for many eukaryotic positive-stranded RNA viruses that CPs are not essential for virus RNA replication. CP genes have been replaced by other genes in a number of expression vectors derived from animal and plant positive-stranded RNA plant viruses, such as alphaviruses, bromoviruses, and tombusviruses (reviewed by Mushegian and Shepherd, 1995; Schlesinger, 1995). In the absence of the capsid protein, levels of accumulation of positive-stranded RNA are sometimes reduced and this has been ascribed to degradation of the RNA in the absence of the protective capsid, e.g., in beet western yellows virus (Reutenauer *et al.,* 1993), cowpea mosaic virus (van Bokhoven *et al.,* 1993), and cucumber mosaic virus (Boccard and Baulcombe, 1993).

In some other cases, there appears to be little or no effect, e.g., brome mosaic virus (French and Ahlquist, 1987; Marsh *et al.,* 1991a), tomato bushy stunt virus (Scholthof *et al.,* 1993), turnip crinkle virus (Hacker *et al.,* 1992). For poliovirus, a large deletion in the capsid-encoding region led to an increase in RNA accumulation (Collis *et al.,* 1992), although there is evidence that poliovirus replication and encapsidation are linked (Pfister *et al.,* 1995).

Plant viruses in the *Alfamovirus* and *Ilarvirus* genera of the *Bromoviridae* family appear to be exceptional in their requirement for CP in both early and late functions. Alfalfa mosaic virus (AlMV) has a tripartite genome, similar to that of brome mosaic virus. RNAs 1 and 2 encode the replication proteins P1 and P2 (equivalent to 1a and 2a), whereas RNA 3 encodes the movement protein and the CP, the latter being translated from a subgenomic RNA. Unlike brome mosaic virus, plants cannot be infected by a mixture of RNA 1, RNA2, and RNA 3; infection requires additionally a small amount of CP or its subgenomic RNA, a process called genome activation (reviewed by Jaspars, 1985). The 3′ terminal 145 nt of the three RNAs are homologous and can be folded into a structure consisting of a series of stem–loops, separated by AUGC motifs, which contains a high affinity CP-binding site (site 1); an additional similar upstream binding site (site 2) is found in RNA 3 (Reusken *et al.,* 1994; Houser-Scott *et al.,* 1994). Mutations in two of the AUGC motifs in binding site 1 of RNA 3 reduced or abolished binding. The RNA-binding site is located near the N terminus of the CP and one or more lysine residues in this region are needed for gene activation (Baer *et al.,* 1994; Yusibov and Loesch-Fries, 1995). When RNA 3 was used to inoculate transgenic plants expressing P1 and P2 from chromosomal expression cassettes (P12 plants), replication of RNA 3 was achieved in the absence of CP in the inoculum (Neeleman *et al.,* 1993). It was suggested that the role of the CP in gene activation may be to stabilize the RNAs in the inoculum to allow translation of the replicase genes and the formation of a replicase complex to occur.

P12 plants could be infected by RNA 3 with deletions in the CP gene; this had little effect on negative-strand synthesis, but reduced the accumulation of positive strands by 100-fold, indicating that the CP also had an additional role in asymmetric positive-strand accumulation (van der Vossen *et al.,* 1994). This was unlikely to be due to stabilization of the RNA by encapsidation because, using infections with chimeric AlMV–tobacco streak ilarvirus (TSV) constructs, it was found that the CP of TSV could encapsidate AlMV RNA, but not stimulate positive-strand synthesis (Reusken *et al.,* 1995). Further analysis indicated that some AlMV CP mutations affected the early (gene activa-

tion) and late (asymmetric positive-strand accumulation) functions in a quantitatively different way and that AUGC motif mutations that abolished CP binding reduced positive-strand but not negative-strand accumulation (van der Vossen *et al.*, 1994). Quadt *et al.* (1991) detected CP in a purified RdRp preparation from AlMV-infected plants, in addition to P1 and P2. De Graaf *et al.* (1995b) showed that synthesis of full-length and subgenomic positive-stranded RNA on a negative-strand AlMV RNA 3 template by an RdRp preparation isolated from P12 plants was strongly stimulated by addition of CP and suggested that this stimulation could explain both gene activation and asymmetric positive-strand accumulation.

The coronaviruses probably also use the nucleocapsid N protein in the control of genomic and subgenomic RNA synthesis (see Section V,D,3).

Binding of a potyvirus CP to the NIb (polymerase) protein was observed in yeast (Hong *et al.*, 1995), but the significance is not known.

G. Replication Complexes and Virus Movement in Plants

As outlined in the Introduction (Section I), plant viruses have evolved specialized proteins to enable them to move from cell to cell in the plant. Generally it has been found that replication in single cells (protoplasts) is not affected when the movement protein gene is deleted. However, two observations suggest that there may be a link between replication and virus movement in the plant. (1) Mutations in the C-terminal region of the 2a brome mosaic virus replication protein had no effect on replication in single cells, but were deleterious for virus spread in the plant (Traynor *et al.*, 1991). (2) Transgenic plants expressing a truncated cucumber mosaic virus 2a protein were resistant to infection by the virus. The major effect was on inhibition of RNA replication, but to a lesser extent virus spread in the plant was also affected (Carr *et al.*, 1994). It is therefore possible that there may be some interactions between replication complexes and the virus-encoded or host proteins that are involved in virus spread through the plant.

III. Host Proteins in RNA Replication

The notion that replication of eukaryotic positive-stranded viral RNA may require host proteins stems from the requirement of host proteins for ssRNA phage replication. Phage Qβ RNA polymerase complex contains, in addition to the phage-encoded polymerase subunit, 30S ribosomal protein S1, and protein synthesis elongation factors EF-

Tu and EF-Ts (Blumenthal and Carmichael, 1979). Another ribosome-associated protein, termed host factor I (HF-I), is required for RNA synthesis on genomic RNA, but not negative-strand RNA, templates.

Various approaches have been used to attempt to obtain evidence for a role of host proteins in eukaryotic viral RNA replication. The first involves purification of solubilized RdRp and searching for host proteins that copurify with the RdRp in multiple purification steps. Copurification of host proteins with RdRps has been found for several viruses, such as brome mosaic virus (BMV) (Quadt and Jaspars, 1990), cowpea mosaic virus (Dorrsers *et al.,* 1984), cucumber mosaic virus (Hayes and Buck, 1990), red clover necrotic mosaic virus (Bates *et al.,* 1995), Sindbis and Semliki Forest viruses (Barton *et al.,* 1991), and turnip yellow mosaic virus (Mouches *et al.,* 1984). All such studies require further evidence that such copurification is not fortuitous. One of the host proteins that copurified with the BMV RdRp was identified as the barley analogue of the p41 subunit of the wheat germ eukaryotic translation initiation factor eIF-3, or a closely related protein (Quadt *et al.,* 1993). The BMV RdRp-associated host protein and the p41 subunit of wheatgerm eIF-3 were found to bind with high affinity and specificity to the BMV 2a polymerase-like protein. Addition of wheatgerm eIF-3 or its p41 subunit to the BMV RdRp gave a threefold stimulation of negative-strand synthesis. Biochemical functions of eIF-3 include stabilization of Met-$tRNA_i^{Met}$ (ternary complex) binding to the 40S ribosomal subunit, mRNA binding to the ribosome, and dissociation of 80S ribosomes into 60S and 40S subunits, and it is thought to play a key role in assembly of the initiation complex (Merrick, 1992; Hannig, 1995). The specific function in protein synthesis of the p41 subunit of wheatgerm eIF-3, which is composed of 10 subunits, is not known.

Another approach has been to search for proteins that bind specifically to terminal sequences of viral RNAs. Specific binding of host proteins to viral terminal sequences has been reported for diverse viruses: alpha-like supergroup, BMV [3′ (+), 3′ (–), and 5′ (+); barley] (Duggal *et al.,* 1994; Duggal and Hall, 1995), cucumber mosaic virus, tobacco mosaic virus [3′ (–); tobacco, spinach] (Hayes *et al.,* 1994b), Sindbis virus [3′ (–); chicken, mosquito] (Pardigon and Strauss, 1992, 1996; Pardigon *et al.,* 1993), rubella, virus [3′ (–); 3′ (+); 5′ (+); simian] (Nakhasi *et al.,* 1990, 1991, 1994); flavi-like supergroup, West Nile virus [3′ (+); hamster] (Blackwell and Brinton, 1995); picorna-like supergroup, hepatitis A [3′ (+), together with internal sites] (Nuesch *et al.,* 1993; Kusov *et al.,* 1996), poliovirus [5′ (+); human] (Najita and Sarnow, 1990), rhinovirus [3′ (+); human] (Todd *et al.,* 1995), corona-like supergroup, mouse hepatitis virus [3′ (–); 5′ (+); human] (Furuya

and Lai, 1993), [3′ (+); murine] (Yu and Leibowitz, 1995), [intergenic (+); murine] (Zhang and Lai, 1995); carmo-like supergroup, red clover necrotic mosaic virus [3′ (–); tobacco, spinach] (Hayes *et al.*, 1994b). In all these cases, it will be important to distinguish between binding required for translational regulation which can require both 5′ (+) and 3 (+) sequences (Gallie, 1991; Ehrenfeld and Gebhard, 1994; Gallie and Kobayashi, 1994; Schmid and Wimmer, 1994; Standart and Jackson, 1994), binding required for replication, and irrelevant binding because the protein recognizes a structure in a viral RNA which is fortuitously related to structures recognized by the protein in its normal function.

Two proteins that bound to the 5′ (+) terminus of rubella virus RNA were shown to be Ro/SS-A-related antigens and it was suggested that they may have a role in translational control of rubella virus RNA (Pogue *et al.*, 1993). A 60-kDa protein that interacted with a rubella virus 3′ (–) sequence (Nakhasi *et al.*, 1991) was also shown to interact with a 3′ (+) stem–loop structure in the genomic RNA, important for the initiation of negative-strand synthesis (Nakhasi *et al.* 1994). This protein has been identified as the simian homologue of human calreticulin (Singh *et al.*, 1994) and the rubella virus RNA-binding activity has been located to the N-terminal region of the protein (Atreya *et al.*, 1995). Calreticulin is a major calcium storage protein found in the lumen of the endoplasmic reticulum in animals and plants, but it may have other functions, such as in the regulation of gene expression (Dedhar, 1994). Ro/SS-A protein, La/SS-B protein, and calreticulin may be present in RNA–protein complexes; autoantibodies to them are found in cases of systemic lupus erythematosus and Sjögren's syndrome (Zhu and Newkirk, 1994).

One of the proteins that bound to the Sindbis virus 3′ (–) terminal sequence has been shown to be the mosquito homologue of the La autoantigen (Pardigon and Strauss, 1996). The La (SS-B) protein is an abundant cellular protein which belongs to the RNP class of RNA binding proteins; it is found in both the nucleus and the cytoplasm. It binds to the 3′-oligouridine stretch found on all newly synthesized RNA polymerase III transcripts and is required for correct transcript termination and release, and also facilitates reinitiation (Maraia *et al.*, 1994). It can unwind both RNA/DNA hybrids and dsRNA (Xiao *et al.*, 1994). It may also be involved in the internal initiation of translation of poliovirus RNA and in poliovirus-infected cells is largely redistributed to the cytoplasm (Meerovitch *et al.*, 1993). It has also been found to bind to several other viral RNAs (van Verooj *et al.*, 1993).

A complex of a cellular protein and the poliovirus $3CD^{pro}$ protein was found to bind to the cloverleaf structure at the 5′ end of poliovirus RNA

(Andino *et al.*, 1990). The cloverleaf structure has been shown to be essential for virus RNA replication (Andino *et al.*, 1993) and this domain functions independently from the internal ribosomal entry site (IRES) required for translation (Rohll *et al.*, 1994). The cellular protein was identified as an N-terminal fragment of EF-1α (Harris *et al.*, 1994). The significance of this complex is not clear in view of the finding that a 3AB–3CD^{pro} complex also binds to the clovercleaf (see Section II,E,4), although it has been found that mouse cells are temperature-sensitive for initiation of poliovirus positive-strand synthesis, implicating a host factor in this process (Shiroki *et al.*, 1993). A putative host factor, a 67-kDa phosphoprotein which possesses autophosphorylation activity and can phosphorylate the α-submit of eIF-2 *in vitro*, has been shown to be able to function in the initiation of poliovirus negative-strand RNA synthesis *in vitro* (Morrow *et al.*, 1985). It has also been suggested that a terminal uridylyltransferase may act as a host factor in the initiation of poliovirus negative-strand synthesis by adding uridine residues to the 3′ poly(A) end of virion RNA, which could anneal back to the poly(A) to form a hairpin primer for the polymerase (Andrews and Baltimore, 1986). dsRNA molecules joined at one end in a hairpin structure have been isolated from poliovirus-infected cells (Young *et al.*, 1985). A model was proposed in which VPg cleaved the hairpin and became covalently attached to the 5′ UMP of the negative strand in a self-catalyzed transesterification reaction (Tobin *et al.*, 1989). Other models were suggested by Lubinski *et al.* (1986). Good evidence for involvement of a host protein in poliovirus replication comes from the use of a combined *in vitro* translation/replication system. Preinitiation complexes isolated after inhibiting the initiation of RNA synthesis with guanidine-HCl were shown to require the addition of soluble cellular factors for initiation of RNA synthesis (Barton *et al.*, 1995).

Recently it has been shown that the human protein Sam68 binds strongly to the poliovirus 3D^{pol} (McBride *et al.*, 1996). In uninfected cells, Sam68 was found to be located mainly in the nucleus. It is known to associate with Src during mitosis (Taylor and Shalloway, 1994; Fumagalli *et al.*, 1994), it has SH-2 and SH-3 binding domains, and it binds to ssRNA and dsRNA (Taylor *et al.*, 1994, 1995). In poliovirus-infected cells, Sam68 was relocated to the cytoplasm where it was bound to the 3D^{pol} protein in membrane-bound replication complexes together with other virus proteins, including 2BC and 2C (McBride *et al.*, 1996).

A host protein may be involved in the resistance to tobacco mosaic virus conferred by the Tm-1 gene in tomato. Expression of this gene causes inhibition of RNA replication (Watanabe *et al.*, 1987) and mutations in resistance-breaking strains of the virus map close to the helicase-

like motif IV in the 126-kDa protein (Meshi *et al.,* 1988). A possible explanation is that the resistance gene could be an allele of a gene encoding an essential host component of the replicase complex. The product of the resistance gene would be a mutant of this host protein, which may be unable to form a functional replicase. The resistance-breaking strains would contain compensating mutations to allow production of an active replicase. Host-specific alterations in brome mosaic virus RNA accumulation, dependent on RNA 1, have also been reported (de Jong and Ahlquist, 1995). A recessive mutation *tom1* in *Arabidopsis thaliana* reduced replication of tobacco mosaic virus RNA to low levels, suggesting that the product of the wild-type gene was required for replication (Ishikawa *et al.,* 1993). Studies of host-range mutants (Kowal and Stollar, 1981) and analysis of the effects of some cis-acting seqences of alphaviruses (Kuhn *et al.,* 1992) also suggest a role of host factors in replication.

Other evidence for roles of host factors come from studies of inhibitors of host transcription, e.g., there are steps in the replication of cowpea mosaic virus (de Varennes *et al.,* 1985), Sindbis virus (Baric *et al.,* 1983a), and tobacco mosaic virus (Dawson, 1978) that are sensitive to actinomycin D. Another study suggested that a host protein may negatively control production of a tobacco mosaic virus subgenomic RNA (Blum *et al.,* 1989). A new approach that should uncover host genes required for the replication of brome mosaic virus has been opened up by the demonstration that this virus can replicate in yeast cells (Janda and Ahlquist, 1993). This opens up the powerful yeast genetic system for the creation of host mutants defective in virus RNA replication, and for the isolation of the genes involved. The ability of brome mosaic virus to replicate in yeast clearly depends on the ability of yeast proteins to substitute for plant proteins in the replication of the virus RNA and implies a degree of conservation. However, several animal and plant genes have been cloned by complementation in yeast, and yeast proteins can often substitute for mammalian proteins, e.g., yeast eIF-3 can substitute for the mammalian factor in a heterologous reconstituted *in vitro* assay system (Naranda *et al.,* 1994). The method may therefore be applicable to a range of animal and plant viruses.

IV. The Role of Membranes in RNA Replication

The genome of animal and plant positive-stranded RNA viruses from a number of different virus supergroups is replicated *in vivo* in membrane-bound complexes. Isolated membrane complexes have been

found to be capable of *in vitro* synthesis of replicative intermediates, elongation and release of genomic-length RNA, and in some cases, initiation of RNA synthesis on endogenous templates which remain bound to the replication complex. Examples include alpha-like virus supergroup, Sindbis virus (Barton *et al.*, 1991), alfalfa mosaic virus (de Graaf *et al.*, 1993), cucumber mosaic virus (Jaspars *et al.*, 1985), foxtail mosaic virus (Rouleau *et al.*, 1993), tobacco mosaic virus (Young and Zaitlin, 1986); picorna-like virus supergroup, poliovirus (Takeda *et al.*, 1986; Bienz *et al.*, 1990); cowpea mosaic virus (Eggen *et al.*, 1988); plum pox virus (Martin *et al.*, 1995); flavi-like virus supergroup, West Nile virus (Grun and Brinton, 1988), Kunjin virus (Chu and Westaway, 1992), dengue virus (Bartholomeusz and Wright, 1993); carmo-like virus supergroup, red clover necrotic mosaic virus (Bates *et al.*, 1995), turnip crinkle virus (Song and Simon, 1994); sobemo-like virus supergroup, velvet tobacco mottle virus (Rohozinski *et al.*, 1986); corona-like virus supergroup, mouse hepatitis virus (Brayton *et al.*, 1982, 1984). In the case of flockhouse virus, it was possible to remove the bound RNA from the membrane-bound complex by nuclease digestion to produce a template-dependent RNA polymerase. Such preparations synthesized only negative-strand RNA (isolated as dsRNA) with a genomic RNA template, but on addition of certain neutral or negatively charged phosphoglycerolipids (PGLs) both negative and positive strands were produced, giving complete replication of the genomic RNA (Wu and Kaesberg, 1991; Wu *et al.*, 1992). It was suggested that initiation of positive-strand synthesis may result from a direct GPL-replicase interaction, analogous to activation of the *E. coli* replication initiator protein dnaA by diphosphatidylglycerol (Sekimizu and Kornberg, 1988), or a change in membrane configuration to mimic similar changes that may occur during replication *in vivo*. Membranes also appear to be important for the replication of poliovirus RNA in a combined translation and replication system (Molla *et al.*, 1991, 1992, 1993; Barton *et al.*, 1995). For some other viruses it has been possible to produce template-dependent RNA polymerases after detergent solubilization and removal of the endogenous RNA (reviewed by de Graaf and Jaspars, 1994). Generally, only the complementary strand has been synthesized by such preparations, although a solubilized polymerase able to catalyze the complete replication of cucumber mosaic virus (CMV) RNA has been described (Hayes and Buck, 1990). *In vitro* replication of CMV RNA in this system was, however, very inefficient with only a small fraction of the template being copied and the ratio of positive to negative strands much lower than that *in vivo*.

Infection of cells with many positive-stranded RNA viruses results in the formation of multiple vesicles or invaginations in the membranes of various organelles. It has been suggested that continuous synthesis of lipid is required for the replication of some viruses. Infection of cells by poliovirus (Guinea and Carrasco, 1990; Maynell *et al.,* 1992) or Semliki Forest virus (Perez *et al.,* 1991) led to increased lipid synthesis; conversely, replication of both these viruses was inhibited by cerulenin, an inhibitor of lipid biosynthesis. Cerulenin, which has several activities apart from inhibition of lipid biosythesis (Odd and Wu, 1993), also inhibited poliovirus replication in an *in vitro* translation/replication system (Molla *et al.,* 1993). It remains uncertain whether increased lipid synthesis in cells infected by these viruses is a requirement for, or a consequence of, virus replication.

Infection of cells with poliovirus results in the formation of numerous vesicles, derived largely from the rough endoplasmic reticulum, which have been compared to the intermediate or transport vesicles involved in cellular protein sorting and secretion (Bienz *et al.,* 1987), although it has recently been shown that the virus-induced membranous structures are bounded by double lipid bilayers (Schlegel *et al.,* 1996). Lysosomes, trans-Golgi, and the trans-Golgi network also contribute to the virus-induced membranous structures (Schlegel *et al.,* 1996). Brefeldin A, an inhibitor of the cellular secretory pathway, inhibits poliovirus replication, possibly by preventing the formation of these vesicles (Maynell *et al.,* 1992). Expression of the 2C or 2BC proteins separately in cells using vaccinia virus vectors induced the formation of similar vesicles, but did not cause an increase in lipid synthesis (Cho *et al.,* 1994; Aldabe and Carrasco, 1995). 2BC has also been reported to induce vesicle formation in yeast (Barco and Carrasco, 1995). The replication complex was located on the surface of the vesicles; proteins 2B, 2C, and 2BC were located exclusively with the complex and may be responsible for its organization, whereas 3D and its precursors were also found in the peripheral cytoplasm (Bienz *et al.,* 1990). Isolated replication complexes were shown to consist of a central replication complex consisting of small, densely packed vesicles, surrounded by larger vesicles in a rosette-like arrangement, which could be disrupted by guanidine, suggesting a role of the 2C protein in the organization of the rosette (Bienz *et al.,* 1992, 1994; Troxler *et al.,* 1992). The larger vesicles may aid in the release of newly synthesized RNA and in its subsequent encapsidation (Pfister *et al.,* 1995). Expression of proteins 2B and 3A separately blocked cellular secretion in the absence of virus infection; evidence indicated that in virus-infected

cells, 3A may function to block the fusion of the virus-induced vesicles with the Golgi membranes (Doedens and Kirkegaard, 1995). The 3A and 2B proteins may attach to membranes via hydrophobic domains (Datta and Dasgupta, 1994; van Kuppeveld *et al.*, 1995); the 2C protein may attach via the hydrophobic side of an amphipathic helix (Paul *et al.*, 1994b). It has been suggested that the 3A protein forms a membrane anchor for VPg (in the form of 3AB) in the initiation of RNA synthesis. Mutations in the hydrophobic domain of 3A affected initiation of RNA synthesis, *in vitro* uridylylation of VPg, and *in vivo* positive-stranded RNA synthesis (Giachetti and Semler, 1991) (see also Section II,E,4).

Large arrays of membranous vesicles, analogous to those seen in poliovirus-infected cells and containing viral RNA and nonstructural proteins, have been detected in cells infected with cowpea mosaic virus, a plant picorna-like virus (Wellinck *et al.*, 1988). Similar structures were observed when the RNA B-encoded 200K or 60K (58K-VPg) protein was expressed in insect cells, and the 60K protein was shown to be associated with the vesicles (van Bokhoven *et al.*, 1992). The cowpea mosaic virus RNA B-encoded 58K protein is the equivalent of the poliovirus 2C protein (see Section II,E,4) and hence may induce the vesicles in a similar way to the 2C protein. Cowpea mosaic virus has no proteins with significant sequence homology to the poliovirus 2B and 3A proteins, but it is possible that the 58K protein might have domains which serve the same functions.

The cytopathic structures formed by viruses in another plant picorna-like virus genus, the potyviruses, are somewhat different and can include formation of pinwheels by the cytoplasmic inclusion (CI) protein (which has the helicase-like domain) and invaginations in the nuclear membrane (Lesemann, 1988). The tobacco etch potyvirus 6-kDa protein expressed in transgenic plants apparently causes the formation of, and localizes to, membranous proliferations associated with the periphery of the nucleus (Restrepo-Hartwig and Carrington, 1994). These may be the sites of virus replication in infected cells. The possible role of the 6-kDa protein (probably equivalent to the poliovirus 3A protein) as a membrane anchor for VPg and its association with other replication proteins was discussed in Section II,E,4. Overall, it appears that there are likely to be similarities and differences between potyvirus and picornavirus replication.

There also appear to be diversities in the membrane localization of replication complexes in the alpha-like supergroup. RNA synthesis of alphaviruses, such as Sindbis and Semliki Forest virus, takes place on membranous structures called type I cytopathic vacuoles (CPVIs), on the surface of which are located the virus replication proteins nsP1,

nsP2, nsP3, and nsP4 (Froshauer *et al.,* 1988; Peränen and Kääriäinen, 1991). CPVIs are modified endosomes and lysosomes, which have characteristic invaginations (spherules), which may represent the attachment sites for the replication complexes. Spherules are also found at the plasma membrane. When nsP1 was expressed alone in cells synchronously, it was located first on the cytoplasmic side of the plasma membrane and then moved to endosomes and later lysosomes (Peränen *et al.,* 1995). Evidence indicated that nsP1 could be acylated, possibly at the plasma membrane, and that it may be attached to membranes by a fatty acid residue; it may act as an anchor for the other replication proteins. A large proportion of the nsP2 protein is normally targeted to the nucleus, but this is irrelevant for replication, because redirection to the cytoplasm by removal of the nuclear localization signal did not affect virus replication (Rikkonen *et al.,* 1994).

Replication complexes of tobacco mosaic virus, a plant alpha-like virus, are associated with cytoplasmic inclusions, called viroplasms, which enlarge during the course of infection to form "X bodies." They are composed of aggregates of tubules, possibly derived from the endoplasmic reticulum, which may be twisted round each other to form ropes, embedded in a ribosome-rich matrix (Esau and Crinshaw, 1967; Saito *et al.,* 1987; Hills *et al.,* 1987). The viroplasms contain the 126-kDa and/or 183-kDa replication proteins, which are associated with the tubules, and are therefore the likely sites of RNA replication. In contrast, cytoplasmic invaginations of the chloroplast outer membrane are the site of RNA synthesis of turnip yellow mosaic virus RNA (reviewed in Garnier *et al.,* 1986); the chloroplast outer membrane is probably also the site of alfalfa mosaic virus RNA replication (de Graaf *et al.,* 1993). Replication of cucumber mosaic virus may be associated with invaginations in the vacuolar membrane (tonoplast) (Hatta and Francki, 1981). In the case of brome mosaic virus, the 1a and 2a replication proteins were initially localized to punctate structures in the cytoplasm, which aggregated into well-defined structures adjacent to the nucleus, containing nascent viral RNA (Restrapo-Hartwig and Ahlquist, 1995). The mechanism of targeting of replication complexes to different regions of the cell is unknown, but it is interesting that replication complexes of brome mosaic virus in yeast were membrane-bound (Quadt *et al.,* 1995). Binding of brome mosaic virus or alfalfa mosaic virus replication proteins to membranes in yeast (Quadt *et al.,* 1995) or plants (de Graaf *et al.,* 1995b) did not require the presence of viral RNA.

In the carmo-like virus supergroup, the formation of multivesicular bodies (MVBs) derived from peroxisome membranes is characteristic of several tombusviruses (artichoke mottled crinkle, cymbidium ringspot,

eggplant mottle crinkle, tomato bushy stunt). In some cases, dsRNA has been located to the MBVs indicating that they are the likely site of replication (Russo *et al.*, 1983; Lupo *et al.*, 1994). However, with another tombusvirus, carnation Italian ringspot virus, MBVs were found to be derived by peripheral vesiculation of mitochondria (Russo *et al.*, 1995). The replication proteins of tombusviruses (see Section II,E,2) are known to be membrane-bound (Lupo *et al.*, 1994; Scholthof *et al.*, 1995). Analysis of hybrids between cymbidium ringspot and carnation Italian ringspot viruses indicated that the subcellular origin of the MBVs was determined by the N-terminal region of the prereadthrough protein encoded by the 5′-proximal ORF (Russo *et al.*, 1995; M. Russo, personal communication).

V. Cis-Acting Nucleotide Sequences Required for RNA Replication

Essential cis-acting sequences will include promoters for negative-strand and positive-strand RNA synthesis. It is also possible that additional sequences might be required for the assembly of replication complexes. Assembly of brome mosaic virus replication complexes in yeast requires the presence of viral RNA (Quadt *et al.*, 1995). However, once complexes have been assembled *in vivo* in yeast or plants, after isolation and removal of the RNA template, they remain competent to initiate negative-strand synthesis on added positive-strand templates (Miller and Hall, 1983; Quadt *et al.*, 1995; Sun and Kao, 1996). It is therefore possible that assembled replication complexes, having copied one template, could be recycled for use on another template. Hence the cis-acting sequences required for replicase assembly and for promoter recognition and initiation of negative-strand synthesis may not be identical, although they are likely to have elements in common. Replication complexes assembled on positive-strand templates could be recycled and modified to recognize negative-strand templates, and evidence suggests that this is the case for the alphaviruses (see Section II,E,1). However, the ability of flockhouse virus to initiate RNA synthesis on negative-strand templates (Ball, 1994) suggests that, for this virus, a functional RNA polymerase can be assembled on a negative-strand template, although whether this occurs during a normal virus infection remains to be shown. Since accumulation of virus RNA *in vivo* depends on expression of replication proteins and encapsidation, as well as the RNA replication process *per se*, methods for identifying cis-acting sequences specifically required for RNA replication have

relied on DI (and in special cases satellite) RNAs, *in vitro* transcription systems, and in the case of viruses with segmented genomes, genomic RNA segments not required for replication.

A. 3′-Terminal Sequences of Positive Strands

1. tRNA-like Sequences in RNAs of Some Plant Alpha-like Viruses

Sequences that can be folded into tRNA-like structures have been found at the 3′-termini of several genera of plant viruses in the alpha-like virus supergroup (reviewed by Florentz and Gierge, 1995). Such termini are substrates for specific aminoacyl-tRNA synthetases and can be aminoacylated with valine (tymoviruses, sunnhemp mosaic tobamovirus), tyrosine (bromoviruses, cucumoviruses), histidine (most tobamoviruses), although aminoacylation is much less efficient than with the cognate canonical tRNA. They can also be adenylated by tRNA nucleotidyltransferases and, after aminoacylation, can bind elongation factors, such as EF1-α. Some RNAs are also substrates for RNase P.

The structural requirements in the tRNA-like sequence of brome mosaic virus (BMV) for RNA replication, aminoacylation and adenylation have been examined in detail (reviewed by David *et al.*, 1992; Duggal *et al.*, 1994). A 134 nt 3′-terminal fragment containing the tRNA-like structure could act as a template for negative-strand synthesis by an isolated RdRp *in vitro*, but large deletions which removed several stem–loops and pseudoknots in the region from nt 135 to 280 upstream of the 3′ terminus reduced accumulation of RNA 3 *in vivo* to undetectable levels, indicating that a larger 3′ region was needed for efficient replication *in vivo* (Lahser *et al.*, 1993). Furthermore, although the 3′-terminal 200 nt of RNAs 1, 2, and 3 are very similar in sequence, reciprocal exchanges led to aberrant replication, again suggesting a requirement for compatibility with upstream sequences (Duggal *et al.*, 1992). In the case of tobacco mosaic virus (TMV), removal of a single pseudoknot structure upstream of the tRNA-like structure reduced RNA replication (Takamatsu *et al.*, 1990). The terminal A residue of the BMV template is not copied or required for infectivity; it is added on by a posttranscriptional mechanism, possibly by the RdRp as with Qβ RNA (Blumenthal and Carmichael, 1979) or by tRNA nucleotidyltransferase (see below). Initiation of RNA synthesis therefore takes place internally on the template (initiating with a 5′-G) and it is noteworthy that the infectivity of RNA transcripts, produced from cDNA clones, is generally tolerant to 3′ extensions of moderate length (Boyer and Haenni, 1994). Mutations in the

terminal C residues of the –CCA terminus of RNA 3 that greatly reduced *in vitro* RNA synthesis were rapidly repaired *in vivo*, probably by nuclease degradation and resynthesis by tRNA nucleotidyltransferase, an enzyme known to be able to repair tRNA CCA ends (Rao *et al.*, 1989), consistent with the suggestion of a telomere-like function for the 3′ ends of genomic RNA molecules (Weiner and Maizels, 1987). Regions of the tRNA-like structure that are important for replication, aminoacylation, and adenylation overlapped, but were distinguishable. Some aminoacylation-defective mutants of RNA 3 were not greatly debilitated in RNA replication, indicating that charging with tyrosine is not essential for replication of this RNA (Dreher *et al.*, 1989). Further experiments indicated a possible correlation between aminoacylation and replication for RNA 1 (Duggal *et al.*, 1994) and RNA 2 (Rao and Hall, 1991). *In vitro* transcription studies with turnip yellow mosaic virus (TYMV) RdRp indicated that 3′ tRNA-like structure contained the promoter for negative-strand synthesis (Morch *et al.*, 1987; Gargouri-Bouzid *et al.*, 1991). Moreover, TYMV RNAs with anticodon loop substitutions that resulted in decreased valylation failed to replicate efficiently (Tsai and Dreher, 1991). In some cases, second-site suppressor mutations that restored both valylation and replication appeared (Tsai and Dreher, 1992). Nevertheless, a requirement of aminoacylation for efficient replication cannot be established unequivocally by such correlations, because sequences needed for aminoacylation may also be required for template recognition by the replicase. The 3′ termini of tobraviruses and furoviruses can be folded into tRNA-like structures that cannot be aminoacylated.

Reciprocal exchanges of 3′ tRNA-like termini have been carried out to determine the template specificity of replicase complexes. It was found that TMV RNA containing a 3′ terminal region of BMV RNA 3 was amplified by TMV replicase, although much less efficiently than the wild-type 3′ end, indicating that the TMV replicase can recognize some feature of the BMV RNA 3 3′-terminal structure (Ishikawa *et al.*, 1991a). However, BMV RNA 3 containing a 3′ terminal region of TMV RNA was not amplifiable by BMV replicase provided by RNA1 and 2, or by TMV RNA, or a mixture of all three. Similarly, replacement of the 3′-terminal tRNA-like structure of TYMV with that of BMV or TMV gave only low viral accumulation in protoplasts and no systemic symptoms on plants (Skuzeski *et al.*, 1996). In contrast, BMV RNA 3 containing a 3′ terminal region of cucumber mosaic virus (CMV) RNA 3 could be amplified by BMV RNA 1 and RNA 2 to give both the hybrid RNA 3 and a subgenomic RNA (Rao and Grantham, 1994), probably reflecting the greater structural similarity of the 3′ end of BMV and

CMV RNAs than BMV and TMV RNAs. BMV RNA 2 containing a 3′-terminal region of CMV RNA 2 could not be amplified by the BMV RNAs, perhaps reflecting the need for interactions with upstream sequences, since only the terminal 186 nt were exchanged. Reciprocal 3′ exchanges between the 3′ ends of the RNAs 3 of BMV and a closely related bromovirus, cowpea chlorotic mottle virus (CCMV), indicated that the 3′ sequences were not the sole or the major determinants of template specificity (Pacha and Ahlquist, 1991). The 3′ sequences of these two viruses are, however, much more closely related to each other than the other examples discussed above. Overall, it may be concluded that 3′ terminal sequences are important determinants of template specificity, with replicase complexes being able to recognize some common structural elements in related viruses.

2. Other 3′ Structural Elements in Alpha-like Viruses

Whereas the bromoviruses and cucumoviruses have 3′ tRNA-like structures, viruses in the other two genera of the *Bromoviridae* family, the alfamoviruses and ilarviruses, have 3′ structures consisting of a series of stem–loop structures (see Section II,F). In an *in vitro* assay with a template-dependent RNA polymerase isolated from plants infected with alfalfa mosaic virus, it was shown that 3′ deletions of up to 133 nt in RNA 3 did not affect its ability to act as a template for negative-strand synthesis (van der Kuyl *et al.*, 1990). However, a 3′ deletion of 163 nt completely abolished its template activity. In an *in vivo* assay in which deletion mutants of RNA 3 were used to infect transgenic plants expressing the P1 and P2 replication proteins, it was found that deletions of 11 to 133 nt from the 3′ end of the RNA-reduced replication to about 1% of that of the full-length RNA 3, whereas a 3′ deletion of 200-nt-reduced replication to undetectable levels (van der Kuyl *et al.*, 1991). It is noteworthy that nucleotides 11 to 127 and nucleotides 133 to 208 of RNA 3 contain independent coat protein (CP) binding sites (Reusken *et al.*, 1994; see Section II,F), the replicase complex in virus-infected plants contains the CP (Quadt *et al.*, 1991) and replicase in P1P2 plants can assemble *in vivo* in the absence of CP. It is possible that the region from nucleotides 11 to 127 is important for the assembly of the replication complex lacking CP and as the promoter for negative-strand synthesis. It is known that CP inhibits the activity of isolated replication complexes in synthesizing negative strands on positive-strand templates, probably by competing for the replicase binding site (Houwing and Jaspars, 1986; Quadt *et al.*, 1991).

Some alpha-like viruses have 3′ poly(A) tails (Table II) and in some cases it has been shown that there is a poly(U) sequence at the 5′ end

of the negative strand, as in potato virus X (Dolja *et al.*, 1987). Deletion of the poly(A) tail from beet necrotic yellow vein virus (BNYVV) RNA 3 caused a great reduction in ability to replicate (in the presence of RNAs 1 and 2) and the poly(A) tail was restored in the progeny, together with a short upstream U-rich sequence not originally present (Jupin *et al.*, 1990a). Sequence evidence suggested that the poly(A) sequence was not restored by recombination with RNA 1 or 2. Addition of poly(A) also occurred during infection with 3′ poly(A)-deficient transcripts of white clover mosaic virus (Guilford *et al.*, 1991). Further deletion analysis with BNYVV RNA 3 showed that a 67-nt sequence upstream of the poly(A), which can be folded into a double-hairpin secondary structure that is conserved in all four of the virus RNAs, was essential for replication and that a further 50-nt upstream contributed to efficient replication (Jupin *et al.*, 1990b; Richards and Tamada, 1992).

Viruses in the animal *Alphavirus* and *Rubivirus* genera of the *Togaviridae* family also have 3′ poly(A) tails. In the alphaviruses, there is a highly conserved 19-nt U-rich sequence upstream of the poly(A) that is required for RNA replication (reviewed by Strauss and Strauss, 1994). All naturally occurring Sindbis and Semliki Forest virus DI RNAs were found to contain a minimum of 50 nt from the 3′ end of the genome, including the conserved 19-nt sequence. Within the 3′ untranslated region of alphaviruses, there are variable numbers of 40–60 nt repeat sequences; these appear to have a host-specific effect on RNA replication. The 19 nt alphavirus conserved sequence is not found at the 3′ end of rubella virus RNA, but there is a sequence 58 nt upstream of the poly(A) that can be folded into a stable stem–loop structure, with a GC-rich stem and a loop composed only of U residues (reviewed by Frey, 1994). There is evidence that this structure may be required for RNA replication and it has been implicated in the the binding of calreticulin (Nakhasi *et al.*, 1994; see Section III).

3. 3′-Terminal Structures in the Picorna-like Viruses

The 3′ termini of all the viruses in the picorna-like virus supergoup are polyadenylated and several lines of evidence suggest that the poly(A) tract is present, not just to protect the RNA from 3′ exonuclease degradation, but also to provide an essential cis-acting sequence for RNA replication. The poly(A) tail has been shown to be required for the infectivity of poliovirus (Spector and Baltimore, 1974; Sarnow, 1989) and cowpea mosaic virus (Eggen *et al.*, 1989a). The poly(A) sequence is transcribed from a poly(U) sequence at the 5′ end of the negative strand. However, in poliovirus, the poly(A) tract has been reported to be longer than the poly(U) tract (Larsen *et al.*, 1980). The additional A

residues may be added on by slippage or by terminal adenylyl transferase activity (Neufeld *et al.*, 1994). Addition of poly(A) tracts occurs during infection with 3′ poly(A)-deficient transcripts of cowpea mosaic virus (Eggen *et al.*, 1989b) and plum pox virus (Riechmann *et al.*, 1990).

Sequences upstream of the 3′ poly(A) are also important for replication, and may form secondary structures that include part of the poly(A) tail. The 3′ 151 nt of cowpea mosaic virus RNA M contains all the 3′-terminal cis-acting elements required for RNA replication (Rohll *et al.*, 1993). The 3′ 65 nt upstream of the poly(A) in both RNAs B and M have a high degree of sequence similarity and can be folded to create a stem–loop, containing four A residues of the poly(A), linked to a Y-shaped structure (Eggen *et al.*, 1989a; Rohll *et al.*, 1993). Mutagenesis showed the importance of both these features for virus RNA replication. Another upstream putative stem–loop structure within the 3′-terminal 151 nt of RNA M was also needed for replication (Rohll *et al.*, 1993), although the sequences of RNAs M and B have little sequence similarity in this region. Nevertheless, replacement of the 3′-terminal 210 nt of RNA M by the 3′ 500 nt of RNA B had only a small effect on replication (van Bokhoven *et al.*, 1993).

The 3′ end of poliovirus RNA, upstream of the poly(A) tract, can be folded into a tRNA-like structure (Pilipenko *et al.*, 1992), although this is less similar to cellular tRNAs than those discussed in Section V,A,1. Further analysis indicated that a pseudoknot, formed between the 3′ untranslated region and sequences upstream of the translational terminator, was important for replication and one of the stem–loop structures involved base-pairing with five A residues of the poly(A) tail (Jacobson *et al.*, 1993). Moreover, an 8-nt insertion which would affect the 5′-proximal of the 3′ stem–loop structures conferred a ts replication phenotype (Sarnow *et al.*, 1986). Further analysis has confirmed the importance for replication of secondary structure in the 3′ untranslated regions of poliovirus and other picornaviruses (Rohll *et al.*, 1995). Mutagenesis of the single 3′-terminal stem–loop formed by a human rhinovirus untranslated region indicated the importance of the loop sequence, the stability of the stem, and its proximity to the poly(A) tract (Rohll *et al.*, 1995). Binding of picornavirus replication complexes to the 3′-terminal sequences may involve both viral and host proteins. Cui *et al.* (1993) reported that the encephalomyocarditis virus (EMC) $3D^{pol}$ protein bound specifically to the 3′ noncoding region of EMC RNA. Further analysis showed that binding was dependent on covalent attachment of the 3′ noncoding region and the poly(A), a U-rich sequence upstream of the poly(A) and part of the poly(A) sequence; a stem–loop structure with a pseudoknot linking the poly(A) to a U-rich

loop was proposed (Cui and Porter, 1995). Evidence also suggested that binding of poliovirus $3CD^{pro}$ protein to the poliovirus 3′ pseudoknot-poly(A) structure in the presence of the 3AB protein occurred via contact points in $3D^{pol}$ (Harris *et al.*, 1994). Host proteins, inducible by virus infection, which bound to the 3′ regions of rhinovirus and poliovirus RNAs, have also been reported (Todd *et al.*, 1995).

4. 3′-Terminal Structures in the Carmo-like Viruses

RNAs of all viruses in the carmo-like virus supergroup, *together* with several satellite and DI RNAs, have –CCC 3′ termini with no poly(A) tail (Russo *et al.*, 1994). Clones of cymbidium ringspot virus (CyRSV) RNA with the terminal CCC deleted, so that the 3′ terminus was –G, were infectious for plants with a delay of 2 to 3 days in appearance of symptoms. Analysis of the progeny RNA showed that the 3′ end had been repaired by the addition of one or more C residues (Dalmay *et al.*, 1993b). Similarly, –GGGG termini were repaired to –GCCC, although –GGCC termini were stable. Artificially added 3′ poly(A) tails were removed *in vivo*. Removal of four nucleotides from the 3′ end led to complete loss of infectivity. A repair mechanism also operated with a CyRSV satellite RNA which had heterogeneous 3′ termini (–C, –CC, –CCC, –CCCA). Up to eight residues could be removed from the longest 3′ terminus without complete loss of infectivity and the molecules were repaired to give termini mostly identical to the wild-type (Dalmay and Rubini, 1995). The satellite is therefore even more tolerant than the genomic RNA to 3′-terminal deletions. Short repeat units were characteristic of the 3′ termini of both the genomic (**GCA GCA** AU **GCA GC** CC) and satellite (**ACA ACA AC** CCA). Turnip crinkle satellite RNA D molecules with 3′-terminal truncations also had their 3′ termini restored *in vivo* to the motif $(C_{1-2})UG(C_{1-3})$, giving ends similar or identical to the wild-type satellite (CCUGCCC) (Carpenter and Simon, 1996). This also probably occurred by a repair mechanism, rather than by recombination with the genomic RNA. Whether all these repairs are carried out by the viral polymerase or a cellular enzyme is not known. However, as with the tRNA-like termini (see Section V,A,1), analogies to telomerase, an enzyme that contains an RNA template, and to the short repeated sequences in the telomeres of cellular chromosomes (Blackburn, 1993) have been made (Dalmay and Rubino, 1995; Carpenter and Simon, 1996).

A CyRSV DI RNA retained a block of 102 nt from the 3′ terminus of the genomic RNA, of which 77 nt were required for replication (Havelda *et al.*, 1995). Similar 3′ regions are retained in DI RNAs of other tombusviruses, such as cucumber necrosis virus and tomato bushy stunt

virus (Knorr *et al.*, 1991; Finnen and Rochon, 1993; White and Morris, 1994; Chang *et al.*, 1995). The 77-nt CyRSV DI RNA 3′ domain could be folded into a structure composed of three hairpins and two non-base-paired regions. Mutational analysis showed that replication competence depended on maintaining the structure of the stems (Havelda and Burgyan, 1995). The ability to form 3′ stem–loop structures is conserved in all the carmo-like viruses. Using an *in vitro* transcription system, it has been shown that the promoter for negative-strand synthesis of a turnip crinkle virus (TCV) satellite RNA (sat-RNA C) is contained within the 3′-terminal 29 nt of the positive strand (Song and Simon, 1995). Structural probing revealed the presence of hairpin structure within this region. Mutagenesis showed that the primary sequence or size of the loop was not important for replication. However mutations that altered the structural integrity of the lower part of the stem strongly reduced template activity *in vitro*. The 3′ 37 nt of sat-RNA C could be joined to an inactive template and the resultant hybrid was competent for transcription *in vitro* by the TCV RdRp. The 3′-terminal sequences of TCV genomic RNA and sat-RNA C are 90% identical and the 3′ genomic RNA could be folded into a similar hairpin structure. However, upstream motifs in the 3′ untranslated region of TCV RNA, close to the end of the coat protein ORF, are also important for RNA replication *in vivo* (Carpenter *et al.*, 1995). It is noteworthy that the TCV RdRp could utilize negative strands of sat-RNA C as templates (Song and Simon 1994). Template activity depended on 5′-proximal sequences, but was insensitive to 3′ deletions, suggesting that the RdRp recognized a 5′ sequence and then scanned the RNA for a 3′ terminus to initiate positive-strand synthesis. It is possible that the RdRp might recognize a stem–loop structure at the 5′ end of the negative strand.

5. 3′-Terminal Structures in Coronavirus RNAs

Analysis of mouse hepatitis virus (MHV) DI RNAs has shown that a 3′-terminal sequence of 436 nt is needed for RNA replication (Kim *et al.*, 1993). This sequence can be folded into a structure containing several hairpins which may correspond to binding sites for host proteins (Yu and Leibowitz, 1995). However, the specific requirements of negative-strand synthesis were investigated using artificial DI RNAs carrying 5′ deletions that prevented complete replication (Lin *et al.*, 1994). It was found that the cis-acting signal for negative-strand synthesis lay in the 55 nt from the 3′ end plus poly(A) tail of the MHV genome, which included the 3′-proximal hairpin structure. No further upstream sequences were required, but DI RNAs which transcribed subgenomic

RNAs synthesized less negative-stranded RNA. It is noteworthy that the poly(U) tract at the 5′ terminus of bovine coronavirus negative strands (8–20 nt) is shorter than the poly(A) tract at the 3′ end of the positive strands (100–130 nt) (Hofmann and Brian, 1991). This is similar to the situation discussed for poliovirus (see Section V,A,3).

6. 3′-Cis-Acting Sequences of Novaviridae RNAs

The cis-acting sequences required for the replication of flockhouse virus (FHV) RNA 2 have been studied using an *in vivo* system in which the templates were transcribed intracellularly from DNA plasmids containing FHV cDNA flanked by a T7 promoter and a ribozyme, using T7 RNA polymerase expressed from a vaccinia virus recombinant; replication proteins were provided by FHV RNA 1. These studies revealed that 50–60 nt at the 3′ end of FHV RNA 2 were required for replication (Ball and Li, 1993; Ball, 1994).

B. 5′-Terminal Sequences of Positive Strands and 3′-Terminal Sequences of Negative Strands

The 5′-terminal regions of positive strands and 3′-terminal regions of negative strands are considered together, because mutations which affect one will necessarily also affect the other, and complementary secondary structures can sometimes be formed for both termini.

The ability of the 5′-terminal 90 nt of poliovirus RNA to fold into a cloverleaf structure, its ability to bind $3CD^{pro}$ in the presence of 3AB or an N-terminal fragment of EF-1α, its requirement for RNA replication, and a model for its role in initiating positive-strand synthesis were discussed in Sections II,E,4 and III. Both the 5′-positive and 3′-negative termini can form cloverleaf structures, but mutagenesis showed that only the positive-strand cloverleaf is functional (Andino *et al.*, 1990). The 5′ 44 nt of RNAs B and M of cowpea mosaic virus, a plant picorna-like virus, have a high degree of sequence similarity and are interchangeable (van Bokhoven *et al.*, 1993), but whether they are structurally and functionally analogous to the poliovirus 5′ cloverleaf structure is not known.

Marsh and Hall (1987) discovered sequences in the 5′ untranslated regions of brome mosaic virus (BMV) RNAs that resemble consensus sequences for the internal control regions (ICR1 and ICR2) of tRNA promoters. In the mature tRNA, ICR1 (box A), and ICR2 (box B) correspond to the D-loop and T-loop respectively (Geiduschek and Kassavetis, 1992). ICR-like sequences have also been found in other bromoviruses, cucumoviruses, tobamoviruses, tobraviruses, tymoviruses, and tobacco

necrosis satellite virus (Marsh *et al.,* 1989). Pogue and Hall (1992) proposed that the 5′-terminal region of BMV RNA 2 could be folded into a stem–loop structure with the ICR2-like motif in the loop and the ICR1-like motif comprising part of the stem. Similar structures were predicted for the 5′ termini of BMV RNAs 1 and 3, RNAs of cucumber mosaic virus (CMV), and cowpea chlorotic mottle virus (CCMV) (Pogue and Hall, 1992) and alfalfa mosaic virus (van der Vossen *et al.,* 1993). Mutational analysis established the importance of the proposed structure and the ICR-like motifs in the replication of BMV RNA 2 (Pogue and Hall, 1992; Pogue *et al.,* 1990, 1992). In particular, it was shown that the structure was important in the 5′-terminal region of the positive strand, but not in the 3′-terminal region of the negative strand. ICR2-like motifs were located within 27-nt repeats in the 5′ untranslated sequence of alfalfa mosaic virus RNA 3, which were shown to be important for the replication of this RNA (van der Vossen *et al.,* 1993). These results led to a model in which a host factor, possibly a transcription factor associated with RNA polymerase III, binds to the ICR-like region at the 5′ terminus of the positive strand of the double-stranded replicative form dsRNA and plays a role in the formation of an initiation complex for positive-stranded RNA synthesis (Pogue and Hall, 1992; Pogue *et al.,* 1994). Support for the model and involvement of ICR-like motifs comes from the finding that binding of some host proteins to the 5′-terminal sequence of the positive strand and the 3′-terminal sequence of the negative strand did not occur in RNA 2 mutants with substitutions in the ICR2-like motif and known to be debilitated in replication (Duggal and Hall, 1995).

It is unlikely that the ICR-like motifs are the only 5′ elements involved in positive-strand RNA replication in these plant alpha-like viruses, since longer 5′ regions (~90 nt) are required for efficient replication of RNA 3 of BMV (French and Ahlquist, 1987), CCMV (Pacha *et al.,* 1990), and CMV (Boccard and Baulcombe, 1993), and ICR-like motifs are not readily discernible in the 5′ region of CMV RNA 3. The 5′ untranslated sequence of tobacco mosaic virus lacks G residues and contains multiple CAA repeats. The sequence is highly conserved and the 5′-terminal 31 nt are almost identical in different strains. Deletion analysis of the 5′-terminal region of the L strain showed that large deletions (nucleotides 9–47 or 25–71) abolished replication, but of approximately 10-nt deletions across the whole region, only a deletion of nucleotides 2–8 abolished replication (Takamatsu *et al.,* 1991). The progeny of *in planta* replication of some TMV subgenomic replicons included a molecule with only 23 nt at the 5′ terminus (Raffo and Dawson, 1991). There is no evidence for the involvement of ICR-like

motifs in this region. The 312-nt 5′ untranslated region of beet necrotic yellow vein virus RNA 3 can be folded into a structure containing several stem–loop structures, which involves long-range base-pairing between different cis-active elements. Chemical and enzymatic probing, and mutational analysis, has provided evidence for the structure and its involvement in the promotion of positive-strand synthesis (Gilmer *et al.*, 1992b, 1993).

The structure of the 3′ end of the negative strand may also be important for replication. Double-stranded RNAs isolated from cells infected by CMV and a satellite (Collmer and Kaper, 1985), potato virus X (Dolja *et al.*, 1987), and Semliki Forest and Sindbis viruses (Wengler *et al.*, 1979, 1982) contained an unpaired G residue at the 3′ terminus of the negative strand. Wu and Kaper (1994) showed that the negative strand of a CMV satellite would only act as a template for an isolated CMV RdRp if it contained this additional nontemplated G. Whether this also applies to replication of the genomic RNAs is not known. It is noteworthy, however, that long 5′-terminal extensions generally render RNA transcripts produced from cDNA clones noninfectious (Boyer and Haenni, 1994). Since a 5′-terminal positive-strand extension would be copied to produce a 3′-terminal negative-strand extension, it may be the unpaired base (rather than the length or sequence of the extension) that is important. This could be required for unwinding an RF structure by helicases that require a 3′ extension (see Section II,B) and possible recognition by the replicase after unwinding.

In the animal alphaviruses, there is a conserved stem–loop structure at the 5′ terminus which is important for replication and this has been proposed to be functional as its complement at the 3′ end of the negative strand (reviewed by Strauss and Strauss, 1994). This sequence is found in some Sindbis virus DI RNAs. However, in some Sindbis virus DI RNAs, the 5′ terminus consists of nt 10 to 75 of a cellular $tRNA^{Asp}$ or a sequence derived from the 5′ end of the subgenomic RNA, suggesting that a structure at the 5′ terminus, rather than a linear sequence, is important for replication. The 5′ terminus of rubella virus RNA can also be folded into a stem–loop structure, with the potential to form a pseudoknot (Frey, 1994).

Analysis of 5′-terminal structures of viruses in other supergroups also indicates the ability to fold into secondary structures and a requirement for RNA replication. Examples include: carmo-like virus supergroup, cucumber necrosis virus (Finnen and Rochon, 1993; Chang *et al.*, 1995), cymbidium ringspot virus (Havelda *et al.*, 1995), tomato bushy stunt virus (Knorr *et al.*, 1991; White and Morris, 1994; Chang *et al.*, 1995); corona-like virus supergroup, mouse hepatitis

virus virus (Kim *et al.*, 1993). The smallest 5′ cis-acting element required for RNA replication appears to be that of RNA 2 of flockhouse virus, which consisted of between 3 and 14 nt, and probably less than 6 nt (Ball and Li, 1993; Ball, 1994).

C. Internal Sequences

Internal cis-acting elements, in either intercistronic or coding regions, which are required for efficient RNA replication have been identified for a number of virus RNAs. In some cases, such sequences may be required to maintain an optimal RNA structure for binding of the replicase complex to promoters at the termini of the positive- or negative-stranded RNAs, or to promote processivity of the replicase during RNA synthesis. In other cases, it is possible that the replicase could bind to internal sequences for a particular purpose, e.g., translational repression, or for an obligatory step in the assembly or modification of RNA complexes. Phage Qβ replicase, lacking the additional host factor, HF1 (see Section III), binds strongly to internal sites on the Qβ RNA positive strand, but weakly if at all to the 3′ terminus. The purpose of this internal binding, which requires the ribosomal S1 protein component of the replicase, is probably to prevent binding of ribosomes upstream of the coat protein cistron, because the replicase is unable to dislodge ribosomes travelling along the RNA in the opposite direction (van Duin, 1988). Addition of HF1, which binds to both an internal and the 3′-terminal region of the RNA, serves to bring the 3′-terminal region in contact with the replicase to allow initiation of negative-strand synthesis (Barrera *et al.*, 1993). Qβ replicase does not bind internally on the RNA negative strand and does not require HF1 for initiation of positive-strand synthesis (van Duin, 1988; Barrera *et al.*, 1993).

A sequence of about 150 nt in the 5′ region of the 244-nt intercistronic region of brome mosaic virus (BMV) RNA 3, which separates the movement protein (3a) and coat protein ORFs, is required in cis for efficient replication of this RNA (French and Ahlquist, 1987). Removal of this region decreased RNA 3 replication to less than 1% of wild-type levels. This decrease was not due to effects on movement protein or coat protein gene expression, because frameshift mutations which abolished synthesis of these proteins had little effect on replication. This intercistronic region contains ICR-like motifs; removal of the motif which best fit the tRNA gene ICR2 consensus sequence reduced RNA replication to 15% of wild-type levels (Pogue *et al.*, 1992), indicating that this sequence contributes to the function of the intercistronic region in RNA 3 replication. In yeast, synthesis of BMV negative

strand RNA 3 required intercistronic sequences, as well as the 3′-terminal sequence (Quadt *et al.*, 1995). *In vitro* synthesis of BMV negative-strand RNA 3 by an isolated RdRp extract did not require the intercistronic sequences; this could indicate that these sequences are required specifically for the assembly of the replication complex or that current RdRp extracts lack essential components and contain only a basal activity. Sequences of the intercistronic regions of RNA 3 of cucumber mosaic virus, which contain an ICR2-like motif (Boccard and Baulcombe, 1993), and alfalfa mosaic virus (van der Vossen *et al.*, 1995), have also been shown to be important for RNA replication. In contrast, replication of cowpea chlorotic mottle virus RNA 3 did not require its intercistronic region and was less susceptible than BMV RNA 3 to large deletions in the RNA (Pacha *et al.*, 1990).

It has also been suggested that some of the BMV 2a protein coding region may be required in cis for replication of RNA 2, since RNA 2 mutants lacking the C-terminal region of the 2a protein could be replicated in trans in the presence of wild-type RNA 2, but mutants with larger deletions, encompassing the central and N-terminal region of 2a, could not (Pogue *et al.*, 1990; Marsh *et al.*, 1991b). In contrast, there do not appear to be any essential cis-acting replication elements in the coding regions of tobacco mosaic virus RNA. Mutants with deletions across the 126-kDa or 183-kDa replication protein ORFs could be replicated in protoplasts in trans by the wild-type proteins; deletion of the C-terminal part of the 183-kDa ORF increased replication (Ogawa *et al.*, 1992). TMV-derived replicons with most of the 126-kDa/183-kDa ORFs removed also replicated well and spread in plants in the presence of a helper virus (Raffo and Dawson, 1991). Furthermore, TMV mutants lacking the movement protein and coat protein ORFs replicated in protoplasts as well as wild-type; a decline in the accumulation of the mutant positive strand relative to that of the wild-type was only observed late in the infection and was ascribed to degradation of the RNA in the absence of the coat protein (Ishikawa *et al.*, 1991b).

The animal alphaviruses have a conserved 51-nt sequence near the beginning of the P123/4 ORF, which can be folded into a structure with two stem–loops. Of 21 silent mutations in this region, 19 resulted in a decrease of virus growth of two to four orders of magnitude (Niesters and Strauss, 1990). Some of the deleterious mutations could potentially disrupt the secondary structure, whereas others would not be expected to. The 51-nt sequence was also found to be necessary for efficient accumulation of DI RNAs, although deletion of the element did not reduce DI RNA replication as much as expected when compared to the effect of mutations in this region in the genomic RNA (Levis *et*

al., 1986; Schlesinger *et al.*, 1987). Internal regions of the DI RNAs derive from other parts of the alphavirus genome and are presumably selected for efficiency of replication and packaging of the RNA (reviewed by Strauss and Strauss, 1994).

A 3′-proximal element of the capsid-coding region (P1) of human rhinovirus 14 RNA was shown to be required for efficient RNA replication (McKnight and Lemon, 1996). Although the P1 region of another picornavirus, poliovirus, was not required for replication (see Section II,F), McKnight and Lemon (1996) suggested that an element within the P2–P3 region of the poliovirus genome, which must undergo translation in cis for RNA replication to proceed (Novak and Kirkegaard, 1994), may be analogous to the human rhinovirus P1 replication element.

Sequences derived from the coding region of polymerase genes have been found to be essential cis-acting elements for the replication of a number of DI RNAs derived from viruses in the carmo-like virus supergroup, such as cucumber necrosis and tomato bushy stunt viruses (Chang *et al.*, 1995) and cymbidium ringspot virus (Havelda *et al.*, 1995). In the corona-like virus supergroup, a 135-nt internal sequence was required for DI RNA replication of one strain of mouse hepatitis virus, but not for another (Kim *et al.*, 1993; Lin and Lai, 1993; van der Most *et al.*, 1994). Although genomic sequences, retained by DI RNAs and required for DI RNA replication, provide useful information, cis-acting elements involved in genomic and DI RNA replication are not necessarily identical, since the selection pressures operating on the two types of RNA, as well as their structures, are different.

D. Sequences Required for Subgenomic RNA Synthesis

1. Alpha-like Virus Supergroup

Subgenomic RNA synthesis in this supergroup occurs by internal initiation on negative-strand templates, first shown *in vitro* by Miller *et al.* (1985) for brome mosaic virus (BMV) using an isolated RdRp and *in vivo* by Gargouri *et al.* (1989) for turnip yellow mosaic virus. Following convention, sequences of subgenomic promoters will be discussed on the positive strand, although it is the negative strand that acts as the template. The BMV subgenomic promoter, for synthesis of the coat protein subgenomic mRNA, has been defined *in vitro* (Marsh *et al.*, 1988) and *in vivo* (French and Ahlquist, 1988). It is contained within the 250-nt intergenic region between the movement protein and coat protein genes, and extends between 74 and 95 nt upstream and 16 nt downstream of the transcriptional initiation site. It consists of three

functional domains. The first of these is the core promoter, which includes the initiation site, 20 nt upstream and about 15 nt downstream; this is sufficient for a low basal level of transcription and determines correct initiation. Immediately upstream of the core promoter is a poly(A) tract, which acts as an activator and possible spacer; similar poly(A) tracts are present in equivalent locations in other bromoviruses. Upstream of the poly(A) is a further enhancer region which contains imperfect direct repeats of sequences in the core promoter. The subgenomic promoter is distinct from, but may overlap at its 5′ extremity, the cis-acting intergenic sequence required for efficient replication of RNA 3 (see Section V,C); mutations in the ICR2-like (box B) motif of the latter reduced RNA 3 synthesis without affecting subgenomic RNA synthesis (Smirnyagina *et al.*, 1994).

The subgenomic promoter of cucumber mosaic virus (CMV) is also located in the 286-nt intergenic region in RNA3 and is contained within a sequence extending about 70 nt upstream and 20 nt downstream of the transcriptional initiation site (Boccard and Baulcombe, 1993). Unlike the BMV subgenomic promoter, this sequence contains an ICR2-like motif. The intergenic region in alfalfa mosaic virus RNA 3 only extends 13 nt upstream of the transcriptional initiation site for the coat protein subgenomic mRNA and the subgenomic promoter extends into the C-terminal end of the movement protein ORF (van der Kuyl *et al.*, 1990, 1991; van der Vossen *et al.*, 1995). The basal AlMV subgenomic promoter was located from −26 to +1, where +1 is the transcriptional start site. The basal level of transcription was increased more than tenfold by extending the upstream sequence to −136 and the downstream sequence to +12. The upstream sequence enhancer sequence was mapped to −136/−94; the downstream element is a U-rich sequence with high homology to the 5′-terminal sequences of AlMV RNAs 1 and 2 (van der Vossen *et al.*, 1995).

The transcriptional initiation site for the 26S subgenomic RNA of the animal alphaviruses lies in the last codon of the P123/4 polyprotein. The basal promoter lies in a "junction" sequence from −19 to +5, which is highly conserved in different alphaviruses (Levis *et al.*, 1990). Full promoter activity (about fivefold greater than the basal promoter) is contained within the −98 to +14 sequence (Raju and Huang, 1991). Sequence comparisons between the animal alphavirus and plant alpha-like virus subgenomic promoters identified a number of conserved motifs (Marsh *et al.*, 1988; French and Ahlquist, 1988; van der Vossen *et al.*, 1995). Mutation of some of these in the AlMV subgenomic promoter led to a decrease in activity (van der Vossen *et al.*, 1995). The poly(A) tract found in bromovirus subgenomic promoters

was not found in the animal alphavirus subgenomic promoters or those of the other plant alpha-like viruses. Mutants of BMV RNA 3 lacking the poly(A) tract synthesize little coat protein, but second-site mutations in the intergenic region can suppress the transcriptional defect (Smirnyagina *et al.,* 1994). One of these mutations was a duplication of the sequence UAUUAUUA immediately 5′ to the deleted poly(A); this sequence had previously been shown to be an important enhancer element in the wild-type subgenomic promoter (Marsh *et al.,* 1988). Hence the BMV subgenomic promoter, like those of other alpha-like viruses, can function in the absence of a poly(A) tract.

Positional effects on the activities of subgenomic promoters have also been noted. For BMV (French and Ahlquist, 1988) and CMV (Boccard and Baulcombe, 1993), the promoter closest to the 3′ end of the positive strand was the most active, whereas with AlMV (van der Vossen *et al.,* 1995) and Sindbis virus (Raju and Huang, 1991), the promoter closest to the 5′ end of the positive strand was the most active. With tobacco mosaic virus (TMV), the level of expression of genes closest to the 3′ end of the positive strand was the highest, but this has been shown to be due to translational control, there being little difference in the levels of subgenomic RNAs produced with promoters at different positions (Culver *et al.,* 1993). Wild-type TMV RNA has two subgenomic promoters, one for the movement protein which is expressed early and another for the coat protein which is expressed late. Expression of the movement protein under the control of the coat protein subgenomic promoter led to late expression, suggesting that the promoter sequence (rather than its position) may control the timing of expression (Lehto *et al.,* 1990).

The promoter for subgenomic RNA synthesis on beet necrotic yellow vein virus RNA 3 differs from those described above in that most of the promoter was located downstream (between + 100 and + 208) of the transcriptional initiation site and only extended to −16 upstream (Balmori *et al.,* 1993).

2. *Carmo-like and Sobemo-like Virus Supergroups*

Although there are fewer studies, it is probable that subgenomic RNAs of viruses in these supergoups are also synthesized by internal initiation on a negative-strand template. Cucumber necrosis virus (tombusvirus) produces two 3′ coterminal subgenomic RNAs, one of 2.1 kb, which directs the synthesis of the coat protein and one of 0.9 kb which directs the synthesis of two proteins, p21 and p20. The core subgenomic promoter for the 0.9 kb RNA was located to a region extending from between −10 and −20 to +6 (Johnston and Rochon, 1995).

The sequence from –11 to +3 was highly conserved in the equivalent regions of different tombusviruses. However, there was very limited similarity between this region, the 5′-terminal sequence of the genomic RNA, and the sequence surrounding the 5′ end of the 2.1 kb subgenomic RNA. This suggests that the two subgenomic promoters may be controlled independently, possibly in a temporal fashion as suggested above for the two tobacco mosaic virus subgenomic promoters (see Section V,D,1).

For several other viruses, there is considerable sequence similarity between the 5′ termini of the genomic and subgenomic RNAs, suggesting that these sequences may be part of the promoters for genomic and subgenomic RNA synthesis. Thirteen of the first fourteen nucleotides of the 5′ termini of the red clover necrotic mosaic virus (dianthovirus) genomic RNA 1 and of the positive strand of the coat protein subgenomic dsRNA are identical (Zavriev *et al.*, 1996). It was hypothesized that the subgenomic promoter consisted of a stem–loop structure composed of nucleotides –53 to +27. Sequence similarity in the 5′-terminal sequences of genomic and subgenomic RNAs has also been found for several luteoviruses (reviewed in Miller *et al.*, 1995). There is also similarity between the 5′-terminal sequences of two subgenomic RNAs of tobacco necrosis virus (necrovirus) (Meulewater *et al.*, 1992); this did not extend upstream of the transcriptional initiation site, suggesting that the two promoters may be controlled independently, consistent with temporal differences in their expression observed *in vivo*. In contrast, the 5′-terminal sequences of two carnation mottle virus (carmovirus) subgenomic RNAs had little similarity, but sequences upstream in the genomic RNA were highly conserved (Carrington and Morris, 1986). In the case of maize chlorotic mottle virus (machlomovirus), a sequence element partially homologous to the 5′-terminal genomic RNA sequence was found upstream of the start of the subgenomic RNA start site (Lommel *et al.*, 1991).

3. *Coronaviruses*

Coronaviruses synthesize a nested set of five to seven 3′-coterminal subgenomic mRNAs for translation of the internal ORFs (reviewed by Lai, 1990). The mechanism of subgenomic RNA synthesis is very different from that of the animal and plant alpha-like viruses, and the plant carmo-like and sobemo-like viruses, discussed above. Every subgenomic RNA has an identical 5′-terminal leader sequence which varies in length (60–90 nt), depending on the coronavirus. The leader sequence is only found at the 5′ end of the genomic RNA, which implies that subgenomic mRNAs are formed by fusion of two noncontiguous

elements. Ultraviolet irradiation inactivation experiments established that the leader and body of the mRNAs were not joined together by a splicing mechanism (Stern and Sefton, 1982). Upstream of each gene are conserved sequence elements, referred to as IS elements. There is sequence homology between sequence elements (present in varying numbers of copies) in the 3′ region of the leader sequence and the IS elements. For mouse hepatitis virus (MHV), every IS contains the sequence (AAUCUAAAC) or a closely related sequence. These IS elements are believed to function as promoters for subgenomic mRNA synthesis (van der Most *et al.*, 1994). For MHV, the smallest mRNAs are generally synthesized in larger amounts that the larger ones, although this is not the case for all coronaviruses (Hiscox *et al.*, 1995).

Several models have been proposed to account for the preceding observations in the synthesis of coronavirus mRNAs (Lai, 1990). One group of models proposes that the virus genomic RNA acts as a template for the synthesis of a full-length negative strand, which in turn acts as a template for the synthesis of both progeny-genomic RNA and the subgenomic RNAs. In the leader-primed hypothesis, the 5′ leader is first synthesized and this then acts as a primer for initiation of subgenomic RNA synthesis at one of the IS elements, either on the same RNA (by looping out of the intervening RNA), or by detachment of the polymerase with its leader and reattachment to another RNA template. The discovery of full-length negative strands and the isolation of replicative intermediates containing nascent subgenomic RNAs with the leader attached was consistent with this model (Baric *et al.*, 1983b). Evidence that the polymerase and leader can detach from one template and attach to another comes from observations that the leader sequences undergo rapid exchanges between two RNA molecules, for example between mRNAs of two MHV strains or between an MHV RNA and a DI RNA (Makino *et al.*, 1986; Makino and Lai, 1989). Also an exogenously added leader can be incorporated into mRNAs in an *in vitro* transcription system (Baker and Lai, 1990) or *in vivo* from a helper virus when supplied as a negative strand containing an IS sequence (Hiscox *et al.*, 1995). Reattachment to new templates could be aided by sequence homology between the 3′ leader region and IS elements. It may also involve host proteins which bind to the IS regions (Zhang and Lai, 1995). Mutations in IS regions which reduced subgenomic RNA synthesis also reduced host protein binding. Further studies showed that the leader sequence can act both in trans and in cis (Zhang *et al.*, 1994) and that leader switching is facilitated by a 9-nt sequence (UUUAUAAAC) located immediately downstream of the leader sequence, possibly at the end of a stem–loop structure (Zhiang

and Lai, 1996). The leader-primed model has some features in common with the model proposed by Carpenter *et al.* (1995) for recombination in turnip-crinkle-virus-associated RNAs, although the former is clearly a more directed process.

The discovery of subgenomic-length negative strands and subgenomic replicative intermediates (Sethna *et al.*, 1989; Sawicki and Sawicki, 1990; Schaad and Baric, 1994) opened up additional possibilities for the synthesis of coronavirus subgenomic RNAs. The suggestion that subgenomic RNAs might be formed by the leader-primed mechanism and then replicate independently seems unlikely in view of the inability of transfected positive-stranded subgenomic RNAs to be replicated in the presence of helper virus (Brian *et al.*, 1994). Sawicki and Sawicki (1990, 1995) have suggested a model in which discontinuous synthesis of subgenomic RNAs occurs during negative-strand synthesis. The polymerase is proposed to pause at an IS element in the positive-strand template and then either continue synthesis or detach with the nascent strand and reattach on the same or different RNA molecules to copy the leader sequence. It was suggested that having copied an IS element, the polymerase might retract in a way analogous to some DNA-dependent RNA polymerases (Kassavetis and Geiduschek, 1993) and suggested for some types of recombination between brome mosaic virus RNAs (Bujarski *et al.*, 1994), and remain associated with the template. The IS sequences at the 3′ end of the nascent strand could then align with complementary sequences on the positive-strand template in the 3′ region of the leader sequence. The newly synthesized subgenomic negative strand could then act as a preferred template for synthesis of positive-strand subgenomic RNA by the same polymerase complex. It has been shown that MHV downstream IS elements have a negative impact on transcription from upstream IS elements, whereas upstream IS elements have little effect on transcription from downstream IS elements (van Marle *et al.*, 1995), consistent with the greater synthesis of the shorter mRNAs. This is nicely explained by the model of Sawicki and Sawicki (1995) because the polymerase would have a choice of detachment or continued synthesis at each IS element encountered and therefore fewer polymerase molecules would reach the IS elements upstream in the positive-strand template. The various models have not yet been unequivocally resolved because other work has suggested that the subgenomic negative strands, apparently present in the cell in double-stranded RNA form, may be a dead-end product (Lin *et al.*, 1994). Double-stranded RNA subgenomic RNAs of the alpha-like viruses are generally considered to be dead-end products, because they lack the upstream promoter sequences. However, further

studies are needed to resolve this issue for coronavirus subgenomic RNAs.

An additional control mechanism must be invoked to account for the predominance of full-length genomic RNA molecules late in infection. It has been suggested by Lai (1990) that the nucleocapsid (N) protein might perform this function, in a manner analogous to the role of the vesicular stomatitis virus nucleocapsid protein in controlling the switch from mRNA synthesis to production of genome-length positive strands, which can then be used as templates for the synthesis of progeny (negative strand) virus RNA molecules (Blumberg *et al.*, 1981). The coronavirus N protein is known to contain an RNA-binding domain and to form high affinity complexes with the 3′ end of the leader sequence (Baric *et al.*, 1988; Nelson and Stohlman, 1993).

VI. Control of Asymmetric Positive- and Negative-Strand Synthesis

During the replication of positive-stranded RNA viruses, a large excess of positive over negative strands is produced. Estimates range from 10:1 (flaviviruses; Chambers *et al.*, 1990), 50–100:1 (coronaviruses; Lai, 1990), 100:1 (brome mosaic virus, French and Ahlquist, 1987), 1000:1 (alfalfa mosaic virus; Nassuth and Bol, 1983). This could be due to down-regulation of negative-strand synthesis and/or up-regulation of positive-strand synthesis. In the alpha-like virus superfamily, it has been shown for Sindbis virus and other alphaviruses (Strauss and Strauss, 1994), alfalfa mosaic virus (van der Vossen *et al.*, 1994), and tobacco mosaic virus (Ishikawa *et al.*, 1991b) that negative-strand synthesis is switched off or greatly reduced a few hours after infection, whereas positive-strand synthesis continues throughout the replication cycle. There are many examples of mutations in virus replication proteins which affect positive- and negative-strand synthesis differently, e.g., in alphaviruses (reviewed by Strauss and Strauss, 1994), bromoviruses (reviewed by Ahlquist, 1992, Duggal *et al.*, 1994), and tobamoviruses (reviewed by Dawson and Lehto, 1990). Hence replication complexes which synthesize negative and positive strands (and subgenomic RNAs) may be different or one may be modified to form the other. It is noteworthy that the phage Qβ replicase complexes for synthesis of positive and negative strands differ. The enzyme for negative-strand synthesis contains the host factor (HF1 protein), whereas that for positive-strand synthesis does not. When HF1 protein is in excess, equal amounts of positive and negative strands are produced. When HF1 is in limiting amount, as in the cell, negative-strand synthesis is

limited and positive-strand synthesis predominates (Blumenthal and Carmichael, 1979).

In the animal alphaviruses, the initial unstable replication complex that synthesizes the negative strand is converted by a process that includes proteolytic cleavage into a stable complex that synthesizes the positive strands (Strauss and Strauss, 1994; Sawicki and Sawicki, 1994; see Section II,E,1). It appears that replication complexes are only formed on the positive-strand templates and the number of replication complexes formed depends on the number of negative strands synthesized. Once negative-strand synthesis has ceased, positive-strand synthesis carries on using the stable replication complex already formed, leading to the observed asymmetry in positive- and negative-strand accumulation.

The model of Pogue and Hall (1992) for the replication of brome mosaic virus also suggests that the replicase complex, having assembled on a positive-strand template and synthesized a negative strand, is then modified to use the negative strand as a template. This does not account for the strand asymmetry, because infection of protoplasts with RNA 1 and 2, which encode the 1a and 2a replication proteins, produced a 1:1 ratio of positive to negative strands. However infection with RNAs 1, 2 and 3 produced a 100:1 ratio of positive to negative strands (Marsh *et al.*, 1991a). RNA 3 therefore controls the strand asymmetry. Mutational analysis showed that the coat protein, which is synthesized from a subgenomic RNA derived from RNA 3, while contributing to the strand asymmetry, was not a major determinant of it. A deletion of the subgenomic core promoter and the first 5′ 43 nt of the subgenomic RNA synthesis reduced the strand asymmetry to 1.8:1. It has been suggested that sequences in the intergenic region of RNA 3 may form a platform on which a positive-strand-synthesizing replicase could be assembled (Duggal *et al.*, 1994). It is noteworthy that deletion of the core subgenomic promoter in RNA 3 resulted in production of an RdRp in yeast with increased ability to synthesize negative strands *in vitro* (Quadt *et al.*, 1995), perhaps again indicating differences between replicase complexes able to synthesize positive and negative strands. Differences in replicase complexes for positive- and negative-strand synthesis have also been invoked for alfalfa mosaic virus, but for this virus the coat protein was shown to be the main activator of positive-strand synthesis (see Sections II,E,1 and II,F). Nevertheless, the shut-off of negative-strand RNA 3 synthesis in P12 transgenic plants appeared to be independent of the coat protein (van der Vossen *et al.*, 1994), a further indication that negative-strand and positive-strand synthesis for this virus may be regulated separately.

Early shut-off of negative-strand synthesis does not occur for all positive-stranded RNA viruses. Negative-stranded RNA synthesis of the flavivirus dengue virus continued throughout the entire replication cycle (Cleaves *et al.*, 1981). Although negative strand-synthesis of coronaviruses peaked at 5–6 hours after infection, some synthesis continued until late in infection (Sawicki and Sawicki, 1986). Viruses in the picorna-like virus supergroup pose a different problem because both positive and negative strands have VPg 5′ termini (see Section II,D). Hence one polyprotein molecule has to be synthesized for every RNA molecule synthesized (unless VPg released from RNAs destined to be translated can be recycled, which seems unlikely for poliovirus if the primer is 3ABpU; see Sections II,D and II,E,4). However poliovirus RNA synthesis is not completely linked to translation throughout the whole replication cycle, because synthesis of both positive and negative strands can continue for some time after addition of inhibitors of protein synthesis (Novak and Kirkegaard, 1994).

VII. Concluding Remarks

It is clear that there are some similarities in RNA replication for all eukaryotic positive-stranded RNA viruses, i.e., the mechanism of polymerization of the nucleotides is probably similar for all. It is noteworthy that all appear to utilize host membranes as a site of replication. Membranes appear to function not just as a way of compartmentalizing virus RNA replication, but also appear to have a central role in the organization and functioning of the replication complex and further studies in this area are needed. Within virus supergroups, similarities are evident between animal and plant viruses, e.g., in the nature and arrangements of replication genes and in sequence similarities of functional domains. However, it is also clear that there has been considerable divergence, even within supergroups. For example, the animal alphaviruses have evolved to encode proteinases which play a central controlling function in the replication cycle, whereas this is not common in the plant alpha-like viruses and even when it occurs, as in the tymoviruses, the strategies that have evolved appear to be significantly different. Some of the divergence could be host-dependent and the increasing interest in the role of host proteins in replication should be fruitful in revealing how different systems have evolved. Even within the plant alpha-like viruses, there are significant differences in regulation mechanisms. An outstanding question which needs to be solved is, for those viruses which appear to show some form of cis-

preferential replication, such as poliovirus (Wimmer *et al.*, 1993; Novak and Kirkegaard, 1994) and turnip yellow mosaic virus (Weiland and Dreher, 1993), whether or not the replication complex assembles on the RNA from which it has been translated. Other possibilities have been discussed by Wimmer *et al.* (1993) and Novak and Kirkegaard (1994). Finally, there are virus supergroups which appear to have no close relatives between animals and plants, such as the animal coronavirus-like supergroup and the plant carmo-like supergroup. Nevertheless, our knowledge of positive-stranded RNA virus replication is still in its infancy and future research may reveal unsuspected similarities. More comprehensive comparisons must await further knowledge of the assembly and structures of replication complexes and how they are modified to utilize and initiate RNA synthesis at different promoters on positive- and negative-strand RNA templates.

Acknowledgments

I am grateful to researchers who sent me reprints and preprints of papers prior to publication. Work on RNA replication in the author's laboratory was funded by the U. K. Biotechnology and Biological Sciences Research Council.

References

Agranovsky, A. A., Koonin, E. V., Boyko, V. P., Maiss, E., Frotschi, R., Lunini, N. A., and Atabekov, J. G. (1994). *Virology* **198**, 311–324.

Ahlquist, P. (1992). *Curr. Opin. Genet. Devel.* **2**, 71–76.

Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. M., Haseloff, J., and Zimmern, D. (1985). *J. Virol.* **53**, 536–542.

Ahola, T., and Kääriänen, L. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 507–511.

Aldabe, R., and Carrasco, L. (1995). *Biochem. Biophys. Res. Comm.* **206**, 64–76.

Andino, R., Rieckhof, G. E., and Baltimore, D. (1990). *Cell* **63**, 369–380.

Andino, R., Rieckhof, G. E., Achacoso, P. L., and Baltimore, D. (1993). *EMBO J.* **12**, 3587–3598.

Andrews, N. C., and Baltimore, D. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 221–225.

Argos, P. (1988). *Nucleic Acids Res.* **16**, 9909–9916.

Atreya, C. D., Singh, N. K., and Nakhasi, H. L. (1995). *J. Virol.* **69**, 3848–3851.

Axelrod, V. D., Brown, E., Priano, C., and Mills, D. R. (1991). *Virology* **184**, 595–608.

Baer, M. L., Houser, F., Loesch-Fries, L. S., and Gehrke, L. (1994). *EMBO J.* **13**, 727–735.

Baker, S. C., and Lai, M. M. C. (1990). *EMBO J.* **9**, 4173–4179.

Ball, L. A. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12443–121447.

Ball, L. A. (1995). *J. Virol.* **69**, 720–727.

Ball, L. A., and Li, Y. (1993). *J. Virol.* **67**, 3544–3551.

Balmori, E., Gilmer, D., Richards, K., Guilley, H., and Jonard, G. (1993). *Biochimie* **75**, 517–521.

Barco, A., and Carrasco, L. (1995). *EMBO J.* **14**, 3349–3364.
Baric, R. S., Carlin, L. J., and Johnston, R. E. (1983a). *J. Virol.* **48**, 200–205.
Baric, R. S., Stohlman, S. A., and Lai, M. M. C. (1983b). *J. Virol.* **48**, 633–640.
Baric, R. S., Nelson, G. W., Fleming, J. O., Deans, R. J., Keck, J. G., Casteel, N., and Stohlman, S. A. (1988). *J. Virol.* **62**, 4280–4287.
Barrera, I., Schuppli, D., Sogo, J. M., and Weber, H. (1993). *J. Mol. Biol.* **232**, 512–521.
Bartholomeusz, A. I., and Wright, P. J. (1993). *Arch. Virol.* **128**, 111–121.
Barton, D. J., Sawicki, S. G., and Sawicki, D. L. (1991). *J. Virol.* **65**, 1496–1506.
Barton, D. J., Black, P. E., and Flanegan, J. B. (1995). *J. Virol.* **69**, 5516–5527.
Barton, D. J., Morasco, B. J., Eisner-Smerage, L., Collins, P. S., Diamond, S. E., Hewlett, J. J., Merchant, M. A., O'Donnell, B. J., and Flanegan, J. B. (1996). *Virology* **217**, 459–469.
Bates, H. J., Farjah, M., Osman, T. A. M., and Buck, K. W. (1995). *J. Gen. Virol.* **76**, 1483–1491.
Baulcombe, D. (1994). *Trends Microbiol.* **2**, 60–62.
Bayliss, C. D., and Smith, G. L. (1996). *J. Virol.* **70**, 794–800.
Beck, D. L., Guiford, P. J., Voot, D. M., Andersen, M. T., and Forster, R. L. (1991). *Virology* **183**, 695–702.
Behrens, S.-E., Tomei, L., and De Francesco, R. (1996). *EMBO J.* **15**, 12–22.
Bienz, K., Egger, D., and Pasamontes, L. (1987). *Virology* **160**, 220–226.
Bienz, K., Egger, D., Troxler, M., and Pasamontes, L. (1990). *J. Virol.* **64**, 1156–1163.
Bienz, K., Egger, D., Pfister, T., and Troxler, M. (1992). *J. Virol.* **66**, 2740–2747.
Bienz, K., Egger, D., and Pfister, T. (1994). *Arch. Virol.* **S9**, 147–157.
Blackburn. E. (1993). *In* "The RNA World" (R. F. Gesteland and J. F. Atkins, eds.), pp. 557–576, Cold Spring Harbor Laboratory Press, Plainview, NY.
Blackwell, J. L., and Brinton, M. A. (1995). *J. Virol.* **69**, 5650–5658.
Blum, H., Gross, H. J., and Beier, H. (1989). *Virology* **169**, 51–61.
Blumberg, B. M., Leppert, M., and Kolakofsky, D. (1981). *Cell* **23**, 837–845.
Blumenthal, T., and Carmichael, G. G. (1979). *Annu. Rev. Biochem.* **48**, 525–548.
Boccard, F., and Baulcombe, D. (1993). *Virology* **193**, 563–578.
Bork, P., and Koonin, E. V. (1993). *Nucleic Acids Res.* **21**, 751–752.
Boyer, J. C., and Haenni, A.-L. (1994). *Virology* **198**, 415–426.
Bramhill, D., and Kornberg, A. (1988). *Cell* **52**, 743–748.
Bransom, K. L., and Dreher, T. W. (1994). *Virology* **184**, 351–358.
Bransom, K. L., Weiland, J. J., and Dreher, T. W. (1991). *Virology* **184**, 351–358.
Brayton, P. R., Stohlman, S. A., and Lai, M. M. C. (1982). *J. Virol.* 42, 847–853.
Brayton, P. R., Lai, M. M. C., Patton, C. D., and Stohlman, S. A. (1984). *Virology* **133**, 197–201.
Brian, D. A., Chang, R-Y., Hofmann, M. A., and Sethna, P. B. (1994). *Arch. Virol.* **S9**, 173–180.
Bruenn, J. A. (1991). *Nucleic Acids Res.* **19**, 217–226.
Brun, J. B., and Brinton, M. A. (1988). *J. Gen. Virol.* **69**, 3121–3127.
Buck, K. W. (1979). *In* "Viruses and Plasmids in Fungi" (P. A. Lemke, ed.), pp. 94–151. Marcel Dekker, New York.
Bujarski, J. J., Nagy, P. D., and Flasinski, S. (1994). *Adv. Virus Res.* **43**, 275–302.
Burroughs, J. N., and Brown, F. (1978). *J. Gen. Virol.* **41**, 443–446.
Caliguiri, L. A., and Tamm, I. (1973). *In* "Selective inhibitors of Viral Functions" (W. Carter, ed.), pp. 257–294. CRC Press, Boca Raton.
Candresse, T. (1993). *In* "Encyclopaedia of Virology" (R. G. Webster and A. Granoff, eds.), vol. 1, pp. 242–248. Academic Press, London.

Candresse, T., Mouches, C., and Bové, J. M. (1986). *Virology* **152**, 322–330.
Carpenter, C. D., and Simon, A. E. (1996). *J. Virol.* **70**, 478–486
Carpenter, C. D., Oh, J-W., Zhang, C., and Simon, A. E. (1995). *J. Mol. Biol.* **245**, 608–622.
Carr, J. P., and Zaitlin, M. (1993). *Semin. Virol.* **4**, 339–347.
Carr, J. P., Galon, A., Palukaitis, P., and Zaitlin, M. (1994). *Virology* **199**, 439–447.
Carrington, J. C., and Morris, T. J. (1986). *Virology* **150**, 196–206.
Cavanagh, D., and MacNaughton, M. R. (1994). *In* "Principles and Practice of Clinical Virology" (A. J. Zuckerman, J. R. Pattison, and J. E. Banatvala, eds.), 3rd ed., pp. 325–335. John Wiley and Sons, New York.
Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990). *Annu. Rev. Microbiol.* **44**, 649–688.
Chambers, T. J., Nestorowicz, A., Amberg, S. E., and Rice, C. M. (1993). *J. Virol.* **67**, 6797–6807.
Chang, Y. C., Borja, M., Scholtof, H. B., Jackson, A. O., and Morris, T. J. (1995). *Virology* **210**, 41–53.
Cho, M. W., Richards, O. C., Dmitrieva, T. M., Agol, V., and Ehrenfeld, E. (1993). *J. Virol.* **67**, 3010–3018.
Cho, M. W., Teterina, N., Egger, D., Bienz, K., and Ehrenfeld, E. (1994). *Virology* **202**, 129–145.
Chu, P. W. G., and Westaway, E. G. (1985). *Virology* **140**, 68–79.
Chu, P. W. G., and Westaway, E. G. (1992). *Arch. Virol.* **125**, 177–191.
Cleaves, G. R., Ryan, T. E., and Schlesinger, R. W. (1981). *Virology* **111**, 73–83.
Coffin, R. S., and Coutts, R. H. A. (1993). *J. Gen. Virol.* **74**, 1475–1483.
Collett, M. S. (1992). *Comp. Immunol. Microbiol. Infect. Dis.* **15**, 145–154.
Collis, P. S., O'Donnell, B. J., Barton, D. J., Rogers, J. A., and Flanegan, J. B. (1992). *J. Virol.* **66**, 6480–6488.
Collmer, C. W., and Kaper, J. M. (1994). *Virology* **145**, 249–259.
Crawford, N. M., and Baltimore, D. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7452–7455.
Cui, T., and Porter, A. (1995). *Nucleic Acids Res.* **23**, 377–382.
Cui, T., Sankar, S., and Porter, A. (1993). *J. Biol. Chem.* **268**, 26093–26098.
Culver, J. N., Lehto, K., Close, S. M., Hilf, M. E., and Dawson, W. O. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2055–2059.
Dalmay, T., Rubino, L., Burgyan, J., Kollar, A., and Russo, M. (1993a). *Virology* **194**, 697–704.
Dalmay, T., Russo, M., and Burgyan, J. (1993b). *Virology* **192**, 551–555.
Dalmay, T., and Rubino, L. (1995). *Virology* **206**, 1092–1098.
Datta, U., and Dasgupta, A. (1994). *J. Virol.* **68**, 4468–4477.
David, C., Gargouri-Bouzid, R., and Haenni, A-L. (1992). *Prog. Nucleic Acid Res. Mol. Biol.* **42**, 157–227.
Dawson, W. D. O. (1978). *Intervirology* **9**, 304–309.
Dawson, W. D. O., and Lehto, K. M. (1990). *Adv. Virus Res.* **38**, 307–342.
Dedhar, S. P. (1994). *Trends Biol. Sci.* **19**, 269–271.
Dé, I., Sawicki, S. G., and Sawicki, D. L. (1996). *J. Virol.* **70**, 2706–2719.
de Graaf, M., and Jaspars, E. M. J. (1994). Annu. Rev. Phytopathol. 32, 311–335.
de Graaf, M., Coscoy, L., and Jaspars, E. M. J. (1993). *Virology* **194**, 878–881.
de Graaf, M., Houwing, C. J., Lukacs, N., and Jaspars, E. M. J. (1995a). *FEBS Lett.* **371**, 219–222.
de Graaf, M., Man In 'T Veld, M. R., and Jaspars, E. M. J. (1995b). *Virology* **208**, 583–589.
de Jong, W., and Ahlquist, P. (1995). *J. Virol.* **69**, 1485–1492.

de Varennes, A., Davies, J. W., Shaw, J. G., and Maule, A. J. (1985). *J. Gen. Virol.* **66**, 817–825.
Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). *Protein Eng.* **3**, 461–467.
Demler, S. A., Rucker, D. G., and de Zoeten, G. A. (1993). *J. Gen. Virol.* **74**, 1–14.
Demler, S. A., Borkhsenious, O. N., Rucker, D. G., and de Zoeten, G. A. (1994). *J. Gen. Virol.* **75**, 997–1007.
Diamond, S. E., and Kirkegaard, K. (1994). *J. Virol.* **68**, 863–876.
Dinant, S., Janda, M., Kroner, P. A., and Ahlquist, P. (1993). *J. Virol.* **67**, 7181–7189.
Dobos, P., and Roberts, T. E. (1983). *Can. J. Microbiol.* **29**, 377–384.
Doedens, J. R., and Kirkegaard, K. (1995). *EMBO J.* **14**, 894–907.
Dolja, V. V., and Carrington, J. C. (1992). *Semin. Virol.* **3**, 315–326.
Dolja, V. V., Grama, D. P., Morozov, S. Y., and Atabekov, J. G. (1987). *FEBS Lett.* **214**, 308–312.
Dolja, V. V., McBride, H. J., and Carrington, J. C. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10208–10212.
Dolja, V. V., Karasev, A. V., and Koonin, E. V. (1994). *Annu. Rev. Phytopathol.* **32**, 261–285.
Domingo, E., and Holland, J. J. (1994). *In* "The Evolutionary Biology of Viruses" (S. S. Morse, ed.), pp. 161–184. Raven Press, New York.
Donald, R. G. K., and Jackson, A. O. (1994). *In* "Encyclopaedia of Virology" (R. G. Webster and A. Granoff, eds.), Vol. 2, pp. 661–664. Academic Press, New York.
Dorssers, L., Van der Krol, S., Van der Meer, J., Van Kammen, A., and Zabel, P. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1951–1955.
Dougherty, W. G., and Semler, B. L. (1993). *Microbiol. Reviews* **57**, 781–822.
Dougherty, W. G., and Parks, T. D. (1995). *Curr. Opin. Cell Biol.* **7**, 399–405.
Dreher, T. W., and Hall, T. C. (1988). *J. Mol. Biol.* **201**, 31–40.
Dreher, T. W., Rao, A. L. N., and Hall, T. C. (1989). *J. Mol. Biol.* **206**, 425–438.
Duggal, R., and Hall, T. C. (1995). *Virology* **214**, 638–641.
Duggal, R., Rao, A. L. N., and Hall, T. C. (1992). *Virology* **187**, 261–270.
Duggal, R., Lahser, F. C., and Hall, T. C. (1994). *Annu. Rev. Phytopathol.* **32**, 287–309.
Dunigan, D. D., and Zaitlin, M. (1990). *J. Biol. Chem.* **265**, 7779–7786
Eagles, R. M., Balmori-Melian, E., Beck, D. L., and Gardner, R. C. (1994). *Eur. J. Biochem.* **224**, 677–684.
Eigen, M., and Briebricher, C. (1988). *In* "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, eds.), Vol. 3, pp. 211–245. CRC Press, Boca Raton.
Eggen, R., and Van Kammen, A. (1988). *In* "RNA Genetics" (P. Ahlquist, J. J. Holland, and E. Domingo, eds.), Vol. 1, pp. 49–69. CRC Press, Boca Raton.
Eggen, R., Kaan, A., Goldbach, R., and Van Kammen, A. (1988). *J. Gen. Virol.* **69**, 2711–2720.
Eggen, R., Verver, J., Wellink, J., Pleij, K., Van Kammen, A., and Goldbach, R. (1989a). *Virology* **173**, 456–464.
Eggen, R., Verver, J., Wellink, J., De Jong, A., Golbach, R., and Van Kammen, A. (1989b). *Virology* **173**, 447–455.
Ehrenfeld, E., and Gebhard, J. G. (1994). *Arch. Virol.* **S9**, 269–277.
Esau, K., and Cronshaw, J. (1967). *J. Cell Biol.* **33**, 665–678.
Fernandez, A., Lain, S., and Garcia, J. A. (1995). *Nucleic Acids Res.* **23**, 1327–1332.
Finnen, R. L., and Rochon, D. M. (1993). *J. Gen. Virol.* **74**, 1715–1720.
Florentz, C., and Gierge, R. (1995). *In* "tRNA: Structure, Biosynthesis, and Function" (D. Soll and U. RajBhandary, eds.), pp. 141–163. American Society for Microbiology, Washington, DC.

Foster, G. D. (1992). *Res. Virol.* **143**, 103–112.
Fraenkel-Conrat, H. (1986). *Crit. Rev. Plant Sci.* **4**, 213–226.
French, R., and Ahlquist, P. (1987). *J. Virol.* **61**, 1457–1465.
French, R., and Ahlquist, P. (1988). *J. Virol.* **62**, 2411–2420.
Frey, T. K. (1994). *Adv. Virus Res.* **44**, 69–160.
Froshauer, S., Kartenbeck, J., and Helenius, A. (1988). *J. Cell Biol.* **107**, 2075–2086.
Fujimura, T., and Wickner, R. B. (1989). *J. Biol. Chem.* **264**, 10872–10877.
Fuller-Pace, F. V. (1994). *Trends Cell Biol.* **4**, 271–274.
Fumagalli, S., Totty, N. F., Hsuan, J. J., and Courtneidge, S. A. (1994). *Nature* **368**, 867–871.
Furiuchi, Y., Muthakrishnan, S., Tomaz, J., and Shatkin, A. J. (1976). *J. Biol. Chem.* **251**, 5043–5053.
Furuya, T., and Lai, M. M. C. (1993). *J. Virol.* **67**, 7215–7222.
Gallie, D. R. (1991). *Genes Dev.* **5**, 2108–2116.
Gallie, D. R., and Kobayashi, M. (1994). *Gene* **142**, 159–165.
Gamez, R., and Leon P. (1988). *In* "The Plant Viruses" (R. Koenig, ed.), Vol. 3, pp. 213–233. Plenum Press, New York.
Gargouri, R., Joshi, R. L., Astier-Manifacier, S., and Haenni, A-L. (1989). *Virology* **171**, 386–393.
Gargouri-Bouzid, R., David, C., and Haenni, A. L. (1991). *FEBS Lett.* **294**, 56–58.
Garnier, M., Mamoun, R., and Bové, J. M. (1980). *Virology* **104**, 357–374.
Garnier, M., Candresse, T., and Bové, J. M. (1986). *Virology* **151**, 100–109.
Geiduschek, E. P., and Kassavetis, G. A. (1992). *In* "Transcriptional Regulation" (S. L. McKnight and K. R. Yamamoto, eds.), Vol. 1, pp. 247–280. Cold Spring Harbor Laboratory Press, Plainview, NY.
Giachetti, C., and Semler, B. L. (1991). *J. Virol.* **65**, 2647–2654.
Gibbs, M. J. (1995). *In* "Molecular Basis of Viral Evolution" (A. J. Gibbs, C. H. Calisher, and F. Garcia-Arenal, eds.), pp. 351–368. Cambridge University Press, Cambridge.
Gieseman-Cookmeyer, D., Kim, K-H., and Lommel, S. A. (1995). *In* "Pathogenesis and Host Specificity in Plant Diseases" (R. P. Singh, U. P. Singh, and K. Kohmoto, eds.), Vol. III, pp. 157–176. Elsevier, Oxford.
Gilmer, D., Bouzoubaa, S., Guilley, H., Richards, K., and Jonard, G. (1992a). *Virology* **189**, 40–47.
Gilmer, D., Bouzoubaa, S., Richards, K., Jonard, G., and Guilley, H. (1992b). *Virology* **190**, 55–67.
Gilmer, D., Allmang, C., Ehresmann C., Guilley, H., Richards, K., Jonard, G., and Ehresmann, B. (1993). *Nucleic Acids Res.* **21**, 1389–1395.
Goldbach, R. W. (1986). *Annu. Rev. Phytopathol.* **24**, 289–310.
Goldbach, R., and de Haan, P. (1994). *In* "The Evolutionary Biology of Viruses" (S. S. Morse, ed.), pp. 105–120. Raven Press, New York.
Goldbach, R., Le Gall, O., and Wellink, J. (1991). *Semin. Virol.* **2**, 19–25.
Goldbach, R., Wellink, J., Verver, J., Van Kammen., A., Kasteel, D., and Van Lent, J. (1994). *Arch. Virol.* **S9**, 87–97.
Gorbalenya, A. E., and Koonin, E. V. (1989). *Nucleic Acids Res.* **17**, 8413–8440.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). *FEBS Lett.* **239**, 16–24.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinow, V. M. (1989). *Nucleic Acids Res.* **17**, 4713–4730.
Gorbalenya, A. E., Koonin, E. V., and Wolf, Y. I. (1990). *FEBS Lett.* **243**, 103–113.
Greene, A. E., and Allison, R. F. (1994). *Science* **263**, 1423–1425.

Gros, C., and Wengler, G. (1996). *Virology* **217**, 367–372.
Gros, C. H., and Shuman, S. (1996). *J. Virol.* **70**, 1706–1713.
Grun, J. B., and Brinton, M. A. (1988). *J. Gen. Virol.* **69**, 3121–3127.
Guilford, P. J., Beck, D. L., and Forster, R. L. S. (1991). *Virology* **182**, 61–67.
Guinea, R., and Carrasco, L. (1990). *EMBO J.* **9**, 2011–2016.
Habili, N., and Symons, R. H. (1989). *Nucleic Acids Res.* **17**, 9543–9555.
Hacker, D. L., Petty, I. T. D., Wei, N., and Morris, T. J. (1992). *Virology* **186**, 1–8.
Hahn, Y. S., Grakoui, C. M., Rice, C. M., Strauss, E. G., and Strauss, J. H. (1989a). *J. Virol.* **63**, 1194–1202.
Hahn, Y. S., Strauss, E. G., and Strauss, J. H. (1989b). *J. Virol.* **63**, 3142–3150.
Hannig, E. M. (1995). *Bioessays* **17**, 915–919.
Harris, K. S., Xiang, W., Alexander, L., Lane, W. S., Paul, A. V., and Wimmer, E. (1994). *J. Biol. Chem.* **269**, 27004–27014.
Haseloff, J., Goelet, P., Zimmern, D., Ahlquist, P., Dasgupta, R., and Kaesberg, P. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4358–4362.
Hatta, T., and Francki, R. I. B. (1981). *J. Gen. Virol.* **53**, 343–346.
Havelda, Z., and Burgyan, J. (1995). *Virology* **214**, 269–272.
Havelda, Z., Dalmay, T., and Burgyan, J. (1995). *J. Gen. Virol.* **76**, 2311–2316.
Hayes, R. J., and Buck, K. W. (1990). *Cell* **63**, 363–368.
Hayes, R. J., and Buck, K. W. (1993). *In* "Molecular Virology: A Practical Approach" (A. J. Davison and R. M. Elliott, eds.), pp. 1–34. IRL Press, Oxford.
Hayes, R. J., Pereira, V. C. A., McQuillin, A., and Buck, K. W. (1994a). *J. Gen. Virol.* **75**, 3177–3184.
Hayes, R. J., Pereira, V. C. A., and Buck, K. W. (1994b). *FEBS Lett.* **352**, 331–334.
Hehn, A., Bouzoubaa, S., Bate, N., Twell, D., Marbach, J., Richards, K., and Jonard, G. (1995). *Virology* **210**, 73–81.
Hellen, C. U. T., and Cooper, J . I. (1987). *J. Gen. Virol.* **68**, 2913–2917.
Heringa, J., and Argos, P. (1994). *In* "The Evolutionary Biology of Viruses" (S. S. Morse, ed.), pp. 87–103. Raven Press, New York.
Hillman, B. I., and Lawrence, D. M. (1995). *In* "Pathogenesis and Host Specificity in Plant Diseases" (R. P. Singh, U. P. Singh and K. Kohmoto, eds.), Vol. III, pp. 35–50. Elsevier, Oxford.
Hills, G. J., Plaskitt, K. A., Young, N. D., Dunigan, D. D., Watts, J. W., Wilson, T. M. A., and Zaitlin, M. (1987). *Virology* **161**, 488–496.
Hirling, H., Scheffner, M., Restle, T., and Stahl, H. (1989). *Nature* **339**, 562–564.
Hiscox, J. A., Mawditt, K. L., Cavanagh, D., and Britton, P. (1995). *J. Virol.* **69**, 6219–6227.
Hodgeman, T. C. (1988). *Nature* **333**, 22–33; 579 (Erratum).
Hofmann, M. A., and Brian, D. A. (1991). *J. Virol.* **65**, 6331–6333.
Hong, Y., Levay, K., Murphy, J. F., Klein, P. G., Shaw, J. G., and Hunt, A. G. (1995). *Virology* **214**, 159–166.
Houser-Scott, F., Baer, M. L., Liem, K. F., Cai, J.-M., and Gehrke, L. (1994). *J. Virol.* **68**, 2194–2205.
Houwing, C. J., and Jaspars, E. M. J. (1986). *FEBS Lett.* **209**, 284–288.
Huisman, M. J., Linthorst, H. J. M., Bol, J. F., and Cornellison, B. J. C. (1988). *J. Gen. Virol.* **69**, 1789–1798.
Inokuchi, Y., and Hirashima, A. (1987). *J. Virol.* **61**, 3946–3949.
Ishikawa, M., Meshi, T., Motoyoshi, F., Takamatsu, N., and Okada, Y. (1986). *Nucleic Acids Res.* **14**, 8291–8305.
Ishikawa, M., Kroner, P., Ahlquist, P., and Meshi, T. (1991a). *J. Virol.* **65**, 3451–3459.

Ishikawa, M., Meshi, T., Ohno, T., and Okada, Y. (1991b). *J. Virol.* **65**, 861–868.
Ishikawa, M., Naito, S., and Ohno, T. (1993). *J. Virol.* **67**, 5328–5338.
Jablonski, S. A., and Morrow, C. D. (1993). *J. Virol.* **65**, 373–381.
Jablonski, S. A., Luo, M., and Morrow, C. D. (1991). *J. Virol.* 65, 4565–4572.
Jacoba-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hiza, A., Hughes, S. H., and Arnold, E. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320–6324.
Jacobson, S. J., Konigs, D. A. M., and Sarnow, P. (1993). *J. Virol.* **67**, 2961–2971.
Jaegle, M., Wellink, J., and Goldbach, R. (1987). *J. Gen. Virol.* **68**, 627–632.
Janda, M., and Ahlquist, P. (1993). *Cell* **72**, 961–970.
Jaramillo, M., Browning, K., Dever, T. E., Blum, S., Trachsel, H., Merrick, W. C., Ravel, J. M., and Sonenberg, N. (1990). *Biochim. Biophys. Acta* **1050**, 134–139.
Jaspars, E. M. J. (1985). *In* "Molecular Plant Virology" (J. W. Davies, ed.), pp. 155–211. CRC Press, Boca Raton.
Jaspars, E. M. J., Gill, D. S., and Symons, R. H. (1985). *Virology* **144**, 410–425.
Jiang, B. M., Monroe, S. S., Koonin, E. V., Stine, S. E., and Glass, R. I. (1993). *Proc Natl. Acad. Sci. U.S.A.* **90**, 10539–10543.
Johnson, K. L., and Sarnow, P. (1991). *J. Virol.* **65**, 4341–4349.
Johnston, J. C., and Rochon, D. M. (1995). *Virology* **214**, 100–109.
Jupin, I., Bouzoubaa, S., Richards, K., Jonard, G., and Guilley, H. (1990a). *Virology* **178**, 281–284.
Jupin, I., Richards, K., Jonard, G., Guilley, H., and Pleij, C. W. A. (1990b). *Virology* **178**, 273–280.
Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F., and Holmes, K. C. (1990). *Nature* **347**, 37–44.
Kadaré, G., Drugeon, G. Savithri, H. S., and Haenni, A-L. (1992). *J. Gen. Virol.* **73**, 493–498.
Kadaré. G., Rozanov, M., and Haenni, A-L. (1995). *J. Gen. Virol.* **76**, 2853–2857.
Kamer, G., and Argos, P. (1984). *Nucleic Acids Res.* **12**, 7269–7282.
Kaper, J. M. (1995). *In* "Pathogenesis and Host Specificity in Plant Diseases" (R. P. Singh, U. P. Singh, and K. Kohmoto, eds.), Vol. III, pp. 373–392. Elsevier, Oxford.
Kao, C. C., and Ahlquist, P. (1992). *J. Virol.* **66**, 7293–7302.
Kao, C. C., Quadt, R., Hershberger, R. P., and Ahlquist, P. (1992). *J. Virol.* **66**, 6322–6329.
Kassavetis, G. A., and Geiduschek, E. P. (1993). *Science* **259**, 944–945.
Kasschau, K. D., and Carrington, J. C. (1995). *Virology* **209**, 268–273.
Kim, J. C., Spence, R. A., Currier, P. A., Lu, X., and Denison, M. R. (1995). *Virology* **206**, 1–8.
Kim, Y. N., Jeong, Y. S., and Makino, S. (1993). *Virology* **197**, 53–63.
Kirkegaard, K., and Baltimore, D. (1986). *Cell* **47**, 433–443.
Klaasen, V. A., Boeshore, M. L., Koonin, E. V., Tian, T., and Falk, B. W. (1995). *Virology* **208**, 99–110.
Knorr, D. A., Mullin, R. H., Hearne, P. Q., and Morris, T. J. (1991). *Virology* **181**, 193–202.
Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992). *Science* **256**, 1783–1790.
Kollar, A., and Burgyan, A. (1994). *Virology* **201**, 169–172.
Koonin, E. V. (1991a). *J. Gen. Virol.* **72**, 2197–2206.
Koonin, E. V. (1991b). *Nature* **352**, 290.
Koonin, E. V. (1993). *J. Gen. Virol.* **74**, 733–740
Koonin, E. V., and Dolja, V. V. (1993). *Crit. Rev. Biochem. Mol. Biol.* **28**, 375–430.
Koonin, E. V., Boyko, V. P., and Dolja, V. V. (1991). *Virology* **181**, 395–398.

Kornberg, A., and Baker, T. A. (1992). "DNA Replication," 2nd Ed. W. H. Freeman and Co., New York.
Kowal, K. J., and Stollar, V. (1981). *Virology* **114**, 140–148.
Kroner, P., Richards, D., Traynor, P., and Ahlquist, P. (1989). *J. Virol.* **63**, 5302–5309.
Kroner, P. A., Young, B. M., and Ahlquist, P. A. (1990). *J. Virol.* **64**, 6110–6120.
Kuhn, R. J., Griffin, D. E., Zhang, H., Niiesters, H. G. M., and Strauss, J. H. (1992). *J. Virol.* **66**, 7121–7127.
Kusov, Y., Weitz, M., Dollenmeier, G., Gauss-Müller, V., and Siegl, G. (1996). *J. Virol.* **70**, 1890–1897.
Laakkonen, P., Hyvonen, M., Peranen, J., and Kääriäinen, L. (1994). *J. Virol.* **68**, 7418–7425.
Lahser, F. C., Marsh, L. E., and Hall, T. C. (1993). *J. Virol.* **67**, 3295–3303.
Lai, M. M. C. (1990). *Annu. Rev. Microbiol.* **44**, 303–333.
Lai, M. M. C. (1992). *Microbiol. Rev.* **56**, 61–79.
Lai, M. M. C. (1995). *In* "Molecular Basis of Viral Evolution" (A. J. Gibbs, C. H. Calisher, and F. Garcia-Arenal, eds.), pp. 119–132. Cambridge University Press, Cambridge.
Lain, S., Riechmann, J. L., Martin, M.T., and Garcia, J. A. (1989). *Gene* **82**, 357–362.
Lain, S., Riechmann, J. L., and Garcia, J. A. (1990). *Nucleic Acids Res.* **18**, 7003–7006.
Lain, S., Martin, M. T., Riechmann, J. L., and Garcia, J. A. (1991). *J. Virol.* **65**, 1–6.
Lama, J., Paul, A. V., Harris, K. S., and Wimmer, E. (1994). *J. Biol. Chem.* **269**, 66–70.
Lambden, P. R., and Clarke, I. N. (1995). *Trends Microbiol.* **3**, 261–265.
Landers, T. A., Blumenthal, T., and Weber, K. (1974). *J. Biol. Chem.* **249**, 5801–5808.
Larsen, G. R., Dorner, A. J., Harris, T. J. R., and Wimmer, E. (1980). *Nucleic Acids Res.* **8**, 1217–1229.
Lastarza, M. W., Lemm, J. A., and Rice, C. M. (1994a). *J. Virol.* **68**, 5781–5791.
Lastarza, M. W., Grakoui, A., and Rice, C. M. (1994b). *Virology* **202**, 224–232.
Lee, C-G., and Hurwitz, J. (1993). *J. Biol. Chem.* **268**, 16822–16830.
Lee, W. M., Monroe, S. S., and Rueckert, R. R. (1993). *J. Virol.* **67**, 2110–2122.
Lehto, K., Grantham, G. L., and Dawson, W. O. (1990). *Virology* **174**, 145–157.
Lemm, J. A., Rumenapf, T., Strauss, E. G., and Strauss, J. H. (1994). *EMBO J.* **13**, 2925–2934.
Lesemann, D. E. (1988). *In* "The Plant Viruses. IV. The Filamentous Plant Viruses," pp. 179–235. Plenum Press, New York.
Levis, R., Weiss, B. G., Tsiang, M., Huang, H., and Schlesinger, S. (1986). *Cell* **44**, 137–145.
Levis, R., Schlesinger, S., and Huang, H.V. (1990). *J. Virol.* **64**, 1726–1733.
Lewis, T. L., Greenberg, H. B., Herrman, J. E., Smith, L. S., and Matsui, S. M. (1994). *J. Virol.* **68**, 77–83.
Li, G., and Rice, C. M. (1989). *J. Virol.* **67**, 5062–5067.
Li, J. P., and Baltimore, D. (1988). *J. Virol.* **62**, 4016–4021.
Li, J. P., and Baltimore, D. (1990). *J. Virol.* **64**, 1102–1107.
Li, X. H., and Carrington, J. C. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 457–461.
Li, Y., and Ball, L. A. (1993). *J. Virol.* **67**, 3854–3960.
Lin, Y-J., and Lai, M. M. C. (1993). *J. Virol.* **67**, 6110–6118.
Lin, Y-J., Liao, C-L., and Lai, M. M. C. (1994). *J. Virol.* **68**, 8131–8140.
Lommel, S. A., Kendall, T. L., Xiong, Z., and Nutter, R. C. (1991). *Virology* **181**, 382–385.
Lomonossoff, G. P. (1995). *Annu. Rev. Phytopathol.* **33**, 323–343.
Lomonossoff, G. P., Shanks, M., and Evans, D. (1985). *Virology* **144**, 351–362.
Longstaff, M., Brigneti, G., Boccard, F., Chapman, S., and Baulcombe, D. (1993). *EMBO J.* **12**, 379–386.

Lu, H-L., Li, X., Cucanati, A., and Wimmer, E. (1995). *J. Virol.* **69**, 7445–7452.
Lubinski, J. M., Kaplan, G., Racaniello, V. R., and Dasgupta, A. (1986). *J. Virol.* **58**, 459–467.
Lucas, W. J. (1995). *Curr. Opin. Cell Biol.* **7**, 673–680.
Lucas, W. J., and Gilbertson, R. L. (1994). *Annu. Rev. Phytopathol.* **32**, 387–411.
Lupo, R., Rubino, L., and Russo, M. (1994). *Arch. Virol.* **138**, 135–142.
Makinen, K., Tamm, T., Naess, V., Truve, E., Puurand, U., Munthe, T., and Saarma. M. (1995). *J. Gen. Virol.* **76**, 2817–2825.
Makino, S., and Lai, M. M. C. (1989). *J. Virol.* **63**, 5285–5292.
Makino, S., Stohlman, S. A., and Lai, M. M. C. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4204–4208.
Maraia, R. J., Kenan, D. J., and Keene, J. D. (1994). *Mol. Cell Biol.* **14**, 2147–2158.
Marsh, L. E., and Hall, T. C. (1987). *Cold Spring Harbor Symp. Quant. Biol.* **52**, 331–341.
Marsh, L. E., Dreher, T. W., and Hall, T. C. (1988). *Nucleic Acids Res.* **16**, 981–995.
Marsh, L. E., Pogue, G. P., and Hall, T. C. (1989). *Virology* **172**, 415–427.
Marsh, L. E., Huntley, C. C., Pogue, G. P., Connell, J. P., and Hall, T. C. (1991a). *Virology* **182**, 76–83.
Marsh, L. E., Pogue, G. P., Szybiak, U., Connell, J. P., and Hall, T. C. (1991b). *J. Gen. Virol.* **72**, 2367–2374.
Martin, M. T., Cervera, M. T., and Garcia, J. A. (1995). *Virus Res.* **37**, 127–137.
Masuta, C., Tanaka, H., Dehara, K., Kuwata, S., Koiwai, A., and Noma, M. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6117–6121.
Matsuura, Y., and Miyamura, T. (1993). *Semin. Virol.* **4**, 297–304.
Maynell, L. A., Kirkegaard, K., and Klymkowsky, M. W. (1992). *J. Virol.* **66**, 1985–1994.
Mayo, M. A., and Fritsch, C. (1994). *FEBS Lett.* **354**, 129–130.
McBride, A. E., Schlegel, A., and Kirkegaard, K. (1996). *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2296–2301.
McKnight, K. L., and Lemon, S. M. (1996). *J. Virol.* **70**, 1941–1952.
Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowicz, F., Kenan, D. J., Chan, E. K. L., Agol, V. I., Keene, J. D., and Sonenberg, N. (1993). *J. Virol.* **67**, 3798–3807.
Merrick, W. C. (1992). *Microbiol. Rev.* **56**, 291–315.
Meshi, T., Motoyoshi, F., Adachi, A., Watanabe, Y., Takamatsu, N., and Okada, Y. (1988). *EMBO J.* **7**, 1575–1581.
Meyer, J., Lundquist, R. E., and Maizel, J. V. (1978). *Virology* **85**, 445–455.
Meulewater, F., Cornelissen, M., and Van Emmelo, J. (1992). *J. Virol.* **66**, 6419–6428.
Mi, S., and Stollar, V. (1990). *Virology* **178**, 429–434.
Mi, S., and Stollar, V. (1991). *Virology* **184**, 423–427.
Mi, S., Durbin, R., Huang, H. V., Rice, C. M., and Stollar, V. (1989). *Virology* **170**, 385–391.
Miller, W. A., and Hall, T. C. (1983). *Virology* **125**, 236–241.
Miller, W. A., Dreher, T., and Hall, T. C. (1985). *Nature* **313**, 68–70.
Miller, W. A., Dinesh-Kumar, S. P., and Paul, C. P. (1995). *Crit. Rev. Plant Sci.* **14**, 179–211.
Mills, D. R., Priano, C., DiMauro, P., and Binderow, B. D. (1988). *J. Mol. Biol.* **205**, 751–764.
Mirzayan, C., and Wimmer, E. (1992). *Virology* **189**, 547–555.
Mirzayan, C., and Wimmer, E. (1994a). *Virology* **199**, 176–187.
Mirzayan, C., and Wimmer, E. (1994b). *In* "Encyclopedia of Virology" (R. G. Webster and A. Granoff, eds.), Vol. 3, pp. 1119–1123. Academic Press, London.
Mizumoto, K., and Kaziro, Y. (1987). *Prog. Nucleic Acids Res. Mol. Biol.* **34**, 1–28.

Mohan, B. R., Dinesh-Kumar, S. P., and Miller, W. A. (1995). *Virology* **212**, 186–195.
Molla, A., Paul, A. V., and Wimmer, E. (1991). *Science* **254**, 1647–1651.
Molla, A., Paul, A. V., and Wimmer, E. (1992). *Dev. Biol. Stand.* **78**, 39–56.
Molla, A., Paul, A. V., and Wimmer, E. (1993). *J. Virol.* **67**, 5932–5938.
Molla, A., Harris, K. S., Paul, A. V., Shin, S. H., Mugavero, J., and Wimmer, E. (1994). *J. Biol. Chem.* **269**, 27015–27020.
Morch, M. D., Joshi, R. L., Denial, T. M., and Haenni, A. L. (1987). *Nucleic Acids Res.* **15**, 4123–4130.
Morrow, C. D., Gibbons, G. F., and Dasgupta, A. (1985). *Cell* **40**, 913–921.
Mouches, C., Candresse, T., and Bove, J. M. (1984). *Virology* **134**, 78–91.
Mueller, E., Gilbert, J., Davenport, G., Brigneti, G., and Baulcombe, D. C. (1995). *Plant J.* **7**, 1001–1013.
Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D., eds. (1995). "Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses," *Arch. Virol.* **Suppl. 10**, 1–586. Springer Verlag, Vienna.
Mushegian, A. R., and Shepherd, R. J. (1995). *Microbiol. Rev.* **59**, 548–578.
Nagy, P. D., Dzianott, A., Ahlquist, P., and Bujarski, J. J. (1995). *J. Virol.* **69**, 2547–2556.
Najita, L., and Sarnow, P. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5846–5850.
Nakhasi, H. L., Roualt, T. A., Haile, D., Liu, T.-Y., and Klausner, R. D. (1990). *New Biol.* **2**, 255–264.
Nakhasi, H. L., Cao, X.-Q., Roualt, T. A., and Liu, T-Y. (1991). *J. Virol.* **65**, 5961–5967.
Nakhasi, H. L., Singh, N. K., Pogue, G. P., Cao, X-Q., Roualt, T. A. (1994). *Arch. Virol.* **S9**, 255–267.
Naranda, T., MacMillan, S. E., and Hershey, J. W. B. (1994). *J. Biol. Chem.* **269**, 32286–32292.
Nassuth, A., and Bol, J. F. (1983). *Virology* **124**, 75–85.
Nassuth, A., Ablas, F., and Bol, J. F. (1981). *Virology* **24**, 75–85.
Natsuaki, T., Mayo, M. A., Jolly, C. A., and Murant, A. F. (1991). *J. Gen. Virol.* **72**, 2183–2189.
Neeleman, L., Van der Vossen, E. A. G., and Bol, J. F. (1993). *Virology* **196**, 883–887.
Nelson, G. W., and Stohlman, S. A. (1993). *J. Gen. Virol.* **74**, 1975–1979.
Nestorowicz, A., Chambers, T. J., and Rice, C. M. (1994). *Virology* **199**, 114–123.
Neufeld, K. L., Richards, K. L., and Ehrenfeld, E. (1991). *Virus Res.* **19**, 173–188.
Neufeld, K. L., Galarza, J. M., Richards, D. F., Summers, D. F., and Ehrenfeld, E. (1994). *J. Virol.* **68**, 5811–5818.
Niesters, H. G. M., and Strauss, J. H. (1990). *J. Virol.* **64**, 1639–1647.
Nilsen, T. W., Wood, D. L., and Baglioni, C. (1981). *Virology* **109**, 82–93.
Novak, J. E., and Kirkegaard, K. (1994). *Genes Develop.* **8**, 1726–1737.
Nuesch, J. P. F., Weitz, M., and Siegl, G. (1993). *Arch. Virol.* **128**, 65–79.
Nutter, R. C., Sheets, K., Panganiban, L. C., and Lommel, S. A. (1989). *Nucleic Acids Res.* **17**, 3163–3177.
Oberg, B. F., and Philipson, L. (1971). *J. Mol. Biol.* **58**, 725–737.
Odd, T., and Wu, H. C. (1993). *J. Biol. Chem.* **268**, 12596–12602.
Ogawa, T., Watanabe, Y., and Okada, Y. (1992). *Virology* **191**, 454–458.
Ollis, D. L., Brick, P., Hamlin, R., Xung, N. G., and Steitz, T. A. (1985). *Nature* **313**, 762–766.
O'Reilly, E. K., Tang, N., Ahlquist, P., and Kao, C. C. (1995). *Virology* **214**, 59–71.
Othman, Y., and Hull, R. (1995). *Virology* **206**, 287–297.
Pacha, R. F., and Ahlquist, P. (1991). *J. Virol.* **65**, 3693–3703.

Pacha, R. F., Allison, R. F., and Ahlquist, P. (1990). *Virology* **174**, 436–443.
Palukaitis, P., Roosinck, M., Dietzgen, R. G., and Francki, R. I. B. (1992). *Adv. Virus Res.* **41**, 281–348.
Pardigon, N., and Strauss, J. H. (1992). *J. Virol.* **66**, 1007–1015.
Pardigon, N., and Strauss, J. H. (1996). *J. Virol.* **70**, 1173–1181.
Pardigon, N., Lenches, E., and Strauss, J. H. (1993). *J. Virol.* **67**, 5003–5011.
Passmore, B. K., Sanger, M., Chin, L-S., Falk, B. W., and Bruening, G. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10168–10172.
Pata, J. D., Schultz, S. C., and Kirkegaard, K. (1995). *RNA* **1**, 466–477.
Paul, A. V., Cao, X., Harris, K. S., Lama, J., and Wimmer, E. (1994a). *J. Biol. Chem.* **269**, 29173–29181.
Paul, A. V., Molla, A., and Wimmer, E. (1994b). *Virology* **199**, 188–199.
Pause, A., and Sonenberg, N. (1992). *EMBO J.* **11**, 2643–2654.
Pause, A., Méthot, N., and Sonenberg, N. (1993). *Mol. Cell Biol.* **13**, 6789–6798.
Peränen, J., and Kääriäinen, L. (1991). *J. Virol.* **64**, 1888–1896.
Peränen, J., Laakkonen, P., Hyvonen, M., and Kääriäinen, L. (1995). *Virology* **208**, 610–620.
Perez, L., Guinea, R., and Carrasco, L. (1991). *Virology* **183**, 74–82.
Peters, S. A., Verver, J., Nollen, E. A. A., van Lent, J. W. M., Wellink, J., and van Kammen, A. (1994). *J. Gen. Virol.* **75**, 3167–3176.
Peters, S. A., Mesnard, J-M., Kooter, I. M., Verver, J., Wellink, J., and Van Kammen, A. (1995). *J. Gen. Virol.* **76**, 1807–1813.
Petty, I. T. D., French, R., Jones, R. W., and Jackson, A. O. (1990). *EMBO J.* **9**, 3453–3457.
Pfister, T., Egger, D., and Bienz, K. (1995). *J. Gen. Virol.* **76**, 63–71.
Pilipenko, E. V., Maslova, S. V., Sinyakov, A. N., and Agol, V. I. (1992). *Nucleic Acids Res.* **20**, 1739–1745.
Plagemann, P. G. W., and Moenning, V. (1992). *Adv. Virus Res.* **41**, 99–192.
Plotch, S. J., and Palant, O. (1995). *J. Virol.* **69**, 7169–7179.
Poch, O., Sauvaget, I., Delarue, M., and Tordo, N. (1989). *EMBO J.* **8**, 3867–3874.
Pogue, G. P., and Hall, T. C. (1992). *J. Virol.* **66**, 674–684.
Pogue, G. P., Marsh, L. E., and Hall, T. C. (1990). *Virology* **178**, 152–160.
Pogue, G. P., Marsh, L. E., Connell, J. P., and Hall, T. C. (1992). *Virology* **188**, 742–753.
Pogue G. P., Cao, X.-Q, Singh, N. K., and Nakhasi, H. L. (1993). *J. Virol.* **67**, 7106–7117.
Pogue, G. P., Huntley, C. C., and Hall, T. C. (1994). *Arch. Virol.* **S9**, 181–194.
Priano, C., Kramer, F. R., and Mills, D. R. (1987). *Cold Spring Harbor Symp. Quant. Biol.* **33**, 373–398.
Purdey, M. A., Tam, A. W., Huang, C.-C., Yarbough, P. O., and Reyes, G. R. (1993). *Semin. Virol.* **4**, 319–326.
Quadt, R., and Jaspars, E. M. J. (1990). *Virology* **178**, 189–194.
Quadt, R., Rosdorff, H. J. M., Hunt, T. W., and Jaspars, E. M. J. (1991). *Virology* **182**, 309–315.
Quadt, R., Kao, C. C., Browning, K. S., Hershberger, R. P., and Ahlquist, P. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1498–1502.
Quadt, R., Ishikawa, M., Janda, M., and Ahlquist, P. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4892–4896.
Raffo, A. J., and Dawson, W. O. (1991). *Virology* **184**, 277–289.
Raju, R., and Huang, H. V. (1991). *J. Virol.* **65**, 2501–2510.
Rao, A. L. N., Dreher, T. W., Marsh, L. E., and Hall, T. C. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5335–5339.
Rao, A. L. N., and Hall, T. C. (1991). *Virology* **180**, 16–22.

Rao, A. L. N., and Grantham, G. L. (1994). *Virology* **204**, 478–481.
Reavy, B., Mayo, M. A., Turnbull-Ross, A. D., and Murant, A. F. (1993). *Arch. Virol.* **131**, 441–446.
Restrapo-Hartwig, M. A., and Ahlquist, P. (1995). Abstr. Am. Soc. Virol. Meeting, Wisconsin, p. 136.
Restrapo-Hartwig, M. A., and Carrington, J. C. (1994). *J. Virol.* **68**, 2388–2397.
Reuer, Q., Kuhn, R. J., and Wimmer, E. (1990). *J. Virol.* **64**, 2967–2975.
Reusken, C. B. E. M., Neeleman, L., and Bol, J. F. (1994). *Nucleic Acids Res.* **22**, 1346–1353.
Reusken, C. B. E. M., Neeleman, L., and Bol, J. F. (1995). *J. Virol.* **69**, 4552–4555.
Reutenauer, A., Ziegler-Graff, V., Lot, H., Scheidecker, D., Guilley, H., Richards, K., and Jonard, G. (1993). *Virology* **195**, 692–699.
Revill, P. A., Davidson, A. D., and Wright, P. J. (1994). *Virology* **202**, 904–911.
Richards, K. E., and Tamada, T. (1992). *Annu. Rev. Phytopathol.* **30**, 291–313.
Richards, O. C., Martin, S. C., Jense, H. G., and Ehrenfeld, E. (1984). *J. Mol. Biol.* **173**, 325–340.
Richards, O. C., Yu, P., Neufeld, K. L., and Ehrenfeld, E. (1992). *J. Biol. Chem.* **267**, 17141–17146.
Riechmann, J. L., Lain, S., and Garcia, J. A. (1990). *Virology* **177**, 710–716.
Riechmann, J. L., Lain, S., and Garcia, J. A. (1992). *J. Gen. Virol.* **73**, 1–16.
Rikkonen, M., Peränen, J., and Kääriäinen, L. (1994). *Arch. Virol.* **S9**, 369–377.
Rodriguez, P. L., and Carrasco, L. (1993). *J. Biol. Chem.* **268**, 8105–8110.
Rohll, J. B., Holness, C. L., Lomonossoff, G. P., and Maule, A. J. (1993). *Virology* **193**, 672–679.
Rohll, J. B., Percy, N., Ley, R., Evans, D. J., Almond, J. W., and Barclay, W. S. (1994). *J. Virol.* **68**, 4384–4391.
Rohll, J. B., Moon, D. H., Evans, D. J., and Almond, J. W. (1995). *J. Virol.* **69**, 7835–7844.
Rohozinski, J., Francki, R. I. B., and Chu, P. W. G. (1986). *Virology* **155**, 27–38.
Roossinck, M. J., Sleat, D., and Palukaitis, P. (1992). *Microbiol. Rev.* **56**, 265–279.
Rothstein, M. A., Richards, O. C., Amin, C., and Ehrenfeld, E. (1988). *Virology* **164**, 301–308.
Rouleau, M., Bancroft, J. B., and Mackie, G. A. (1993). *Virology* **197**, 695–703.
Rozanov, M. N., Koonin, E. V., and Gorbalenya, A. E. (1992). *J. Gen. Virol.* **73**, 2129–2134.
Rozen, F., Edery, I., Meerovitch, K., Dever, T. E., Merrick, W, C., and Sonenberg, N. (1990). *Mol. Cell. Biol.* **10**, 1134–1144.
Russo, M., Di Franco, S., and Martelli, G. P. (1983). *J. Ultrastruct. Res.* **82**, 52–63.
Russo, M., Burgyan, J., and Martelli, G. P. (1994). *Adv. Virus Res.* **44**, 381–428.
Russo, M., Burgyan, J., and Rubino, L. (1995). Abstr. Fourth Internat. Symp. Positive Strand RNA Viruses, Utrecht, P6-58.
Saito, T., Hosokawa, D., Meshi, T., and Okada, Y. (1987). *Virology* **160**, 477–481.
Sankar, S., and Porter, A. G. (1992). *J. Biol. Chem.* **267**, 10168–10176.
Saraste, M., Sibbald, P. R., and Wittinghofer, A. (1990). *Trends Biochem. Sci.* **15**, 430–434.
Sarnow, P. (1989). *J. Virol.* **63**, 467–470.
Sarnow, P., Bernstein, H. D., and Baltimore, D. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 571–575.
Sawicki, S. G., and Sawicki, D. L. (1986). *J. Virol.* **57**, 328–334.
Sawicki, S. G., and Sawicki, D. L. (1990). *J. Virol.* **64**, 1050–1056.
Sawicki, D. L., and Sawicki, S. G. (1994). *Arch. Virol.* **S9**, 393–405.
Sawicki, S. G., and Sawicki, D. L. (1995). *In* "Corona- and Related Viruses" (P. J. Talbot and G. A. Levy, eds.), pp. 499–506. Plenum Press, New York.

Schaad, M. C., and Baric, R. S. (1994). *J. Virol.* **68**, 8169–8179.
Scheffner, M., Knippers, R., and Stahl, H. (1989). *Cell* **57**, 955–963.
Sheidel, L. M., Durbin, R. K., and Stollar, V. (1989). *Virology* **173**, 408–414.
Scheidel, L. M., and Stollar, V. (1991). *Virology* **181**, 490–499.
Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., and Sänger, H. L. (1993a). *J. Biol. Chem.* **263**, 11851–11857.
Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., and Sänger, H. L. (1993b). *J. Biol. Chem.* **263**, 11858–11867.
Schlegel, A., Giddings, T. H., Ladinsky, M. S., and Kirkegaard, K. (1996). *J. Virol.*, in press.
Schlesinger, S. (1995). *Mol. Biotech.* **3**, 155–165.
Schlesinger, S., Levis, R., Weiss, B. G., Tsiang, M., and Huang, H. (1987). *UCLA Symp. Mol. Cell Biol. New Ser.* **54**, 241–250.
Schmid, M., and Wimmer, E. (1994). *Arch. Virol.* **S9**, 279–289.
Scholthof, H. B., Morris, T. J., and Jackson, A. O. (1993). *Mol. Plant-Microbe Interact.* **6**, 309–322.
Scholthof, K.-B. G., Scholthof, H. B., and Jackson, A. O. (1995). *Virology* **208**, 365–369.
Schulz, G. E. (1992). *Curr. Opin. Struct. Biol.* **2**, 61–67.
Schultz, S. C., Hansen, J. L., and Long, A. (1995). Abstr. Fourth Internat. Symp. Positive Strand RNA Viruses, Utrecht, p. 13.
Shuman, S. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11798–11802.
Shuman, S. (1993). *J. Biol. Chem.* **268**, 11798–11802.
Sekimizu, K., and Kornberg, A. (1988). *J. Biol. Chem.* **263**, 7131–7135.
Sethna, P. B., Hung, S.-L., and Brian, D. A. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5626–5630.
Shatkin, A. J., and Kozak, M. (1983). *In* "The Reoviridae" (W. K. Joklik, ed.), pp. 79–106. Plenum Press, New York.
Shen, P., Kaniewska, M. B., Smith, C., and Beachy, R. N. (1993). *Virology* **193**, 621–630.
Shuman, S., and Moss, B. (1990). *Methods Enzymol.* **181**, 170–180.
Shirako, Y., and Wilson, T. M. A. (1993). *Virology* **195**, 16–32.
Shirako, Y., and Strauss, J. H. (1994). *J. Virol.* **68**, 1874–1885.
Shiroki, K., Kato, H., Koike, S., Odaka, T., and Nomoto, A. (1993). *J. Virol.* **67**, 3989–3996.
Shuman, S. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10935–10939.
Simon, A. E., and Bujarski, J. J. (1994). *Annu. Rev. Phytopathol.* **32**, 337–362.
Singh, N. K., Atreya, C. A., and Nakhasi, H. L. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12770–12774.
Skuzeski, J. M., Bozarth, C. S., and Dreher, T. W. (1996). *J. Virol.* **70**, 2107–2115.
Smirnyagina, E., Hsu, Y.-H., Chua, N., and Ahlquist, P. (1994). *Virology* **198**, 427–436.
Snijder, E. J., and Horzinek, M. C. (1993). *J. Gen. Virol.* **74**, 2305–2316.
Snijder, E. J., Wassenaar, A. L. M., Spaan, W. J. M., and Gorbalenya, A. E. (1995). *J. Biol. Chem.* **270**, 16671–16676.
Solovyev, A. G., Novikov, V. K., Merits, A., Savenkov, E. I., Zelenina, D. A., Tyulkina, L. G., and Morozov, S. Y. (1994). *J. Gen. Virol.* **75**, 259–267.
Song, C., and Simon, A. E. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8792–8796.
Song, C., and Simon, A. E. (1995). *J. Virol.* **69**, 4020–4028.
Sosnovtsev, S., and Green, K. Y. (1995). *Virology* **210**, 383–390.
Soumounou, Y., and Laliberte, J-F. (1994). *J. Gen. Virol.* **75**, 2567–2573.
Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B. C. (1993). *Nature* **364**, 593–599.
Spector, D. H., and Baltimore, D. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2983–2987.
Standart, N., and Jackson, R. J. (1994). *Biochimie* **76**, 867–879.

Stern, D. F., and Sefton, B. M. (1982). *J. Virol.* **42**, 755–759.
Strauss, J. H., and Strauss, E. G. (1988). *Annu. Rev. Microbiol.* **42**, 657–683.
Strauss, J. H., and Strauss, E. G. (1994). *Microbiol. Rev.* **58**, 491–562.
Sun, J., and Kao, C. C. (1996). *Virology*, in press.
Suzich, J. A., Tamura, J. K., Palmer-Hill, F., Warrener, P., Grakoui, A., Rice, C. M., Feinstone, S. M., and Collett, M. S. (1993). *J. Virol.* **67**, 6153–6158.
Takamatsu, N., Watanabe, Y., Meshi, T., and Okada, Y. (1990). *J. Virol.* **64**, 3686–3693.
Takamatsu, N., Watanabe, Y., Iwasaka, T., Shiba, T., Meshi, T., and Okada, Y. (1991). *J. Virol.* **65**, 1619–1622.
Takeda, N., Kuhn, R. J., Yang, C. F., Takegami, T., and Wimmer, E. (1986). *J. Virol.* **60**, 43–53.
Tamura, J. K., Warrener, P., and Collett (1993). *Virology* **193**, 1–10.
Taylor, S. J., and Shalloway, D. (1994). *Nature* **368**, 867–871.
Taylor, S. J., Anafi, M., Pawson, T., and Shalloway, D. (1995). *J. Biol. Chem.* **270**, 10120–10124.
Teterina, N. L., Kean, K. M., Gorbalenya, A. E., Agol, V. L., and Girard, M. (1992). *J. Gen. Virol.* **73**, 1977–1986.
Tobin, G. J., Young, D. C., and Flanegan, J. B. (1989). *Cell* **59**, 511–519.
Todd, S., Nguyen, J. H. C., and Semler, B. L. (1995). *J. Virol.* **69**, 3605–3614.
Tolskaya, E. A., Romanova, L. I., Kolesnikova, M. S., Gmyl, A. P., Gorbalenya, A. E., and Agol, V. I. (1994). *J. Mol. Biol.* **236**, 1310–1323.
Toyoda, H., Yang, C. F., Takeda, N., Nomoto, A., and Wimmer, E. (1987). *J. Virol.* **61**, 2816–2822.
Traut, T. W. (1994). *Eur. J. Biochem.* **222**, 9–19.
Traynor, P., and Ahlquist, P. (1990). *J. Virol.* **64**, 69–77.
Traynor, P., Young, B. M., and Ahlquist, P. (1991). *J. Virol.* **65**, 2807–2815.
Troxler, M., Egger, D., Pfister, T., and Bienz, K. (1992). *Virology* **191**, 687–697.
Tsai, C-H, and Dreher, T. W. (1991). *J. Virol.* **65**, 3060–3067.
Tsai, C-H, and Dreher, T. W. (1992). *J. Virol.* **66**, 5190–5199.
Turnbull-Ross, A. D., Mayo, M. A., Reavy, B., and Murant, A. F. (1993). *J. Gen. Virol.* **74**, 555–561.
van Bokhoven, H., Van Lent, J. W. M., Custers, R., Vlak, J. M., Wellinck, J., and Van Kammen, A. (1992). *J. Gen. Virol.* **73**, 2775–2784.
van Bokhoven, H., Le Gall, O., Kasteel, D., Verver, J., Wellinck, J., and Van Kammen, A. (1993). *Virology* **195**, 377–386.
van der Kuyl, A. C., Langereis, K., Houwing, C. J., Jaspars, E. M. J., and Bol, J. F. (1990). *Virology* **176**, 346–354.
van der Kuyl, A. C., Neeleman, L., and Bol, J. F. (1991). *Virology* **183**, 687–694.
van der Meer, J., Dorssers, L., Van Kammen, A., and Zabel, P. (1984). *Virology* **132**, 413–425.
van der Most, R. G., De Groot, R. J., and Spaan, W. J. M. (1994). *J. Virol.* **68**, 3656–3666.
van der Vossen, E. A. G., Neeleman, L., and Bol, J. F. (1993). *Nucleic Acids Res.* **21**, 1361–1367.
van der Vossen, E. A. G., Neeleman, L., and Bol, J. F. (1994). *Virology* **202**, 891–903.
van der Vossen, E. A. G., Notenboom, T., and Bol, J. F. (1995). *Virology* **212**, 663–672.
van der Werf, S., Bradley, J., Wimmer, E., Studier, F. W., and Dunn, J. J. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2330–2334.
van Duin, J. (1988). *In* "The Bacteriophages" (R. Calandar, ed.), pp. 117–167. Plenum Press, New York.
van Dyke, T. A., and Flanegan, J. B. (1980). *J. Virol.* **35**, 733–740.

van Etten, J. L., Burbank, D. E., Cuppels, D. A., Lane, L. C., and Vidaver, A. K. (1980). *J. Virol.* **33**, 769–773.
van Kuppeveld, F. J. M., Galama, J. M. D., Zoll, J., and Melchers, W. J. G. (1995). *J. Virol.* **69**, 7782–7790.
van Marle, G., Luytjes, W., Van der Most, R. G., Van Der Straaten, T., and Spaan, W. J. M. (1995). *J. Virol.* **69**, 7851–7856.
van Verooj, W. J., Slobbe, R. L., and Pruijn G. J. M. (1993). *Mol. Biol. Rep.* **18**, 113–119.
Verchot, J., and Carrington, J. C. (1995). *J. Virol.* **69**, 3668–3674.
Volloch, V. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1208–1212.
Volloch, V., Schweitzer, B., and Rits, S. (1987). *J. Cell Biol.* **105**, 137–143.
Vos, P., Jaegle, M., Wellink, J., Van Kammen, A., and Goldbach, R. (1988). *Virology* **165**, 33–41.
Waigmann, E., and Zambryski, P. (1994). *Curr. Biol.* **4**, 713–716.
Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). *EMBO J.* **1**, 945–951.
Wang, H.-L., O'Rear, J., and Stollar, V. (1996). *Virology* **217**, 527–531.
Wang, Y-F., Sawicki, S. G., and Sawicki, D. L. (1994). *J. Virol.* **68**, 6466–6475.
Ward, C. W. (1994). *Res. Virol.* **144**, 419–453.
Warrener, P., and Collett, M. C. (1995). *J. Virol.* **69**, 1720–1726.
Warrener, P., Tamura, J. K., and Collett, M. C. (1993). *J. Virol.* **67**, 989–996.
Watanabe, Y., Kishibayashi, N., Motoyoshi, F., and Okada, Y. (1987). *Virology* **161**, 527–532.
Watanabe, Y., Ogawa, T., Takahashi, H., Ishida, I., Takeuchi, Y., Yamamoto, M., and Okada, Y. (1995). *FEBS Lett.* **372**, 165–168.
Wei, N., Hacker, D. L., and Morris, T. J. (1992). *Virology* **190**, 346–355.
Weiland, J. J., and Dreher, T. W. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6095–6099.
Weiner, B. M., and Bradley, M. K. (1991). *J. Virol.* **65**, 4973–4984.
Weiner, A. M., and Maizels, N. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7383–7387.
Wellink, J., Van Lent, J., and Goldbach, R. (1988). *J. Gen. Virol.* **69**, 751–755.
Wengler, G., and Wengler, G. (1993). *Virology* **197**, 265–273.
Wengler, G., Wengler, G., and Gross, H. J. (1979). *Nature* 282, 754–756.
Wengler, G., Boege, U., Wengler, G., Bischoff, H., and Wahn, K. (1982). *Virology* **118**, 401–410.
White, K. A., and Morris, T. J. (1994). *J. Virol.* **68**, 14–24.
White, K. A., Skuzeski, J. M., Li, W., Wei, N., and Morris, T. J. (1995). *Virology* **211**, 525–534.
Willcocks, M. M., Brown, T. D. K., Madeley, C. R., and Carter, M. J. (1994). *J. Gen. Virol.* **75**, 1785–1788.
Wimmer, E., Hellen, C. U. T., and Cao, X. (1993). *Annu. Rev. Genet.* **27**, 353–436.
Witherell, G. W., Gott, J. M., and Uhlenbeck, O. C. (1991). *Prog. Nucleic Acid Res. Mol. Biol.* **40**, 185–220.
Wu, G., and Kaper, J. M. (1994). *J. Mol. Biol.* **238**, 655–657.
Wu, S.-X., and Kaesberg, P. (1991). *Virology* **183**, 392–396.
Wu, S.-X., Ahlquist, P., and Kaesberg, P. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11136–11140.
Young, N. D., and Zaitlin, M. (1986). *Plant Mol. Biol.* **6**, 455–465.
Young, D. C., Tuschall, D. M., and Flanegan, J. B. (1985). *J. Virol.* **54**, 256–264.
Yu, W., and Leibowitz, J. L. (1995). *J. Virol.* **69**, 2016–2023.
Yu, S. F., Benton, P., Bovee, M., Sessions, L., and Lloyd, R. E. (1995). *J. Virol.* **69**, 247–252.
Yusibov, V. M., and Loesch-Fries, L. S. (1995). *Virology* **208**, 405–407.

Xiang, W., Cuconati, A., Paul, A. V., Cao, X., and Wimmer, E. (1995a). *RNA* **1**, 892–904.
Xiang, W., Harris, K. S., Alexander, L., and Wimmer, E. (1995b). *J. Virol.* **69**, 3658–3667.
Xiao, Q., Sharp, T. S., Jeffrey, I. W., James, M. C., Pruijn, G. J. M., van Venrooij, W. J., and Clemens, M. J. (1994). *Nucleic Acids Res.* **22**, 2512–2518.
Zavriev, S. K., Hickey, C. M., and Lommel, S. A. (1996). *Virology* **216**, 407–410.
Zerfass, K., and Beier, H. (1992). *EMBO J.* **11**, 4167–4173.
Zhang, X., and Lai, M. M. C. (1995). *J. Virol.* **69**, 1637–1644.
Zhang, X., and Lai, M. M. C. (1996). *J. Virol.* **70**, 705–711.
Zhang, X., Liao, C.-L., and Lai, M. M. C. (1994). *J. Virol.* **68**, 4738–4746.
Zhu, J., and Newkirk, M. M. (1994). *Clin. Invest. Med.* **17**, 196–205.
Ziegler, A., Natsuaki, T., Mayo, M. A., Jolly, C. A., and Murant, A. F. (1992). *J. Gen. Virol.* **73**, 3213–3218.
Zimmern, D. (1988). *In* "RNA Genetics" (J. Holland, E. Domingo, and P. Ahlquist, eds.), Vol. 2, pp. 211–240. CRC Press, Boca Raton.

This Page Intentionally Left Blank

ADVANCES IN VIRUS RESEARCH, VOL. 47

TRANS-ACTIVATION OF CELLULAR GENES BY HEPATITIS B VIRUS PROTEINS: A POSSIBLE MECHANISM OF HEPATOCARCINOGENESIS

Wolfgang H. Caselmann

Department of Medicine II
Klinikum Grosshadern
Ludwig-Maximilians-Universität
D-81366 Munich, Germany
and
Department of Virus Research
Max-Planck-Institut für Biochemie
D-82152 Martinsried, Germany

I. Introduction

In 1965 Blumberg *et al.* (1965) identified a new antigen in leukemic serum of Australian aborigines. This "Australia antigen" turned out to be the surface antigen (HBsAg) of the hepatitis B virus (HBV; Prince, 1968). Two years later the complete human HBV particle could be isolated from patient's serum and visualized by electron microscopy (Dane *et al.*, 1970). Since this time, intensive research has successfully elucidated HBV gene expression and viral replication and a variety of addi-

tional hepatotropic DNA viruses (hepadnaviruses) have been identified. Up to now, however, the molecular pathomechanisms underlying HBV-associated chronic liver disease, and especially hepatocarcinogenesis, are only poorly understood.

A. Hepadnaviruses

The human HBV and several related species-specific animal hepatitis viruses share common features which are the basis for classification of these viruses as hepadnaviruses. The woodchuck hepatitis virus (WHV; Summers *et al.,* 1978), the ground squirrel hepatitis virus (GSHV; Marion *et al.,* 1980), and the tree squirrel hepatitis virus (TSHV; Feitelson *et al.,* 1986) represent mammalian hepadnaviruses, while the duck hepatitis B virus (DHBV; Mason *et al.,* 1980) and the heron hepatitis virus (HeHBV; Sprengel *et al.,* 1988) are summarized in the group of avian hepadnaviruses. Similarities include virion and DNA size, structure, and genetic organization as well as the replication mechanism including reverse transcription of a pregenomic RNA transcript of more than genome length (see Section II,D). Phylogenetically, hepadnaviruses seem to be related to retroviruses due to similarities in gene number, function, and organization (Miller and Robinson, 1986). Mammalian and avian hepadnaviruses represent two separate genera within *Hepadnaviridae*, because there are differences in nucleotide (nt) sequence homology, gene number and organization, polypeptide size and virion structure, antigenic cross-reaction, and host range among both groups (Werner *et al.,* 1979; Gerlich *et al.,* 1980; Mandart *et al.,* 1984). For example, avian hepadnaviruses lack the X open reading frame (ORF). Hepadnaviruses and other known causative agents of viral hepatitis like hepatitis viruses A, C, D, E, and G belong to completely different virus families.

B. Epidemiology of HBV Infection

More than 200 million people are chronic carriers of HBV. More than 2 million of them die annually from the sequelae of HBV-associated liver disease. This is approximately 100,000 deaths by fulminant, 500,000 by acute, and 400,000 by complications of chronic hepatitis B per year. About 700,000 patients die annually from HBV-induced liver cirrhosis and approximately 600,000 from primary hepatocellular carcinoma (HCC; Szmuness, 1978). According to estimates by the World Health Organization there will be 400 million HBV carriers in the world by the year 2000, despite the existence of an effective HBV vaccine.

In Western Europe and the United States of America the spread of HBV is usually horizontal, by blood products or mucosal contact. In endemic areas like Southeast Asia or Equatorial Africa vertical transmission perinatally from an HBV-infected mother to the newborn child is more common (Chalmers and Alter, 1971). After an incubation period of 35–150 days HBV infection becomes clinically apparent in less than 50% of all infected individuals. In general, complete remission occurs in 90–95% of them within 3–4 months (Bamber *et al.,* 1983). Chronic hepatitis B evolves if HBV infection persists more than 6 months. Up to 85% of chronic hepatitis B cases displaying severe histological changes progress to cirrhosis; 2 to 15% of these will develop HCCs, depending on regional factors (Popper *et al.,* 1987).

II. The Hepatitis B Virus

A. Morphology

The 42-nm infectious virion consists of a 27-nm icosahedric nucleocapsid of 180 units of the hepatitis B virus core antigen (HBcAg; Onodera *et al.,* 1982; see Section II,E,2) and a 7-nm lipoprotein bilayer derived from the endoplasmic reticulum (ER) membrane of the host (Patzer *et al.,* 1986). Three differently glycosylated HBV surface proteins of varying sizes, LHBs, MHBs, and SHBs (see Section II,E,1), are inserted in this lipoprotein bilayer. The viral DNA, a virus-encoded RNA-dependent DNA polymerase, a protein kinase activity which phosphorylates HBcAg (Albin and Robinson, 1980), and a genome-bound protein covalently linked to the 5′-end of the *L*(–) strand of HBV DNA (see Section II,B) are contained in the nucleocapsid (Fig. 1; Tiollais *et al.,* 1985; Blum *et al.,* 1989: Kann *et al.,* 1993). Nine different complete HBV DNA sequences have been isolated and characterized to date (Pasek *et al.,* 1979; Werner *et al.,* 1979, Valenzuela *et al.,* 1980; Fujiyama *et al.,* 1993; Ono *et al.,* 1983; Gan *et al.,* 1984; Kobayashi and Koike, 1984; Bichko *et al.,* 1985). In the initial phase of infection up to 10^{10} particles/ml are detectable in serum (Dane *et al.,* 1970). Besides intact infectious virions (Dane *et al.,* 1970), approximately 22-nm spheric and filamentous structures are found in sera of infected patients in 1000- to 100,000-fold excess. These particles consist exclusively of hepatitis B surface antigen (HBsAg) and do not contain viral DNA. The small spherical HBsAg particles display a buoyant density of about 1.18 g/ml in CsC1 (Gerin *et al.,* 1983). The 27-nm virion core particles obtained after treatment with nonionic detergents like

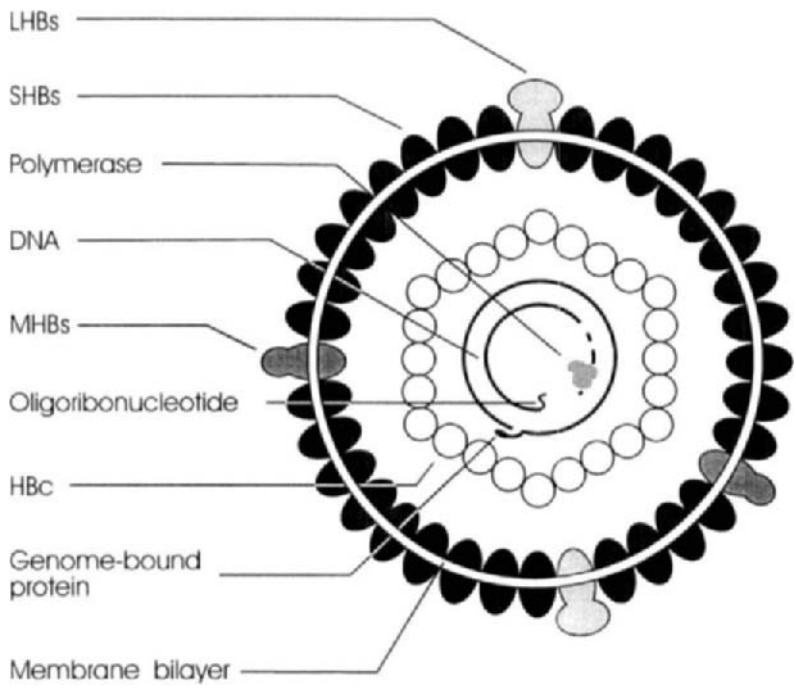


FIG. 1. Schematic structure of HBV particle. LHBs, large hepatitis B surface protein; MHBs, middle hepatitis B surface protein; SHBs, small hepatitis B surface protein; HBc, core protein; polymerase, virus-encoded RNA-dependent DNA polymerase; DNA, 3.2-kb partially double-stranded HBV DNA genome; genome-bound protein, covalently linked protein; membrane bilayer, lipoprotein bilayer of cellular origin.

Nonidet P-40 have a buoyant density of about 1.36 g/ml in CsC1 as compared to 1.22 g/ml for intact virions (Kaplan *et al.*, 1976).

B. Genomic Organization

Due to its complex organization the complete genetic HBV information is encoded in about 3200 base pairs (bp). The large [*L*(–)] strand of the partially double-stranded genome carries five overlapping ORFs—preS1/S2 (surface), preC/C (core), P (polymerase), X, and ORF 5—which do not contain introns and encode at least seven proteins (Fig. 2). For ORF 5 within the X gene as well as for a couple of short ORFs encoded by the small [*S*(+)] strand (ORF 6; Kaneko and Miller, 1988), neither a gene product nor a distinct function has been proven so far (Miller, 1988; Miller *et al.*, 1989). Circulation of the genome is achieved

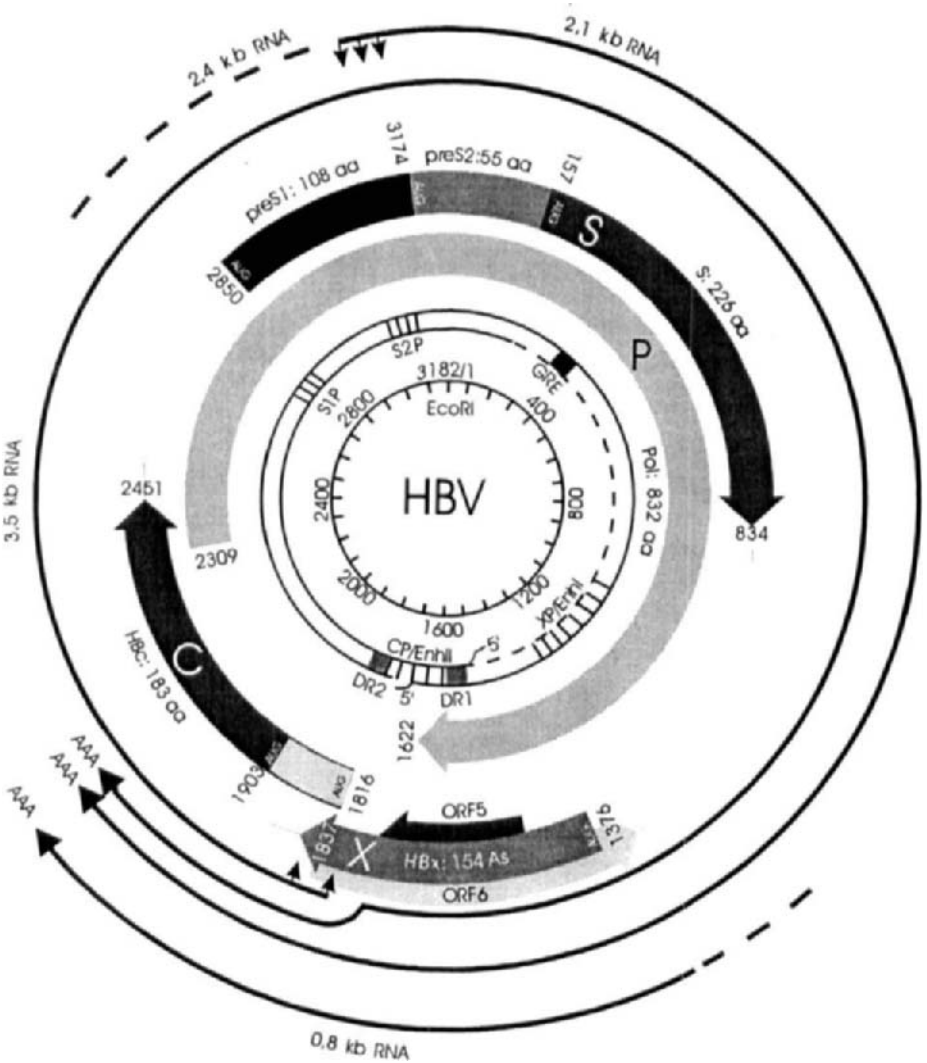


FIG. 2. Genomic organization of HBV (subtype *ayw*). Inner circle numbering of nucleotide positions according to Galibert *et al.* (1979). The partially double-stranded circle represents the viral genome, showing direct repeats 1 and 2 (DR1 and DR2), the S1, S2, X, and C promoters (S1P, S2P, XP, and CP, respectively), and the enhancers (EnhI and EnhII, respectively) as well as the glucocorticoid-responsive element (GRE). The genome-bound protein and the oligoribonucleotide are indicated at the 5′ end of the *L*(–) strand and at the 5′ end of the *S*(+) strand, respectively. The ORF for preS1/S2/S, preC/C, P, X, and ORFs S and 6 are indicated with the encoded polypeptides, if shown experimentally. The outer circle depicts the unspliced viral transcripts with their heterologous initiation sites and the common polyadenylation site.

by base pairing of the free 5′ end of the *S*(+) strand with the nicked 5′ end of the *L*(–) strand in a region of approximately 220 nt flanked by two 11-bp direct repeats, DR1 and DR2 (Sattler and Robinson, 1979). A polypeptide most likely encoded by the amino-terminal region of the HBV P gene (Bartenschlager and Schaller, 1988) is covalently linked to the 5′ end of the *L*(–) strand and serves as protein primer for *L*(–) strand DNA synthesis during replication. A basic protein bearing the HBV DNA polymerase and RNase H activity is linked to the 3′ end of the *S*(+) strand (Bavand *et al.*, 1989; Toh *et al.*, 1983). A 19-nt capped ribonucleotide is covalently attached to the 5′ end of the *S*(+) strand, most likely serving as a primer for the *S*(+) strand DNA synthesis (see Section II,D).

The HBV genome carries four promoters: the S1 promoter (S1P in Fig. 2; Siddiqui *et al.*, 1986), the S2 promoter (S2P; Cattaneo *et al.*, 1983; Siddiqui *et al.*, 1986; De Medina *et al.*, 1988), the C promoter (CP; Honigwachs *et al.*, 1989), the X promoter (XP; Treinin and Laub, 1987), and a polyadenylation signal within the C gene approximately 20 bp upstream of the 3′ end of the *L*(–) strand (Cattaneo *et al.*, 1984). In addition, two enhancers, EnhI and EnhII (Shaul *et al.*, 1985; Shaul and Ben-Levy, 1987; Yee, 1989), and a glucocorticoid-responsive element, GRE (Tur-Kaspar *et al.*, 1986), regulate gene expression (Fig. 2; see Section II,C).

EnhI (nt 1135–1254) has binding sites for various liver-specific (Ori and Shaul, 1995; Ori *et al.*, 1994) and ubiquitous DNA-binding transcription factors (Trujillo *et al.*, 1991; Yuh and Ting, 1990; Jameel and Siddiqui, 1986). Deletion mapping revealed that EnhI consists of a 5′-basal enhancer module and a 3′-accessory module which increases the basal enhancer activity in an orientation- and distance-dependent manner (Guo *et al.*, 1991). EnhI stimulates gene transcription from the S2 and C promoters as well as from the X promoter via the X-responsive element XRE (Faktor and Shaul, 1990). In its native position within the HBV genome EnhI functions less efficiently than if it is placed immediately upstream of the C promoter. This weak activity is due to repression of EnhI by a cellular trans-acting factor, which binds to the NF-1 site located between EnhI and the C promoter (Spandau and Lee, 1992). This factor seems to be related to NF-1, but not to be NF-1 itself (Patel *et al.*, 1989).

EnhII is localized at nt position 1627–1774 (Yee, 1989) and partially overlaps the C promoter (Y. Wang *et al.*, 1990). It consists of two functional subunits, element A (nt 1627–1687) and element B (nt 1688–1774). Element A seems to be responsible for the liver specificity of EnhII by binding liver-specific transcription factors; element B is

essential for its transcriptional stimulatory function (Yuh and Ting, 1990). A negative regulating element (NRE; nt 1613–1636) represses EnhII activity in differentiated cells, possibly by masking transcription factor binding sites (Lo and Ting, 1994). It has marginal inhibitory effect if isolated, but strong repressive potential in the presence of a functional EnhII.

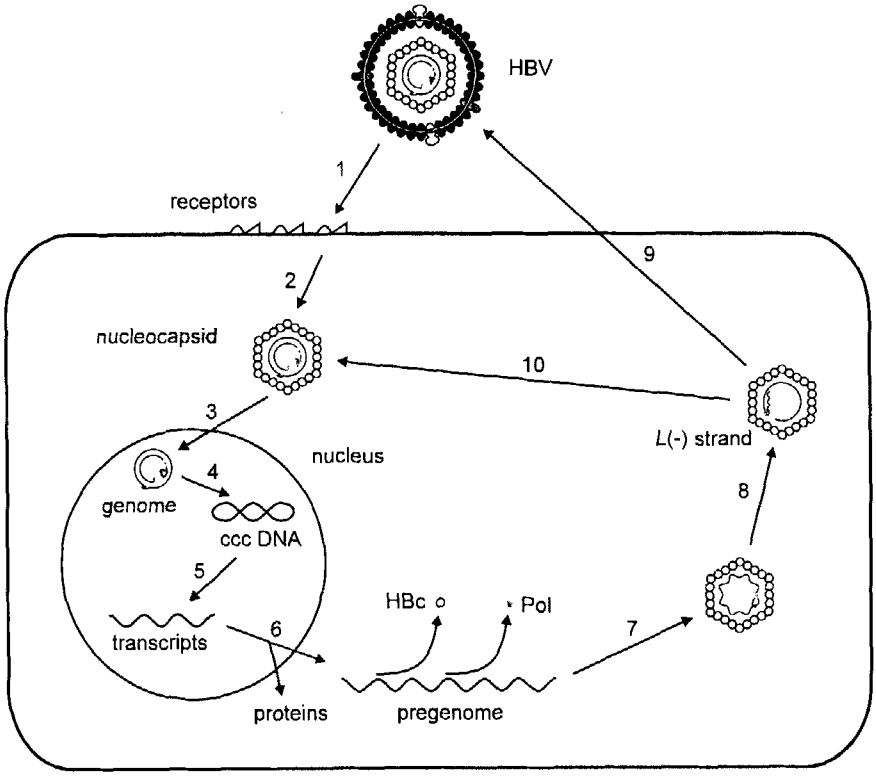
A binding site for hepatocyte nuclear factor 1 within the S1 promoter (nt 2717–2739) is probably responsible for the liver specificity of the S1 promoter (H. K. Chang *et al.*, 1989). At nt position 2967–2986 within the S2 promoter which lacks a TATA box there is a binding site for nuclear factor 1 (Shaul *et al.*, 1986), and a repressor/activator (nt 3079–3134) which stimulates HBV expression in HeLa cells and represses it in liver cells (Bulla and Siddiqui, 1989), as well as an enhancer-like sequence (nt 3155–30) with homologies to an AP-1 binding site, to the cyclic AMP-responsive element (CRE), and to the serum-responsive element (SRE). The GRE (nt 341–370) stimulates HBV expression from the S2 promoter three- to fivefold after glucocorticoid induction (Tur-Kaspar et al., 1986).

C. Transcription

Transcription of covalently closed circular (ccc) HBV DNA (see Section II,D) is performed by the cellular DNA-dependent RNA polymerase II. On the one hand, it creates an RNA template for HBV replication; on the other, it leads to the synthesis of all viral transcripts from which the HBV proteins are translated (Fig. 3). All major transcripts have (–) strand polarity and terminate approximately 20 nt downstream of the conserved polyadenylation site with a poly(A) tail of about 100 nt (nt 1888–TATAA–1993; Cattaneo *et al.*, 1983; Simonsen and Levinson, 1983; Cattaneo *et al.*, 1984).

1. C-mRNA

The 3.4-kb C-mRNA initiates within the preC region, approximately 6 nt upstream of the DR1 sequence, and terminates within the C gene with a terminal redundance of 63 nt. The polyadenylation signal is read through when it is reached the first time (Fig. 2). The C promoter regulates the initiation of the C-mRNA (Gough, 1983; Will *et al.*, 1987). The mechanism regulating the DNA-dependent RNA polymerase II readthrough is unknown. This RNA represents the template for the synthesis of HBcAg as well as for the synthesis of the viral polymerase by internal initiation of translation (L. J. Chang *et al.*, 1989; Schlicht *et al.*, 1989). In addition, the C-mRNA is the template for reverse transcrip-



tion (see Section II,D) by the viral polymerase after this pregenome has been packed in nucleocapsids. For this the presence of a 94-bp packaging signal ε at the 5′ end of the RNA is essential (Junker-Niepmann *et al.*, 1990; Knaus and Nassal, 1993; Pollack and Ganem, 1993).

2. *PreC-mRNA*

The initiation of the slightly longer 3.5-kb preC-mRNA transcript is also regulated by the C promoter. The transcription initiation site, which is located upstream of the preC start codon, is used much less frequently than the initiation site for the C-mRNA. The 3.5-kb transcript is also terminally redundant. It serves as template for the synthesis of the hepatitis B e antigen (HBeAg) that undergoes proteolytic processing before secretion (Schlicht *et al.*, 1989; Fig. 2).

3. PreS1-mRNA

The 2.4-kb preS1-mRNA transcript initiates approximately 38 nt upstream of the start codon of the preS/S gene (Will *et al.,* 1987). It is the template for the synthesis of the large hepatitis B surface protein. There is no evidence for internal initiation of translation of the middle and/or small hepatitis B surface proteins (see Section II,E,1). Due to the liver specificity of the S1 promoter (H. K. Chang *et al.,* 1989), the 2.4-kb preS1-mRNA transcripts are only transcribed in hepatocytes.

4. PreS2/S-mRNA

Three preS2/S-mRNA transcripts of approximately 2.1 kb initiate at three different initiation sites around the start codon of the preS2 region. The transcription is regulated by the S2 promoter within the preS1 region which lacks a TATA box and tissue specificity. It shares some homologies with the simian virus 40 (SV 40) late promoter (Cattaneo *et al.,* 1983). PreS2/S transcripts are the templates for the synthesis of middle and small hepatitis B surface proteins (Standring *et al.,* 1984; Ou and Rutter, 1985; Siddiqui *et al.,* 1986; Yaginuma *et al.,* 1987) which are translated in approximately equal amounts. A post-transcriptional regulating element (PRE) on HBV S transcripts regulates the level of preS/S-specific transcripts by activating their transport from the nucleus to the cytoplasm (Huang and Yen, 1994). Huang and Liang (1993) showed that this element can partially substitute for the *rev*-responsive element of human immunodeficiency virus 1 (HIV-1) and concluded that PRE exhibits *rev* protein/*rev*-responsive element-like functions by inhibiting splicing. However, most recent data suggest that PRE functions in cis to allow the export of nuclear transcripts that do not interact efficiently with the splicing pathway and are normally not exported well from the nucleus (Huang and Yen, 1995).

5. X-mRNA

X-mRNA transcripts are so far not detectable in liver tissues infected with HBV (Cattaneo *et al.,* 1984; Will *et al.,* 1987; Su *et al.,* 1989). A 0.8-kb X-mRNA transcript was demonstrated after transfection of cell lines with HBV DNA (Gough *et al.,* 1983; Simonsen and Levinson, 1983; Saito *et al.,* 1986; Siddiqui *et al.,* 1987; Zelent *et al.,* 1987, Koike *et al.,* 1988; Arii *et al.,* 1992; Runkel *et al.,* 1993; Yaginuma *et al.,* 1993) and in HBX-transgenic mice (Lee *et al.,* 1990; Kim *et al.,* 1991; Balsano *et al.,* 1993, 1994; Koike *et al.,* 1994a).

6. Spliced Transcripts

In liver tissues of chronically infected patients and in the hepatoblastoma cell line HepG2 (ATCC HB8065) spliced HBV transcripts have been described (P. J. Chen *et al.*, 1989; Suzuki *et al.*, 1989; Su *et al.*, 1989; Wu *et al.*, 1991). Single splicing would give rise to a 2.2-kb transcript encoding a *C*-terminally truncated HBcAg or an *N*-terminally truncated polymerase. The splice donor site is located at nt position 2445, the splice acceptor site at nt position 487. Double splicing results in a 2.0-kb transcript which encodes an HBc/HBs fusion protein or an *N*-terminally truncated polymerase. The splice donor sites for this event are at nt 2067 and 2740, the corresponding splice acceptor sites at nt positions 2350 and 282, respectively. Since mutations of the splice donor and acceptor sites do not influence viral replication (Su *et al.*, 1989; Wu *et al.*, 1991), spliced HBV transcripts do not seem to be essential for the viral life and replication cycle. However, encapsidation of the singly spliced 2.2-kb HBV RNA leads to the secretion of circulating defective HBV particles in patients with chronic hepatitis (Terre *et al.*, 1991). These data suggest a role of these defective particles in viral persistence. *In vitro* transfection of HuH7 cells with cloned cDNA corresponding to the singly spliced 2.2-kb transcript leads to intracellular accumulation of the hepatitis B core protein and increased secretion of hepatitis B e antigen (see Section II,E,2; Rosmorduc *et al.*, 1995).

D. Replication

Current models of HBV replication are largely derived from animal models. The most prominent feature of HBV replication is reverse transcription of pregenomic C-mRNA, which is also typical for retroviruses.

Circulating HBV particles are most likely taken up by the hepatocyte by specific binding of a viral surface protein to an unknown cellular receptor (Acs and Price, 1990; Rigg and Schaller, 1992; Fig. 3, [1]). Polymerized human serum albumin (Pontisso *et al.*, 1989), endonexin II (Hertogs *et al.*, 1993), and apolipoprotein H (Mehdi *et al.*, 1994) have been suggested possible candidates to facilitate virus entry. After the lipoprotein envelope is removed (Fig. 3, [2]) and the nucleocapsid is uncoated (Fig. 3, [3]), the partially double-stranded HBV genome is transported to the nucleus (Guidotti *et al.*, 1994). A DNA-dependent DNA polymerase fills up the *S*(+) strand after removal of the oligoribonucleotide primer (Köck and Schlicht, 1993). Covalently closed circular (ccc) DNA is produced by ligation of DNA ends (Fig. 3, [4]) after re-

moval of the protein primer from the *L*(–) strand and elimination of its terminal redundancy. This supercoiled DNA is the template for transcription and translation (Fig. 3, [5,6]) as well as pregenomic C-mRNA (Fig. 3, [6]; Büscher *et al.,* 1985; Enders *et al.,* 1985). During reverse transcription the nucleic acid is packaged in capsids together with the RNA-dependent DNA polymerase (Fig. 3, [7]; Enders *et al.,* 1987; Pollack and Ganem, 1994). An internal tyrosine residue at the *N* terminus of the HBV polymerase constitutes the first reaction partner for *L*(–) strand synthesis by reverse transcription, which is primed by a protein encoded by the 5′ region of the P genes and covalently linked to the 3′ end of DR I (Fig. 3, [8]; Mack *et al.,* 1988; Zoulim and Seeger, 1994). In parallel, the RNase H function of the HBV polymerase degrades the pregenomic C-mRNA, with exception of a 5′ oligoribonucleotide, which serves as primer for the *S*(+) strand synthesis after translocation to DR II (Wang and Seeger, 1992). After circulation of the genome, the synthesis of viral DNA stops before *S*(+) strand completion (Fig. 3, [9]), when the nucleocapsids enter the endoplasmic reticulum (ER). From there they are released in the circulation as enveloped viruses. Alternatively, the viral genome is transported to the nucleus (Fig. 3, [10]), where it initiates another round of replication (Huovila *et al.,* 1992).

Detailed reviews of hepadnaviral replication are given elsewhere (Seeger *et al.,* 1986; Ganem and Varmus, 1987; Will *et al.,* 1987; Schirmacher *et al.,* 1993; Nassal and Schaller, 1993; Ganem *et al.,* 1994).

E. Viral Proteins

At least seven different viral proteins are synthesized from HBV RNA: the large, middle, and small hepatitis B surface proteins, the hepatitis B core and e proteins, the RNA-dependent DNA polymerase with reverse transcriptase and RNase H activity, and the genome-bound protein have essential functions during HBV infection in humans. There is only indirect evidence for the production of a hepatitis B x protein, and so far no experimental proof at all for the existence of products encoded by ORF 5 or ORF 6.

1. Hepatitis B Surface Proteins

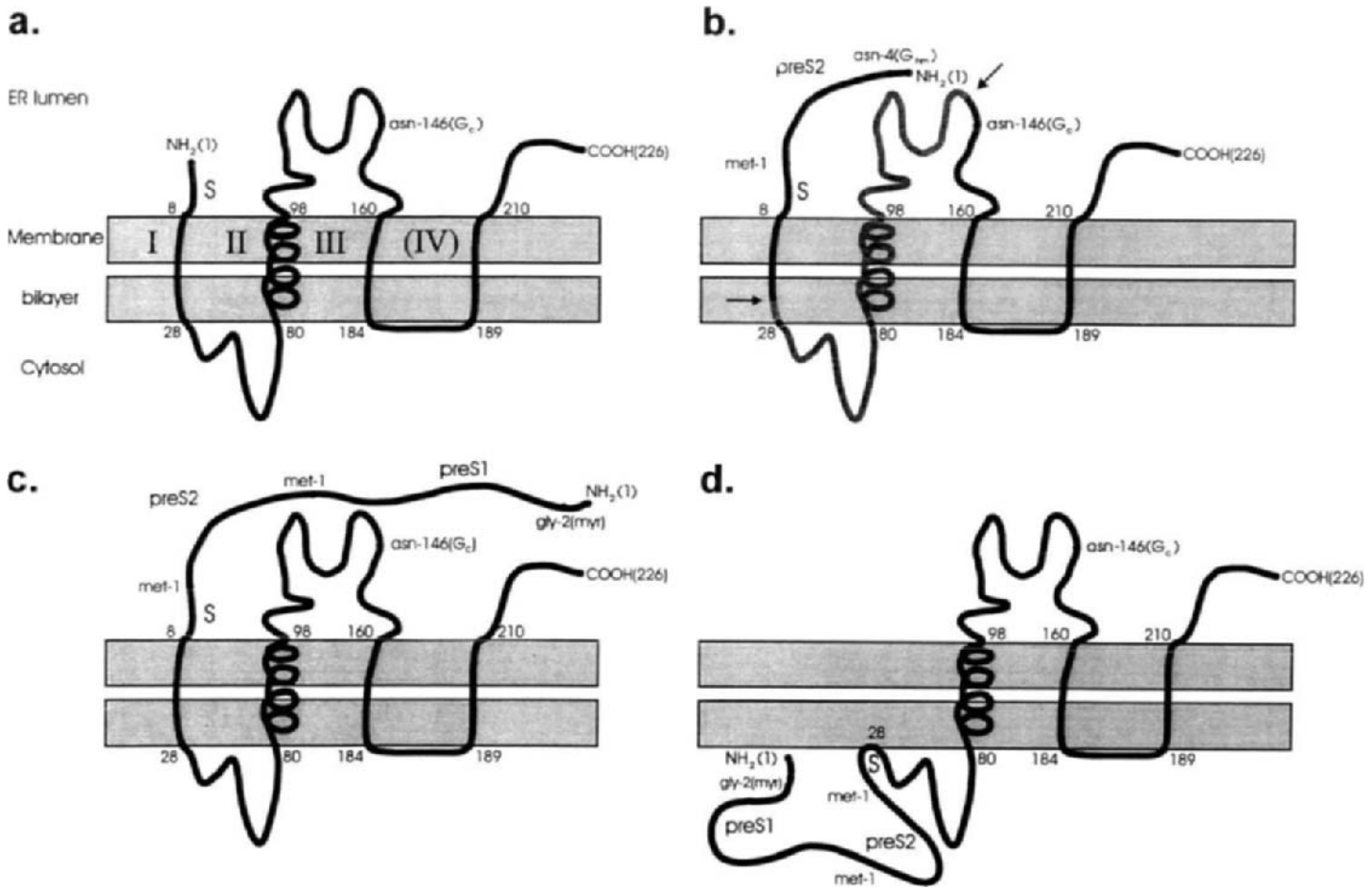
Three hepatitis B surface proteins of different size and modification are the constituents of the HBV envelope (Fig. 1). They mediate the adsorption of the virions to the hepatocyte surface (Fig. 3) and are believed to regulate virus assembly and amplification of supercoiled DNA (Neurath *et al.,* 1988; Summers, 1988). They are highly antigenic

and stimulate the production of virus-neutralizing antibodies (Krugman and Giles, 1973; Hollinger *et al.*, 1974). Therefore, they are used for the production of effective plasma-derived or recombinant antiviral vaccines (Krugman and Giles, 1973; McAleer *et al.*, 1984).

The small or major hepatitis B surface protein (SHBs; Fig. 4a) consists of 226 amino acids of the S region. It exists in an unglycosylated form ($p24^s$) and a monoglycosylated form ($gp27^s$) with a complex glycan residue at amino acid asparagine-146. It accounts for approximately 90% of the viral surface proteins in patients' sera. Both differently glycosylated proteins are equally represented in the viral envelope of viral particles (Heermann *et al.*, 1984). After aggregation of SHBs molecules, secretion of 22-nm SHBs particles occurs via the Golgi apparatus. One fourth of each particle consist of lipids derived from the ER membrane (Patzer *et al.*, 1986). The orientation of SHBs in the ER membranes displays the *N* and *C* termini to the ER lumen (Eble *et al.*, 1986, 1987; Bruss and Ganem, 1991; Fig. 4a). The S domain comprises three hydrophobic regions, two of which represent ER signal sequences (I, amino acids 11–28; II, amino acids 80–98) that initiate the translocation of *N*-terminal sequences and allow the anchoring of the protein within the ER membrane. The *C*-terminal region of the protein (amino acids 169–210) represents a third (and possibly fourth) transmembrane region of the third lipophilic domain. The ER regions (I and II) could have α-helical structure, which is unlikely for the entire ER III/IV domain. Since cells which synthesize SHBs exclusively can form regular 22-nm particles, it is likely that the S region contains the complete information for virus assembly (Liu *et al.*, 1982). Between amino acids 120 and 148 there is the major antigenic B-cell epitope of SHBs (Guerrero *et al.*, 1990).

The middle hepatitis B surface protein (MHBs) consists of 281 amino acids encoded by the preS2 and S region (Fig. 4b). It is found in two differentially glycosylated forms: $gp33^s$ is glycosylated at amino acid asparagine-4 of the preS2 region with a mannose-rich glycan residue with terminal sialinic acid. In addition, $gp36^s$ is glycosylated at amino acid asparagine-146 within the S region with a complex glycan residue. The glycosylation signal in the preS2 region is highly con-

FIG. 4. Two-dimensional model of hepatitis B virus surface proteins: gly, glycin; met, methionin; asn, asparagin; myr, myristyl residue; G_c, complex glycosylated glycan; G_{hm}, high-mannose glycosylation. Arabic numbers reflect amino acid positions; roman numbers reflect transmembrane domains. (a) SHBs, (b) MHBs, trans-activity-on region is depicted in gray and marked with arrows, (c) LHBs, (d) LHBs with alternative membrane topology of *N* terminus (modified according to Bruss *et al.*, 1994).


a.
ER lumen
NH2(1)
S
asn-146(Gc)
COOH(226)
8
98
160
210
Membrane
I
II
III
(IV)
bilayer
28
80
184
189
Cytosol
b.
preS2
asn-4(Gpre)
NH2(1)
asn-146(Gc)
met-1
S
COOH(226)
8
98
160
210
28
80
184
189
c.
preS2
met-1
preS1
NH2(1)
gly-2(myr)
asn-146(Gc)
met-1
S
COOH(226)
8
98
160
210
28
80
184
189
d.
asn-146(Gc)
COOH(226)
98
160
210
28
NH2(1)
S
80
184
189
gly-2(myr)
met-1
preS1
preS2
met-1

served in all hepadnaviruses. Therefore, it is speculated that it may have a function in virus adsorption to the hepatocyte membrane. A binding site for *in vitro* polymerized human serum albumin was identified at amino acids 13–29 of the preS2 region (Pontisso *et al.*, 1989). An immunogenic epitope is formed by amino acids 1–26 of the preS2 region. Furthermore, amino acids 47–55 carry recognition sites for the cellular proteases trypsin, chymotrypsin, and V8.

The large hepatitis B surface protein (LHBs) accounts for only 1–5% of all surface proteins, but represents up to 20% of viral envelope proteins. It exists in an unglycosylated form ($p39^s$) and a monoglycosylated form ($gp42^s$), which carries a complex glycan residue at amino acid asparagine-146 of the S region (Fig. 4c). LHBs has a virus-specific main determinant *a* which comprises amino acid 2–13 of the preS1 region as well as subtype-specific determinants *d/y* and *w/r*, which are mutually exclusive (LeBouvier *et al.*, 1972). This results in the four major HBV subtypes *adw, adr, ayw*, and *ayr*, which vary with respect to their geographic distribution (Mazzur *et al.*, 1974). Additional subtype-specific determinants *q, x*, and *g* and unusual combinations like *adyr* or *adywr* have been reported (Courouce *et al.*, 1976). In Europe, the United States of America, and Africa the predominant subtypes are *adw* and *ayw*, while subtype *adr* is frequently found in Southeast Asia. Subtype *ayr* is the least common worldwide.

Depending on the subtype, LHBs consists of 389–400 amino acids. The 108–119 residue preS1 domain appears to mask the preS2 and S region. Therefore, glycosylation at amino acid asparagine-4 of the preS2 region (Fig. 4c; asparagine-112 for HBV subtype *ayw*) does not take place. After hydrolysis of the *N*-terminal amino acid met-1 the preS1 domain can be myristylated at gly-2 (Persing *et al.*, 1987). The myristyl residue may help to anchor LHBs in the lipoprotein bilayer of the virus envelope. LHBs is cosecreted with nucleocapsids from the ER (Ueda *et al.*, 1991; Kuroki *et al.*, 1989). When LHBs is overexpressed as compared to the smaller hepatitis B surface proteins, it inhibits the secretion of the latter (Chisari *et al.*, 1986; Molnar-Kimber and Jarocki-Witek, 1988). Amino acids 21–47 of the preS1 region allow specific binding of HBV to the human hepatocyte membrane (Pontisso *et al.*, 1989; Petit *et al.*, 1992; Klingmüller and Schaller, 1993), suggesting a potential receptor function of LHBs. LHBs and MHBs are believed to have a similar topology to SHBs (see the preceding description; Bruss *et al.*, 1994). The accessibility of preS-specifc epitopes of both proteins in seceted particles and the glycosylation of MHBs at asparagine-4 support this notion (Heermann *et al.*,, 1984; Kuroki *et al.*, 1989). However, the lack of glycosylation of LHBs at asparagine-4 in the preS2

region may suggest an altered topology comprising cytoplasmic disposition of the *N* terminus (Ostapchuk *et al.*, 1994; Fig. 4d). Prange and Streek (1995) were recently able to demonstrate that posttranslational alteration of LHBs transmembrane localization occurs, resulting in two topologically different populations of LHBs maintained upon secretion of empty envelope particles. The sequences possibly accounting for this effect were mapped to amino acids 70–107 of LHBs.

2. *Hepatitis B Virus Core and E Proteins*

The hepatitis B virus core protein (HBcAg) is translated from the C-mRNA (see Section II,C,1). The 183 amino acid $p21^c$ has an arginine-rich basic *C* terminus with signal sequence function (Yeh *et al.*, 1990). The viral capsid, which takes up the pregenomic C-mRNA after entry into the cytoplasm, is made up of 180 units of HBcAg. The icosahedric 3-D structure of recombinant capsids was recently clarified by cryo-electronmicroscopy (Crowther *et al.*, 1994). Core particles possess an endogenous protein kinase activity which phosphorylates $p21^c$ at its *C*-terminal domain (Albin and Robinson, 1980; Feitelson *et al.*, 1982). Accordingly, phosphorylated HBcAg was detectable in the cytoplasm and unphosphorylated HBcAg in the nucleus of infected hepatocytes (Yeh *et al.*, 1993). Its localization is probably regulated in a cell-cycle-dependent manner; the relevance of the phosphorylation has not been completely clarified (Guidotti *et al.*, 1994; Kann and Gerlich, 1994). Core particles are strongly immunogenic. Antibodies against HBcAg are the first markers of acute infection (anti-HBc-IgM) and remain detectable as anti-HBc-IgG life-long (Cohen, 1978). The cellular immune response against HBcAg is considered a major pathogenic mechanism of HBV-induced liver damage (Mondelli *et al.*, 1982; Ferrari *et al.*, 1988; see Section IV,C).

The hepatitis B virus e antigen (HBeAg) is produced by proteolytic cleavage of a $p25^e$ precursor protein which is translated from the start codon of the preC region. A signal peptide encoded by the preC domain mediates the transport of the precursor protein to the ER (Ou *et al.*, 1986; Weimer *et al.*, 1987; Garcia *et al.*, 1988; Carlier *et al.*, 1995). After *N*- and *C*-terminal proteolytic processing to $p15–18^e$, HBeAg is secreted into the blood or medium. It can be used as sufficient marker for active virus replication (Krugman *et al.*, 1979; Takahashi *et al.*, 1983; Bruss and Gerlich, 1988).

3. *Hepatitis B Virus Polymerase*

The hepatitis B viral polymerase (*pol*) is translated from pregenomic C-mRNA (Ou *et al.*, 1990; see Section II,C,1). Theoretically, the P ORF

encodes an 832-amino-acid protein with a calculated molecular mass of about 93 kDa. *Pol* consists of at least three functionally different regions: the *N* terminus encoding a genome-bound protein linked to the 5′ end of the *L*(–) strand and serving as primer for reverse transcription (Bartenschlager and Schaller, 1988; Bosch *et al.*, 1988; see Section II,D). These sequences are separated by a spacer from the RNA-dependent DNA polymerase-encoding sequences. The 3′ region of *pol* displays RNase H activity (Toh *et al.*, 1983; Radziwill *et al.*, 1990).

These data have been obtained after overexpression of the P gene in bacterial or eukaryotic systems (Bartenschlager *et al.*, 1992; McGlynn *et al.*, 1992; Tavis *et al.*, 1993), since the *pol* protein could not be shown directly in infected hepatocytes. The presence of anti-*pol* antibodies in patients' sera (Stemler *et al.*, 1988) suggests that the *pol* protein is also synthesized *in vivo*. Polymerase activity is detectable in visions, in which the *pol* protein can be demonstrated by anti-*pol* antibodies (Bavand *et al.*, 1989; Mack *et al.*, 1988). It is essential for virus replication, as shown after transfection of hepatoma cell lines with a P gene mutant (Yaginuma *et al.*, 1987).

4. The Hepatitis B Virus x Protein

The hepatitis B virus x protein (HBx) consists of 154 amino acids resulting in a calculated molecular mass of 16.5 kDa. Translation usually starts at the first of three AUGs at the 5′ end of the X mRNA. Possibly due to its short half-life (Schek *et al.*, 1991), HBx has not been directly detected in patients' sera, but circulating anti-HBx antibodies suggest its expression (Moriarty *et al.*, 1985; Kay *et al.*, 1985; Elfassi *et al.*, 1986; Pfaff *et al.*,, 1987). HBx is detectable after overexpression in prokaryotic (Kay *et al.*, 1985; Meyers *et al.*, 1986; Elfassi *et al.*, 1986; Jameel *et al.*, 1990; M. L. Chen *et al.*, 1988) or eukaryotic expression systems (Klein *et al.*, 1991; Spandau *et al.*, 1991) or after *in vitro* translation (Pfaff *et al.*, 1987; Lin and Lo, 1989; Wu *et al.*, 1990). The codon usage of the X gene (Miller and Robinson, 1986) that is characteristic for eukaryotes suggests that the X gene may be of cellular origin. However, so far no significant sequence homology of HBx with cellular proteins has been reported.

The importance of HBx for viral replication is controversial. No virions were produced after injection of replication-competent WHV DNA-harboring mutations within the X gene into the liver of woodchucks (H. S. Chen *et al.*, 1993; Zoulim *et al.*, 1994), suggesting an essential role of HBx for the WHV infection. After transfection of cell lines with mutated X DNA neither viral transcription nor translation nor HBV replication was altered (Blum *et al.*, 1992). These data indi-

cate that HBx may play a less crucial role in human HBV infection. Nakatake *et al.* (1993), however, detected decreased levels of secreted HBsAg and HBeAg when hepatoma cells were transfected with X DNA comprising a stop mutation within the X gene.

Takada and Koike (1990) have described a sequence homology between HBx and Kunitz-type serine protease inhibitors which was mainly based on HBx amino acids 67-glycine-proline-cysteine-69, which are highly conserved within the active center of these protease inhibitors, and HBx region 135-glycine-glycine-cysteine-arginine-138. In all Kunitz-type serine protease inhibitors five highly conserved cysteins are found. However, none of these are present in HBx. In addition, taking its cleavability by serine proteases, its intracellular localization, and its physicochemical properties into account, the evidence that HBx represents a Kunitz-type serine protease inhibitor is only weak.

An intrinsic protein serine/threonine kinase function with autophosphorylating capacity has been claimed by Wu *et al.* (1990) for HBx. After purification of HBx from total cell extracts of recombinant *E. coli* by cell filtration and reverse phase high-pressure liquid chromatography the authors observed autophosphorylation and phosphorylation of histone H1. These experiments were not reproducible (Wu *et al.*, 1990, published erratum), which can most likely be explained by contaminating *E. coli* kinases. Similar results by Maguire *et al.* (1991) showing phosphorylation of histone H1 by immunoprecipitated HBx may have encountered the same problem with contaminating *E. coli* kinases. Only recently, a dinucleotide kinase activity has been demonstrated (De Medina and Shaul, 1994). Also in this case, contaminating *E. coli* kinases and dinucleotide kinases were not excluded as contributions to the observed activity. The trans-activating function of HBx and its possible relation to HBV-associated hepatocarcinogenesis will be discussed seperately below.

III. Hepatocellular Carcinoma

A. Epidemiology of HBV-Associated HCC

Epidemiological data suggest a causal relation between HBV infection and HCC development. There is a good geographic correlation between HBsAg prevalence and HCC incidence worldwide. In HBV-endemic areas like southern and Equatorial Africa as well as in Southeast Asia and Taiwan, where 5–20% of the population are chronic HBsAg carriers, 20–150 cases of HCC per 100,000 inhabitants are

newly diagnosed per year (Szmuness, 1978). In contrast, the HCC incidence in low HBV-endemic areas like Germany, in which the HBsAg prevalence is between 0.1 and 0.5, only 4 HCC cases are newly registered per 100,000 inhabitants a year.

In a large epidemiological study that was conducted in 22,707 Taiwanese male government employees aged 40–59 years, Beasley (1988) demonstrated prospectively an approximately 98-fold elevated relative risk for HBsAg-positive carriers to develop HCC as compared to HBsAg-negative male controls. Markers of persistent or past HBV infection are detectable in serum or liver tissue of more than 60% of HCC patients (Hadziyannis *et al.*, 1983; Thung, 1979). The clinical course of HBV infection with subsequent development of HCC is well documented (Omata *et al.*, 1982). Family clustering of HCC development in HBV-infected children of HBsAg-positive mothers has been reported (Tong, 1979).

B. Molecular Findings

HBV DNA is integrated in the hepatocyte genome in more than 90% of HBsAg-positive HCCs. The preferred viral site of recombination is the region between DR I and DR II. Rearrangements like deletions, duplications, inversions, and translocation of both the integrated HBV DNA and cellular DNA (Nagaya *et al.*, 1987; Tokino *et al.*, 1987; Tokino and Matsubara, 1991; Mizusawa *et al.*, 1985; Koike *et al.*, 1983; Meyer *et al.*, 1992a) often occur during repeated integrations.

The genomic integration of HBV DNA is considered to be crucial for hepatocarcinogenesis. Since HBV does not carry an acutely transforming oncogene, HBV DNA integration could on theoretical grounds alter cellular gene expression by cis and/or trans mechanisms.

1. Cis (In-)Activation of Cellular Genes

Insertional mutagenesis by HBV as a result of directed integration of HBV DNA in cellular oncogenes or tumor suppressor genes has been studied extensively. In HCCs, allelic losses have been documented on chromosomes 1 (Simon *et al.*, 1991; Ding *et al.*, 1991; Ding and Habib, 1995), 4 (Zhang *et al.*, 1990; Walker *et al.*, 1991; Buetow *et al.*, 1989), 5 (Ding *et al.*, 1991; Ding *et al.*, 1993), 8 (Emi *et al.*, 1992; Slagle *et al.*, 1991), 10 (Fujimori *et al.*, 1991), 11 (Rogler *et al.*, 1985; Fujimori *et al.*, 1991), 13 (Walker *et al.*, 1991), 16 (Slagle *et al.*, 1991; Zhang *et al.*, 1990; Fujimori *et al.*, 1991), 17 (Slagle *et al.*, 1991; Fujimori *et al.*, 1991; Bressac *et al.*, 1991; Ding *et al.*, 1991; Scorsone *et al.*, 1992; Kress *et al.*, 1992; Hosono *et al.*, 1993; Nishida *et al.*, 1993), and 22 (Takahashi *et al.*, 1993). However, none of these losses has yet led to identification of

any HCC-specific tumor suppressor genes. The regions 11p13–15 bearing the Wilms tumor suppressor gene, 13q14 comprising the retinoblastoma gene, and 17p13 containing the p53 tumor suppressor gene were studied in detail. Allelic deletion of 11p13–14 was associated with HBV integration in one HCC (Rogler *et al.,* 1985) and detectable in 6 of 14 HBV-associated HCCs (Wang and Rogler, 1988). Allelic deletions at 13q12–qter was described in 5 of 10 HBV-associated HCCs (Wang and Rogler, 1988). Other authors (Walker *et al.,* 1991; Fujimori *et al.,* 1991) described HCC allele losses on 11p and 13q irrespective of the HBV status. Aberrations in the p53 tumor suppressor gene seem to be common genetic changes in human cancers, including HCC (Levine *et al.,* 1991; Hollstein *et al.,* 1991). Loss of heterozygosity at the 17p13 locus seems to occur irrespective of ethnic variation (Bressac *et al.,* 1991; Ding *et al.,* 1991; Slagle *et al.,* 1991; Emi *et al.,* 1992) in chronic HBV infection (Fujimori *et al.,* 1991) and cirrhosis (Walker *et al.,* 1991). Specific G to T point mutations within codon 249 of p53 leading to an arginine-to-serine substitution seem to occur more frequently in patients exposed to aflatoxin B1, but do not seem to be an HBV-specific event (Hosono *et al.,* 1993; Shieh *et al.,* 1993; Gerbes and Caselmann, 1993; see Section III,B,1).

Insertional activation of expression of cellular genes possibly relevant for proliferation or malignant hepatocyte transformation has been documented in only a few cases. Wang *et al.* (1993) identified integration of HBV DNA in an intron of the human cyclin A gene leading to a deletion of cellular sequences responsible for cyclin A degradation (J. Wang *et al.,* 1992). This led to impaired cyclin A degradation and severe alteration of cell cycle regulation. A second example is the integration of HBV DNA in cellular DNA homologous to the coding sequences for the tyrosine protein kinase domain of epidermal growth factor receptor (Zhang *et al.,* 1992). A third integration was demonstrated in a cellular gene similar to the v-*erb* A oncogene, which encodes a previously unknown retinoid acid receptor (Dejean *et al.,* 1986; De Thé *et al.,* 1987). The most recent HBV DNA integration characterized occurred in the cellular mevalonate kinase gene causing insertional activation of mevalonate kinase in the hepatocarcinoma cell line PLC/PRF/5 and transcription of HBV-mevalonate kinase fusion transcripts (Graef *et al.,* 1994; Graef *et al.,* 1995). Overexpression of mevalonate kinase, which is the key enzyme of the cholesterol biosynthesis pathway and is indirectly involved in the farnesylation of growth-related proteins like the *ras* oncoprotein, may contribute to the dysregulation of hepatocyte growth.

Insertional (in-)activation of cellular genes by integrated DNA has only been observed in rare cases, although approximately 200 HBV

DNA integrations have been cloned. Therefore, cis-activation of cellular genes appears not to be the general mechanism of HBV-associated hepatocarcinogenesis in humans. Cis-mechanisms seem to be particularly important in WHV-associated hepatocarcinogenesis in woodchucks, which develop HCCs within 2 years in nearly 100% of cases (Popper *et al.*, 1987). Specific integration of WHV DNA into N-*myc*-1 and N-*myc*-2 sequences are common (Möröy *et al.*, 1986; Fourel *et al.*, 1990; Hansen *et al.*, 1993). Insertional activation of *myc* genes, mainly the intronless N-*myc*-2 oncogene, can be found in more than 50% of infected woodchucks. Cloning analysis of single integration sites in four woodchuck tumors carrying wild-type *myc* alleles led to the identification of an additional preferred integration locus (*win*) on the long arm of the woodchuck X chromosome, from which abundant production of N-*myc*-2 transcripts can be stimulated over a distance up to 180 kb (Fourel *et al.*, 1994).

7. Trans-Activation of Cellular Gene Expression

In the absence of preferred cellular integration sites, integrated HBV DNA could still exert an effect on cellular gene expression by transmechanisms. Transcriptional trans-activators that can cause malignant transformation have been shown in many other DNA tumor viruses, such as adenoviruses, human papilloma viruses type 16 and 18, or the Epstein-Barr virus, which encode E1A (Raychaudhury *et al.*, 1991), E6, E7 (Phelps *et al.*, 1988; Scheffner *et al.*, 1990), and EBNA-2 proteins (Cohen and Kieff, 1991) and are common in retroviruses (see Section I,A). A trans-activating function of HBx was postulated by Miller and Robinson (1986). Shortly afterward, the trans-activation by HBx was experimentally shown by Twu and Schloemer (1987). Wollersheim *et al.* (1988) and Zahm *et al.* (1988) could demonstrate that the trans-activating potential of the X gene is retained also in the integrated state.

IV. HBV Trans-Activators in Hepatocarcinogenesis

A. Hepatitis B Virus x Protein

1. Trans-Activation by HBx

Commonly used trans-activation assays are based on cotransfection of cells or cell lines with test and reporter plasmids. When the bacterial chloramphenicol acetyltransferase (*cat*) reporter gene is used, which is not found in eukaryotic cells, differently acetylated forms of radio-

actively labeled ^{14}C-chloramphenicol (1-, 3-, and 1,3-acetyl-^{14}C-chloramphenicol) can be separated by thin-layer chromatography (Gorman *et al.*, 1982) and quantified by an isotope scanner. Alternatively, a commercially available *cat* ELISA (Boehringer Mannheim, Mannheim, Germany) can be used for direct quantification of the *cat* enzyme with polyclonal anti-*cat* antibodies. An alternative sensitive and rapid method to determine trans-activation by test DNA is the luciferase (*luc*) assay (De Wet *et al.*, 1987), which uses *luc* cDNA derived from *Photinus pyralis* as reporter plasmid. The luciferase catalyzes oxidation of luciferin in the presence of ATP and Mg^{2+}, resulting in the production of oxyluciferin, AMP, and CO_2 as well as light emission, which can be quantified in a luminometer. If appropriate controls are used and the transfection efficiency is standardized, acetylation of chloramphenycol, production of the *cat* enzyme, or the light emission during oxyluciferin production correlates positively with the trans-activation exerted by the test DNA. The transferability of these *in vitro* results to the stimulation of cellular gene expression *in vivo* is presently under investigation.

In most experiments transient transfections of cell cultures with trans-activator DNA have been used to determine the trans-activatory potential of a derived protein. Only in rare cases has trans-activation by HBx been demonstrated in stably transfected cell lines carrying integrated HBV DNA (Twu *et al.*, 1989) or after scrape loading of trans-activator proteins (S. Urban, 1995, personal communication). There is a single report in which the trans-activating potential of HBx was shown *in vivo* (Balsano *et al.*, 1993, 1994). Transgenic mice expressing the X gene under the control of the human antithrombin III regulatory sequences were crossbred with transgenics carrying either the *cat* or the β-galactosidase (*lac*Z) reporter gene under the control of the HIV-1 LTR, which is a known target of trans-activation by HBx (see Section IV,A,2). Expression of HBx in the liver of these double transgenic mice stimulated the expression of either the *cat* or *lac*Z reporter gene by a factor of 2.5 to 6.0. No significantly enhanced reporter gene expression was observed in tissues like spleen, brain, and heart which do not express X gene sequences.

The X gene can encode two or possibly three polypeptides of different lengths from a single mRNA using alternate translation initiation from any of three in-frame AUG codons (Kwee *et al.*, 1992; Zheng and Yen, 1994). The two *N*-terminally truncated HBx proteins have sizes of about 8.0 and 6.6 kDa. Each of these defined HBx proteins is able individually to trans-activate RNA polymerase III–transcribed promoters. In contrast, combinations of different HBx species were required

by RNA polymerase II–transcribed promoters, while activation of an NF-κB-dependent promoter was exclusively achieved by a full-length HBx protein.

2. *Target Sequences of Trans-Activation*

Initial trans-activation experiments were performed with a reporter construct, in which the reporter gene was driven by the heterologous SV40 early promoter/enhancer (Twu and Schloemer, 1987; Wollersheim *et al.*, 1988; Zahm *et al.*, 1988). In the following years various autologous (Table I), heterologous viral (Table II), and cellular regulatory

TABLE I

HBV REGULATORY SEQUENCES AS TARGETS OF HBX-MEDIATED TRANS-ACTIVATION[a]

Target sequence	Cell line	Trans-activation	Reference
CP/EnhI/II	CV-1	–	Jameel and Siddiqui (1986)
	HepG2	+	Koike *et al.* (1989)
		–	Spandau and Lee (1988)
	huH7	+	Spandau and Lee (1988)
		–	Koike *et al.* (1989)
	Jurkat	+	Siddiqui *et al.* (1989), Cross *et al.* (1993)
	PLC/PRF/5	+	Colgrove *et al.* (1989)
S1P/EnhI/II	HepG2	+	Koike *et al.* (1989)
	huH7	+	Rossner *et al.* (1990)
		–	Koike *et al.* (1989)
		–	Zhou *et al.* (1990)
S2P/EnhI/II	CV-1	+	Colgrove *et al.* (1989)
	HepG2	+	Koike *et al.* (1989)
	huH7	+	Rossner *et al.* (1990)
		–	Koike *et al.* (1989)
	Jurkat	+	Cross *et al.* (1993)
XP/EnhI	CCl13	+	Wollersheim *et al.* (1988)
	HeLa	+	Siddiqui *et al.* (1989), Faktor *et al.* (1990), Balsano *et al.* (1991), Renner *et al.* (1995)
	HepG2	+	Levrero *et al.* (1990)
	PLC/PRF/5	+	Unger and Shaul (1990)
	Vero	–	Twu and Schloemer (1987)

[a] Modified according to Rossner (1992).

TABLE II

HETEROLOGOUS VIRAL REGULATORY SEQUENCES AS TARGETS OF HBx-MEDIATED TRANS-ACTIVATION[a]

Target sequence	Cell line	Trans-activation	Reference
HIV-1 LTR	CV-1	+	Siddiqui *et al.* (1989)
	HeLa	+	Twu and Robinson (1989)
		–	Seto *et al.* (1990)
	HepG2	+	Ritter *et al.* (1990)
		–	Seto *et al.* (1988)
	huH7	–	Twu and Robinson (1989)
	Jurkat	+	Twu and Robinson (1989)
	NIH3T3	–	Seto *et al.* (1989)
	PLC/PRF/5	–	Seto *et al.* (1989)
	Vero	+	Seto *et al.* (1989), Jameel *et al.* (1990), Cross *et al.* (1993)
	HepG2	+	Levrero *et al.* (1990)
HSV TK	CCl13	+	Zahm *et al.* (1988)
	CV-1	+	Cross *et al.* (1993)
	HepG2	–	Twu *et al.* (1989)
	Jurkat	+	Cross *et al.* (1993)
		–	Seto *et al.* (1988)
HTLV-I LTR	CCl13	+	Zahm *et al.* (1988)
	CV-1	–	Cross *et al.* (1993)
	HepG2	+	Wollersheim *et al.* (1988), Siddiqui *et al.* (1989)
		–	Twu and Robinson (1989)
	Jurkat	–	Cross *et al.* (1993)
	Vero	–	Twu and Schloemer (1987)
MMTV LTR	CCl13	+	Wollersheim *et al.* (1988)
		+	Zahm *et al.* (1988)
RSV LTR	CCl13	+	Zahm *et al.* (1988)
	CV-1	+	Colgrove *et al.* (1989)
	HeLa	+	Spandau and Lee (1988)
	HepG2	+	Balsano *et al.* (1991)
	Jurkat	+	Cross *et al.* (1993)
		–	Seto *et al.* (1989)
	NIH3T3	+	Seto *et al.* (1989)
	PLC/PRF/5	+	Ritter *et al.* (1991)
	Vero	+	Cross *et al.* (1993)
		–	Twu and Schloemer (1987)
SV40 eP	HepG2	–	Siddiqui *et al.* (1989)
	NIH3T3	+	Koike *et al.* (1989)
	PLC/PRF/5	–	Faktor *et al.* (1990)

(continued)

TABLE II *(continued)*

Target sequence	Cell line	Trans-activation	Reference
SV40 eP/EnH	CCl13	+	Renner *et al.* (1995)
	COS-1	–	Spandau and Lee (1988)
	CV-1	+	Wollersheim *et al.* (1988)
	HeLa	+	Aufiero and Schneider (1990)
	HepG2	+	Arii *et al.* (1992)
	Jurkat	+	Twu *et al.* (1989), Colgrove *et al.* (1989), Spandau and Lee (1988), Ritter *et al.* (1991), Cross *et al.* (1993)
	NIH3T3	+	Koike *et al.* (1989)
	PLC/PRF/5	+	Zahm *et al.* (1988)
	Vero	+	Siddiqui *et al.* (1989)
		–	Zahm *et al.* (1988)
CMV/Enh/P	CV-1	+	Cross *et al.* (1993)

[a] Modified according to Rossner (1992).

sequences (Table III) as well as isolated binding sites for nuclear transcription factors (Table IV) have been identified as target sequences of HBx-mediated trans-activation. These experiments were performed to identify common sequence motifs susceptible to HBx-mediated trans-activation. However, as summarized in Tables I–IV, the efforts could not satisfactorily achieve this goal and some data are still contradictory due to the different test conditions used. There are unsolved discrepancies like the HBx-mediated trans-activation of the HBV C promoter/enhancer I/II in HepG2 and huH7 cells (Koike *et al.*, 1989; Spandau and Lee, 1988) or of the S1 or S2 promoter/enhancer I/II in huH7 cells (Rossner *et al.*, 1990; Koike *et al.*, 1989; Zouh *et al.*, 1990). There are discordant findings concerning the herpes simplex virus (HSV) thymidine kinase (TK) promoter in Jurkat cells (Cross *et al.*, 1993; Seto *et al.*, 1988), the human T-cell lymphotropic virus type I (HTLV-I) LTR in HepG2 cells (Wollersheim *et al.*, 1988; Siddiqui *et al.*, 1989; Twu and Robinson, 1989), the HIV-1 LTR in HepG2 cells (Ritter *et al.*, 1991; Seto *et al.*, 1988) and HeLa cells (Twu and Robinson, 1989; Seto *et al.*, 1989), the Rous sarcoma virus (RSV) LTR in Jurkat (Cross *et al.*, 1993; Seto *et al.*, 1989) and Vero cells (Cross *et al.*, 1993; Twu and Schloemer, 1987), and the SV40 early promoter/enhancer in Vero cells (Siddiqui *et al.*, 1989; Zahm *et al.*, 1988). In part, these discrepancies are due to varying experimental conditions like molar ratio of test and reporter plasmids, cultivation conditions of cell lines, or quality of

TABLE III

CELLULAR REGULATORY SEQUENCES AS TARGETS OF HBx-MEDIATED TRANS-ACTIVATION[a]

Target sequence	Cell line	Trans-activation	Reference
β-actin	CCl13	–	Levrero *et al.* (1990), Zahm *et al.* (1988)
	CV-1	+	Cross *et al.* (1993)
	HepG2	–	Siddiqui *et al.* (1989)
	H938	–	Twu *et al.* (1989)
	Jurkat	+	Cross *et al.* (1993)
Serum albumin	HepG2	–	Rossner *et al.* (1990)
c-*fos*	HeLa	+	Meyer *et al.* (1992b), Avantaggiati *et al.* (1992, 1993)
α-globin	PLC/PRF/5	–	Faktor *et al.* (1990), Unger and Shaul (1990)
H-2K/MHC 1	HepG2	+	Zhou *et al.* (1990)
HLA DR/MHC II	HepG2, huH7	+	Hu *et al.* (1990)
β-interferon	HepG2	–	Zhou *et al.* (1990)
	Vero	+	Twu and Schloemer (1988)
Metallothionein	CCl13	+	Zahm *et al.* (1988)
	CV-1, Jurkat	+	Cross *et al.* (1993)
c-*myc*	HepG2	+	Balsano *et al.* (1991)
	HeLa	+	Koike *et al.* (1989), Avantagiatti *et al.* (1992, 1993)
	NIH3T3	+	Koike *et al.* (1989)
$tRNA^{Ala}$	CCH3	+	Aufiero *et al.* (1990)

[a] Modified according to Rossner (1992).

DNAs used in different experiments. Furthermore, most authors failed to show the presence of HBx protein or RNA directly in their trans-activation assays.

With respect to the cellular target elements identified *in vitro* (Table III) the promoters of c-*myc*, c-*fos*, and c-*jun* may also be relevant *in vivo*; these genes are known to be involved in cell proliferation and transformation. While c-*myc* is strongly expressed in the human hepatocarcinoma cell line huH-4 (Kekulé *et al.*, 1990), c-*jun* is overexpressed in HCCs or liver tissue of HBV-infected patients as compared to normal liver (Twu *et al.*, 1993).

Analysis of distinct transcription-factor binding sites allowed to identify binding motifs for AP-1, AP-2, NF-κB, and SP-1 as target sequences of the pleiotropic HBx-mediated trans-activation (Table IV). All binding motifs do not have apparent sequence homology and therefore argue against a direct interaction of HBx with these DNA se-

TABLE IV

TRANSCRIPTION FACTOR BINDING SITES AS TARGETS OF HBx-MEDIATED TRANS-ACTIVATION[a]

Target sequence	Cell line	Trans-activation	Reference
AP-1	CCl13	+	Kekulé *et al.* (1993)
	CV-1	+	Seto *et al.* (1990)
	HepG2	+	Schlüter and Caselmann (1995)
AP-2	CCl13	+	Kekulé *et al.* (1993)
	CV-1	+	Seto *et al.* (1990)
	HepG2	–	Twu *et al.* (1989)
AP-3	CV-1	–	Seto *et al.* (1990)
NF-κB	CCl13	+	Kekulé *et al.* (1993)
	Cos7	+	Meyer *et al.* (1992b)
	HepG2	+	Twu *et al.* (1989), Schlüter and Caselmann (1995)
	HeLa	+	Meyer *et al.* (1992b)
	PLC/PRF/5	+	Faktor *et al.* (1990)
SP-1	Jurkat	+	Seto *et al.* (1990)

[a] Modified according to Rossner (1992).

quences. Some of the experiments showing the increased nuclear appearance of NF-κB or AP-1 have been performed as electromobility shift analysis. In these experiments the binding of distinct transcription factors to their radioactively labeled oligonucleotide binding sites led to specific retardation of these DNA/protein complexes during gel electrophoresis (Meyer *et al.*, 1992b; Schlüter and Caselmann, 1995).

3. The Functional HBx Trans-Activator Domain

Detailed deletion, insertion, and point mutation analysis of the X gene has led to the identification of X gene regions essential or dispensable for the HBx trans-activating function. Approximately 12 amino acids of the *C* terminus (Levrero *et al.*, 1990; Unger and Shaul, 1990; Balsano *et al.*, 1991; Ritter *et al.*, 1991; Renner *et al.*, 1995) and at least 50 amino acids of the *N* terminus of the X gene (Ritter *et al.*, 1991, Arii *et al.*, 1992; Renner *et al.*, 1995) can be deleted without abolishing the trans-activating function of HBx. Point mutation and insertion experiments point out the relevance of amino acids 67–69 and 110–139 for trans-activation by HBx (Arii *et al.*, 1992; Runkel *et al.*, 1993). The detailed compilation is summarized in Table V. There are also some minor contradictory findings, which may again be explained by the different test systems that varied in the target sequences and transfected cell lines used, as described above.

TABLE V

ESSENTIAL AND DISPENSABLE HBx DOMAINS FOR TRANS-ACTIVATION[a]

Mutation	Target sequence	Cell line	Trans-activation	Reference
N-terminal deletions				
Δ2-9	SV40 eP/Enh	HepG2, Vero	+	Ritter *et al.* (1991)
Δ2-9	RSV LTR	HepG2, Vero	+	Ritter *et al.* (1991)
Δ2-10	SV40 eP/Enh	CCl13, HepG2	+	Renner *et al.* (1995)
Δ2-30	CCl13, HepG2			Renner *et al.* (1995)
Δ2-30	SV40 eP/Enh	HepG2, Vero	+	Ritter *et al.* (1991)
Δ2-30	RSV LTR	HepG2, Vero	+	Ritter *et al.* (1991)
Δ2-30	HBV Enh/XP	PLC/PRF/5	–	Faktor and Shaul (1990)
Δ2-30	HBV Enh/SV40 eP	PLC/PRF/5	–	Faktor and Shaul (1990)
Δ2-48	SV40 eP/Enh	CCl13, HepG2	+	Renner *et al.* (1995)
Δ2-52	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
Δ2-48, Δ111-123	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ2-50	pHECx2CAT	HepG2	+	Murakami *et al.* (1994a)
Δ2-70	SV40 eP/Enh, MMTV LTR	HepG2, A549	–	Kim *et al.* (1993)
Δ2-71	pHECx2CAT	HepG2	–	Murakami *et al.* (1994a)
Δ2-73	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ2-105	pHECx2CAT	HepG2	–	Murakami *et al.* (1994a)
Internal deletions				
Δ5-27	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
Δ5-62	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ5-87	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ11-60	SV40 eP/Enh	HepG2, Vero	–	Ritter *et al.* (1991)
Δ11-60	RSV LTR	HepG2, Vero	–	Ritter *et al.* (1991)
Δ32-66	SV40 eP/Enh	HepG2, Vero	–	Ritter *et al.* (1991)
Δ49-90	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ49-136	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ49-143	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ118-133	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ118-143	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
C-terminal deletions				
Δ68-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ73-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ83-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ87-154	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ91-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ92-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ104-154	HBV Enh/SV40 eP	HepG2	–	Levrero *et al.* (1990)
Δ104-154	HIV-1 LTR	HepG2	–	Levrero *et al.* (1990)

(continued)

TABLE V *(continued)*

Mutation	Target sequence	Cell line	Trans-activation	Reference
Δ105-154	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ110-154	HBV-E site	HepG2	–	Unger and Shaul (1990)
Δ110-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ110-154	HBV Enh/XP	HepG2	–	Unger and Shaul (1990)
Δ114-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ118-154	HBV Enh/SV40 eP	HepG2	+	Levrero *et al.* (1990)
Δ118-154	HIV1-LTR	HepG2	+	Levrero *et al.* (1990)
Δ118-154	HIV1-LTR	HepG2	+	Levrero *et al.* (1990)
Δ118-154	HBV Enh/XP	HepG2, HeLa	+	Balsano *et al.* (1991)
Δ118-154	c-*myc* p	HepG2, HeLa	+	Balsano *et al.* (1991)
Δ118-154	SV40 eP/Enh	CCl13	–	Caselmann *et al.* (1990)
Δ118-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ122-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ130-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ114-154	SV40 eP/Enh	HepG2	–	Takada and Koike (1990)
Δ134-154	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
Δ137-154	SV40 eP/Enh, MMTV LTR	HepG2, A549	–	Kim *et al.* (1993)
Δ137-154	SV40 eP/Enh	CCl13, HepG2	–	Renner *et al.* (1995)
Δ143-154	HBV Enh/SV40 eP	HepG2	+	Levrero *et al.* (1990)
Δ143-154	HIV1-LTR	HepG2	+	Levrero *et al.* (1990)
Δ143-154	HBV Enh/XP	HepG2, HeLa	+	Balsano *et al.* (1991)
Δ143-154	c-*myc* promoter	HepG2, HeLa	+	Balsano *et al.* (1991)
Δ144-154	HBV Enh/XP	HepG2	+	Unger and Shaul (1990)
Δ144-154	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
Δ144-154	SV40 eP/Enh	CCl13	+	Renner *et al.* (1995)
Δ148-154	SV40 eP/Enh	HepG2, Vero	+	Ritter *et al.* (1991)
Δ148-154	RSV LTR	HepG2, Vero	+	Ritter *et al.* (1991)
Δ149-154	pHECx2CAT	HepG2	+	Murakami *et al.* (1994a)
Δ137-154	pHECx2CAT	HepG2	–	Murakami *et al.* (1994a)
Point mutations				
31 S→D	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
39 S→R	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
43 S→R	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
46 P→N	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
49 H→I	SV40 eP/Enh	HepG2	+/–	Arii *et al.* (1992)
52 H→G	SV40 eP/Enh	HepG2	+/–	Arii *et al.* (1992)
58 L→R	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
61 C→D	SV40 eP/Enh	HepG2	+/–	Arii *et al.* (1992)
64 S→R	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
69 C→A	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
74 T→K	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)

TABLE V *(continued)*

Mutation	Target sequence	Cell line	Trans-activation	Reference
82 T→R	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
107 D→V	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
111 Y→G	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
114 D→A	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
126 E→F	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
132 F→D	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
134 L→R	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
137 C→S	SV40 eP/Enh	HepG2	+/–	Arii *et al.* (1992)
138 R→G	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
139 H→P	SV40 eP/Enh	HepG2	+/–	Arii *et al.* (1992)
140 K→A	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
Double-point mutations				
13 RD→LV	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
52-HL→DI	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
155 LR→SC	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
67-GP→RA	SV40 eP/Enh	HepG2	–	Arii *et al.* (1992)
72-RF→VD	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
94 AK→LT	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
121-EE→LQ	SV40 eP/Enh	HepG2	+	Arii *et al.* (1992)
121 EE→VV	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
138 RH→TY	SV40 eP/Enh, MMTV LTR	HepG2, A549	+	Kim *et al.* (1993)
Insertional mutations				
5-L**RP**C-6	AP-1	HepG2	+	Runkel *et al.* (1993)
24-E**RP**S-25	AP-1	HepG2	+	Runkel *et al.* (1993)
36-T**RP**L-37	AP-1	HepG2	+	Runkel *et al.* (1993)
45-V**KP**P-46	AP-1	HepG2	+	Runkel *et al.* (1993)
58-L**RP**P-59	AP-1	HepG2	–	Runkel *et al.* (1993)
68-P**RP**C-69	AP-1	HepG2	–	Runkel *et al.* (1993)
88-I**RP**L-6	AP-1	HepG2	+	Runkel *et al.* (1993)
99-G**RP**L-100	AP-1	HepG2	+	Runkel *et al.* (1993)
110-A**RP**Y-111	AP-1	HepG2	–	Runkel *et al.* (1993)
119-D**RP**W-120	AP-1	HepG2	–	Runkel *et al.* (1993)
128-R**RP**L-129	AP-1	HepG2	–	Runkel *et al.* (1993)
138-R**RP**H-139	AP-1	HepG2	–	Runkel *et al.* (1993)
147-P**RP**C-148	AP-1	HepG2	+	Runkel *et al.* (1993)

[a] Kindly prepared by S. Urban and M. Renner, 1995.

4. Mechanism of Trans-Activation

Trans-activators of many DNA tumor viruses either bind directly to specific DNA sequences or exert their trans-activating potential by interacting with the transcriptional machinery, cellular signal transduction pathways, or unrelated cellular elements. So far, a direct and specific interaction between HBx and distinct cellular DNA sequences has not been shown by several groups (Jameel *et al.*, 1990; Twu *et al.*, 1990; Wu *et al.*, 1990; Maguire *et al.*, 1991; Ritter *et al.*, 1991). Therefore, HBx protein–protein interactions are more likely.

a. Binding of HBx to transcription factors and interaction with the transcriptional machinery. Maguire *et al.* (1991) were able to show binding of HBx to the CREB and ATF-2 transcription factors, which are part of the leucin zipper family. In electromobility shift analysis both transcription factors complexed with HBx bound to a labeled HBV enhancer I fragment (Williams and Andrisani, 1995). In the absence of HBx only weak binding was observed. Nevertheless, this explains only a limited set of identified target sequences (see Section IV,A,3), because only a few of them contain binding sites for CREB or ATF-2.

Qadri *et al.* (1995) were able to demonstrate that HBx interacts with the basal transcription factor TATA-binding protein (TBP) *in vitro*. Deletion analysis suggested that a domain consisting of 71 amino acids in the highly conserved carboxyl terminus of TBP is necessary for this interaction with HBx amino acids 110–143.

The screening of a λgt11 cDNA library of the HepG2 cell line with a recombinant HBx protein led to the identification of proteins that are able to bind HBx. One of these proteins was homologous to a submit shared by RNA polymerases (RP). The HBx trans-activation domain and the central region of human RP B5 were necessary for specific binding as shown *in vitro* with deletion mutants and in transfected HepG2 cells (Cheong *et al.*, 1995). These results support the notion that interaction of HBx and RP B5 can facilitate trans-activation and that RP B5 contains a domain that is able to communicate with transcriptional regulators.

Furthermore, Haviv *et al.* (1995) suggested that the effect of HBx on transcription involves a coactivation process, because HBx cannot activate certain DNA elements like AP-1 in the absence of the corresponding cellular activators. HBx preferentially coactivates potent activators, especially those with acidic activation domains like the *N*-terminal 65 amino acids of *p53* or the portion *C*-terminal to the bzip domain of *fos*. Weak activators with nonacidic activation domains like the *N*-terminal

portion of *jun* are not potentiated by HBx. These data provide a tempting concept to explain the molecular basis for the pleiotropism of the HBx effect on transcription. However, they would favor a nuclear localization of HBx, which is in discordance with the cytosolic HBx appearance demonstrated by Schek *et al.* (1991).

Biochemical analysis of bacterially produced HBx revealed a hydrolytic activity specific for adenine nucleotides (ATPase activity), which is not DNA-dependent (De Medina *et al.,* 1994). For maximum activity HBx amino acids 88–119 were required. The role of ATPases in transcription complexes is only poorly understood. ATP hydrolysis is required in at least two steps of the activation of the preinitiation complex: ATP hydrolysis facilitates unwinding of the DNA template at the site of transcription initiation (W. Wang *et al.,* 1992) and is required for phosphorylation of the RNA polymerase II, which is necessary for activation of the preinitiation complex (Chesnut *et al.,* 1992). Alternatively, the HBx ATPase activity may be involved in phosphorylation of an enhancer-binding protein and thereby stimulate transcription (De Medina *et al.,* 1994).

By means of two-hybrid protein interaction screening, HBx interaction with the *C*-terminal portion of a previously unknown human proteasome alpha-subunit was found (Fischer *et al.,* 1995). Insertion of two amino acids behind amino acid 128 abolished the interaction. This region is known to be essential for the trans-activatory potential of HBx (see Table V). HBx may therefore interfere with degradative processes by binding to a specific proteasome alpha-subunit. This may lead to an increased half-life of different transcription factors and other nuclear regulatory proteins responsible for transcriptional activation. However, considering the numerical superiority of proteasomes in relation to the limited number of HBx molecules, this concept is less attractive as a general mechanism of HBx action.

b. Interaction of HBx with the protein kinase C signal transduction pathway. HBx can use the protein kinase C (PKC) tumor promoter signaling pathway to mediate transcriptional activation as shown by an increase of the endogenous PKC activator *sn*-1,2-diacylglycerol (Kekulé *et al.,* 1992; Fig. 5). Various PKC inhibitors like H7, sangivamycin, sphingosine, GF109203X, or staurosporine and PKC depletion by prolonged stimulation of cells with 12-*O*-tetra-decanoyl-phorbol-13-acetate (TPA) abolished trans-activation by HBx via AP-1, thereby emphasizing the relevance of this pathway (Kekulé *et al.,* 1993). However, Lucito and Schneider (1992) as well as Murakami *et al.* (1994b) could independently not confirm this requirement for PKC. This may be ex-

plained by the delayed exposure of cells to the PKC inhibitors 6 hours after transfection, which may not have been early enough to inhibit PKC activation.

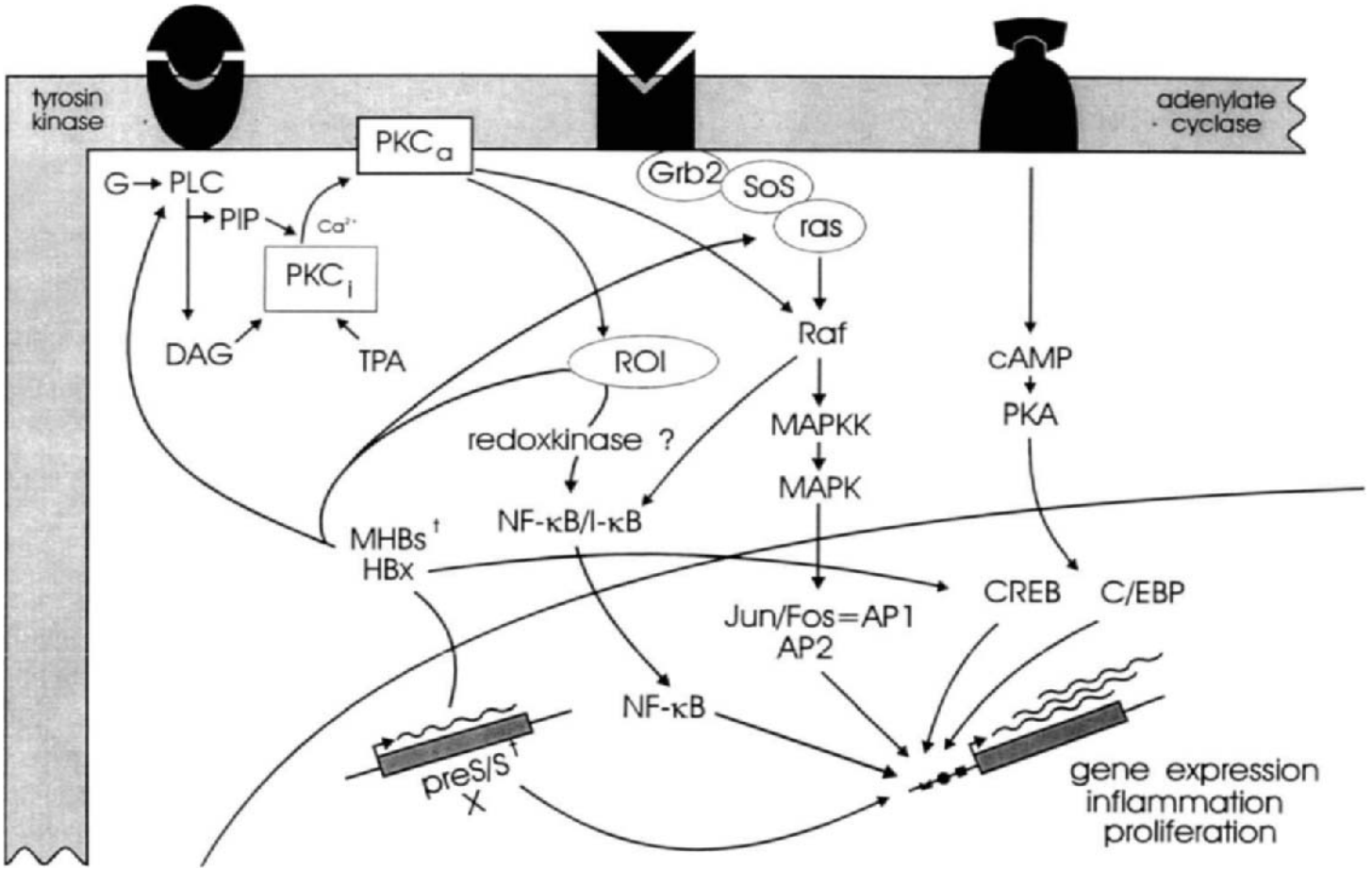
c. Involvement of reactive oxygen intermediates in HBx-mediated trans-activation. Using the antioxidant pyrrolidine dithiocarbamate (PDTC) in 60 μM concentration, which is known to be a selective inhibitor of NF-κB-controlled transcriptional activation (Staal *et al.,* 1990; Schreck *et al.,* 1991, 1992), HBx-induced *cat* gene expression from NF-κB-controlled regulatory elements was almost completely blocked in a dose-dependent manner (Meyer *et al.,* 1992b). Reporter gene expression from the cellular c-*fos* promoter which is not controlled by NF-κB binding elements was not affected by antioxidant exposure. Electromobility shift analysis of nuclear extracts prepared from COS7 cells transfected with X DNA confirmed the specific nuclear translocation of NF-κB that could also be inhibited by PDTC. These data show striking similarity to the redox-regulated HTLV-I trans-activator *tax* and prompted us to postulate also an involvement of reactive oxygen intermediates (ROI) in HBx-mediated trans-activation via NF-κB (Meyer *et al.,* 1992b; Fig. 5).

d. Interaction of HBx with the ras–raf–mitogen-activated protein kinase pathway. Benn and Schneider (1994) were able to demonstrate that HBx induces the cytoplasmic signaling cascade of *ras, raf,* mitogen-activated protein kinase kinase (MAPKK), and mitogen-activated protein kinase (MAPK; Fig 5). In their experiments HBx elevated GTP-bound *ras,* activated and phosphorylated *raf,* as well as phosphorylated and activated MAPK. In HeLa and F9 cells HBx was able to increase the activity of transfected c-*jun* (Natoli *et al.,* 1994b). By use of dominant Ha-*ras* and *raf*-1 negative mutants the same group showed that both proteins are required for HBx-induced activation of c-*jun* transcriptional activity.

e. Interaction of HBx with cellular proteins unrelated to the basal transcriptional machinery. HBx and p53 could be coimmunoprecipitated from extracts of HBV-associated HCCs (Feitelson *et al.,* 1993). Evidence that p53 can directly bind to HBx was also recently presented

---

FIG. 5. Putative signal transduction pathways of trans-activation by HBV proteins. G, G protein; Grb2, adaptor molecule; I-κB, inhibitor-kappa B; MAPKK, mitogen-activated protein kinase kinase; MAPK, mitogen-activated protein kinase; NF-κB, nuclear factor-kappa B; PLC, phospholipase C; PIP, phosphatidylinositol-diphosphate; DAG, diacylglycerol; PKA, protein kinase A; $PKC_{i/a}$, (in-)active form of protein kinase C; TPA, 12-O-tetra-decanoyl-phorbol-13-acetate; ROI, reactive oxygen intermediates; SoS, nucleotide-exchange factor. [Modified from Caselmann (1995).]


tyrosin kinase
adenylate cyclase
PKC$_a$
G → PLC
PIP
Ca^{2+}
PKC$_i$
DAG
TPA
Grb2
SoS
ras
Raf
ROI
redoxkinase ?
MAPKK
MAPK
cAMP
PKA
MHBs^t
HBx
NF-κB/I-κB
CREB
C/EBP
Jun/Fos=AP1
AP2
NF-κB
preS/S^t
X
gene expression
inflammation
proliferation

by Wang *et al.* (1994). This interaction of HBx with p53 did not prevent p53 from specific DNA binding. Most recent data showed that HBx is able to inhibit p53-mediated trans-activation of a reporter gene containing a p53 response element as well as to repress p53-stimulated transcription *in vitro*. Interaction of HBx with p53 did not prevent the binding of the TATA box-binding protein subunit of TFIID and the p62 subunit of TFIIH to the activation domain of p53. Therefore, Truant *et al.* (1995) speculated that HBx may contain a repression domain which may repress transcription by direct interaction with DNA-bound p53.

Using the two-hybrid system, Lee *et al.* (1995) were able to identify another cellular protein that can interact with HBx. This protein is a human homologue of the UV-damaged DNA-binding protein (UV-DDB) isolated from a monkey cell cDNA library. It is presumably involved in DNA repair. An inhibitory effect of HBx by interaction with the cellular DNA repair system could encourage HBV-associated cell transformation. However, the relevance of these different interactions to hepatocarcinogenesis *in vivo* remains to be elucidated. At present the experimental data do not allow definition of an exclusive mechanism mediating transcriptional trans-activation by HBx.

B. Carboxy-Terminally Truncated Hepatitis B Virus Surface Proteins

1. The Functional Trans-Activator Domain

A second trans-activating function unrelated to HBx is encoded in the preS/S region of HBV. Respective trans-activator sequences have been cloned from the human hepatocarcinoma cell line huH-4, which contained 5.6 kb of rearranged HBV DNA (Kekulé *et al.*, 1990) and from surgically resected human HCC tissue that comprised 2.0 bp of integrated HBV DNA (Caselmann *et al.*, 1990). In both cases the trans-activating function could be independently assigned to the preS/S region, because subclones lacking functional X gene sequences exerted a trans-activating effect in cotransfection-assays (Kekulé *et al.*, 1990; Caselmann *et al.*, 1990). In contrast to HBx, preS/S-encoded trans-activators require carboxy-terminal truncation to gain their trans-activating function (Caselmann *et al.*, 1990; Kekulé *et al.*, 1990). While the preS1 region seems to be dispensable for the trans-activating effect (Meyer *et al.*, 1992b; Natoli *et al.*, 1992), deletion or mutation analysis of preS2/S sequences reveals that gene truncation has to occur within a defined region of the S gene (trans-activity-on region; Fig. 4b) between S amino acids 22 and 139 (Lauer *et al.*, 1992). According to recent inves-

tigations even the presence of preS2 amino acids 1–54 may be sufficient to exert the trans-activating effect (Hildt *et al.,* 1995). The trans-activating potential originally described in integrated HBV DNA is also retained in nonintegrated wild-type HBV DNA after artificial truncation.

2. Structural Characterization

The low trans-activator gene expression in transient expression systems and the lack of antibodies specifically recognizing truncated MHBs^t proteins in the presence of full-length MHBs established the need to produce larger quantities of recombinant MHBs^t trans-activator protein. MHBs^{t167} (comprising 167 *N*-terminal amino acids of MHBs), synthesized in the vaccinia virus system (Meyer *et al.,* 1992b; Schlüter and Caselmann, 1995), is glycosylated at asparagine-4 of the preS2 domain with a mannose-rich glycan. MHBs^{t76} (comprising 76 *N*-terminal amino acids of MHBs), expressed in *Spodoptera frugiperda* (Sf9) insect cells or in *E. coli,* lacks this glycosylation but retains its trans-activating activity. Cotransfection of CCl13 cells with an MHBs^{t76} expression vector in which the glycosylation site had been mutated confirmed that glycosylation is not required for the trans-activating function (Hildt *et al.,* 1993). Double immunofluorescence staining and subcellular fractionation revealed that MHBs^{t167} and MHBs^{t76} are, in contrast to MHBs, retained in the ER and not secreted into the cell culture medium. Treatment with 0.1-*M* Na_2CO_3 at pH 11.5 showed that MHBs^{t167} is tightly associated with membranes and suggested that it may be an integral part of the ER membrane (Meyer *et al.,* 1992b). The role of ER retention for the trans-activating function is presently under investigation (Hildt *et al.,* 1995; Schlüter *et al.,* personal communication). So far, there is evidence that artificially truncated MHBs^t proteins not associated with membranes can also represent functional trans-activators (Hildt *et al.,* 1995). These shorter trans-activator proteins may also have lower stability. In liver tissues or hepatocarcinoma cell lines these MHBs^t proteins or integrated HBV DNA sequences truncated in a way to give rise to non-membrane-associated MHBs^t trans-activators have not been identified to date.

3. Target Sequences of Trans-Activation

Differently truncated MHBs^t proteins have been used to identify target sequences of MHBs^t-mediated trans-activation. Like HBx, heterologous viral and cellular regulatory sequences as well as isolated transcription-factor binding sites have been shown to represent target sequences in eukaryotic and eubacterial systems (Table VI). Apart from the viral SV40 early promoter/enhancer that was originally used

TABLE VI

TARGET SEQUENCES FOR MHBs^T MEDIATED TRANS-ACTIVATION

Target sequence	Cell line	Trans-activator	Reference
SV40 eP/Enh	CCl13	MHBs^{t79}	Kekulé *et al.* (1990)
	PLC/PRF/5, HepG2, HeLa	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)
	Sf9, CCl13	MHBs^{t76}	Hildt *et al.* (1993)
HIV-1 LTR	CCl13	MHBs^{t167}	Meyer *et al.* (1992b)
HSV TK	CCl13, HeLa	MHBs^{t76}	Lauer *et al.* (1994)
HTLV-1 LTR	HeLa, CCl13	MHBs^{t167}	Meyer *et al.* (1992b)
RSV LTR	HeLa	MHBs^{t76}	Lauer *et al.* (1994)
c-*fos*	HeLa	MHBs^{t167}	Meyer *et al.* (1992b)
	HeLa	MHBs^{t76}	Lauer *et al.* (1994)
	HeLa, HepG2, PLC/PRF/5	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)
IL-6	HeLa, CCl13	MHBs^{t167}	Meyer *et al.* (1992b)
c-*myc*	CCl13	MHBs^{t79}	Kekulé *et al.* (1990)
	CV-1	MHBs^{t76}	Lauer *et al.* (1994)
	PLC/PRF/5, HepG2, HeLa	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)
c-Ha-*ras*	HeLa, CCl13	MHBs^{t167}	Meyer *et al.* (1992b)
AP-1	Sf9, CCl13	MHBs^{t76}	Hildt *et al.* (1993)
	HeLa	MHBs^{t76}	Lauer *et al.* (1994)
AP-2	CCl13	MHBs^{t76}	Lauer *et al.* (1994)
NF-κB	CCl13	MHBs^{t76}	Lauer *et al.* (1994)
	HeLa	MHBs^{t167}	Meyer *et al.* (1992b)
	PLC/PRF/5, HepG2, HeLa	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)
SP1	HeLa	MHBs^{t167}	M. Meyer *et al.*, personal communication
SRE	PLC/PRF/5, HepG2, HeLa	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)
TRE	PLC/PRF/5, HepG2, HeLa	MHBs^{t207}, MHBs^{t87}	Natoli *et al.* (1992)

as test target (Caselmann *et al.*, 1990; Kekulé *et al.*, 1990; Natoli *et al.*, 1992, Hildt *et al.*, 1993), the promoter sequences of the oncogenes c-*fos* (Meyer *et al.*, 1992b; Natoli *et al.*, 1992; Lauer *et al.*, 1994), c-*myc* (Kekulé *et al.*, 1990; Natoli *et al.*, 1992; Lauer *et al.*, 1994), and c-Ha-*ras* (Meyer *et al.*, 1992b) as well as the regulatory element of the main mediator of the hepatic acute-phase reaction, interleukin-6 (IL-6; Meyer *et al.*, 1992b), are cellular targets of trans-activation by MHBs^t. Al-

though the various target sequences have not been compared systematically, targets of MHBs^t-mediated trans-activation do not appear to differ significantly from HBx targets. Therefore, signal transduction pathways used to mediate trans-activation by MHBs^t may also be similar.

4. Mechanism of Trans-Activation

As in the case of HBx, MHBs^t-mediated trans-activation seems to be pleiotropic. MHBs^{t167} and other MHBs^t trans-activators do not seem to bind DNA directly, but exert their effect through transcription factors like AP-1 (Hildt *et al.*, 1993; Lauer *et al.*, 1994, Schlüter and Caselmann, 1995), AP-2 (Lauer *et al.*, 1994), NF-κB (Meyer *et al.*, 1992b; Natoli *et al.*, 1992; Lauer *et al.*, 1994), and others (Table VI). MHBs^{t167} can stimulate nuclear appearance of the redox-regulated nuclear transcription factor NF-κB and activate transcription of various NF-κB-controlled reporter genes *in vitro* (Meyer *et al.*, 1992b; Schlüter and Caselmann, 1995). The antioxidants *N*-acetyl-L-cysteine (NAC) and PDTC can potently suppress trans-activation at 30-m*M* and 60-μM concentrations, respectively (Meyer *et al.*, 1992b). The effects are selective, since antioxidants do not interfere with MHBs^{t167}-induced activation of the c-*fos* promoter, which is not regulated by NF-κB, nor with the basal activity of several other reporter genes lacking functional NF-κB binding motifs. These data suggest indirectly that NF-κB-mediated trans-activation by MHBs^t may involve ROI as in the case of HBx and therefore be dependent on a prooxidant state of the cell (Meyer *et al.*, 1992b). In analogy to HBx, the relevance of PKC- and mitogen-activated signal transduction pathways (see Section IV,A,4) is presently being investigated.

C. Experimental Evidence for Oncogenic Potential of HBV Trans-Activators

Transgenic mouse models have been used to link HBV trans-activator expression with hepatocarcinogenesis. HBx-transgenic mice which expressed the transgene from its homologous promoter and enhancer developed hepatocyte dysplasia resulting in the formation of multifocal areas of altered hepatocytes, benign liver adenomas, and HCCs within about 13 months (Kim *et al.*, 1991; Koike *et al.*, 1994a). These data are in discordance with previous results, which did not show HCC formation in transgenic mice carrying the X gene under the control of the heterologous human α_1-antitrypsin promoter (Lee *et al.*, 1990). Also, no HCC development was observed in X/C-transgenic mice, in which the X

and C genes were expressed under the control of their autologous enhancer/promoter elements (K. Reifenberg and H. J. Schlicht, 1996, in press). However, increased 8-hydroxy-2'-deoxy-guanosine DNA adduct formation reflecting ROI-induced DNA damage was detectable in liver tissues of these 590-day-old X/C-transgenic mice as compared to nontransgenic and 60-day-old X/C-transgenic mice (R. Gehrke *et al.*, 1995, personal communication). This suggests that HBx may cause increased DNA adduct formation by ROI involvement also *in vivo*.

Immortalized fetal mouse hepatocytes FMH202 (Seifer *et al.*, 1991; Seifer and Gerlich, 1992) and NIH3T3 fibroblasts can successfully be transformed by transfected X gene sequences (Shirakata *et al.*, 1989). The transforming potential of HBx is less pronounced in fetal mouse hepatocytes than that of the total HBV genome (Höhne *et al.*, 1990). Using the same FMH202 cells, integrated HBV X and truncated preS/S sequences cloned from the huH-4 cell line (Kekulé *et al.*, 1990) displayed soft agar growth and nude mouse tumorigenicity after stable transfection. However, rearrangements of transfected HBV DNA had taken place. Expression was shut off during nude mouse passage in one of two cases (B. Luber, 1995, personal communication). Koike *et al.* (1994b) demonstrated increased DNA synthesis in NIH3T3 cells after induction of HBx overexpression with dexamethasone. Following X gene expression most cells in the G_0/G_1 phase moved into the S phase within 24 hours and the cell cycle progressed further towards 48 hours. Therefore, the authors speculated that HBx may play a role in hepatocyte transformation by driving deregulated cell cycle progression.

Neither transgenic mice nor appropriate cell culture systems are presently available to test the transforming capacity of truncated $MHBs^t$ sequences. PreS1/S2/S transgenic mice leading to overexpression and cellular accumulation of LHBs developed regenerative hyperplasia and neoplasia in accordance with HCC (Chisari *et al.*, 1989). When these mice were exposed to aflatoxin or diethylnitrosamine, which are strong hepatocarcinogens in rats but are effective only upon neonatal exposure in the mouse (Diwan *et al.*, 1986; Vesselinovitch *et al.*, 1972), adenomas and HCCs could be induced more frequently in LHBs-overexpressing mice than in nontransgenic controls (Sell *et al.*, 1991). Similar data had been obtained by Dunsford *et al.* (1990) using an HBS-transgenic mouse system. Since overexpression of the preS1/S2/S gene led to a marked reduction of detectable hepatitis B virus surface antigen in the mouse serum and intracellular accumulation of the envelope proteins (Chisari *et al.*, 1986), similar mechanisms as observed with ER-retained truncated $MHBS^t$ (see Section IV,B,2) may be effective in LHBs-producing transgenic mice.

HeLa cell lines expressing constitutively and inducibly truncated preS2/S sequences, respectively, are presently under construction (R. Gehrke *et al.,* 1995, personal communication). So far, only epidemiological evidence supports a role of trans-activators in hepatocarcinogenesis. Eight of 26 (31%) HCC tissues or hepatoma cell lines contained at least one insert comprising truncated preS/S sequences, 18 of 26 (69%) X gene sequences, and 21 of 26 (81%) preS/S and/or X gene sequences (Schlüter *et al.,* 1994). Transcription of X- and preS/S-specific RNA and translation of HBx and $MHBs^t$ was shown for four integrated X gene and three preS/S sequences. By all these HBV integrations reporter gene expression was stimulated fivefold to tenfold from the c-*fos* protooncogene promoter, indicating that trans-activator sequences are frequently found in HCC cell lines/tissues and are functional in model systems in all cases investigated so far. In most cases viral–cellular fusion proteins were translated. There was evidence for mutual enhancement when both trans-activators were coexpressed.

D. The Trans-Hypothesis of Hepatocarcinogenesis

Integration of HBV DNA into the hepatocyte genome can take place at any time during infection. This can result in single or multiple HBV DNA insertions at various cellular sites, causing rearrangements of both viral and cellular DNA. In rare cases, insertional mutagenesis leads to overexpression of cellular genes favoring proliferation, inflammation, or transformation. Deletions of tumor suppressor gene sequences may occur, but viral sequences also may get lost. The HBV does not exert a direct cytopathogenic effect. Liver damage is considered to be mediated by cytotoxic T cells that recognize viral antigens like HBcAg together with major histocompatibility complex class I antigens. When the HBcAg coding sequences that are pivotal for virus elimination are deleted during integration, the HBV-infected hepatocyte may escape the host's immune response (Chisari and Ferrari, 1995; Jung *et al.,* 1994; Meuer and Moebius, 1994; Naumov and Eddleston, 1994; Milich *et al.,* 1993). Subsequently or in parallel, viral trans-activators HBx or $MHBs^t$ may be expressed in the multifactor process of hepatocarcinogenesis. This may lead to increased ROI formation favoring DNA adduct formation, lipid peroxidation, and protein damage. PKC-dependent pathways or cellular signal cascades comprising mitogen-activated protein kinases may be used to stimulate cellular gene expression and thereby contribute to HBV-associated hepatocarcinogenesis (Fig. 5).

V. Future Research Directions and Clinical Perspectives

In the future, major emphasis will be laid on the demonstration of HBx and truncated MHBs^t trans-activators *in vivo* and under physiological conditions. HBV trans-activator proteins should be demonstrated in peripheral blood and liver tissues of patients with HBV-associated liver disease. The link between trans-activation by HBV proteins and malignant transformation needs to be established more conclusively. Especially for MHBs^t trans-activators, there are no appropriate experimental systems available yet. If further evidence for a relevant function of HBV trans-activators in hepatocarcinogenesis can be provided and practicable detection systems can be used, the monitoring of HBV trans-activators in liver tissues or sera of HBV-infected patients may become a method for preventive HCC screening and then supplement established procedures like alpha-fetoprotein and ultrasound follow-up of patients with chronic liver disease.

Understanding of the mechanism underlying trans-activation by HBV proteins is essential and requires further experimental clarification. Only a detailed knowledge of the different signal transduction pathways may allow specific interference with single mediators of the cascade in the future. Thereby it may become possible to inhibit proliferation or even transformation in chronically infected patients and to lay the grounds for experimental prevention or even therapy of HBV-associated HCC.

Acknowledgments

The author's research is supported by a Heisenberg fellowship (CA113/10-1) of the Deutsche Forschungsgemeinschaft. We wish to thank Dr. Stephan Urban and Dr. Matthias Renner for the preparation of Table V, Mr. Matthias Weber for typing the manuscript, and Mr. Matthias Wissnet for drawing the figures.

References

Acs, G., and Price, P. M. (1990). *Prog. Liver Dis.* **9**, 379–389.

Albin, C., and Robinson, W. S. (1980). *J. Virol* **34**, 297–302.

Arii, M., Takada, S., and Koike, K. (1992). *Oncogene* **7**, 397–403.

Aufiero, B., and Schneider, R. J. (1990). *EMBO J.* **9**, 497–504.

Avantaggiati, M. L., Balsano, C., Natoli, G., De Marzio, E., Will, H., Elfassi, E., and Levrero, M. (1992). *Arch. Virol.* **4**, 57–61.

Avantaggiati, M. L., Natoli, G., Balsano, C., Chirillo, P., Artini, M., De Marzio, E., Collepardo, D., and Levrero, M. (1993). *Oncogene* **8**, 1567–1574.

Balsano, C., Avantaggiati, M. L., Natoli, G., DeMarzio, E., Will, H., Perricaudet, M., and Levrero, M. (1991). *Biochem. Biophys. Res. Comm.* **176**, 985–992.
Balsano, C., Billet, O., Bennoun, M., Cavard, C., Zider, A., Grimber, G., Natoli, G., Briand, P., and Levrero, M. (1993). *Arch. Virol.* (Suppl.) **8**, 63–71.
Balsano, C., Billet, O., Bennoun, M., Cavard, C., Zider, A., Grimber, G., Natoli, G., Briand, P., and Levrero, M. (1994). *J. Hepatol.* **21**, 103–109.
Bamber, M., Thomas, H. C., Bannister, B., and Sherlock, S. (1983). *Gut* **24**, 561–564.
Bartenschlager, R., and Schaller, H. (1988). *EMBO J.* **7**, 4185–4192.
Bartenschlager, R., Kuhn, C., and Schaller, H. (1992). *Nucleic Acids Res.* **20**, 195–202.
Bavand, M., Feitelson, M., and Laub, O. (1989). *J. Virol.* **63**, 1019–1021.
Beasley, R. P. (1988). *Cancer* **61**, 1942–1956.
Benn, J., and Schneider, R. T. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10350–10354.
Bichko, V., Pushko P., Dreilina, D., Pumpen, P., and Gren, E., (1985). *FEBS Lett.* **185**, 208–212.
Block, T. M., Lu, X., Platt, F. M., Foster, G. R., Gerlich, W. H.,Blumberg, B. S., and Dwek, R. A. (1994). *Proc. Natl. Acad Sci. U.S.A.* **91**, 2235–2239.
Blum, H. E., Gerok, W., and Vyas, G. N. (1989). *Trends Genet.* **5**, 154–158.
Blum, H. E., Zhang, Z. S., Galun, E., von Weizsäcker, F., Garner, B., Liang, T. J., and Wands, J. R. (1992). *J. Virol.* **66**, 1223–1227.
Blumberg, B. S., Alter, H. J., and Visnich, S. (1965). *JAMA* **191**, 541–546.
Bosch, V., Bartenschlager, R., Radziwill, G., and Schaller, H. (1988). *Virology* **166**, 475–485.
Bressac, B., Kew, M., Wands, J., and Ozturk, M. (1991). *Nature* **350**, 429–431.
Bruss, V., and Ganem, D. (1991). *J. Virol.* **65**, 3813–3820.
Bruss, V., and Gerlich, W. H. (1988). *Virology* **163**, 268–275.
Bruss, V., Lu, X., Thomssen, R., and Gerlich, W. H. (1994). *EMBO J.* **13**, 2273–2279.
Büscher, M., Reiser, W., Will, H., and Schaller, H. (1985). *Cell* **40**, 717–724.
Buetow, K. H., Murray, J. C, Israel, J. L., London, W. T., Smith, M., and Kew, M. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8852–8856.
Bulla, G. A., and Siddiqui, A. (1989). *Virology* **170**, 251–260.
Carlier, D., Jean-Jean, O., Fouillot, N., Will, H., and Rossignol, J. M. (1995). *J. Gen. Virol.* **76**, 1041–1045.
Caselmann, W. H. (1994). *Antiviral Res.* **24**, 121–129.
Caselmann, W. H. (1995). *J. Hepatol.* **22 (Suppl. 1)**, 34–37.
Caselmann, W. H., Meyer, M., Kekulé, A. S., Lauer, U., Hofschneider, P. H., and Koshy, R. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2970–2974.
Cattaneo, R., Will, H., Hernandez, N., and Schafer, H. (1983). *Nature* **305**, 336–338.
Cattaneo, R., Will, H., and Schaller, H. (1984). *EMBO J.* **3**, 2191–2196.
Chalmers, T. C., and Alter, H. J. (1971). *N. Engl. J. Med.* **285**, 613–617.
Chang, H. K., Wang, B. Y., Yuh, C. H., Wei, C. L., and Ting, L. P. (1989). *Mol. Cell. Biol.* **9**, 5189–5197.
Chang, L. J., Pryciak, P., Ganem, D., and Varmus, H. E. (1989). *Nature* **337**, 364–368.
Chen, H. S., Kaneko, S., Girones, R., Anderson, R. W., Hornbuckle, W. E., Tennant, B. C., Coté, P. J., Gerin, J. L., Purcell, R. H., and Miller, R. H. (1993). *J. Virol.* **67**, 1218–1223.
Chen, M. L., Lee, Y. H. W., and Lo, S. J. (1988). *Gene* **62**, 315–321.
Chen, P. J., Chen, C. R., Sung, J. L., and Chen, D. S. (1989). *J. Virol.* **63**, 4165–4171.
Cheong, J. H., Yi, M. K., Lin, Y., and Murakami, S. (1995). *EMBO J.* **14**, 143–150.
Chesnut, J. D., Stephens, J. H., and Dahmus, M. E. (1992). *J. Biol. Chem.* **267**, 10500–10506.
Chisari, F. V., and Ferrari, C. (1995). *Annu. Rev. Immunol.* **13**, 29–60.

Chisari, F. V., Filippi, P., McLachlan, A., Milich, D. R., Riggs, M., Lee, S. Palmiter, R. D., Pinkert, C. A., and Brinster, R. L. (1986). *J. Virol.* **60**, 880–887.

Chisari, F. V., Klopchin, K., Moriyama, T., Pasquinelli, C., Dunsford, H. A., Sell, S., Pinkert, C. A., Brinster, R. L., and Palmiter, R. D. (1989). *Cell* **59**, 1145–1156.

Cohen, B. J. (1978). *J. Med. Virol.* **3**, 141–149.

Cohen, J. I., and Kieff, E. (1991). *J. Virol.* **65**, 5880–5885.

Colgrove, R., Simon, G., and Ganem, D. (1989). *J. Virol.* **63**, 4019–4026.

Courouce, A. M., Holland, P. V., Muller, P. V., and Soulier, J. P. (1976). *Bibl. Haematol.* **42**, 1–158.

Cross, J. C., Wen, P., and Rutter, W. J. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8078–8082.

Crowther, R. A., Kiselev, N. A., Böttcher, B., Berrimn, J. A., Borisova, G. P., Ose, V., and Pumpens, P. (1994). *Cell* **77**, 943–950.

Dane, D. S., Cameron, C. H., and Briggs, H. (1970). *Lancet* **1**, 695–698.

Dejean, A., Bougueleret, L., Grzeschik, K. H., and Tiollais, P. (1986). *Nature* **322**, 70–72.

De Medina, T., and Shaul, Y. (1994). *FEBS Lett.* **351**, 423–426.

De Medina, R., Faktor, O., and Shaul, Y. (1988). *Mol. Cell. Biol.* **8**, 2449–2455.

De Medina, T., Haviv, I., Noiman, S., and Shaul, Y. (1994). *Virology* **202**, 401–407.

De Thé, H., Marchio, A., Tiollais, P., and Dejean, A. (1987). *Nature* **330**, 667–670.

De Wet, J. R., Wood, K. V., De Luca, M., Helinski, D. R., and Subramani, S. (1987) *Mol. Cell. Biol.* **7**, 725–737.

Ding, S. F., and Habib, N. A. (1995). *J. Hepatol.* **22**, 230–238.

Ding, S. F., Habib, N. A., Dooley, J. S., Wood, C. B., Bowles, L., and Delhanty, J. D. A. (1991). *Br. J. Cancer* **64**, 1083–1087.

Ding, S. F., Delhanty, J. D. A., Dooley, J. S., Bowles, L., Wood, C. B., and Habib, N. A. (1993). *Cancer Det. Prev.* **17**, 405–409.

Diwan, B. A., Rice, J. M., Oshima, M., and Ward, J. M. (1986). *Carcinogenesis* **7**, 215–220.

Dunsford, H. A., Sell, S., and Chisari, F. V. (1990). *Cancer Res.* **50**, 3400–3407.

Eble, B., Lingappa, V., and Ganem, D. (1986). *Mol. Cell. Biol.* **6**, 1454–1463.

Eble, B. E., McRae, D. R., Lingappa, V. R., and Ganem, D. (1987). *Mol. Cell. Biol.* **7**, 3591–3601.

Elfassi, E., Haseltine, W. A., and Dienstag, J. L. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2219–2222.

Emi, M., Fujiwara, Y., Nakajima, T., Tsuchiya, E., Tsuda, H., and Hirohashi, S. (1992). *Cancer Det. Prev.* **52**, 5368–5372.

Enders, G. H., Ganem, D., and Varmus, H. (1985). *Cell* **42**, 297–308.

Enders, G. H., Ganem, D., and Varmus, H. E. (1987). *J. Virol.* **61**, 35–41.

Faktor, O., and Shaul, Y. (1990). *Oncogene* **5**, 867–872.

Faktor, O., Budlovsky, S., Ben-Levy, R., and Shaul, Y. (1990). *J. Virol.* **64**, 1861–1863.

Feitelson, M. A., Marion, P. L., Robinson, P. S. (1982). *J. Virol.* **43**, 687–696.

Feitelson, M. A., Millman, I., Hakbherr, T., Simmonds, H., and Blumberg, B. S. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2233–2237.

Feitelson, M. A., Zhu, M., Duan, L. X., and London, W. T. (1993). *Oncogene* **8**, 1009–1017.

Ferrari, C., Penna, A., Degli, A., and Ficcadori, F. (1988). *J. Hepatol.* **7**, 21–33.

Fischer, M., Runkel, L., and Schaller, H. (1995). *Virus Genes* **10**, 99–102.

Fourel, G., Trepo, C., Bougueleret, L., Henglein, B., Ponzetto, A., Tiollais, P., and Buendia, M. A. (1990). *Nature* **347**, 294–298.

Fourel, G., Couturier, J., Wei, Y., Apiou, F., Tiollais, P., and Buendia, M. A. (1994). *EMBO J.* **13**, 2526–2534.

Fujimori, M., Tokino, T., Hino, O., Kitagawa, T., Imamura, T., and Okamoto, E. (1991). *Cancer Res.* **51**, 89–93.

Fujiyama, A., Miyanohara, A., Nozaki, C., Yoneyama, T., Ohtomo, N., Matsubara, K., and Matsubara, K. (1983). *Nucleic Acids Res.* **11**, 4601–4610.

Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P., and Charnay, P. (1979). *Nature* **281**, 464–650.

Gan, R. B., Chu, M. J., Shen, L. P., and Li, Z. P. (1984). *Acta Biochim. Biophys. Sinica* **16**, 316–319.

Ganem, D., and Varmus, H. E. (1987). *Annu. Rev. Biochem.* **56**, 651–693.

Ganem, D., Pollack, J. R., and Tavis, J. (1994). *Infect. Agents Dis.* **3**, 85–93.

Garcia, P. D., Ou, J. H., Rutter, W. J., and Walter, P. (1988). *J. Cell. Biol.* **106**, 1093–1104.

Gerbes, A. L., and Caselmann, W. H. (1993). *J. Hepatol.* **19**, 312–315.

Gerin, J. L., Tennant, B. C., Ponzetto, A, Purcell, R. H., and Tyeryar, F. J. (1983). *Prog. Clin. Biol. Res.* **143**, 23–28.

Gerlach, K. K., and Schloemer, R. H. (1992). *Virology* **189**, 59–66.

Gerlich, W. H., Feitelson, M. A., Marion, P. L., and Robinson, W. S. (1980). *J. Virol.* **36**, 787–795.

Gorman, C. M., Moffat, L. F., Howard, B. H. (1982). *Mol. Cell. Biol.* **2**, 1044–1051.

Gough, N. M. (198Z). *J. Mol. Viol.* **165**, 683–699.

Graef, E., Caselmann, W. H., Wells, J., and Koshy, R. (1994). *Oncogene* **9**, 81–87.

Graef, E., Caselmann, W. H., Hofschneider, P. H., and Koshy, R. (1995). *Virology* **208**, 696–703.

Guerrero E., Swenson, P. D., Hu, P. S., and Peterson, D. L. (1990). *Mol. Immunol.* **27**, 435–441.

Guidotti, L. G., Martinez, V., Loh, Y.-T., Rogler, C. E., and Chisari, F. V (1994). *J. Virol.* **68**, 5469–5475.

Guo, W., Bell, K. D., and Ou, J. H. (1991). *J. Virol.* **65**, 6686–6692.

Hadziyannis, S. J., Lieberman, H. M., Karvountzis, G. G., and Shafritz, D. A. (1983). *Hepatology* **3**, 656–662.

Hansen, L. J., Tennant, B. C., Seeger, C., and Ganem, D. (1993). *Mol. Cell. Biol.* **13**, 659–667.

Haviv, I, Vaizel, D., and Shaul, Y. (1995). *Mol. Cell. Biol.* **15**, 1079–1085.

Heermann, K. H., Goldmann, U., Schwartz, W., Seyffarth, T., Baumgarten, H., and Gerlich, W. H. (1984). *J. Virol.* **52**, 396–402.

Hertogs, K., Leenders, W. P. J., Depla, E., De Bruin, W. C. C., Meheus, L., Raymackers, J., Moshage, H., and Yap, S. H. (1993). *Virology* **197**, 549–557.

Hildt, E., Urban, S., Lauer, U., Hofschneider, P. H., and Kekulé, A. S. (1993). *Oncogene* **8**, 3359–3367.

Hildt, E., Urban, S., and Hofschneider, P. H. (1995). *Oncogene* **11**, 2055–2066.

Höhne, M., Schaefer, S., Seifer, M., Feitelson, M. A., Paul, D., and Gerlich, W. H. (1990). *EMBO J.* **9**, 1137–1142.

Hollinger, F. B., Werch, J., and Melnick, J. L. (1974). *N. Engl. J. Med.* **290**, 1104–1109.

Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. (1991). *Science* **253**, 49–53.

Honigwachs, J., Faktor, O., Dikstein, R., Shaul, Y., and Laub, O. (1989). *J. Virol.* **63**, 919–924.

Hosono, S., Chou, M.-J., Lee, C.-S., and Shih, C. (1993). *Oncogene* **8**, 491–496.

Hu, K. Q., Vierling, J. M., and Siddiqui, A. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7140–7144.

Huang, J., and Liang, T. J. (1993). *Mol. Cell. Biol.* **13**, 7476–7486.

Huang, Z. M., and Yen, T. S. B. (1994). *J. Virol.* **68**, 3193–3199.
Huang, Z. M., and Yen, T. S. B. (1995). *Mol. Cell. Biol.* **15**, 3864–3869.
Huovila, A. P. J., Eder, A. M., and Fuller, S. D. (1992). *J. Cell. Biol.* **118**, 1305–1320.
Jameel, S., and Siddiqui, A. (1986). *Mol. Cell. Biol.* **6**, 710–715.
Jameel, S., Siddiqui, A., Maguire, H. F., and Rao, K. V. S. (1990). *J. Virol.* **64**, 3963–3966.
Johnson, J. L., Raney, A. K., and McLachlan, A. (1995). *Virology* **208**, 147–158.
Jung, M. C., Stemler, M., Weimer, T., Spengler, U., Döhrmann, J., Hoffmann, R., Eichenlaub, D., Eisenburg, J., Paumgartner, G., Riethmüller, G., Will, H., and Pape, G. R. (I991). *Hepatology* **13**, 637–643.
Jung, M. C., Diepolder, H. M., and Pape, G. R. (1994). *Eur. J. Clin. Invest.* **24**, 641–650.
Junker-Niepmann, M., Bartenschlager, R., and Schaller, H. (1990). *EMBO J.* **9**, 3389–3396.
Kann, M., and Gerlich, W. H. (1994). *J. Virol.* **68**, 7993–8000.
Kann, M., Thomssen, R., Kochel, H. G., and Gerlich, W. H. (1993). *Arch. Virol.* (Suppl.) **8**, 53–62.
Kaplan, P. M., Ford, E. C., Purcell, R. H., and Gerin, J. L. (1976). *J. Virol.* **17**, 885–893.
Kaneko, S., and Miller, R. H. (1988). *J. Virol.* **62**, 3979–3984.
Kawamoto, S., Ueda, K., Mita, E., and Matsubara, K. (1994). *J. Virol. Method.* **49**, 113–128.
Kay, A., Mandart, E., Trepo, C., and Galibert, F. (1985). *EMBO J.* **4**, 1287–1292.
Kekulé, A. S., Lauer, U., Meyer, M., Caselmann, W. H., Hofschneider, P. H., and Koshy, R. (1990). *Nature* **343**, 457–461.
Kekulé, A. S., Lauer, U., Weiss, L, Hofschneider, P. H., and Koshy, R. (1992). *Arch. Virol.* (Suppl.) **4**, 63–64.
Kekulé, A. S., Lauer, U., Weiss, L., Luber, B., and Hofschneider, P. H. (1993). *Nature* **361**, 742–745.
Kim, C. M., Koike, K., Saito, I., Miyamura, T., and Jay, G. (1991). *Nature* **351**, 317–320.
Kim, Y. H., Kang, S. K., and Lee, Y. I. (1993). *Biochem. Biophys. Res. Comm.* **197**, 894–903.
Klein, R., Schröder, C. H., and Zentgraf, H. (1991). *Virus Genes* **5**, 157–174.
Klingmüller, U., and Schaller, H. (1993). *J. Virol.* **67**, 7414–7422.
Knaus, T., and Nassal, M. (1993). *Nucleic Acids Res.* **21**, 3967–3975.
Kobayashi, M., and Koike, K. (1984). *Gene* **30**, 227–232.
Köck, J., and Schlicht, H.-J. (1993). *J. Virol.* **67**, 4867–4874.
Koike, K., Kobayashi, M., Mizusawa, H., Yoshida, E., Yaginuma, K., and Taira, M. (1983). *Nucleic Acids Res.* **11**, 5391–5402.
Koike, K., Akatsuka, T., and Miyamura, T. (1988). *Virology* **163**, 233–235.
Koike, K., Shirakata, Y., Yaginuma, K., Arii, M., Takada, S., Nakamura, I., Hayashi, Y., Kawada, M., and Kobayashi, M. (1989). *Mol. Biol. Med.* **6**, 151–160.
Koike, K., Moriya, K., Ino, S., Yotsuyanagi, H., Endo, Y., Miyamura, T., and Kurokawa, K. (1994a). *Hepatology* **19**, 810–819.
Koike, K., Moriya, K., Yotsuyanagi, H., Lino, S., and Kurokawa, K. (1994b). *J. Clin. Invest.* **94**, 44–49.
Kress, S., Jahn, U.-R., Buchman, A., Bannasch, P., and Schwarz, M. (1992). *Cancer Res.* **52**, 3220–3223.
Krugman, S., and Giles, J. P. (1973). *N. Engl. J. Med.* **288**, 755–760.
Krugman, S., Overby, L. R., Mushahwar, I. K., Ling, C. M., Frosner, G. G., and Deinhardt, F. (1979). *N. Engl. J. Med.* **300**, 101–106.
Kuroki, K., Russnak, R., and Ganem, D. (1989). *Mol. Cell. Biol.* **9**, 4459–4466.
Kwee, L., Lucito, R., Aufiero, B., Schneider, R. J. (1992). *J. Virol.* **66**, 4382–4389.
Lauer, U., Weiss, L., Lipp, M., Hofschneider, P. H., and Kekulé, A. S. (1992). *J. Virol.* **66**, 5284–5289.

Lauer, U., Weiss, L., Lipp, M., Hofschneider, P. H., and Kekulé, A. S. (1994). *Hepatology* **19**, 23–31.

LeBouvier, G. L., McCollum, R. W., Hierholzer, W. J., Irwin, G. R., Krugman, S., and Giles, J. P. (1972). *J. Am. Med. Assoc.* **222**, 928–930.

Lee, T. H., Finegold, M. J., Shen, R. F., DeMayo, J. L., Woo, S. L. C., and Butel, J. S (1990). *J. Virol.* **64**, 5939–5947.

Lee, T. H., Elledge, S. J., and Butel, J. S. (1995) *J. Virol.* **69**, 1107–1114.

Levine, A. J., Momand, J., and Finlay, C. A. (1991). *Nature* **351**, 453–456.

Levrero, M., Balsano, C., Natoli, G., Avantaggiati, L., and Elfassi, E. (1990). *J. Virol.* **64**, 2082–2086.

Liang, T. J., Jeffers, L. J., Reddy, K. R., De Medina, M., Parker, I. T., Cheinquer, H., Idrovo, V., Rabassa, A., and Schiff, E. R. (1993). *Hepatology* **18**, 1326–1333.

Lin, M. H., and Lo, S. J. (1989). *Biochem. Biophys. Res. Comm.* **164**, 14–21.

Liu, C. C., Yansura, D., and Levinson, A. (1982). *DNA* **1**, 213–221.

Lo, W. Y., and Ting, L. P. (1994). *J. Virol.* **68**, 1758–1764.

Luber, B., Bürgelt, E., Fromental, C., Kanno, M., and Koch, W. (1991). *Virology* **184**, 808–813.

Luber, B., Lauer, U., Weiss, L., Höhne, M., Hofschneider, P. H., and Kekulé, A. S. (1993). *Res. Virol.* **144**, 311–321.

Lucito, R., and Schneider, R. J. (1992). *J. Virol.* **66**, 983–991.

Mack, D. H., Bloch, W., Nrapendra, N., and Sninsky, J. J. (1988). *J. Virol.* **62**, 4786–4790.

Maguire, H. F., Hoeffler, T. P., and Siddiqui, A. (1991). *Science* **252**, 842–844.

Mandart, E., Kay, A., and Galibert, F. (1984). *J. Virol.* **49**, 782–792.

Marion, P. L., Oshiro, L. S., Regnery, D. C., Scullard, G. H., and Robinson, W. S. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2941–2945.

Mason, W. S., Seal, G., and Summers, J. (1980). *J. Virol.* **36**, 829–836.

Mazzur, S., Burgert, S., and Blumberg, B. S. (1974). *Nature* **247**, 38–40.

McAleer, W. J., Buynak, E. B., Maigetter, R. Z., Wampler, D. E., Miller, W. J., and Hilleman, M. R. (1984). *Nature* **307**, 178–180.

McGlynn, B., Reutener, S., Matter, A., Wildner, G., Will, H., and Lydon, N. B. (1992). *Gen. Virol.* **73**, 1515–1519).

Mehdi, H., Kaplan, M. J., Anlar, F. Y., Yang, X., Bayer, R., Sutherland, K., and Peeples, M. E. (1994). *J. Virol.* **68**, 2415–2424.

Meuer, S. C., and Moebius, U. (1994). *Semin. Virol.* **5**, 289–295.

Meyer, M., Wiedorn, K. H., Hofschneider, P. H., Kohsy, R., and Caselmann, W. H. (1992a). *Hepatology* **15**, 665–671.

Meyer, M., Caselmann, W. H., Schlüter, V., Schreck, R., Hofschneider, P. H., and Baeuerle, P. A. (1992b). *EMBO J.* **11**, 2991–3001.

Meyers, M. L., Vitvitski-Trepo, L. V., Nath, N., and Sninsky, J. J. (1986). *J. Virol.* **57**, 101–109.

Milich, D. R., Jones, J., Hughes, J., and Maruyama, T. (1993). *J. Immunotherapy* **14**, 226–233.

Miller, R. H. (1988). *Science* **239**, 1420–1422.

Miller, R. H., and Robinson, W. S. (1986). *Proc. Natl. Acad Sci. U.S.A.* **83**, 2531–2535.

Miller, R. H., Kaneko, S., Chung, C. T., Girones, R., and Purcell, R. H. (1989). *Hepatology* **9**, 322–327.

Mizusawa, H., Taira, M., Yaginuma, K., Kobayashi, M., Yoshida, E., and Koike, K. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 208–212.

Molnar-Kimber, K. L., and Jarocki-Witek, V. (1988). *J. Virol.* **62**, 407–416.

Mondelli, M. U., Mieli-Vergani, G., Alberti, A., Vergani, D., Portmann, B., Eddleson, A. L., and Williams, R. (1982). *J. Immunol.* **129**, 2773–2778.

Moriarty, A. M., Alexander, H., Lerner, R. A., and Thornton, G. B. (1985). *Science* **227**, 429–433.
Möröy, T., Marchio, A., Etiemble, J., Trepo, C., Tiollais, P., and Buendia, M. A. (1986). *Nature* **324**, 276–279.
Murakarni, S., Cheong, J., and Kaneko, S. (1994a). *J. Biol. Chem.* **269**, 15118–15123.
Murakami, S., Cheong, J., Ohno, S., Matsushima, K., and Kaneko, S. (1994b). *Virology* **199**, 243–246.
Nagaya, T., Nakamura, T., Tokino, T., Tsurimoto, T., Imai, M., Mayumi, T., Kamino, K., Yamamura, K., and Matsuhara, K. (1987). *Genes Dev.* **1**, 773–782.
Nakatake H., Chisaka, O., Yamamoto, S., Matsubara, K., and Koshy, R. (1993). *Virology* **195**, 305–314.
Nassal, M., and Schaller, H. (1993). *Trends Microbiol.* **222**, 221–228.
Natoli, G., Avantaggiati, M. L., Balsano, C., De Marzio, E., Collepardo, D., Elfassi, E., and Levrero, M. (1992). *Virology* **187**, 663–670.
Natoli, G., Avantaggiati, M. L., Chirillo, P., Costanzo, A., Artini, M., Balsano, C., and Levrero, M. (1994a). *Mol. Cell. Biol.* **2**, 989–998.
Natoli, G., Avantaggiati, M. L., Chirillo, P., Puri, P. L., Ianni, A., Balsano, C., and Levrero, M. (1994b). *Oncogene* **9**, 2837–2843.
Naumov, N. V., and Eddleston, A. L. (1994). *Gut* **35**, 1013–1017.
Neurath, A. R., Kent, S. B., Strick, N., and Parker, K. (1988). *Ann. Inst. Pasteur Virol.* **139**, 13–38.
Nishida, N., Fukuda, Y., Kokyuru, H., Toguchida, J., Yandell, D. W., and Ikenega, M. (1993). *Cancer Res.* **53**, 368–372.
Omata, M., Mori, J., Yokusuka, O., Iwama, S., Ito, Y., and Okuda, K. (1982). *Liver* **2**, 125–132.
Ono, Y., Onda, H., Sasada, R., Igarahi, K., Sugino, Y., and Nishioka, K. (1983). *Nucleic Acids Res.* **11**, 1747–1757.
Onodera, S., Ohori, H., Yamaki, M., and Ishida, N. (1982). *J. Med. Virol.* **10**, 147–155.
Ori, A., and Shaul, Y. (1995). *Virology* **207**, 98–106.
Ori, A., Atzmony, D., Hviv, J., and Shaul, Y. (1994). *Virology* **204**, 600–608.
Ostapchuk, P., Hearing, P., and Ganem, D. (1994). *EMBO J.* **13**, 1048–1057.
Ou, J.-H., and Rutter, W. J. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 83–87.
Ou, J.-H., Laub, O., and Rutter, W. J. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1578–1582.
Ou, J.-H., Bao, H., Shih, C., and Tahara, S. M. (1990). *J. Virol.* **64**, 4578–4581.
Pasek, M., Goto, T., Gilbert, W., Zink, B., Schaller, H., McKay, P., Leadbetter, G., and Murray, K. (1979). *Nature* **282**, 575–579.
Patel, N. U., Jameel, S., Isom, H., and Siddiqui, A. (1989). *J. Virol.* **63**, 5293–5301.
Patzer, E. J., Nakamura, G. R., Simonsen, C. C., Levinson, A. D., and Brands, R. (1986). *J. Virol.* **58**, 884–892.
Persing, D. H., Varmus, H. E., and Ganem, D. (1987). *J. Virol.* **61**, 1672–1677.
Petit, M. A., Capel, F., Dubanchet, S., and Mabit, H. (1992). *Virology* **187**, 211–222.
Pfaff, E., Salfeld, J., Gmelin, K., Schafer, H., and Theilmann, L. (1987). *Virology* **158**, 456–460.
Phelps, W. C., Yee, C. L., Münger, S., and Howley, P. M. (1988). *Cell* **53**, 539–547.
Pollack, J. R., and Ganem, D. (1993). *J. Virol.* **67**, 3254–3263.
Pollack, J., and Ganem, D. (1994). *J. Virol.* **68**, 5579–5587.
Popper, H., Shafritz, D. A., and Hoofnagle, J. H. (1987). *Hepatology* **7**, 764–772.
Pontisso, P., Petit, M. A., Bankowski, M. J., and Peeples, M. E. (1989). *J. Virol.* **63**, 981–988.
Prange, R., and Streeck, R. E. (1995). *EMBO J.* **14**, 247–256.

Prince, A. M. (1968). *Proc. Nutl. Acad. Sci. U.S.A.* **60**, 814–821.
Qadri, I., Maguires H. F., and Siddiqui, A. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1003–1007.
Radziwill, G., Tucker, W., and Schaller, H. (1990). *J. Virol.* **64**, 613–620.
Raychaudhuri, P., Bagchi, S., Devoto, S. H., Kraus, V. B., Moran, E., and Nevins, J. R. (1991). *Genes Develop.* **5**, 1200–1211.
Reifenberg, K., Löhler, J., Pudollek, H. P., Schmitteckert, E., Spindler, G., Köck, J., and Schlicht, H. J. (1996). *Hepatol.*, in press.
Renner, M., Haniel, A., Bürgelt, E., Hofschneider, P. H., and Koch, W. (1995). *J. Hepatol.* **23**, 53–65.
Rigg, R. J., and Schaller, H. (1992). *J. Virol.* **66**, 2829–2836.
Ritter, S. E., Whitten, T. M., Quets, A. T., and Schloemer, R. H. (1991). *Virology* **182**, 841–845.
Rogler, C. E., Sherman, M., Su, C. Y., Sharritz, D. A., Summers, J., and Shows, T. B. (1985). *Science* **230**, 319–322.
Rosmorduc, O., Petit, M. A., Pol, S., Capel, F., Bortolotti, F., Berthelot, P., Bréchot, C., and Kremsdorf, D. (1995). *Hepatology* **21**, 10–19.
Rossner, M. T. (1992). *J. Med. Virol.* **36**, 101–117.
Rossner, M. T., Jackson, R. J., and Murray, K. (1990). *Proc. R. Soc. Lond. B* **241**, 51–58.
Runkel, L., Fischer, M., and Schaller, H. (1993). *Virology* **197**, 529–536.
Saito, I., Oya, Y., and Shimojo, H. (1986). *J. Virol.* **58**, 554–560.
Sattler, F., and Robinson, W. F. (1979). *J. Virol.* **32**, 226–233.
Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J., and Howley, P. M. (1990). *Cell* **63**, 1129–1136.
Schek, N., Bartenschlager, R., Kuhn, C., and Schaller, H. (1991). *Oncogene* **6**, 1735–1744.
Schirmacher, P., Rogler, C. E., and Dienes, H. P. (1993). *Virchows Archiv B Cell Pathol.* **63**, 71–89.
Schlicht, H. J., Radziwill, G., and Schafer, H. (1989). *Cell* **56**, 85–92.
Schlüter, V., and Caselmann, W. H. (1995). *Method. Mol. Genetics* **7**, 152–166.
Schlüter, V., Meyer, M., Hofschneider, P. H., Koshy, R., and Caselmann, W. H. (1994). *Oncogene* **9**, 3335–3344.
Schreck, R., Riber, P., and Baeuerle, P. A. (1991). *EMBO J.* **10**, 2247–2258.
Schreck, R., Meier, B., Mannel, D., Droge, W., and Baeuerle, P. A. (1992). *J. Exp. Med.* **10**, 2247–2258.
Scorsone, K. A., Zhou, Y.-Z., Butel, J. S., and Slagle, B. L. (1992). *Cancer Res.* **53**, 368–372.
Seeger, C., Ganem, D., and Vamus, H. E. (1986). *Science* **232**, 477–484.
Seifer, M., and Gerlich, W. H. (1992). *Arch. Virol.* **126**, 119–128.
Seifer, M, Höhne, M., Schaefer, S., and Gerlich, W. H. (1991). *J. Hepatol.* **13** (Suppl. 4), 561–565.
Sell, S., Hunt, J. M., Dunsford, H. A., and Chisari, F. V. (1991). *Cancer Res.* **51**, 1278–1285.
Seto, E., Yen, T. S., Peterlin, B. M., and Ou, J. H. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8286–8290.
Seto, E, Zhou, D. X., Peterlin, B. M., and Benedict Yen, T. S. (1989). *Virology* **173**, 764–766.
Seto, E., Mitchell, P. J., and Yen, T. S. (1990). *Nature* **344**, 72–74.
Shaul, Y., and Ben-Levy, R. (1987). *EMBO J.* **6**, 1913–1920.
Shaul, Y., Rutter, W. J. and Laub, O. (1985). *EMBO J.* **4**, 427–430.
Shaul, Y., Ben-Levy, R., and De Medina, T. (1986). *EMBO J.* **5**, 1967–1971.
Shieh, Y. S., Nguyen, C., Vocal, M. V., and Chu, H.-W. (1993). *Int. J. Cancer* **54**, 558–562.

Shirakata, Y., Kawada, M., Fujiki, Y., Sano, H., Oda, M., Yaginuma, K., Kobayashi, M., and Koike, M. (1989). *Jpn. J. Cancer Res.* **80**, 617–621.
Siddiqui, A., Jameel, S., and Mapoles, J. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 566–570.
Siddiqui, A., Jameel, S., and Mapoles, J. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2513–2517.
Siddiqui, A., Gaynor, R., Srinivasan, A., Mapoles, J., and Farr, R. W. (1989). *Virology* **169**, 479–484.
Simon, D., Knowles, B. B., and Weith, A. (1991). *Oncogene* **6**, 765–770.
Simonsen, C. C., and Levinson, A. D. (1983). *Mol. Cell. Biol.* **3**, 2250–2258.
Slagle, B. L., Zhou, Y.-Z., and Butel, J. S. (1991). *Cancer Res.* **51**, 49–54.
Spandau, D. F., and Lee, C. H. (1988). *J. Virol.* **62**, 427–434.
Spandau, D. F., and Lee, C. H. (1992) *J. Gen. Virol.* **73**, 131–137.
Spandau, D. F., Wang, H. G. H., Fraser, M. J., and Lee, C. H. (1991). *Virology* **185**, 938–941.
Sprengel, R., Kaleta, E. F., and Will, H. (1988). *J. Virol.* **62**, 3832–3839.
Staal, F. J. T., Roederer, M., and Herzenberg, L. A. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9943–9947.
Standring, D. N., Rutter, W. J., Varmus, H. E., and Ganem, D. (1984). *J. Virol.* **50**, 563–571.
Stemler, M., Hess, J., Braun, R., Will, H., and Schroeder, C. H. (1988). *J. Gen. Virol.* **69**, 689–693.
Su, T. S., Lai, C. J., Huang, J. L., Lin, L. H., Yauk, Y. K., Chang, C., and Han, S. H. (1989). *J. Virol.* **63**, 4011–4018.
Summers, J. (1988). *Cancer* **61**, 1957–1962.
Summers, J., Smolec, J., and Snyder, R. (1978). Proc. Natl. Acad. Sci. U.S.A. **75**, 4533–4537.
Suzuki, T., Masui, N., Kajino, K., Saito, I., and Miyamura, T. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8422–8426.
Szmuness, W. (1978). *Prog. Med. Virol.* **24**, 40–69.
Takada, S., and Koike, K. (1989) *Oncogene* **4**, 189–193.
Takada, S., and Koike, K. (1990). *Jpn. J. Cancer Res.* **81**, 1191–1194.
Takahashi, K., Machida, A., Funatsu, G., Nomura, M, Usuda, S., Aoyagi, S., Tachibana, K., Miyamoto, H., Imai, M., Nakamura, T., Miyakawa, Y., and Mayumi, M. (1983). *J. Immunol.* **130**, 2903–2907.
Takahashi, K., Kudo, J., Ishibashi, H., Hirata, Y., and Niho, Y. (1993). *Hepatology* **17**, 794–799.
Takahashi, H., Fujimoto, J., Hanada, S., and Isselbacher, K. J. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1470–1474.
Tavis, J. E., and Ganem, D. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4107–4111.
Terre, S., Petit, M. A, and Bréchot, C. (1991). *J. Virol.* **65**, 5539–5543
Thung, S. N. (1979). *Lab. Invest.* **41**, 101–105.
Tiollais, P., Pourcel, C., and Dejean, A. (1985). *Nature* **317**, 489–495.
Toh, H., Hayashida, H., and Miyata, T. (1983). *Nature* **305**, 827–829.
Tokino, T. and Matsubara, K. (1991). *J. Virol.* **65**, 6761–6764.
Tokino, T., Fukushige, S., Nakamura, T., Nagaya, T., Murotsu, T., Shiga, K., Aoki, N., and Matsubara, K. (1987). *J. Virol.* **61**, 3848–3854.
Tong, M. J. (1979). *Cancer* **44**, 2338–2334.
Treinin, M., and Laub, O. (1987). *Mol. Cell. Biol.* **7**, 545–548.
Truant, R., Antunovic, J., Greenblatt, J., Prives, C., and Cromlish, J. A. (1995). *J. Virol.* **69**, 1851–1859.

Trujillo, M. A., Letovsky, J., Maguire, H. F., Babrera, M. L., and Sidiqui, A. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3797–3801.
Tur-Kaspa, J., Burk, R. D., Shaul, Y., and Shafritz, D. A. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1627–1631.
Twu, J. S., and Robinson, W. S. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2046–2050.
Twu, J. S., and Schloemer, R. H. (1987). *J. Virol.* **61**, 3448–3453.
Twu, J. S., and Schloemer, R. H. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 252–256.
Twu, J. S., Chu, K., Robinson, W. S. (1989). *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5168–5172.
Twu, J. S., Wu, J. Y., and Robinson, W. S. (1990). *Virology* **177**, 406–410.
Twu, J. S., Lai, M. Y., Chem, D. S., and Robinson, W. S. (1993). *Virology* **192**, 346–350.
Ueda, K., Tsurimoto, T., and Matsubara, K. (1991). *J. Virol.* **65**, 3521–3529.
Unger, T., and Shaul, Y. (1990). *EMBO J.* **9**, 1889–1895.
Valenzuela, P., Quiroga, M., Zaldivar, J., Gray, P., and Rutter, W. J. (1980). *In* "Animal Virus Genetics" (B. N. Fields, R. Jaenisch, and C. F. Fox, eds.), pp. 57–70. Academic Press, New York.
Vesselinovitch, S. D., Mihailovich, N., Wogan, G. N., Lombard, L. S., and Kao, K. V. N. (1972). *Cancer Res.* **32**, 2289–2291.
Walker, G. J., Hayward, N. K., Falvey, S., and Cooksley, W. G. E. (1991). *Cancer Res.* **51**, 4367–4370.
Wang, G. H., and Seeger, C. (1992). *Cell* **71**, 663–670.
Wang, H. P., and Rogler, C. E. (1988). *Cytogenet. Cell. Genet.* **48**, 72–78.
Wang, J., Chenivesse, X., Henglein, B., and Bréchot, C. (1990). *Nature* **343**, 555–557.
Wang, J., Zindy, F., Chenivesse, X., Lamas, E., Henglein, B., and Bréchot, C. (1992). *Oncogene* **7**, 1653–1656.
Wang, W., Carey, M., and Gralla, J. D. (1992). *Science* **255**, 450–453.
Wang, X. W., Forrester, K., Yeh, H., Feitelson, M. A., Gu, J. R., and Harris, C. C. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2230–2234.
Wang, Y., Chen, P., Wu, X., Sun, A., Wang, H., Zhu, Y., and Li, Z. (1990). *J. Virol.* **64**, 3977–3981.
Weimer, T., Salfeld, J., and Will, H. (1987). *J. Virol.* **61**, 3109–3113.
Werner, B. G., Smolec, J. M., Snyder, R., and Summers, J. (1979). *J. Virol.* **32**, 314–322.
Will, H., Reiser, W., Weimer, T., Pfaff, E., Büscher, M., Sprengel, R., Cattaneo, R., and Schaller, H. (1987). *J. Virol.* **61**, 904–911.
Williams, J. S., and Andrisani, O. M. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3819–3823.
Wollersheim, M., Debelka, U., and Hofschneider, P. H. (1988). *Oncogene* **3**, 545–552.
Wu, H. L., Chen, P. J., Tu, S. J., Lin, M. H., Lai, M. Y., and Chen, D. S. (1991). *J. Virol.* **65**, 1680–1686.
Wu, J. Y., Zhou, Z. Y., Judd, A., Cartwright, C. A., and Robinson, W. S. (1990) *Cell* **63**, 687–695 [published erratum (1993). *Cell* **75**].
Yaginuma, K., Shirakata, Y., Kobayashi, M., and Koike, K. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2678–2682.
Yaginuma, K., Nakamura, I., Takada, S., and Koike, K. (1993). *J. Virol.* **67**, 2559–2565.
Yee, J. K. (1989). *Science* **246**, 658–661.
Yeh, C. T., Liaw, Y. F., and Ou, J. H. (1990). *J. Virol.* **64**, 6141–6147.
Yeh, C. T., Wong, S. W., Fung, Y.-K., and Ou, J.-H. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6459–6463.
Yuh, C. H., and Ting, L. P. (1990). *J. Virol.* **64**, 4281–4287.
Zahm, P., Hofschneider, P. H., and Koshy, R. (1988). *Oncogene* **3**, 169–177.
Zelent, A., Sells, M. A., Price, P. M., Mohamad, A., Acs, G., and Christman, J. K. (1987). *J. Virol.* **61**, 1108–1115.

Zhang, W., Hirohashi, S., Tsuda, H., Shimosato, Y., Yokota, J., and Terada, M. (1990). *Jpn. J. Cancer Res.* **81**, 108–111.

Zhang, X. K., Egan, J. O., Huang, D. P., Sun, Z. L., Chien, V. K. Y., and Chiu, J. F. (1992). *Biochem. Biophys. Res. Comm.* **188**, 344–351.

Zheng, Y. W., and Yen, T. S. B. (1994). *J. Biol. Chem.* **269**, 8857–8862.

Zhou, D. X., Taraboulos, A., Ou, J. H., and Yen, T. S. B. (1990). *J. Virol.* **64**, 4025–4028.

Zoulim, F., Saputelli, J., and Seeger C. (1994). *J. Virol.* **68**, 2026–2030.

Zoulim, F., and Seeger, C. (1994). *J. Virol.* **68**, 6–13.

ADVANCES IN VIRUS RESEARCH, VOL. 47

AUTONOMOUS PARVOVIRUS AND DENSOVIRUS GENE VECTORS

Joe Corsini,* Boris Afanasiev,*‡ Ian H. Maxwell,† and Jonathan O. Carlson*

*Department of Microbiology
Colorado State University
Fort Collins, Colorado 80523
†University of Colorado Cancer Center and Division of Medical Oncology
University of Colorado Health Sciences Center
Denver, Colorado 80262
‡Engelhardt Institute of Molecular Biology
Moscow, Russia

I. Introduction

Viruses have evolved as efficient natural vehicles of genetic information and they are logical systems to develop as gene vectors. Transfer of heterologous DNA by an animal virus (SV40) was first demonstrated in the 1970s (Sambrook, 1978; Mulligan *et al.*, 1979) and seminal work developing systems for retrovirus-mediated gene transfer took place in the early 1980s (Miller *et al.*, 1983; Mann *et al.*, 1983; Joyner *et al.*, 1983).

Since this early work, the notion of gene replacement therapy has become more tantalizing as the mechanisms which generate genetic disorders have become better understood. It is now apparent that a number of genetic diseases are caused by single-gene defects that could theoretically be cured if a functional gene were introduced into a defective tissue. The concept of curing other diseases, such as cancer and

infectious disease, with genetic therapy has also emerged, and numerous anticancer and antiviral genes are being engineered or isolated. As a result, a number of virus families are now being explored as potential delivery vehicles for application in gene therapy. These include retroviruses (Culver *et al.,* 1992; Cournoyer and Caskey, 1993; Kay *et al.,* 1993; Lin *et al.,* 1994), adenoviruses (Berkner, 1988; Rosenfeld *et al.,* 1992; Welsh *et al.,* 1994), herpesviruses (Breakfield and DeLuca, 1991; Smith *et al.,* 1994), polyomaviruses (Oppenheim *et al.,* 1986; Oppenheim and Peleg, 1989), positive-strand RNA viruses (alphaviruses and picornaviruses) (Choi *et al.,* 1991; Liljestrom and Garoff, 1991; Bredenbeek and Rice, 1992; Hahn *et al.,* 1992; Bredenbeek *et al.,* 1993; Porter *et al.,* 1995), and negative-strand RNA viruses (orthomyxoviruses and paramyxoviruses) (Park *et al.,* 1991; Collins *et al.,* 1991). Several reviews have dealt with viral vectors (Anderson, 1984, 1992; Breakfield and DeLuca, 1991; Muzyczka, 1992; Roemer and Friedman, 1992; Bredenbeek and Rice, 1992; Miller, 1992; Schlesinger, 1993; Mulligan, 1993; Kotin, 1994).

In addition to their application to gene therapy, viral vectors are powerful tools for probing basic processes of viruses and the cells they infect. Retrovirus vectors, e.g., have been used as traps to identify cellular promoters (Melchner *et al.,* 1990), and we have used parvovirus vectors to answer basic questions about the encapsidation of autonomous parvovirus genomes (Corsini *et al.,* 1995) and the determinants of autonomous parvovirus tissue tropisms (Maxwell *et al.,* 1995). Herpesvirus vectors are being developed to study gene function in neurons (Breakfield and DeLuca, 1991), and adeno-associated virus (AAV) has been used to produce recombinant AAV/B19 virus to facilitate study of parvovirus B19 (Srivastava *et al.,* 1989). Thus, though gene therapy applications have fueled the development of viral vectors, they are also proving to be of considerable value to more basic studies.

Further possibilities exist in the engineering of invertebrate viruses for study and genetic manipulation of invertebrate species. At this time, several insect viruses are being developed toward this end (Lucknow and Summers, 1988; Carlson *et al.,* 1995; Iatrou and Meidinger, 1989). Another interesting possibility for use of insect viruses as gene delivery vehicles is in insect control strategies that utilize a virus to deliver insect-specific toxin genes for pest control of target insects (Cory *et al.,* 1994; Belloncik, 1990; Tal and Attathom, 1993).

Several members of the family *Parvoviridae* are being evaluated for their potential as gene delivery vehicles. Transfer of a foreign gene by a parvovirus was first demonstrated with adeno-associated virus (Hermonat and Muzyczka, 1984; Tratschin *et al.,* 1984a); development of the AAV gene delivery system has progressed to the point where AAV

vectors may soon be used in clinical trials (reviewed by Carter, 1990a; Muzyczka, 1992; Kotin, 1994). Until recently, however, there has been little effort to exploit the gene delivery potential of autonomous parvoviruses or the parvoviruses of arthropods, the densoviruses. Although delivery systems based on both types of parvovirus are relatively inchoate, there have been significant advances in their development as gene delivery vectors and they offer many exciting possibilities, if some of the technical hurdles can be overcome.

II. Parvoviruses

Parvoviruses are small nonenveloped viruses with linear, single-stranded DNA genomes of approximately 4000–6000 nucleotides. Historically, the family has been divided into three genera: dependoviruses (adeno-associated parvoviruses) and autonomous parvoviruses, which both infect mammalian and avian species, and densoviruses, which infect arthropods. All are stable in the environment, and most are resistant to short-term heating and mild extraction with organic solvents. The known capsid structures of parvoviruses have been reviewed (Agbandje *et al.,* 1995). Tsao *et al.* (1991) have obtained a 2.8-Å X-ray crystal structure for the canine parvovirus (CPV), an autonomous parvovirus, showing it to be icosahedral with canyons encircling the fivefold axes and spikes at the threefold axes; unlike the rhinoviruses, the canyons of CPV do not appear to be the sites of receptor binding (Strassheim *et al.,* 1994; Chapman and Rossman, 1993). The crystal structure of human parvovirus B19 has been resolved to 8.0 Å; it appears to lack the prominent spikes that are present on the threefold axes of CPV (Agbandje *et al.,* 1994). Cellular receptors for parvoviruses have not been identified, with the exception of B19, whose receptor was shown to be the glycolipid globoside (P antigen) present on erythroid cells (Brown *et al.,* 1993).

Parvovirus replication occurs in the nucleus, requiring viral nonstructural proteins, cellular DNA polymerases, and other as yet poorly characterized nuclear factors (Berns, 1990; Astell, 1990; Tattersall and Cotmore, 1990; Tattersall and Gardiner, 1990; Cotmore and Tattersall, 1995). Replication origins exist in the terminal 100–200 nucleotides of the linear genome (Rhode, 1977; Hauswirth and Berns, 1977; Faust and Ward, 1979; Rhode and Klassen, 1982; Senapathy *et al.,* 1984) and there is evidence that sequences internal to the termini of the rodent parvovirus minute virus of mice (MVM) might contribute to replication efficiency in some cell types (Salvino *et al.,* 1991; Tam and Astell, 1993).

The genomic termini contain palindromic sequences able to form hairpin structures which act as replication origins (Astell, 1990; Tattersall and Cotmore, 1990; Cotmore and Tattersall, 1995). A simplified model of parvovirus DNA replication is shown in Fig. 1. For AAV, a double-stranded monomeric replicative form (RF) molecule is synthesized by self-priming from the 3′ palindrome. This RF molecule is then nicked by site-specific endonuclease activities of the Rep proteins, and in a process dubbed "hairpin transfer" the termini are replicated (Im and Muzyczka, 1990; Berns, 1990; Muzyczka, 1992). The situation with MVM appears to be more complicated, requiring a dimeric intermediate to explain the observed patterns of RF resolution and genome formation (Astell *et al.*, 1985; Tyson *et al.*, 1990; Cotmore *et al.*, 1993; Cotmore and Tattersall, 1994; Liu *et al.*, 1994). As with AAV, the major nonstructural protein (NS1) is the enzyme which nicks RF to allow hairpin transfer (Nuesch *et al.*, 1992). The terminal palindromes also contain packaging signals (Rhode, 1978; Faust and Ward, 1979; McLaughlin *et al.*, 1988; Samulski *et al.*, 1989) which mediate concomitant generation of single-stranded progeny DNA and encapsidation (Rhode, 1976; Richards *et al.*, 1977; Muller and Siegl, 1983a,b; Myers and Carter, 1980; Tratschin *et al.*, 1984b; Hermonat *et al.*, 1984; Cotmore and Tattersall, 1989; Willwand and Kaaden, 1990; Willwand and Hirt, 1991, 1993). As discussed below, presence of the replication and packaging signals in the termini allows replacement of nearly the entire genome with foreign DNA.

A. Adeno-Associated Virus

Adeno-associated virus (AAV) is so named because it was discovered as a contaminant in stocks of adenovirus (Atchison *et al.*, 1965). Productive AAV infection requires a helper virus, either herpesvirus or adenovirus [vaccinia virus has also been reported to "help" AAV (Schlehofer *et al.*, 1986)]. The actions of helper adenovirus on AAV are complicated and not completely understood; it is clear, however, that E1A, E1B, E2A, E4, and the VA transcripts contribute helper functions (Carter, 1990b). Certain experimental conditions allow limited helper-free replication, showing that the dependent condition of AAV is not absolute (Schlehofer *et al.*, 1986; Yakobson *et al.*, 1989). AAV seems to be nonpathogenic (Siegl *et al.*, 1985), although recently AAV-2 has been shown to induce fetal death in mice (Botquin *et al.*, 1994). Interestingly, AAV integrates into the host genome in the absence of helper virus (Cheung *et al.*, 1980; Carter *et al.*, 1990a; Muzyczka, 1992). Integration has been reported to occur preferentially in a specific 100-nucleotide

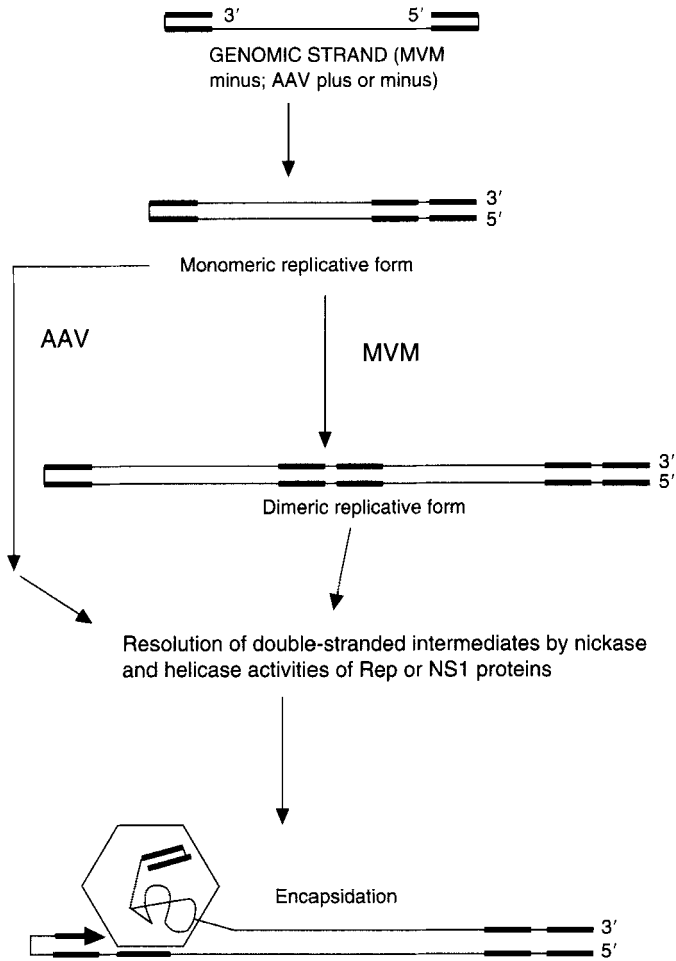


FIG. 1. Simplified model of parvovirus replication. Black bars represent palindromic sequences. AAV is adeno-associated virus and MVM is the minute virus of mice. A single-stranded genome is converted in the nucleus of an infected cell to a monomeric double-stranded form by cellular DNA polymerases. In the case of AAV, this monomer is nicked and unwound by the Rep proteins, after (or perhaps during) which single strands are thought to be threaded into preformed capsids. In the case of MVM, the dimeric molecule appears to be the substrate for NS1 nicking and helicase activities.

target region of chromosome 19 of several human cell lines (Kotin and Berns, 1989; Samulski *et al.*, 1991; Kotin *et al.*, 1992). Target specificity of this 100-nucleotide region was maintained after its insertion into Epstein-Barr-based plasmids (Giraud *et al.*, 1994) and Rep has been

shown to nick the target sequence and mediate its asymmetric replication *in vitro* (Urcelay *et al.*, 1995).

The AAV genomic organization and transcription strategy is shown in Fig. 2A. Upon entry into the nucleus, genomic plus or minus strands are converted to double-stranded intermediates which are presumably the templates for transcription (as well as replication). Three promoters regulate synthesis of at least seven RNAs (Green and Roeder, 1980; Hermonat *et al.*, 1984; Carter *et al.*, 1990b). Four of these originate from the left end of the genome, from promoters at map units 5 or 19, and encode the Rep proteins, Rep 78, Rep 68, Rep 52, and Rep 40. The Rep proteins are intimately involved with DNA replication, although the specific functions of all have yet to be defined. The three capsid proteins of AAV are translated from the two spliced transcripts that originate from the promoter at map unit 40, P40, with VP1 originating from one transcript and VP2 and VP3 originating from different start codons on the other (Becerra *et al.*, 1988; Cassinotti *et al.*, 1989). The start codon for VP1 is located between two alternative splice acceptors; use of the second acceptor eliminates this ATG and generates the more abundant mRNA encoding VP2 and VP3. A nonstandard start codon (ACG) upstream of the AUG for VP3 is utilized for translation of VP2 (Becerra *et al.*, 1988).

B. Autonomous Parvoviruses

The autonomous parvoviruses fall into distinct evolutionary groups (Truyen *et al.*, 1995). The rodent parvoviruses include H-1, minute virus of mice, Kilham's rat virus, LuIII, and several new isolates (Ball-Goodrich and Johnson, 1994; Besselsen *et al.*, 1994). The feline parvovirus subgroup includes feline panleukopenia virus, mink enteritis virus, raccoon parvovirus, and canine parvovirus (Truyen *et al.*, 1995). The rodent, feline, and porcine parvoviruses have similar genomic organization (Bergeron *et al.*, 1993), while the more distantly related bovine parvovirus (BPV), B19 (human), and Aleutian disease virus (mink) have unique genomic organizations. The autonomous parvoviruses exhibit a range of encapsidation patterns. H-1 and MVM encapsidate predominantly the minus strand, bovine parvovirus encapsidates the plus strand approximately 10% of the time, and LuIII and B19 efficiently encapsidate either strand (Shull *et al.*, 1988; Bates *et al.*, 1984; Muller and Siegl, 1983a,b; Cotmore and Tattersall, 1984; Corsini *et al.*, 1995). The genomic organization and transcription strat-

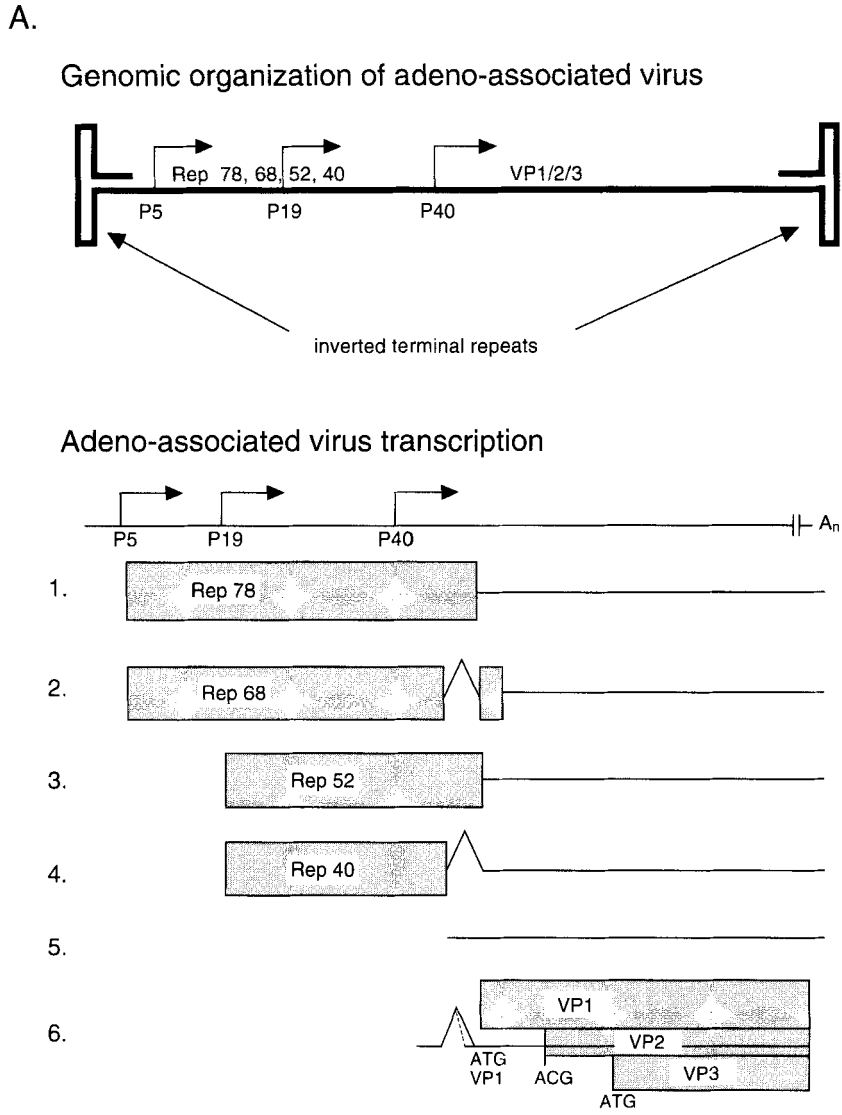
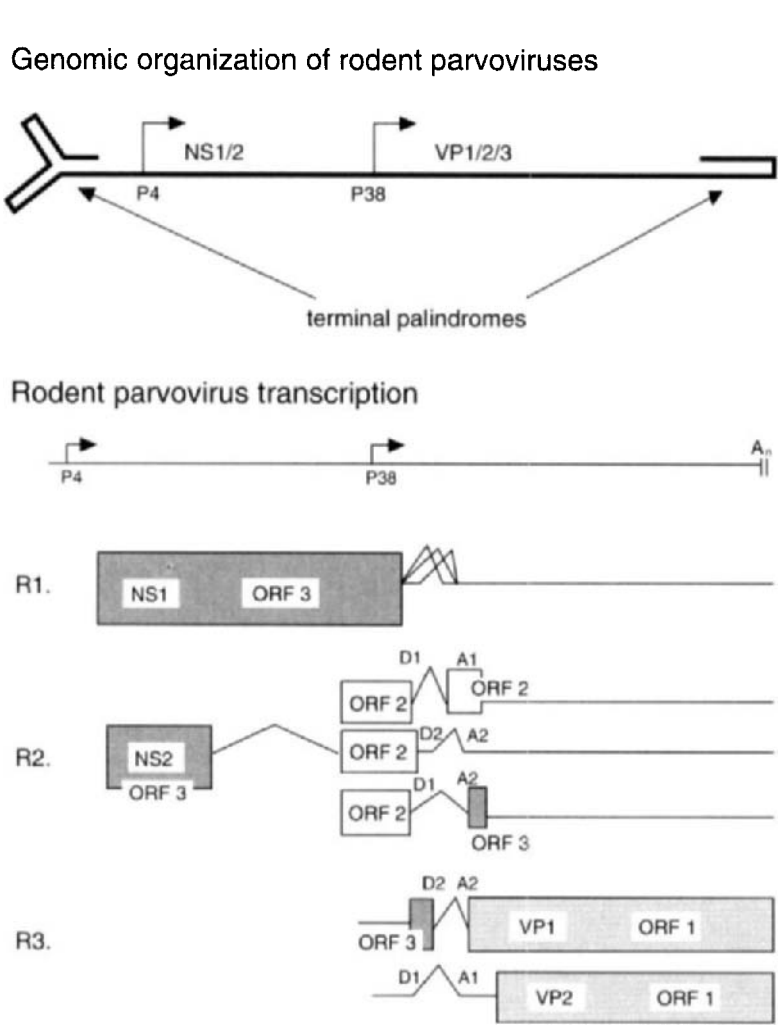


FIG. 2. (A) Genomic organization of adeno-associated virus. Four Rep proteins, Rep 78, Rep 68, Rep 52, and Rep 40, are produced from spliced mRNAs initiating at promoters P5 or P19. The three coat proteins, VP1, VP2, and VP3, are produced from two differentially spliced mRNA's initiating at promoter P40. (B) Genomic organization of the rodent parvoviruses (Pintel *et al.*, 1995). Two nonstructural proteins, NS1 and NS2, are produced from spliced mRNAs initiating from the P4 promoter. Two coat proteins, VP1



and VP2, are produced from alternatively spliced mRNA's initiating at the P38 promoter. (C) Genomic organization of two densoviruses, *Junonia coenia* DNV (JcDNV, top panel) and *Aedes aegypti* DNV (AeDNV, bottom panel). The structural proteins of JcDNV are encoded by one strand and the nonstructural by the other. Pr indicates the positions of the JcDNV promoters. All proteins of the AeDNV are encoded on one strand. P7 indicates the promoter at map unit 7 and P61 indicates the promoter at map unit 61.

FIG. 2. *(continued)*

C.

Genomic organization of densoviruses

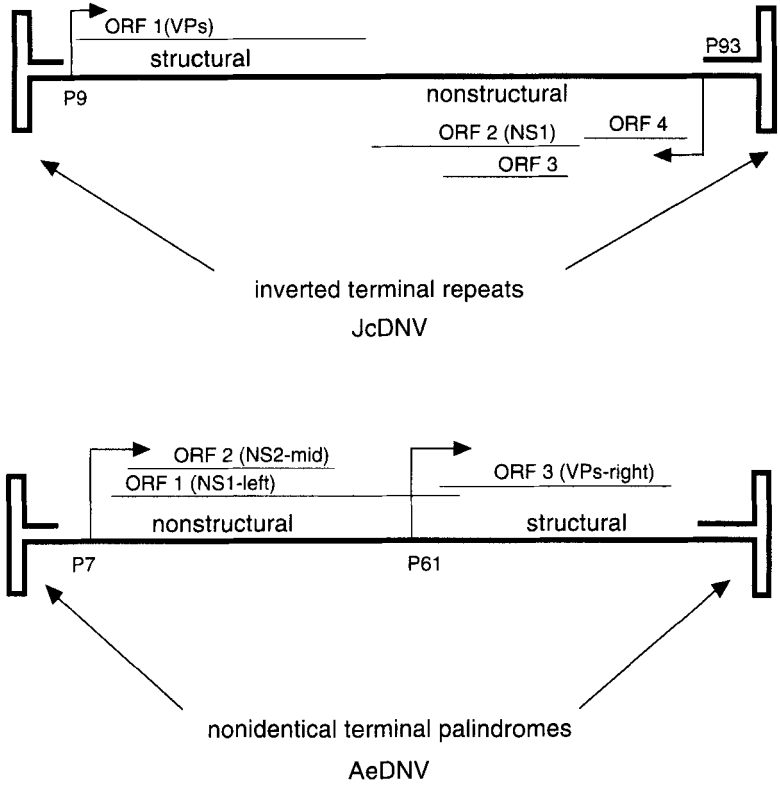


FIG. 2. *(continued)*

egy of the rodent parvoviruses is shown in Fig 2B. The P4 promoter is moderately strong and constitutive, whereas the P38 promoter is only active in the presence of NS1, the major gene product from the P4 transcription unit (Rhode, 1985; Doerig *et al.,* 1988; Doerig *et al.,* 1990; Christensen *et al.,* 1995). The nonstructural proteins of other autonomous parvoviruses also upregulate their capsid gene promoter (Doerig *et al.,* 1987; Clemens and Carlson, 1989; Doerig *et al.,* 1990; Christensen *et al.,* 1993). The P4 transcripts are alternatively spliced to generate mRNAs encoding the nonstructural proteins NS1 and NS2 (Pintel *et al.,* 1983; Ben-Asher and Aloni, 1984; Morgan and Ward, 1986; Cotmore

and Tattersall, 1986a; Jongeneel *et al.*, 1986; Cotmore, 1990), and there are apparently sequences in the downstream small intron (see Fig. 2B) that affect the efficiency with which the large intron is spliced (Zhao *et al.*, 1994). The P38 transcripts are also spliced using alternative splice donors and acceptors to generate mRNAs encoding the minor (VP1) and major (VP2) capsid proteins (Pintel *et al.*, 1983; Ben-Asher and Aloni, 1984; Labniec-Pintel and Pintel, 1986; Jongeneel *et al.*, 1986; Cotmore, 1990; Pintel *et al.*, 1995).

NS1 is a multifunctional protein containing helicase, endonuclease activities, and concomitant ATPase activities that are involved with replication and resolution of the replicative intermediate (Rhode, 1985, 1989; Tullis *et al.*, 1988; Cotmore and Tattersall, 1992; Wilson *et al.*, 1991; Nuesch *et al.*, 1992; Cotmore *et al.*, 1993). NS1 is at least transiently attached covalently to the 5′ terminus of the packaged genome, a feature that may have some bearing on the encapsidation process (Cotmore and Tattersall, 1988). This protein has also been shown to be toxic to cultured mammalian cells (Caillet-Fauquet *et al.*, 1990), a property which has frustrated attempts to generate NS1-producing cell lines. NS2 is a minor nonstructural protein shown, for MVM, to be necessary for efficient infection of murine A9 cells but not of human NB324K cells (Naeger *et al.*, 1990) and, for H-1, to be required for efficient infection of rat cells and newborn rats (Li and Rhode, 1991). Thus, NS2 seems to be involved in the cell type and species specificity of parvoviruses. As with AAV, autonomous parvovirus particles contain three capsid proteins; VP1 and VP2 are translated from alternatively spliced mRNAs, while VP3 is a proteolytic cleavage product of VP2 (Tattersall *et al.*, 1976).

C. Densoviruses

Parvoviruses isolated from arthropods (mostly insects) have been conventionally called densoviruses (shortened from "densonucleosis viruses," which was originally used because of the observation that cells infected with these viruses develop characteristic densely stained nuclei). We now know that densoviruses and vertebrate parvoviruses have very similar morphogenesis except that densoviruses (DNVs) usually also cause hypertrophy of infected nuclei (Siegl, 1984; Tijssen *et al.*, 1990). About 25 DNVs have been isolated to date, most from insects and several from crustaceans (Tijssen and Bergoin, 1995). The subfamily *Densovirinae* is divided into three genera: *Densovirus* (lepidopteran DNVs), *Contravirus* (*Aedes* DNVs), and *Iterovirus* (*Bombyx*

parvovirus) (Murphy *et al.,* 1995). The host range of densoviruses varies with species. For example, GmDNV infects only *Galleria mellonella,* while JcDNV infects *Junonia coenia, Bombyx mori, Lymantria dispar, Mamestra brassicae, Spodoptera littoralis* and *Spodoptera exigua,* and others (Kawase *et al.,* 1990). The *Aedes aegypti* virus (AeDNV) infects numerous mosquito species from the genera *Aedes, Culex,* and *Culiseta* (Lebedeva *et al.,* 1973). At the tissue level, some of the densoviruses, such as GmDNV and AeDNV, are polytropic. Others, such as BmDNV or *Sibine fusca* DNV, infect only the midgut tissues (Kawase *et al.,* 1990).

Because Section III treats the densoviruses and contraviruses, this introduction will focus on these genera. Sequencing of *Junonia coenia* (JcDNV) (Dumas *et al.,* 1992) and *Galleria mellonella* (GmDNV) (Gene Bank accession L32896) revealed genomes about 6000 nucleotides in length. The genomic organization of these viruses is shown in Fig. 2C. They contain inverted terminal repeats (ITR) of approximately 540 nucleotides. The first 136 nucleotides (GmDNV) of the ITR can be folded into a typical T-shaped hairpin structure. Genes for nonstructural and structural proteins are located on complementary strands (Fig. 2C). Three ORFs originating on one strand of the genome encode nonstructural proteins (NS1, NS2, NS3) and one ORF on the other strand putatively encodes four structural proteins (VP1, VP2, VP3, VP4) with approximate sizes of 88, 70, 60, and 45 kDa. Two promoters (p9 and p93) are thought to drive the transcription of two mRNAs, one for nonstructural proteins and the other for capsid proteins (Tal and Attathom, 1993). These viruses package both strands with equal frequency. Preliminary evidence indicates that the viruses of these genera, like AAV, are able to integrate into host genome under appropriate conditions (Dumas *et al.,* 1992).

Densoviruses of the genus *Contravirus* were isolated from mosquitos of the family *Culicidae* (reviewed in Buchatsky, 1989). Because of the small size of their genome (~4000 nt), the name *Brevidensovirus* was proposed recently (Tijssen and Bergoin, 1995) and seems appropriate. The genomic sequences of *Aedes aegypti* virus (AeDNV) and *Aedes albopictus* virus (AaPV or AaDNV) have been determined (Afanasiev *et al.,* 1991; Boublik *et al.,* 1994). The genomic organization of these virues is shown in Fig. 2C. In contrast to the JcDNV and GmDNV, all proteins are encoded on one strand. The terminal sequences (left ~ 140 nt and right ~ 170 nt) of the genome of these viruses are mostly unique and can assume the T-shaped secondary structures which act as replication origins (authors' unpublished data). Two virion proteins (38 and 40

kDa) have been identified (Afanasiev *et al.*, 1991) and sequence analysis reveals ORFs encoding two putative nonstructural proteins. These viruses package predominantly (85–90%) minus strand.

Parvoviruses carry out their infectious cycle in the cell nucleus, so most virally encoded proteins (VP and NS1) are localized in the nucleus (Cotmore and Tattersall, 1986b; Nuesch *et al.*, 1992; Yang *et al.*, 1992). The presence of a sequence in NS1 and capsid proteins of AeDNV and JcDNV similar to the mammalian parvovirus nuclear localization signal suggests that these proteins are translocated to the nucleus; indeed, VPs of JcDNV and AeDNV appear to be translocated to the nucleus (Giraud *et al.*, 1992; Afanasiev *et al.*, 1994). The NS1 gene of densoviruses contains two sequence motifs in common with mammalian parvoviruses: the helicase protein motif, including NTP-binding domain (Anton and Lane, 1986; Astell *et al.*, 1987; Koonin, 1993), and a recently described motif in the rolling circle replication initiator protein (Ilyina and Koonin, 1992; Nuesch *et al.*, 1995; Tijssen and Bergoin, 1995). Similar to NS1 of mammalian parvoviruses, the NS1 proteins of two densoviruses have been shown to mediate excision of a cloned genome from plasmid and to trans-activate viral promoters (Giraud *et al.*, 1992; Afanasiev *et al.*, 1994). The NS2 proteins of the densoviruses have not been characterized. Densoviruses generally contain two to four virion proteins (Kelly *et al.*, 1980; Afanasiev *et al.*, 1991; Nakagaki and Kawase, 1980), although a putative parvovirus isolated from shrimp may have only one virion protein (Bonami *et al.*, 1995). Densovirus capsid proteins are encoded by one ORF as a nested set of overlapping molecules and it is not definitively known whether translation requires alternative splicing, internal initiation of the ribosome, or, as occurs with mammalian parvoviruses, proteolytic processing of the virion proteins. *In vitro* translation experiments with GmDNV RNA suggest that nested nonstructural and structural proteins are expressed from one mRNA via multiple internal initiations of the ribosome (Tal and Attathom, 1993), but it is not known whether this is a common characteristic of densoviruses.

Available sequence analyses (Bando *et al.*, 1987, 1990; Afanasiev *et al.*, 1991; Dumas *et al.*, 1992; Boublik *et al.*, 1994) show that the densoviruses are at least as heterogeneous as the mammalian parvoviruses, although they are similar in their general functional organization. Preliminary data suggest differences in the mechanistic details of function between densoviruses and mammalian parvoviruses that probably reflect differences in the biology of invertebrate cells compared to mammalian cells. For example, transcriptional activators used in mammalian cells are probably quite different from those in

insect cells (Courey and Tjian, 1988; Hoey *et al.,* 1993); consequently, the upstream sequence motifs common to promoter regions of mammalian parvoviruses are not observed in the genome of DNVs. Instead, sequences located downstream of the TATA box might be used to specify and/or modulate the expression in the DNV genome; thus, the consensus for a proposed arthropod initiator element, A/C/T CAGT (Hultmark *et al.,* 1986; Cherbas and Cherbas, 1993), occurs close (downstream) to densovirus TATA boxes. Another difference appears to be lack of splicing in the insect parvoviruses. Although RNA splicing commonly occurs in insect cells, preliminary data suggest that splicing is not required by GmDNV (Tal and Attathom, 1993). Interestingly, another group of insect DNA viruses, the *Baculoviridae,* does not appear to utilize splicing (Bilimoria, 1991).

D. Oncolytic Effects of Parvoviruses

Since autonomous parvovirus vectors have potential application in cancer therapy, it is worth summarizing the anticancer (antioncogenic) characteristics of wild-type parvoviruses. There have been numerous demonstrations that parvoviruses can protect against tumorigenesis in animals. Infection of neonatal hamsters with H1 virus reduced the subsequent incidence of spontaneous tumors in surviving animals by up to 20 times (Toolan, 1967). The protective effect was greater in those animals that had experienced more severe infection and impaired development of teeth, palate, and skull. It was later shown that such hamsters resisted induction of tumorigenesis by both chemical carcinogens and oncogenic viruses; the incidence of dimethylbenzanthracene-induced (Toolan *et al.,* 1982) or adenovirus 12–induced (Toolan and Ledinko, 1968) tumors was 40% of that in control animals. Furthermore, concomitant parvovirus infection proved strongly inhibitory to the growth of transplanted tumors in animals (Guetta *et al.,* 1986; Dupressoir *et al.,* 1989; Yang, 1987). Protection against tumor growth from intraperitoneally injected Ehrlich ascites cells was elicited in 90% of mice receiving concomitant intramuscular injection of the fibrotropic strain of minute virus of mice (MVMp), while the lymphotropic strain of MVM (MVMi), to which these ascites cells are not susceptible, did not inhibit tumor growth (Guetta *et al.,* 1986). The MVMp-injected mice were also protected against subsequent reinjection of the tumor cells. These observations provided evidence for a direct interaction between the virus and the tumor cells *in vivo,* although additional involvement of the immune system could not be excluded. More recently, parvovirus H-1 was shown to protect nude mice from developing tumors

arising from injection of SV40-transformed human mammary epithelial cells which were susceptible to H-1 killing *in vitro* (Dupressoir *et al.*, 1989). Moreover, intravenous injection of H-1 was effective in causing complete regression in two of four animals that had already developed 0.5-cm tumors following injection of these cells.

The oncolytic effect of parvoviruses is not limited to rodents; tumorigenesis in dogs by canine transmissible venereal sarcoma cells was inhibited by both virulent and vaccine strains of CPV (Yang, 1987), the former being more effective. The oncolytic property of parvoviruses has led to one report of a "last-ditch" attempt to treat human osteosarcoma patients with the parvovirus H-1, which was administered to two 13-year-old girls with advanced disease; no tumor regression was observed, although H-1 viremia was established (Toolan *et al.*, 1965). It is possible that treatment might have been more successful had the virus been administered earlier.

Oncolytic effects have also been described for AAV. A substantial inhibition of adenovirus 12 tumorigenesis by AAV was observed in newborn hamsters; in those experiments, AAV and its defective interfering particles both effectively reduced tumors (Maza and Carter, 1981). Such examples probably involve interference with the oncogenic functions of adenoviruses (or herpes simplex virus) during reciprocal interactions of AAV with these helpers (Rommelaere and Tattersall, 1990). Protection by AAV in humans against cervical and prostate carcinoma can be circumstantially inferred from the lower frequency of antibodies against AAV in such patients compared with normal subjects (Rommelaere and Tattersall, 1990). Selective inhibition of *in vitro* proliferation of freshly isolated human melanoma and carcinoma cells by AAV has also been reported (Bantel-Schaal, 1990).

Although the mechanisms of parvovirus-mediated oncosuppression are not known, it has frequently been observed that transformation of cultured cell lines by chemical or physical agents, or oncogenic viruses, is accompanied by markedly increased cytotoxicity upon infection by autonomous parvoviruses (Guetta *et al.*, 1990; Cornelis *et al.*, 1988a,b; Faisst *et al.*, 1989; Mousset *et al.*, 1994). Observations that cell killing correlated with the level of expression of nonstructural proteins, particularly NS1, rather than with virus production, led to the conclusion that intracellular accumulation of these proteins was toxic (Van Hille *et al.*, 1989; Caillet-Fauquet *et al.*, 1990; Becquart *et al.*, 1993). It therefore seems likely that preferential expression and toxicity of nonstructural proteins in cancer cells may contribute to the observed oncolytic effects of parvoviruses.

III. Parvovirus Gene Vectors

As noted in Section I, many virus families are being engineered into gene delivery vehicles. In general, the principles behind this technology are based on genetic complementation wherein a section of the viral genome is replaced with heterologous sequences and the deleted region is provided in trans on a separate piece of nucleic acid. Similarly, the general approach to producing recombinant parvoviruses utilizes a helper genome that provides nonstructural and structural proteins for the transducing genome, which has had some or all of its coding sequences replaced by a foreign gene (Fig. 3). The helper can be either a wild-type viral genome or a defective viral genome that cannot replicate or be packaged. In most cases, a defective genome is used and steps are taken to minimize recombination, which can result in contamination of recombinant virus stocks with infectious virus. As noted in Section II, the replication and packaging signals are contained within the terminal palindromes, so the transducing genome typically consists of a foreign gene bounded by the parvovirus termini. Recombinant virus is generated by cotransfection of the two defective viral genomes into producer cells and incubation for 3–5 days. After this time, recombinant "transducing virus" is harvested from cell culture medium combined with freeze–thaw cell lysates, clarified by centrifugation or filtration, and stored at –20°C.

A. Adeno-Associated Virus Vectors

Since the focus of this review is autonomous parvovirus and densovirus gene vectors, we will deal only briefly with AAV vectors. For more extensive reviews see Muzyczka (1992) and Kotin (1994). There are several features of AAV that make it attractive as a gene delivery vehicle. Because AAV is able to integrate into the host cell genome, vectors based on this virus are able to deliver genes permanently to infected cells. As mentioned previously, this integration is targeted to a small region of chromosome 19 (Kotin and Berns, 1989; Samulski *et al.*, 1991; Kotin *et al.*, 1992) and it appears that this characteristic will be useful in overcoming positional effects of random integration. However, there are several aspects of AAV biology that warrant closer study. First, targeted integration seems to require the Rep proteins to maintain specificity (Muzyczka, 1992; Urcelay *et al.*, 1995). This is potentially a problem, not only because retention of the *rep* region would severely limit capacity for foreign DNA but also because the *rep* gene products

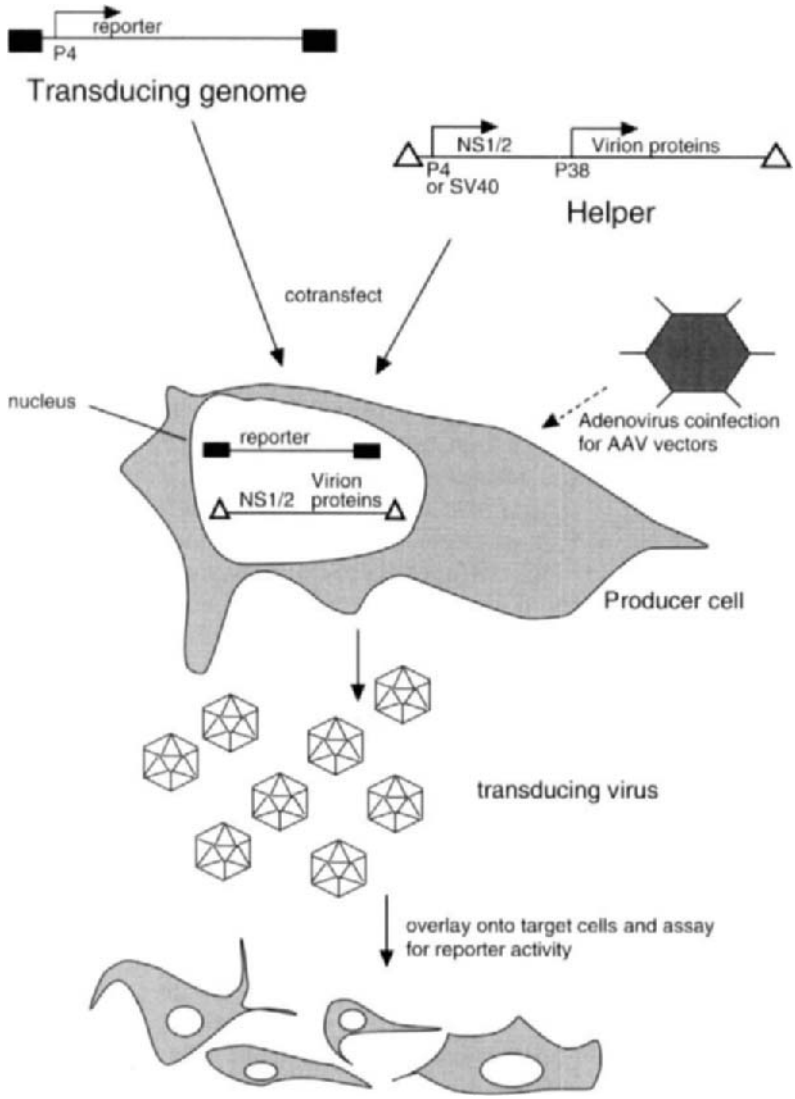


FIG. 3. General strategy for producing recombinant parvovirus (transducing virus). Two plasmids, one containing a helper genome and the other containing a transducing genome carrying the gene of interest, are cotransfected into producer cells. In the case of adeno-associated virus, coinfection with adenovirus is necessary for production of AAV transducing virus. The helper provides proteins required for replication and packaging of the transducing genome. Producer cells are typically incubated for 3–5 days, after which cells and medium are frozen and thawed three times to release cell-associated transducing virus. The lysate containing transducing virus (termed transducing medium) is clarified by centrifugation and then stored frozen for analysis.

exert variable effects on their own and heterologous promoters depending on the target cell line (McLaughlin *et al.,* 1988; Tratschin *et al.,* 1986; Mendelson *et al.,* 1988; Labow and Berns, 1988; Labow *et al.,* 1987; Antoni *et al.,* 1991; Rittner *et al.,* 1992; Yang *et al.,* 1995). Second, AAV-2 and vectors based on AAV-2 have been observed to inhibit embryogenesis in mice (Botquin *et al.,* 1994) (interestingly, the antiembryogenic effect occurred in the absence of *rep* gene sequences), so it is possible that these vectors might interfere with differentiation of some progenitor cells. Third, episomal copies of AAV–*neo* transducing genomes have been observed in transduced cell lines, indicating that the intracellular state of the virus is not always predictable (McLaughlin *et al.,* 1988). Fourth, it has been noted that the region of chromosome 19 into which AAV integrates (in the presence of *Rep*), q13.3–qter, displays a high frequency of crossover events and rearrangements that may lead to chronic B-cell leukemia; consequently, integration into this region may at some point present itself as a safety issue (Mulligan, 1993). Fifth, many targets for gene therapy are nondividing cells, and although AAV vectors seem to integrate into nondividing cells, there is some controversy as to the efficiency with which this occurs (Russell *et al.,* 1994; Podsakoff *et al.,* 1994; Halbert *et al.,* 1995; Alexander *et al.,* 1994).

The general strategy for producing recombinant AAV is shown in Fig. 3. Infectious clones were derived by Samulski *et al.* (1982) and Laughlin *et al.* (1983). Production of recombinant AAV requires the presence of adenovirus (or herpes simplex virus), which is included during the cotransfection step in Fig. 3. Adenovirus is then removed from the transducing stocks by heat inactivation and/or CsCl banding. Stocks of AAV transducing virus are quite stable and can be concentrated by high-speed centrifugation (Samulski *et al.,* 1989).

The original AAV gene delivery vectors demonstrated the ability to deliver the neomycin phosphotransferase (*neo*) gene or the chloramphenicol acetyl transferase gene (CAT) to cultured human cells (Hermonat and Muzyczka, 1984; Tratschin *et al.,* 1984a, 1985). Since the early studies, numerous genes have been inserted between the AAV termini and delivered to cultured cell lines. The AAV/hygromycin phosphotransferase recombinants have been used to derive hygromycin B–resistant cell lines (Samulski *et al.,* 1989). A recombinant AAV encoding HIV antisense expression units has been shown to inhibit HIV replication in human $CD4^+$ lymphoid cells (Chatterjee *et al.,* 1992). A γ-globin gene has been delivered to erythroid cells (Walsh *et al.,* 1992) and a β-globin gene to cultured epidermoid carcinoma (KB) and erythroid leukemia lines (Dixit *et al.,* 1991; Einerhand *et al.,* 1995). AAV vectors containing placental alkaline phosphatase or *β-geo* (a *lac*Z–neomycin phospho-

transferase gene fusion) have been used to transduce several human cell lines and primary human foreskin fibroblasts (Alexander *et al.*, 1994).

Two examples of potential targets for gene therapy with AAV vectors which are currently of considerable interest are respiratory epithelium and the hematopoietic system. The packaging capacity of AAV is just sufficient to accommodate the 4.7-kb cDNA encoding the cystic fibrosis transmembrane protein (CFTR) which is defective in cystic fibrosis patients. Flotte *et al.* (1993) administered an AAV–CFTR recombinant to rabbit lung *in vivo* and observed expression of CFTR RNA and protein. Expression decreased with time but was still detectable at 6 months. However, Halbert *et al.* (1995) showed that *in vitro* transduction of primary airway (nasal polyp) epithelium cells by AAV occurred only inefficiently, calling into question the usefulness of AAV vectors for cystic fibrosis therapy.

Diseases such as hemoglobinopathies and immune deficiencies make hematopoietic stem cells desirable targets for corrective gene therapy. Several recent studies have employed AAV vectors to transduce human $CD34^+$ cell populations (thought to contain primitive hematopoietic stem cells) derived from bone marrow or umbilical cord blood (Zhou *et al.*, 1994; Goodman *et al.*, 1994; Miller *et al.*, 1994). We have also observed low levels of AAV-delivered luciferase in $CD34^+$-enriched cells 2–6 days posttransduction, although we have not yet determined that this expression is occurring in the $CD34^+$ population (unpublished results). It should be noted that, because recombinant AAV genomes can persist inside the transduced cell without integration or expression (Russell *et al.*, 1994; Halbert *et al.*, 1995), reverse transcriptase PCR (rather than DNA-based PCR) is required for detection of transduction events that lead to expression of the transgene. Miller *et al.* (1994) used this method to demonstrate that, following transduction of $CD34^+$ cells with AAV-γ-globin vectors, approximately 30% of BFU-e colonies expressed γ-globin RNA. A modest increase in fetal hemoglobin content, presumably derived from the transgene, was also observed. Thus, it appears that at least some hematopoietic progenitors can be transduced with AAV vectors.

There is clearly potential for the use of AAV as a gene delivery vehicle in therapeutic situations. It remains to be seen, however, whether clinically useful results can be obtained from the relatively low-titer stocks currently available.

B. Autonomous Parvovirus Vectors

Autonomous parvovirus vectors developed thus far are based on either LuIII or MVM, which are closely related rodent parvoviruses.

For several reasons, they are attractive as delivery vehicles for some gene therapy applications. First, since there is no evidence for integration into cellular DNA, they may be useful in situations where only transient expression of a gene is desired (such as toxin genes or therapeutic ribozymes). Second, the rodent parvoviruses efficiently infect a variety of human cell lines in culture. Third, there is no evidence for pathogenicity of rodent parvoviruses in humans. Fourth, they exhibit oncolytic characteristics which may make them useful in cancer therapy. Finally, like AAV, the virions are stable, facilitating handling and concentration of recombinant virus. The basic strategy for producing recombinant autonomous parvoviruses is shown in Fig. 3. It proceeds essentially as that for producing recombinant AAV, except that there is no requirement for adenovirus.

1. Parvovirus LuIII Vectors

LuIII is closely related to the rodent parvoviruses, MVM and H-1 (Diffoot *et al.*, 1993). Although it was originally isolated from a human lung cell line, its natural host is unknown; LuIII can, however, cause experimental disease in newborn and fetal hamsters (Soike *et al.*, 1976). Like MVM and H-1, LuIII infects a number of human cell lines including NB, 293, and HeLa cells. We have found that, although the hamster can be experimentally infected with LuIII, hamster kidney cells (BHK21) are not lysed by the virus but instead become persisently infected for at least eight passages (authors' unpublished observations).

With a view toward eventual application in cancer therapy, we have used the LuIII infectious clone pGLu883 (Diffoot *et al.*, 1993) to develop a system for encapsidation of foreign genes into LuIII virions. The prototype transducing genomes contained either the *lac*Z or the luciferase gene inserted downstream from the P4 promoter in such a way as to retain the LuIII polyadenylation signal (Fig. 4; Maxwell *et al.*, 1993b). Two prototype helper plasmids were used, one based entirely on the LuIII genome (pGLuΔBE; Hanson and Rhode, 1991) and the other, pSVLu, containing the SV40 early promoter in place of the P4 promoter (Fig. 4). β-Galactosidase and luciferase reporter genes maintained on transducing genomes have been packaged into LuIII capsids and transduced into recipient cells of several types (see following discussion). Transduction by the recombinant luciferase virus was inhibited by antiserum against LuIII, demonstrating that transfer of the luciferase gene was particle-mediated (Maxwell *et al.*, 1993b).

2. Minute Virus of Mice Vectors

MVM was discovered as a contaminant in stocks of murine adenovirus (Crawford, 1966). There are two well-characterized strains of

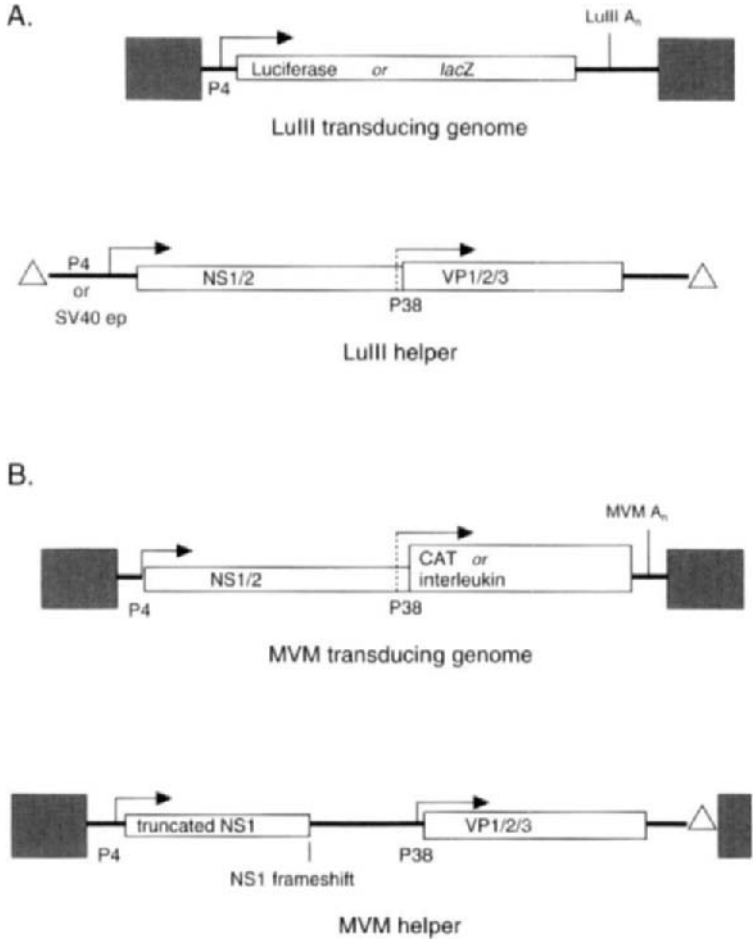


FIG. 4. (A) The prototype viral genomes used to generate LuIII transducing virus (Maxwell *et al.*, 1993a). The transducing genome consisted of either the luciferase or the *lacZ* gene inserted between the LuIII termini downstream from the P4 promoter of LuIII. The helper genome consisted of the entire coding sequence of LuIII controlled by either LuIII P4 or the SV40 early promoter; the terminal palindromes were removed, eliminating ability to replicate and be packaged. P4 and P38 indicate promoters at map unit 4 and 38, respectively. The delta symbol represents a deletion in the terminal palindrome; A_n, polyadenylation signal. (B) The viral genomes used to generate MVM transducing virus (Russell *et al.*, 1992; Dupont *et al.*, 1994). A_n represents the polyadenylation signal. The transducing genome consisted of either the chloramphenicol transacetylase (CAT) or interleukin (2 or 4) gene inserted downstream from the MVM P38 promoter. These constructs retain both termini and the NS1 coding sequences. The helper genome retained the NS1 coding sequences but did not produce functional NS1 due to a frameshift mutation in the NS1 reading frame; it was also unable to replicate due to a deletion in the right (by convention, 5′ end of the minus strand) palindrome.

MVM, MVMp, and MVMi. MVMp exhibits tropism for murine fibroblastoid lineages, while MVMi exhibits a tropism for murine T-lymphoid cells. Russell *et al.* (1992) inserted cDNAs for interleukin 2 or 4 downstream from the P38 promoter in place of the structural protein coding region of an MVMp infectious clone, pMM984, in a manner that left the nonstructural genes undisturbed (Fig. 4). These transducing genomes were packaged by cotransfection with either the infectious clone or an NS1 defective helper construct. Since the NS1 gene was present on this transducing genome, it was expected to replicate in the absence of a helper genome, in theory generating more template for transcription of the interleukin transgene. Expression of both NS1 (due to direct toxicity and immune stimulation) and an interleukin gene was predicted to kill tumor cells more effectively than either one alone. It was observed that IL-2 was efficiently expressed in SV40-transformed fibroblasts but not in their nontransformed analogues. This apparently reflects the tropism of parvoviruses for transformed cell types and shows that, at least in culture, expression of foreign genes from the viral promoters requires factors present in transformed cell types. In subsequent work, the bacterial chloramphenicol acetyl transferase gene (CAT) was inserted downstream from P38 in a manner analogous to the IL-2 and IL-4 genes already mentioned (Dupont *et al.*, 1994). The transducing genome was replicated and expressed CAT to significantly higher levels in transformed cell lines than in their nontransformed analogues. Thus, it appears that these MVM transducing genomes may exhibit at least some degree of specificity for tumor cells.

3. Titers of Transducing Virus

Two methods have been used to estimate titers of LuIII transducing virus. One is direct detection of β-galactosidase in cells transduced with a *lac*Z-bearing recombinant (Maxwell *et al.*, 1993b), and the other an infected cell (*in situ*) hybridization assay that relies on coinfection with wild-type LuIII to amplify the transducing genome (Maxwell and Maxwell, 1994). Using the hybridization assay, estimated titers of transducing units[1] have been in the range 10^4–10^5 per milliliter for the

[1]A transducing unit is defined as a transduction event detected by either the infected cell hybridization (ICH) assay (luciferase virus) or X-gal histochemistry (*lac*Z virus). It represents the entry of a transducing virus into a cell and its subsequent replication (ICH assay) or expression (X-gal assay). Estimates of transducing units vary with cell type, so we have routinely used NB324K cells as the standard target line. As discussed in text, the number of transducing units is substantially less than the number of DNA-containing transducing virus particles.

luciferase virus and 10^3–10^4 per milliliter for the *lac*Z transducing virus. Using the β-galactosidase histochemical assay, we have obtained estimates of 10^3–10^5 transducing units per milliliter, depending on transducing virus stock. Reported titers of MVM–IL2 vectors were quite low [less than 10^2 transducing units per milliliter as indicated by an infected center hybridization assay (Russell *et al.*, 1992)], and those of MVM–CAT virus were approximately 10^4 transducing units per milliliter (Dupont *et al.*, 1994). The titers of AAV transducing viruses have been estimated by several assays on a variety of cell lines. Using an infected cell hybridization assay or DNA replication assays, estimates of 10^4–10^6 transducing units per milliliter have been obtained (Hermonat and Muzyczka, 1984; Tratschin *et al.*, 1985; McLaughlin *et al.*, 1988). Using the number of drug-resistant colonies as an indicator of transduction, estimates of 10^3–10^5 transducing units per milliliter were obtained with AAV–neomycin or AAV–hygromycin transducing viruses (Samulski *et al.*, 1989; McLaughlin *et al.*, 1988).

Thus, judging from titers of LuIII, MVM, and AAV transducing virus, it appears that a level of 10^6 transducing units per milliliter of crude (nonconcentrated) virus stock is the practical limit of production capacity for parvovirus vectors in DNA cotransfection schemes used to date. In our experience, the efficiency of transfecting NB324K producer cells (via electroporation) has ranged from 15–50%, so it may be possible to increase the titers somewhat by improving the transfection efficiency; even so, crude titers in the range of 10^6–10^7 transducing units per milliliter are quite low compared to the 10^9 (or more) plaque-forming units per milliliter generated in culture medium during a productive infection of NB324K cells with LuIII. Increasing the yield of transducing virus is a primary goal of future research.

A factor contributing to the generally low titers of recombinant parvovirus stocks may be a decrease in infectivity, reflected in an increase in the observed particle ratio (defined as the ratio of total DNA-containing particles to transducing units) relative to the parental wild-type virus. Particle ratios in the range 300–3000:1 have been reported for wild-type MVM and H1 (Tattersall, 1972; Paradiso, 1981) and a value of about 800:1 for LuIII (Majaniemi and Siegl, 1984). This compares with our observed particle ratios of 5000–10,000:1 for LuIII–luciferase transducing virus stocks [as measured by comparison of DNase-resistant DNA content with the yields of transducing units determined in the biological assays described above (Maxwell *et al.*, 1993b)]. For several other LuIII recombinants the particle ratios appear to be even higher (unpublished observations). Particle ratios appear to be less with wild-type AAV, ranging from 20:1 to 100:1 (cited by

Tratschin *et al.*, 1985, and determined by McLaughlin *et al.*, 1988), which may allow for the lower PRs reported for AAV recombinants. Estimates of the infectivity of AAV recombinants transducing the *neo* gene, made by comparing the packaged DNA content with infectious units determined as G418-resistant colonies on various cell types, have been quite variable. Although early reports implied a high particle ratio (Tratschin *et al.*, 1985), this was probably caused by inhibitory effects of *rep* gene expression, as shown by McLaughlin *et al.* (1988). However, even for *rep*-minus vectors, the particle ratio determined by this method has been variable, ranging from 10–100:1 (McLaughlin *et al.*, 1988; Flotte *et al.*, 1995) to 500–1000:1 (Samulski *et al.*, 1989; Halbert *et al.*, 1995). This variation may be partly due to the use of different cell lines or culture conditions. It should be noted that particle ratio values are likely to vary depending on the target cell type, with some cells, such as primary cultures or nondividing cells (Russell *et al.*, 1994; Halbert *et al.*, 1995), probably giving higher particle ratios.

With regard to the low infectivity of some LuIII recombinants, one can speculate that this might result from a failure of the packaged recombinant DNA to interact properly with the capsid. This notion is supported by observation of at least partial ordering of the nucleic acid in the CPV full capsid (Tsao *et al.*, 1991). Better understanding of these interactions may enable the design of more efficient transducing recombinants. It is at present unclear why the particle ratio is apparently higher for the autonomous parvoviruses than for their AAV counterparts. Perhaps some aspect of assembly and maturation of AAV is more tolerant of foreign sequences.

4. *Generation of Infectious Virus by Recombination during Production of Transducing Virus*

Most packaging systems designed to produce transducing virus have generated infectious virus via recombination between the complementing helper and transducing genomes (Mulligan, 1993; Miller, 1992; Oppenheim and Peleg, 1989; Breakfield and DeLuca, 1991; Smith *et al.*, 1994; Bredenbeek *et al.*, 1993). The same has been observed in the parvovirus systems. Infectious AAV was formed with the early AAV packaging system (Hermonat and Muzyczkza, 1984), but, more recently, stocks of recombinant virus free of wild-type AAV have been generated by removing all regions of homology between the helper and transducing genomes (Samulski *et al.*, 1989). When producing MVM–P38IL2 vectors with a defective helper, infectious MVM was generated at low levels and was amplified upon serial passage (Russell *et al.*, 1992). It is not surprising that infectious virus was produced from

those constructs, since the transducing and helper genomes shared long regions of homology (see Fig. 4). In recent studies with a similar MVM P38–CAT transducing genome and defective helper, wild-type contamination ranged from 63–85% of the total virus particles in a given stock (Dupont *et al.*, 1994).

The LuIII helper and transducing genomes also recombined to form infectious LuIII virus (Maxwell *et al.*, 1993b). Removal of homology from only the 5′ ends reduced the frequency of regenerating plaque-forming virus to two in six stocks, and removal of homology from both the 5′ and 3′ ends may eliminate the generation of plaque-forming LuIII, although at some expense to the yield of transducing virus (authors' unpublished observations). These experiments suggest that the primary mechanism of regenerating plaque-forming LuIII is via homologous recombination, although this has not been confirmed by sequencing of the resultant viruses.

As an aside, the presence of infectious virus might be advantageous in some therapeutic situations, such as *in vivo* treatment of metastatic disease. In such a situation, it would be unlikely that practically administered levels of recombinant virus (perhaps 10^8–10^9 transducing units) would be therapeutically effective. Coinfecting the defective virus (carrying the antitumor gene) with a wild-type virus might allow amplification of the defective virus *in vivo* as well as exploit the natural oncolytic tendencies of the wild-type virus. This sort of approach might be clinically acceptable, since no pathology has been observed in humans inoculated with rodent parvoviruses.

5. *Packaging Cell Lines*

Stable cell lines able to package a transducing genome and release recombinant virus are attractive because of the potential to generate higher titers and the relative ease of producing transducing virus. The generation of such packaging cell lines involves introduction and stable expression of genes for both virion and nonstructural proteins, as well as the transducing genome. HeLa cell lines containing an integrated AAV helper virus genome have been reported, although yields were disappointingly low (10^3–10^4 particles/ml) (Vincent *et al.*, 1990). More recently, Yang *et al.* (1994) have generated derivatives of 293 cells containing the *rep* gene under control of the murine metallothionein-I promoter. These cells produced 5×10^7 transducing units per milliliter from a transfected AAV genome containing a *rep* frameshift. This cell line retained Rep inducibility for at least one year of continuous passage and was resistent to cytotoxic effects of the Rep proteins. Holscher *et al.* (1994) reported HeLa cell lines containing *rep* under

control of the mouse mammary tumor virus long terminal repeat. Rep produced in these cell lines mediated RF formation but was for some reason defective in single-strand progeny formation. Clark *et al.* (1995) recently obtained HeLa cell lines containing both helper and transducing genomes. These lines were capable of producing yields of AAV transducing virus (upon infection with adenovirus) which were comparable to those obtained from transient transfections of helper and transducing genome. These types of cell lines may be very useful for scaling up transducing virus production.

Attempts to generate packaging cell lines based on LuIII genomes have met with difficulties because of toxic properties (Caillet-Fauquet *et al.,* 1990; Momeda *et al.,* 1994; Mousset *et al.,* 1994; Vanacker and Rommelaere, 1995) of the essential viral protein, NS1 (Corsini 1994). We have inserted transcription units consisting of an inducible promoter [*Drosophila* heat-shock protein 70, human metallothionein, engineered human metallothionein (McNeall *et al.,* 1989), or tetracycline response elements (Gossen and Bujard, 1992)] upstream from the NS1–2 gene and placed these units into Epstein-Barr virus–based plasmids containing the hygromycin marker. Although these promoters were inducible in transient transfection assays, they did not tightly control NS1 expression in 293, HeLa, or NB324K cells. Numerous hygromycin-resistant cell lines were generated but none produced significant levels of NS1. *Drosophila* heat-shock 70–NS1 Epstein-Barr plasmids rescued from one cell line contained a deletion in the nonstructural region of the LuIII genome, suggesting that loss of NS1 function led to survival of colonies. Thus, we have concluded that basal NS1 expression during transient transfection interfered with derivation of stable cell lines able to support LuIII nonstructural genes. Caillet-Fauquet *et al.* (1990) have reported isolation of stable NB-E cell packaging lines able to produce detectable levels of MVM NS1 from the mouse mammary tumor virus long terminal repeat upon induction; however, there has been no report of using these cells for packaging of MVM recombinants.

6. Regulation of Heterologous Promoters

Heterologous promoters have been inserted into AAV vectors and used to transcribe foreign genes. Walsh *et al.* (1992) and Einerhand *et al.* (1995) constructed such vectors containing γ- or β-globin genes under control of their own regulatory elements and showed expression in erythroleukemia cells. The SV40 early promoter has been used to express neomycin or hygromycin phosphotransferase genes in AAV vectors (Hermonat and Muzyczka, 1984; McLaughlin *et al.,* 1988; Samulski *et al.,* 1989). Both the murine sarcoma virus long terminal

repeat (LTR) and the cytomegalovirus early promoter have been used by Lebkowski *et al.* (1988) to express CAT and neomycin phosphotransferase genes from AAV vectors in a variety of cell types.

Expression from heterologous promoters in context of LuIII recombinant genomes has been examined (Maxwell *et al.*, 1996). Since LuIII vectors can transduce various cell types (although with differing efficiency; see the following discussion), it will be important to impose cell-specific transcriptional control in situations where specificity of targeted expression is required. To investigate the possible influence of LuIII on regulation of such promoters, we generated several recombinants with regulated promoters substituted for P4 (Maxwell *et al.*, 1996). In the first construct, P4 was replaced by a chimeric promoter containing an enhancer sequence (from an alpha-1 protease inhibitor gene) known to direct preferential transcription in transfected hepatic cells. When this construct (pGLu-A1Pi-LUC) was transfected into LuIII-permissive cells together with a helper plasmid supplying the viral nonstructural and capsid proteins, excision and amplification of the recombinant genome was observed, confirming that these processes did not require the P4 region (nucleotides 148–265). Transducing virions were produced by these cotransfected cells, as shown by expression of the luciferase reporter in recipient HepG2 and HeLa cells. Correcting for relative infection efficiency of the two cell types, the HepG2 hepatoma cells expressed the reporter 10 to 20 times more efficiently than HeLa cells, indicating that the chimeric promoter retained specificity for hepatic cells in the context of the LuIII transducing genome. In other LuIII recombinants, the luciferase reporter was linked with chimeric promoters containing response elements for either GAL4 (Ma and Ptashne, 1987) or a tetracycline-regulated transactivator (tTA) (Gossen and Bujard, 1992). Luciferase expression was strongly activated when these viruses were used to infect cells containing a cognate trans-activator (GAL4 or tTA) introduced by transfection (Maxwell *et al.*, 1996). These results confirmed that appropriate transcriptional regulation could be achieved for genes transduced by an autonomous parvovirus vector. Such vectors therefore show promise for the delivery of therapeutic genes in situations requiring cell-specific expression, e.g., in targeting cancer cells with toxin genes. It is worth noting that preliminary results using an infected cell hybridization assay (Maxwell and Maxwell, 1994) suggested that the levels of infectious virus generated from the transcriptionally regulated recombinants were lower than those from the prototype LuIII–luciferase genome (pGLuP4luc1). These results again raise the possibility that the production efficiency and/or infectivity of LuIII recombinant virions may be influenced by the particular DNA sequences included.

7. *Transduction Range of MVM and LuIII Transducing Viruses*

The transduction range of MVM vectors has been explored by Dupont *et al.* (1994), who transduced fibroblast, epithelial, and white blood cell lineages with an MVM vector containing CAT driven by P38. Transformed and nontransformed cell types were compared to examine oncotropism of the MVM vectors. In general, transformed cells replicated the MVM–CAT transducing genome and expressed CAT to a much higher degree than the nontransformed analogues. In particular, transformed cells of fibroblast and macrophage lineage expressed transduced CAT at levels 500–1000 times and 1000–7000 times higher (respectively) than did their nontransformed analogues. In contrast, Burkitt's lymphoma lines expressed only low levels of CAT upon transduction with the MVM–CAT virus, perhaps due to inefficient entry into B cells. These results suggest that the MVM vectors may be useful for preferentially targeting tumor cells with P38 controlled antitumor genes.

Data concerning the range of cells transduced by LuIII vectors are shown in Table I. In contrast to the MVM-based vectors just described, recombinant viruses used in these studies expressed reporter genes

TABLE I

TYPICAL TRANSDUCTIONS OF A VARIETY OF CELL TYPES WITH THE LUIII–LUCIFERASE TRANSDUCING VIRUS (PGLUP4LUC1)[a]

Cell line[b]	Transduced luciferase activity (light units/ml of virus stock)
HeLa (human cervical carcinoma)	$0.7–2 \times 10^6$
293 (human fetal kidney, adenovirus transformed)	1×10^6
NB324K (human newborn kidney, SV40 transformed)	1×10^6
HepG2 (human hepatoma)	$10^4 - 10^5$
OVCARIII (human ovarian carcinoma)	6×10^4
T47D (human breast carcinoma)	4×10^4
WI38 (human fetal diploid fibroblast)	2×10^4
NeAl (human B-cell, EBV transformed)	2×10^3
GM4025 (human B-cell, EBV transformed)	$1–2 \times 10^3$
Raji (Burkitt's lymphoma)	$1–2 \times 10^3$

[a] Maxwell *et al.* (1993a).

[b] Each cell line was overlaid with transducing virus, incubated 24 hours, and then assayed for luciferase activity.

from the native P4 promoter of LuIII (Fig. 4). Such recombinants transduced the luciferase gene into a variety of transformed cell types, as well as into one nontransformed human diploid fibroblast line, WI38. Not surprisingly, different cell types were transduced with different efficiencies. For example, human cervical carcinoma (HeLa), human fetal kidney transformed with adenovirus (293), and human newborn kidney transformed with SV40 (NB324K) cells were all efficiently transduced, whereas Burkitt's lymphoma and other lymphocyte lines were poorly transduced. This suggests that in their current form, neither MVM nor LuIII vectors will be useful for delivering genes to cells originating from B lymphocytes. Experiments using a LuIII–*lac*Z virus have generally yielded results similar to those obtained with LuIII–luciferase virus.

8. Pseudotyping of the LuIII–Luciferase Genome

Generation of recombinant virions with coat proteins from alternative parvoviruses may offer a means of varying the transducing host range with respect to tissue and/or species. To test this possibility, the LuIII–luciferase genome was packaged with replication-defective helper constructs from H1, MVMp (prototype), or MVMi (immunosuppressive) (Maxwell *et al.,* 1993b). These experiments (Fig. 5) relied on the fact that heterologous NS proteins are able to mediate replication from the LuIII origins, avoiding the need to substitute heterologous capsid genes (MVM or H-1) into a LuIII helper. Efficient production of LuIII–lucifierase transducing virions from pGLuP4luc1 by cotransfection with each of these alternative helpers was observed. Transduction by the H1-pseudotyped virus was neutralized by anti-H1 but not anti-LuIII serum, and vice versa for the virus generated with the homologous LuIII helper.

Packaging a recombinant LuIII genome with capsids of MVMp and MVMi enabled us to determine whether such recombinants would retain the respective MVM host range. MVMp is tropic for fibroblasts, whereas MVMi replicates in murine T-lymphoid cells but not in the A9 fibroblast cell line (a derivative of mouse L cells). This tropism is conferred by two amino acid differences in the VP2 capsid coding region (Ball-Goodrich and Tattersall, 1992) which are thought to be involved in interaction with a host-cell-specific determinant in establishing productive infection. In previous host-range studies (Antonietti *et al.,* 1988; Gardiner and Tattersall, 1988a), the infecting viruses retained the homologous combination of capsid and its encoding DNA, so a direct, sequence-dependent involvement of the DNA in determining

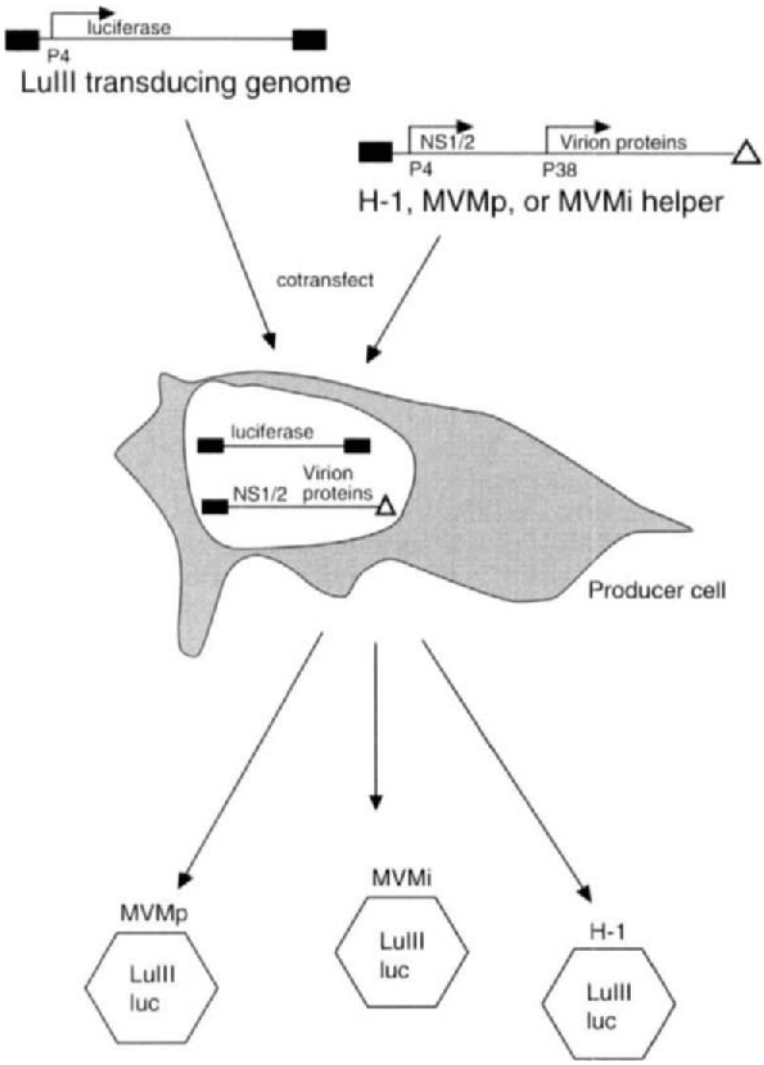


FIG. 5. Pseudotyping of the LuIII–luciferase transducing genome with each of three rodent parvoviruses, H-1, fibrotropic strain of MVM (MVMp), or lymphotropic strain of MVM (MVMi). The strategy is to cotransfect the LuIII transducing genome with a heterologous helper, and then collect the transducing medium for analysis. P4 is the rodent parvoviral promoter at map unit 4 and P38 the promoter at map unit 38. The delta symbol represents a deletion in the terminal palindrome; luc, luciferase.

tropism could not be excluded. Infection of A9 fibroblasts with GLuP4luc1/MVMp resulted in efficient expression of the transduced luciferase reporter, whereas infection with GLuP4luc1/MVMi gave virtually no expression in these cells. Both pseudotyped viruses expressed luciferase efficiently on infection of human NB324K cells, which are permissive for both MVMp and MVMi. This result supported the notion of a specific involvement of capsid protein in establishing efficient transcription in infected cells (Gardiner and Tattersall, 1988b) and excluded direct involvement of DNA encoding the VP, at least in the initial stages of productive infection.

It was of interest to determine whether a recombinant LuIII genome could be pseudotyped by capsid proteins from more distantly related parvoviruses such as the feline panleukopenia virus (FPV) or the canine parvovirus (CPV). These viruses are very closely related to one another but show considerably less homology to LuIII than does MVM. In these experiments, recombinant viruses were produced by triple cotransfection of NB324K cells with transducing genome (pGLuP4luc1), an NS expression plasmid, and a plasmid expressing capsid proteins of LuIII, FPV, or CPV (using the respective P38 promoters). In each case, transducing virus was generated that expressed luciferase in the Crandell feline kidney cell line, which is permissive for both FPV and CPV. In contrast, the FPV and CPV pseudotypes were devoid of transducing activity in human NB324K cells (Spitzer *et al.*, 1996).

9. Persistence of Expression from LuIII–Luciferase Transducing Virus

Available evidence indicates that MVM and LuIII do not integrate into the host genome (Richards and Armentrout, 1979; Majaniemi and Siegl, 1984; Ron and Tal, 1985), so expression from these vectors might be expected to be transient. However, experiments with both CMV–luciferase and P4–luciferase transducing viruses (authors' unpublished observations) have shown that total luciferase expression in transduced NB324K cells increased in an exponentially growing bulk culture of transduced NB324K cells during a 3-week time course. This may stem from some sort of integration event or possibly episomal maintenance of the transducing genome. It is also conceivable that recombinant genomes are sequestered inside the cell [as seems to occur with AAV (Russell *et al.*, 1994)], and then "recruited" for expression at a later time. This phenomenon requires further characterization to assess whether the LuIII or MVM vectors can be considered strictly transient gene delivery vehicles.

10. Use of LuIII–Luciferase Transducing Virus to Study Genome Encapsidation Patterns

LuIII encapsidates either strand of its genome with similar efficiency (Bates *et al.,* 1984; Muller and Siegl, 1983a,b). In contrast, MVM and H-1, which are closely related to LuIII, encapsidate predominantly the minus strand when grown in the same cell type (Bates *et al.,* 1984). It was postulated that a unique AT-rich region near the right end of the LuIII genome is responsible for symmetric encapsidation of plus and minus strands by LuIII. To address this hypothesis, recombinant LuIII–luciferase genomes containing or lacking the AT-rich sequence (AT) were packaged into LuIII capsids (Corsini *et al.,* 1995). Hybridization of strand-specific probes to DNA from these virions revealed that either strand of the genome was packaged regardless of the presence of AT. In addition, encapsidation of both strands of the AT^+ LuIII–luciferase genome into MVM and H-1 pseudotyped virions was observed, suggesting that MVM and H-1 proteins are not responsible for the minus-strand packaging bias of these two viruses. Alignment of the published LuIII and MVMp sequences shows that AT exists as an insertion into an element that, in MVM, binds cellular proteins, and it was speculated that in LuIII, AT disrupts binding of these cellular proteins, allowing encapsidation of either strand (Corsini *et al.,* 1995).

C. Densovirus Vectors

Manipulation of insect genomes (other than *Drosophila*) has proven difficult because the gene transfer tools developed for *Drosophila* have not been readily transferable to other genera. Because of this, there is interest in developing expression and gene delivery systems specific for medically (mosquitos and ticks), agriculturally (silkworm and medfly), and aquaculturally (shrimp) important arthropod species. Densoviral promoter elements have been used for expression of foreign genes in insect cells, and strategies for packaging and delivering heterologous genes have been reported for at least two insect densoviruses, JcDNV and AeDNV (Giraud *et al.,* 1992; Afanasiev *et al.,* 1994). Densoviruses are still in the early stages of characterization, and it is possible that some or many of them will be of the integrating sort, a characteristic which would make them extremely useful as tools for manipulation of the insect host genome.

Another potential application of wild-type and recombinant densoviruses is in insect control schemes. In Colombia, a DNV of the lepidop-

teran *Sibine fusca* was effectively used to control infestations of palm trees by *S. fusca*, and in Ivory Coast, *Casophilia extranea* DNV was used to control the lepidopteran pest *C. extranea* in palm groves (Belloncik, 1990). Tal and Attathom (1993) have explored the insecticidal potential of GmDNV in the host wax moth, concluding that third instar larvae are the most susceptible. Recombinant densoviruses carrying toxin genes might be useful in certain insect control strategies, especially in cases where expedient killing of target insects is desired. Such a strategy has been successfully utilized with baculoviruses (Lucknow and Summers, 1988; Cory *et al.*, 1994). Although attractive in concept, there are many technical matters, such as specificity of toxins and production of large quantities of transducing virus for field application, that require attention before these sorts of strategies become feasible.

To date, packaging systems able to generate transducing virus have been developed for AeDNV and JcDNV. The strategy for generating DNV transducing virus is essentially the same as that for the mammalian parvoviruses (Fig. 3).

1. Junonia Coenia DNV Vectors

The genomes of JcDNV and GmDNV have an unusual organization, with structural mRNAs synthesized from one strand and nonstructural mRNAs from the other (Dumas *et al.*, 1992) (Fig. 2C). As mentioned, JcDNV has a very wide host range, making it a good candidate for development into a broad-spectrum gene vector for lepidopteran insects. The complete JcDNV genome has been cloned into pBR322 and generates wild-type virus when transfected into susceptible *Spodoptera littoralis* larvae (Jourdan *et al.*, 1990). This infectious clone has been used to generate *lac*Z–virion protein (VP) gene fusion constructs which are efficiently expressed (Giraud *et al.*, 1992). A recombinant genome containing a *lac*Z fusion in the VP gene and deletions in the nonstructural regions produces β-galactosidase efficiently and can be packaged into virions (by cotransfection with either the infectious clone or a helper containing deletions in the terminal palindromes) able to transfer β-galactosidase activity to fresh cells (Giraud *et al.*, 1992). These results, though promising, are preliminary and further studies are required to determine such parameters as transducing virus titers, transduction efficiency, transduction range, and recombination frequencies.

2. Aedes Aegypti DNV Vectors

The DNV of *Aedes aegypti* mosquitoes was isolated from fourth instar larvae of a laboratory culture of *Aedes aegypti* (Lebedeva *et al.*, 1973). It is infectious for five species of *Aedes* as well as several *Culex*

and *Culiseta* species. Many of its physical characteristics have been reported (Buchatsky, 1989). The genome was cloned and sequenced by Afanasiev *et al.* (1991) and contains two leftward ORFs and one rightward ORF. Fusion of the *lac*Z gene into both leftward ORFs (nonstructural protein) and the right ORF (virion protein) showed that proteins are generated from all three ORFs. Recombinant genomes containing the *lac*Z gene fused into the right ORF have been cotransfected along with an infectious clone (Afanasiev *et al.,* 1994), and medium collected from such a cotransfection was able to transfer the *lac*Z gene to fresh cells. Inhibition of this transfer by anti-AeDNV serum has demonstrated that this transfer is virion-mediated. The numbers of cells transduced by this *lac*Z virus were generally not high, but in some cases up to 2×10^4 target cells were transduced (Afanasiev *et al.,* 1994). As with other parvovirus packaging systems, efficiency of encapsidation decreases as the genome size is increased above 100% of the wild-type genome. Up to 108% of the genome has been packaged, but the efficiency is significantly less than that for a 102% genome. As for JcDNV, further studies are required to determine such parameters as particle ratios and recombination frequencies.

3. Potential Vectors Based on Other Densoviruses

The first steps toward the development of the GmDNV (wax moth) into a gene delivery vehicle have been taken. Tal and Attathom (1993) have placed the chloramphenicol transacetylase (CAT) gene under control of nonstructural and structural promoters, then injected calcium precipitates of these plasmids into *G. mellonella* larvae. Elevated levels of CAT were observed in injected larvae. By analogy with the other densoviruses, it seems very likely that these constructs, which retain the termini, will be packagable by cotransfection with the infectious clone or defective helpers. Genomic clones of AaDNV (mosquito) and BmDNV (silkworm) exist (Boublik *et al.,* 1994; Bando *et al.,* 1990), so it should be possible to derive transducing virus packaging systems based on these viruses as well.

IV. Summary and Prospects

A. Autonomous Parvovirus-Mediated Gene Delivery

Table II summarizes the state of each parvovirus vector system, its potential applications, and its current limitations. Autonomous parvovirus gene vectors are still in the early stages of development. Recombinant MVM and LuIII vectors bearing reporter genes have been used

TABLE II

SUMMARY OF THE PARVOVIRUS GENE VECTOR SYSTEMS, THEIR POTENTIAL APPLICATIONS, AND THE CURRENT OBSTACLES IMPEDING THEIR DEVELOPMENT

Vector	Host	General applications	Current obstacles
Adeno-associated virus	Mammalian	1. Integrating vector, site specific in some cases: integrated genes seem to be stable so should be useful for permanent delivery of genes. 2. Broad target cell range: useful at this point for *ex vivo* gene delivery.	1. Inefficient integration into nondividing cells (may be enhanced by DNA-damaging agents). 2. The requirement of Rep for targeted integration. 3. Relatively low numbers of transducing units (high and variable particle to infectivity ratios).
Autonomous parvoviruses	Mammalian	1. Probably nonintegrating: potentially useful for transient delivery of genes. 2. Broad target cell range (pseudotyping may allow directed delivery *in vivo*) potentially useful for transfer of antitumor genes. 3. NS1-containing vectors may be useful because of the inherent oncolytic characteristics of NS1.	1. Relatively low numbers of transducing units (high and variable particle to infectivity ratio). 2. Possible persistence of expression in transduced cells requires further characterization.
Densoviruses	Invertebrate	1. Numerous insect species infected by a variety of densoviruses: insecticidals and delivery of insecticidal genes. 2. Genes can be expressed from densoviral promoters and delivered in densoviral virions: genetic engineering of insects.	1. Fundamental aspects of the viruses, such as splicing strategies and infectious titers, require characterization. 2. Many fundamental aspects of the gene delivery system, such as particle to infectivity ratios, transduction range, formation of infectious virus, require characterization.

to transduce various transformed or untransformed cell lines. Reasonably high titers have been obtained, but there is much room for improvement. Wild-type virus resulting from recombination of defective DNA constructs has been a problem, but appears to be controllable by minimizing homology between helper and transducing genomes. Activity of several foreign transcriptional control elements in context of the recombinant LuIII transducing virion has been characterized and, in

general, genes cloned downstream from heterologous promoters are properly regulated. To extend the utility of the LuIII vectors, LuIII–luciferase transducing genomes have been pseudotyped with virions from heterologous parvoviruses. The results suggest that packaging LuIII transducing genomes into heterologous virions can effectively alter the specificity of transduction and may allow delivery of genes to a wider variety of cell types. When specific tissues are to be targeted, it should be possible to restrict expression of the transgene to certain cell or tissue types with the use of transcriptional control elements.

As mentioned in Section I, complementation systems developed to package recombinant LuIII (or MVM) transducing genomes should also be useful for other basic studies of parvovirus biology. For example, the complementation system provides the opportunity for identification of LuIII sequences that might be required for optimal formation of infectious virus particles. Basic questions about contributions of the individual coat proteins to host-range specificity might be addressed by packaging transducing genomes into chimeric virions containing VP1 from one parvovirus and VP2 from another (Maxwell *et al.,* 1995).

Retention of the NS region in the MVM vectors provides an intriguing means of exploiting the natural oncolytic tendencies of the parvoviruses. Current versions of this system express the cytotoxic NS1 protein from the P4 promoter and, although it has been shown that these MVM vectors do not replicate or express well in some primary cell types, detectable expression of the transgene, though significantly lower than in transformed analogues, does occur in nontransformed cell types (Dupont *et al.,* 1994). Thus, it remains to be seen whether the relative levels of both NS1 and the transgene will differ sufficiently for targeted therapeutic effects.

A number of other potential applications for gene delivery with autonomous parvovirus vectors exist. Since MVM–CAT vectors enter and are expressed in macrophages, it may be possible to deliver therapeutic genes to macrophages infected with *Mycobacterium tuberculosis* or human immunodeficiency virus. Such genes might encode abzymes that degrade specific components of the disease agent, or antisense RNA which interferes with gene expression. Still another as yet unexplored application of autonomous parvovirus gene delivery vehicles might be delivery of genes encoding growth factors which promote healing of damaged tissues. The list of potential applications may increase or decrease in length as more becomes understood about autonomous parvoviruses; nevertheless, the autonomous parvoviruses have shown promise as gene delivery vehicles and will likely find a place in the repertoire of gene therapy tools.

B. Densovirus-Mediated Gene Delivery

Development of densovirus gene vectors is in its infancy, and there remains much work before these systems are well-characterized. Initial results with several densoviruses indicate that they and their promoters will be useful for introducing and expressing foreign genes in insect cells and probably live insects. The β-galactosidase gene has been inserted into genomes of both JcDNV and AeDNV, and then packaged into particles able to deliver *lacZ* to target cells in culture, indicating that it will also be possible to utilize these vectors for gene delivery. One intriguing prospect is that of integration by DNVs. It is not known at this time whether any of the insect parvoviruses integrate into the host genome; there is some evidence for integration of JcDNV genome in insect cell lines (Dumas *et al.*, 1992; Bergoin *et al.*, 1995), but the results are preliminary and await confirmation. If the densoviruses turn out to integrate, they will be extremely useful in the derivation of transgenic insects, which, though now commonplace with *Drosophila*, has proven difficult with other insect species.

Despite their promise, the vector systems require further characterization. For example, basic splicing patterns of the viruses, if any, require study to facilitate manipulation of both helper and transducing genomes. Potential toxic effects of the nonstructural proteins require study in order to predict their effects on long-term expression (i.e., in genetically transformed insects). Since the DNVs are lytic in their insect hosts, packaging systems that utilize defective helpers which eliminate infectious virus will be necessary before they are useful as *in vivo* gene delivery vehicles. Titers of transducing virus require systematic determination, as do particle to infectivity ratios. With more comprehensive characterization, however, we foresee considerable potential for use of these viruses as gene delivery vehicles.

References

Afanasiev, B. N., Galyov, E. E., Buchatsky, L. P., and Kozlov, Y. V. (1991). Nucleotide sequence and genomic organization of *Aedes* densonucleosis virus. *Virology* **185**, 323–336.

Afanasiev, B. N., Kozlov, Y. V., Carlson, J. O., and Beaty, B. J. (1994). Densovirus of *Aedes aegypti* as an expression vector in mosquito cells. *Exp. Parisitol.* **79**, 322–339.

Agbandje, M., Kajigaya, S., McKenna, R., Young, N. S., and Rossman, M. G. (1994). The structure of human parvovirus B19 at 8 Å resolution. *Virology* **203**, 106–115.

Agbanje, M., Parrish, C. R., and Rossman, M. G. (1995). The structure of parvoviruses. *Semin. Virol.* **6**, 299–309.

Alexander, I. E., Russel, D. W., and Miller, A. D. (1994). DNA-damaging agents greatly

increase the transduction of non-dividing cells by adeno-associated virus vectors. *J. Virol.* **68**, 8282–8287.

Anderson, W. F. (1984). Prospects for human gene therapy. *Science* **226**, 401–409.

Anderson, W. F. (1992). Human gene therapy. *Science* **256**, 808–813.

Anton, I. A., and Lane, D. P. (1986). Non-structural protein 1 of parvoviruses: Homology to purine nucleotide using proteins and early proteins of papovaviruses. *Nucleic Acids Res.* **14**, 7813.

Antoni, B. A., Rabson, A. B., Miller, I. L., Trempe, J. P., Chejanovsky, N., and Carter, B. J. (1991). Adeno-associated virus rep protein inhibits human immunodeficiency virus production in human cells. *J. Virol.* **59**, 284–291.

Antonietti, J.-P., Sahli, B., Beard, P., and Hirt, B. (1988). Characterization of the cell type-specific determinant in the genome of minute virus of mice. *J. Virol.* **62**, 552–557.

Astell, C. R. (1990). Terminal hairpins of parvovirus genomes and their role in DNA replication. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 4. CRC Press, Boca Raton, Florida.

Astell, C. R., Chow, M. B., and Ward, D. C. (1985). Sequence analysis of the termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling hairpin model for autonomous parvovirus DNA replication. *J. Virol.* **54**, 171–177.

Astell, C. R., Mol, C. D., and Anderson, W. F. (1987). Structural and functional homology of parvovirus and papovavirus polypeptides. *J. Gen. Virol.* **68**, 885–893.

Atchison, R. W., Casto, B. C., and Hammon, W. McD. (1965). Adeno-associated defective virus particles. *Science* **149**, 754–756.

Ball-Goodrich, L. J., and Johnson, E. (1994). Molecular characterization of a newly recognized mouse parvovirus. *J. Virol.* **68**, 6476–6486.

Ball-Goodrich, L. J., and Tattersall, P. (1992). Two amino acid subsitutions within the capsid are coordinately required for aquisition of fibrotropism by the lymphotropic strain of minute virus of mice. *J. Virol.* **66**, 3415–3423.

Bando, H., Kusuda, J., Gojobori, T., Maruyama, T., and Kawase, S. (1987). Organization and nucleotide sequence of a densovirus genome imply a host-dependent evolution of parvoviruses. *J. Virol.* **61**, 553–560.

Bando, H., Choi, H., Ito, Y., and Kawase, S. (1990). Terminal structure of a densovirus implies a hairpin transfer replication which is similar to the model for AAV. *Virology* **179**, 57–63.

Bantel-Schaal, U. (1990). Adeno-associated parvoviruses inhibit growth of cells derived from malignant human tumors. *Int. J. Cancer* **45**, 190–194.

Bates, R. C., Snyder, C., Bannerjee, P. T., and Mitra, S. (1984). Autonomous parvovirus LuIII encapsidates equal amounts of plus and minus DNA strands. *J. Virol.* **49**, 319–324.

Becerra, S. P., Koczot, F., Fabisch, P., and Rose, J. A. (1988). Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. *J. Virol.* **62**, 2745–2754.

Becquart, P., Vanacker, J.-M., Duponchel, N., Begue, A., and Rommelaere, J. (1993). Expression of the non-structural proteins of parvovirus MVMp from recombinant retroviruses: Predominant role of the parvoviral NS-1 product in host cell disturbance. *Res. Virol.* **144**, 465–470.

Belloncik, S. (1990). Potential use of densonucleosis viruses as biological control agents of insect pests. *In* "Handbook of parvoviruses" (P. Tijssen, ed.), Vol. 2, Ch. 19. CRC Press, Boca Raton, Florida.

Ben-Asher, E., and Aloni, Y. (1984). Transcription of minute virus of mice, an autonomous parvovirus, may be regulated by attenuation. *J. Virol.* **52**, 266–276.

Bergeron, J., Menezes, and Tijssen, P. (1993). Genomic organization and mapping of transcription and translation products of the NADL-2 strain of porcine parvovirus. *Virology* **197**, 86–98.

Bergoin, M., Jousset, F.-X., Jourdan, M., Giraud, C., Rolling, F., Li, Y., Romane, C., Yuan, S., and Bossin, H. (1995). The genome of the *Junonia coenia* densovirus as vector for stable transformation of insect cells. *In* "Proceedings of the First International Workshop on Transgenesis of Invertebrates of Medical, Agricultural, and Aquacultural Importance."

Berkner, K. L. (1988). Development of adenovirus vectors for the expression of heterologous genes. *Biotechniques* **6**, 616–629.

Berns, K. (1990). Parvovirus replication. *Microbiol. Rev.* **54**, 316–329.

Besselsen, D., Riley, L., and Pintel, D. (1994). Characterization of newly recognized rodent parvoviruses by sequence analysis, serotyping, and in vitro host cell range. 13th American Society for Virology Meeting, Madison.

Bilimoria, S. L. (1991). The biology of nuclear polyhedrosis viruses. *In* "Viruses of Invertebrates" (Kurstak, ed.), p. 35. Marcel Dekker, New York.

Bonami, J. R., Trumper, B., Mari, J., Brehlin, M., and Lightner, D. V. (1990). Purification and characterization of the infectious hypodermal and hematopoietic virus of penaeid shrimps. *J. Gen. Virol.* **71**, 2657–2664.

Bonami, J.-R., Mari, J., Poulos, B. T., and Lightner, D. V. (1995). Characterization of hepatopancreatic parvo-like virus, a second unusual parvovirus pathogenic for penaeid shrimps. *J. Gen. Virol.* **76**, 813–817.

Botquin, V., Cid-Arregui, A., and Schlehofer, J. R. (1994). Adeno-associated virus type 2 interferes with early development of mouse embryos. *J. Gen. Virol.* **75**, 2655–2662.

Boublik, Y., Jousset, F.-X., and Bergoin, M. (1994). Complete nucleotide sequence and genomic organization of the *Aedes albopictus* parvovirus (AaPV) pathogenic for *Aedes aegypti* larvae. *Virology* **200**, 752–763.

Breakfield, X. O., and DeLuca, N. A. (1991). Herpes simplex virus for gene delivery to neurons. *New Biologist* **3**, 203–218.

Bredenbeek, P. J., and Rice, C. M. (1992). Animal RNA virus expression systems. *Semin. Virol.* **3**, 297–310.

Bredenbeek, P. J., Frolov, I., Rice, C. M., and Schlesinger, S. (1993). Sindbis virus expression vectors: Packaging of RNA replicons by using defective helper RNA's. *J. Virol.* **67**, 6439–6446.

Brown, K. E., Anderson, S. M., and Young, N. S. (1993). Erythrocyte P antigen: Cellular receptor for B19 parvovirus. *Science* **262**, 114–117.

Buchatsky, L. P. (1989). Denosonucleosis virus of blood sucking mosquitos. *Dis. Aquat. Org.* **6**, 145–150.

Caillet-Fauquet, P., Perros, M., Brandenburger, A., Spegelaere, P., and Rommelaere, J. (1990). Programmed killing of human cells by means of an inducible clone of parvoviral genes encoding non-structural proteins. *EMBO J.* **9**, 2989–2995.

Carlson, J., Olson, K., Higgs, S., and Beaty, B. (1995). Molecular genetic manipulation of mosquito vectors. *Annu. Rev. Entomol.* **40**, 359–388.

Carter, B. J. (1990a). Parvoviruses as vectors. *In* "Handbook of parvoviruses" (P. Tijssen, ed.), Vol. 2, Ch. 18, CRC Press, Boca Raton, Florida.

Carter, B. J. (1990b). Adeno-associated virus helper functions. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 13. CRC Press, Boca Raton, Florida.

Carter, B. J., Mendelson, E., and Trempe, J. P. (1990a). AAV DNA replication, integration, and genetics. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 11. CRC Press, Boca Raton, Florida.

Carter, B. J., Trempe, J. P., and Mendelson, E. (1990b). Adeno-associated virus gene expression and regulation. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 12, CRC Press, Boca Raton, Florida.

Cassinotti, P., Weitz, M., and Tratschin, J.-D. (1989). Organization of the adeno-associated virus (AAV) capsid gene: Mapping of a minor spliced RNA coding for the virus capsid protein 1. *Virology* **167**, 176–184.

Chapman, M. S., and Rossman, M. G. (1993). Structure, sequence and function correlations among parvoviruses. *Virology* **194**, 491–508.

Chatterjee, S., Johnson, P. R., and Wong, K. K. (1992). Dual-target inhibition of HIV-1 in vitro by means of an adeno-associated virus antisense vector. *Science* **258**, 1485–1488.

Cherbas, L., and Cherbas, P. (1993). The arthropod initiator: The cap site consensus plays an important role in transcription. *Ins. Biochem. Mol. Biol.* **23**, 81–90.

Cheung, A. K. M., Hoggan, M. D., Hauswirth, W. W., and Berns, K. I. (1980). Integration of adeno-associated virus genome into cellular DNA in latently infected Detroit 6 cells. *J. Virol.* **33**, 739–748.

Choi, W. S., Pal-Gosh, R., and Morrow, C. D. (1991). Expression of human immunodeficiency virus type 1 (HIV-1) *gag, pol,* and *env* proteins from chimeric HIV-1–poliovirus minireplicons. *J. Virol.* **65**, 2875–2883.

Christensen, J. T., Storgaard, T., Viuff, B., Aasted, B., and Alexandersen, S. (1993). Comparison of promoter activity in Aleutian mink disease parvovirus, minute virus of mice, and canine parvovirus: Possible role of weak promoters in the pathogenesis of Aleutian disease parvovirus infection. *J. Virol.* **67**, 1877–1886.

Christensen, J., Cotmore, S. F., and Tattersall, P. (1995). Minute virus of mice transcriptional activator protein NS1 binds directly to the transactivation region of the viral P38 promoter in a strictly ATP-dependent manner. *J. Virol.* **69**, 5422–5430.

Clark, K. R., Voulgaropoulou, F., Fraley, D. M., and Johnson, P. R. (1995). Cell lines for production of recombinant adeno-associated virus. *Hum. Gene Ther.* **6**, 1329–1341.

Clemens, D., and Carlson, J. (1989). Regulated expression of the feline panleukopenia virus P38 promoter on extra chromosomal FPV/EBV chimeric plasmids. *J. Virol.* **63**, 2737–2745.

Collins, P. L., Mink, M. A., and Stec, D. S. (1991). Rescue of synthetic analogues of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9663–9667.

Cornelis, J. J., Becquart, P., Duponchel, N., Salome, N., Avalosse, B. L., Namba, N., and Rommelaere, J. (1988a). Transformation of human fibroblasts by ionizing radiation, a chemical carcinogen, or simian virus 40 correlates with an increase in susceptibility to the autonomous parvoviruses H-1 virus and minute virus of mice. *J. Virol.* **62**, 1679–1686.

Cornelis, J. J., Spruyt, N., Spegelaere, P., Guetta, E., Darawshi, T., Cotmore, S. F., Tal, J., and Rommelaere, J. (1988b). Sensitization of transformed rat fibroblasts to killing by parvovirus minute virus of mice correlates with an increase in viral gene expression. *J. Virol.* **62**, 3438–3444.

Corsini, J. A. (1994). Development of Parvovirus LuIII Gene Vectors. Ph.D. dissertation, Colorado State University.

Corsini, J., Carlson, J. O., Maxwell, F., and Maxwell, I. H. (1995). Symmetric-strand packaging of recombinant parvovirus LuIII genomes that retain only the terminal regions. *J. Virol.* **69**, 2692–2696.

Cory, J. S., Hirst, M. L., Williams, T., Hails, R. S., Goulson, D., Green, B. M., Carty, T. M., Possee, R. D., Cayley, P. J., and Bishop, D. H. L. (1994). Field trial of a genetically improved baculovirus insecticide. *Nature* **370**, 138–140.

Cotmore, S. F. (1990). Gene expression in the autonomous parvoviruses. *In* "Handbook of the parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 9. CRC Press, Boca Raton, Florida.

Cotmore, S. F., and Tattersall, P. (1984). Characterization and molecular cloning of a human parvovirus genome. *Science* **226**, 1161–1165.

Cotmore, S. F., and Tattersall, P. (1986a). Organization of nonstructural genes of the autonomous parvovirus minute virus of mice. *J. Virol.* **58**, 724–732.

Cotmore, S. F., and Tattersall, P. (1986b). The NS-1 polypeptide of autonomous parvovirus MVM is a nuclear phosphoprotein. *Virus Res.* **4**, 243–250.

Cotmore, S. F., and Tattersall, P. (1988). The NS1 polypeptide of minute virus of mice is covalently attatched to the 5′ termini of duplex replicative-form DNA and progeny single strands. *J. Virol.* **62**, 851–860.

Cotmore, S. F., and Tattersall, P. (1989). A genome-linked copy of the NS-1 polypeptide is located on the outside of infectious parvovirus particles. *J. Virol.* **63**, 3902–3911.

Cotmore, S. F., and Tattersall, P. (1992). In vivo resolution of circular plasmids containing concatemer junction fragments from minute virus of mice DNA and their subsequent replication as linear molecules. *J. Virol.* **66**, 420–431.

Cotmore, S. F., and Tattersall, P. (1994). An asymetric nucleotide in the parvoviral 3′ hairpin directs segregation of a single active origin of DNA replication. *EMBO J.* **13**, 4145–4152.

Cotmore, S. F., and Tattersall, P. (1995). DNA replication in the parvoviruses. *Semin. Virol.* **6**, 271–281.

Cotmore, S. F., Nuesch, J. P. F., and Tattersall, P. (1993). Asymmetric resolution of a parvovirus palindrome in vitro. *J. Virol.* **67**, 1579–1589.

Courey, A. J., and Tjian R. (1988). Analysis of Sp1 in vivo reveals multiple transcription domains, including a novel glutamine-rich activation motif. *Cell* **55**, 887–898.

Cournoyer, D., and Caskey, C. (1993). Gene therapy of the immune system. *Annu. Rev. Immunol.* **11**, 297–329.

Crawford, L. V. (1966). A minute virus of mice. *Virology* **29**, 605–612.

Culver, K. W., Ram, Z., Wallbridge, S., Ishi, H., Oldfield, E. H., and Blaese, R. M. (1992). In vivo transfer with retroviral vector-producer cells for treatment of experimental brain tumors. *Science* **256**, 1550–1552.

Dedieu, J.-F., Ronco, J., van der Werf, S., Hogle, J. M., Henin, Y., and Girard, M. (1992). Poliovirus chimas expressing sequences from the principal neutralization domain of human immunodeficiency virus type 1. *J. Virol.* **66**, 3161–3167.

Diffoot, N., Shull, B. C., Chen, K. C., Stout, E. R., Lederman, M., and Bates, R. C. (1989). Identical ends are not required for the equal encapsidation of plus- and minus-strand parvovirus LuIII DNA. *J. Virol.* **63**, 3180–3184.

Diffoot, N., Chen, K. C., Bates, R. C., and Lederman, M. (1993). The complete nucleotide sequence of parvovirus LuIII and localization of a unique sequence possibly responsible for its encapsidation pattern. *Virology* **192**, 339–345.

Dixit, M., Webb, M. S., Smart, W. C., and Ohi, S. (1991). Construction and expression of a recombinant adeno-associated virus that harbours a human β-globin encoding cDNA. *Gene* **104**, 253–257.

Doerig, C., Beard, P., and Hirt, B. (1987). A transcriptional promoter of the human parvovirus B19 active in vitro and in vivo. *Virology* **157**, 539–542.

Doerig, C., Hirt, B., Beard, P., and Antonietti, J.-P. (1988). Minute virus of mice nonstructural protein is necessary and sufficient for transactivation of the viral P39 promoter. *J. Gen. Virol.* **69**, 2563–2573.

Doerig, C., Hirt, B., Antonietti, J.-P., and Beard, P. (1990). Nonstructural proteins of parvovirus B19 and minute virus of mice controls transcription. *J. Virol.* **64**, 387–396.

Dumas, B., Jourdan, M., Pascaud, A.-M., and Bergoin, M. (1992). Complete nucleotide sequence of the cloned infectious genome of *Junonia coenia* densovirus reveals an organization unique among parvoviruses. *Virology* **191**, 202–222.

Dupont, F., Tenenbaum, L., Guo, L.-P., Spegelaere, P., Zeicher, M., and Rommelaere, J. (1994). Use of an autonomous parvovirus vector for selective transfer of a foreign gene into transformed human cells of different tissue origins and its expression therein. *J. Virol.* **68**, 1397–1406.

Dupressoir, T., Vanacker, J.-M., Cornelis, J. J., Duponchel, N., and Rommelaere, J. (1989). Inhibition by parvovirus H-1 of the formation of tumors in nude mice and colonies in vitro by transformed human mammary epithelial cells. *Cancer Res.* **49**, 3203–3208.

Einerhand, M. P. W., Antoniou, A., Zolotukhin, S., Muzyczka, N., Berns, K. I., Grosveld, F., and Valerio, D. (1995). Regulated high-level human β-globin expression in erythroid cells following recombinant adeno-asociated virus-mediated gene transfer. *Gene Ther.* **2**, 336–343.

Faisst, S., Schlehofer, J. R., and zur Hausen, H. (1989). Transformation of human cells by oncogenic viruses supports permissiveness for parvovirus H-1 propagation. *J. Virol.* **63**, 2152–2158.

Faust, E., and Ward, D. C. (1979). Incomplete genomes of the parvovirus minute virus of mice: Selective conservation of genome termini, including the origin for DNA replication. *J. Virol.* **32**, 276–292.

Flotte, T. R., Afione, S. A., Conrad, C., McGrath, S. A., Solow, R., Oka, H., Zeitlin, P. L., Guggino, W. B., and Carter, B. J. (1993). Stable in vivo expression of the cyctic fibrosis transmembrane conductance regulator with an adenoassociated virus vector. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10613–10617.

Flotte. T. R., Barraza-Oritz, X., Solow, R., Afione, S. A., Carter, B. J., and Guggino, W. B. (1995). An improved system for packaged recombinant adeno-associated virus vectors capable of in vivo transduction. *Gene Ther.* **2**, 29–37.

Gardiner, E. M., and Tattersall, P. (1988a). Mapping of the fibrotropic and lymphotropic host range determinants of the parvovirus minute virus of mice. *J. Virol.* **62**, 2605–2613.

Gardiner, E. M., and Tattersall, P. (1988b). Evidence that developmentally regulated control of gene expression by a parvoviral allotropic determinant is particle mediated. *J. Virol.* **62**, 1713–1722.

Giraud, C., Devauchelle, G., and Bergoin, M. (1992). The densovirus of *Junonia coenia* (JcDNV) as an insect cell expresion vector. *Virology* **186**, 207–218.

Giraud, C., Linden, R. M., and Berns, K. I. (1994). Site-specific integration by AAV: Structure of recombinants formed with an EBV-based vector. 13th Annual American Society of Virology Meeting, Madison. [Abstract]

Goodman, S., Xiao, X., Donahue, R. E., Moulton, A., Miller, J., Walsh, C., Young, N. S., Samulski, R. J., and Nienhuis, A. W. (1994). Recombinant adeno-associated virus-mediated gene transfer into hematopoietic progenitor cells. *Blood* **84**, 1492–1500.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5547–5551.

Green, M. R., and Roeder, R. G. (1980). Transcripts of adeno-associated virus genome: Mapping of major RNAs. *J. Virol.* **36**, 79–92.

Guetta, E., Graziani, Y., and Tal, J. (1986). Suppression of Ehrlich's ascites tumors in mice by minute virus of mice. *J. Natl. Cancer Inst.* **76**, 1177–1179.

Guetta, E., Mincberg, M., Mousset, S., Bertinchamps, C., Rommelaere, J., and Tal, J. (1990). Selective killing of transformed rat cells by minute virus of mice does not require infectious virus production. *J. Virol.* **64**, 458–462.

Halbert, C. L., Alexander, I. E., Wolgamot, G. M., and Miller, A. D. (1995). Adeno-associated virus vectors transduce primary cells much less efficiently than immortalized cells. *J. Virol.* **69**, 1473–1479.

Hanson, N. D., and Rhode, S. L. (1991). Parvovirus NS1 stimulates P4 expression by interaction with the terminal repeats and through DNA amplification. *J. Virol.* **65**, 4325–4333.

Hauswirth, W. W., and Berns, K. I. (1977). Origin and termination of adeno-associated virus DNA replication. *Virology* **79**, 488–499.

Hermonat, P. L., and Muzyczka, N. (1984). Use of adeno-associated virus as a mammalian DNA cloning vector: Transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6466–6470.

Hermonat, P. L., Labow, M. A., Wright, R., Berns, K., and Muzyczka, N. (1984). Genetics of adeno-associated virus: Isolation and preliminary characterization of adeno-associated virus type 2 mutants. *J. Virol.* **51**, 329–339.

Hoey, T., Weinzierl, R. O. J., Gill, G., Chen, J. L., Dynlacht, B. D., and Tjian, R. (1993). Molecular cloning and functional analysis of *Drosophila* TAF110 reveal properties expected of coactivators. *Cell* **72**, 247–260.

Holscher, C., Horer, M., Kleinschmidt, J. A., Zentgraf, H., Burkle, A., and Heilbronn, R. (1994). Cell lines inducibly expressing the adeno-associated virus (AAV) *rep* gene: Requirements for productive replication of *rep*-negative AAV mutants. *J. Virol.* **68**, 7169–7177.

Hultmark, D., Klemenz, R., and Gehring, W. J. (1986). Translational and transcriptional control elements in the untranslated leader of the heat shock gene hsp22. *Cell* **44**, 429–438.

Iatrou, K., and Meidinger, R. G. (1989). Bombyx mori nuclear polyhedrosis virus-based vectors for expressing passenger genes in silkmoth cells under viral or cellular promoter. *Gene* **75**, 59–71.

Ilyina T. V., and Koonin E.V. (1992). Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucariotes and archaebacteria. *Nucleic Acids Res.* **20**, 3279–3285.

Im, D.-S., and Muzyczka, N. (1990). The AAV origin binding protein *rep*68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell* **61**, 447–457.

Jongeneel, C. V., Sahli, R., McMaster, G. K., and Hirt, B. (1986). A precise map of the splice junction in the RNA's of the minute virus of mice, an autonomous parvovirus. *J. Virol.* **59**, 564–573.

Jourdan, M., Jousset, F.-X., Gervais, M., Skory, S., Bergoin, M., and Dumas, B. (1990). Cloning of the genome of a densovirus and rescue of infectious virions from recombinant plasmid in the insect host *Spodoptera littoralis*. *Virology* **179**, 403–409.

Joyner, A., Keller, G., Phillips, R. A., and Bernstein, A. (1983). Retrovirus transfer of a bacterial gene into haematopoietic progenitor cells. *Nature* **305**, 556–558.

Kawase, S., Garzon, S., Su, D.-M., and Tijssen, P. (1990). Insect parvovirus diseases. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 2, Ch. 16. CRC Press, Boca Raton, Florida.

Kay, M. A., Rothenberg, S., Landen, C. N., Bellinger, D. A., Leland, F., Toman, C., Finegold, M., Thompson, A. R., Read, M. S., Brinkhous, K. M., and Woo, S. L. C. (1993). In vivo therapy of hemophilia B: Sustained partial correction in factor IX deficient dogs. *Science* **262**, 117–119.

Kelly, D. C., Moore, N. F., Spilling, C. R., Barwise, A. H., and Walker, I. O. (1980). Densonucleosis virus structural proteins. *J. Virol.* **36**, 224–235.

Koonin, E. V. (1993). A common set of conserved motifs in a vast variety of putative nuclei

acid-dependent ATPases including MCM proteins involved in the initiation of eucaryotic DNA replication. *Nucleic Acid Res.* **21**, 2541–2547.

Kotin, R. M. (1994). Prospects for the use of adeno-associated virus for human gene therapy. *Hum. Gene Ther.* **5**, 793–801.

Kotin, R., and Berns, K. (1989). Organization of adeno-associated virus in latently infected Detroit 6 cells. *Virology* **170**, 460–467.

Kotin, R. M., Linden, R. M., and Berns, K. I. (1992). Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by nonhomologous recombination. *EMBO J.* **11**, 5071–5078.

Labieniec-Pintel, L., and Pintel, D. (1986). The minute virus of mice P39 transcription unit can encode both capsid proteins. *J. Virol.* **57**, 1163–1167.

Labow, M., and Berns, K. I. (1988). The adeno-associated virus *rep* gene inhibits replication of an adeno-associated virus/simian virus 40 hybrid genome in cos-7 cells. *J. Virol.* **62**, 1705–1712.

Labow, M., Graf, L. H., and Berns, K. I. (1987). Adeno-associated virus gene expression inhibits cellular transformation by heterologous genes. *Mol. Cell. Biol.* **7**, 1320–1325.

Laughlin, C., Tratschin, J.-D., Coon, H., and Carter, B. J. (1983). Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene* **23**, 65–73.

Lebedeva, O. P., Kuznetsova, M. A., Zelenko, A. P., and Gudz-Gorban, A. P. (1973). Investigation of a virus disease of densonucleosis type in a laboratory culture of *Aedes aegypti*. *Acta. Virol.* **17**, 253–256.

Lebkowski, J. S., McNally, M. M., Okarma, T. B., and Lerch, L. B. (1988). Adeno-associated virus: A vector system for efficient introduction into a variety of mammalian cell types. *Mol. Cell. Biol.* **8**, 3988–3996.

Legendre, D., and Rommelaere, J. (1992). Terminal regions of the NS-1 protein of the parvovirus minute virus of mice are involved in cytotoxicity and promoter transinhibition. *J. Virol.* **66**, 5705–5713.

Li, X., and Rhode, S. L. (1991). Nonstructural protein NS2 of parvovirus H-1 is required for efficient viral protein synthesis and virus production in rat cells in vivo and in vitro. *Virology* **184**, 117–130.

Liljestrom, P., and Garoff, H. (1991). A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Biotechnology* **9**, 1356–1361.

Lin, S., Gaiano, N., Culp, P., Burns, J. C., Friedmann, T., Yee, J.-K., and Hopkins, N. (1994). Integration and germ line transmission of a pseudotyped retroviral vector in zebrafish. *Science* **265**, 666–669.

Liu, Q., Yong, C. B., and Astell, C. R. (1994). In vitro resolution of the minute virus of mice (MVM) genome supports the modified rolling hairpin model for MVM replication. *Virology* **201**, 251–262.

Lucknow, V. A., and Summers, M. D. (1988). Trends in the development of baculovirus expression vectors. *Biotechnology* **6**, 47–55.

Ma, J., and Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847–853.

Majaniemi, I., and Siegl, G. (1984). Early events in the replication of parvovirus LuIII. *Arch. Virol.* **81**, 285–302.

Mann, R., Mulligan, R. C., and Baltimore, D. (1983). Construction of a retrovirus packaging mutant and its use to produce helper free defective retrovirus. *Cell* **33**, 153–159.

Maxwell I. H., and Maxwell, F. (1994). A modified plaque assay and infected cell hybridization assay for wild-type and recombinant LuIII autonomous parvovirus. *Biotechniques* **16**, 876–881.

Maxwell, I. H., Long, C. J., Carlson, J. O., Rhode, S. L., and Maxwell, F. (1993a). Encapsi-

dation of a recombinant LuIII parvovirus genome by H1 virus and the fibrotropic or lymphotropic strains on minute virus of mice. *J. Gen. Virol.* **74**, 1175–1179.

Maxwell, I. H., Maxwell, F., Rhode, S. L., Corsini, J., and Carlson, J. O. (1993b). Recombinant LuIII autonomous parvovirus as a transient transducing vector for human cells. *Hum. Gene Ther.* **4**, 441–450.

Maxwell, I. H., Spitzer, A. L., Maxwell F., and Pintel, D. (1995). The capsid determinant of fibrotropism for the MVMp strain of minute virus of mice functions via VP2, and not VP1. *J. Virol.* **69**, 5829–5832.

Maxwell, I. H., Spitzer, A. L., Long, C. J., and Maxwell, F. (1996). Autonomous parvovirus transduction of a gene under control of tissue-specific or inducible promoters. *Hum. Gene Ther.*, in press.

Maza, de la L. M., and Carter, B. J. (1981). Inhibition of adenovirus oncogenicity in hamsters by adeno-associated virus DNA. *J. Natl. Cancer Inst.* **67**, 1323–1326.

McLaughlin, S. K., Collis, P., Hermonat, P. L., and Muzyczka, N. (1988). Adeno-associated virus general transduction vectors: Analysis of proviral structures. *J. Virol.* **62**, 1963–1973.

McNeall, J., Sanchez, A., Gray, P. P., Chesterman, C. N., and Sleigh, M. J. (1989). Hyper-inducible gene expression from a metallothionein promoter containing additional metal-responsive elements. *Gene* **76**, 81–88.

Melchner, H. von, Reddy, S., and Ruley, H. E. (1990). Isolation of cellular promoters using a retrovirus promoter trap. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3733–3737.

Mendelson, E., Smith, M. G., Miller, I. L., and Carter, B. J. (1988). Effect of a viral *rep* gene on transformation of cells by an adeno-associated virus vector. *Virology* **166**, 612–615.

Miller, A. D. (1992). Human gene therapy comes of age. *Nature* **357**, 455–460.

Miller, A. D., and Rosman, G. J. (1989). Improved retroviral vectors for gene transfer and expression. *Biotechniques* **7**, 980–990.

Miller, A. D., Jolly, D. J., Friedmenn, T., and Verma, I. (1983). A transmissable retrovirus expressing human hypoxanthine phosphotransferase (HRPT): Gene transfer into cells obtained from humans deficient in HRPT. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4709–4713.

Miller, J. L., Donahue, R. E., Sellers, S. E., Samulski, R. J., Young, N. S., and Niehaus, A. W. (1994). Recombinant adeno-associated virus (rAAV)-mediated expression of a human γ-globin gene in human progenitor-derived erythroid cells. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10183–10187.

Momeda, M., Wong, S., Kawase, M., Young, N.S., and Kajigaya, S. (1994). A putative nucleoside triphosphate binding domain in the non-structural protein of B19 parvovirus is required for cytotoxicity. *J. Virol.* **68**, 8443–8446.

Morgan, W. R., and Ward, D. C. (1986). Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs. *J. Virol.* **60**, 1170–1174.

Mousset, S., Ouadrhiri, Y., Caillet-Fauquet, P., and Rommelaere, J. (1994). The cytotoxicity of the autonomous parvovirus minute virus of mice nonstructural proteins in FR3T3 rat cells depends on oncogene expression. *J. Virol.* **68**, 6446–6453.

Muller, D.-E., and Siegl, G. (1983a). Maturation of parvovirus LuIII in a subcellular system. I. Optimal conditions for in vitro synthesis and encapsidation of viral DNA. *J. Gen. Virol.* **64**, 1043–1054.

Muller, D.-E., and Siegl, G. (1983b). Maturation of parvovirus LuIII in a subcellular system. II. Isolation and characterization of nucleoprotein intermediates. *J. Gen. Virol.* **64**, 1055–1067.

Mulligan, R. C. (1993). The basic science of gene therapy. *Science* **260**, 926–932.

Mulligan, R. C., Howard, B. H., and Berg, P. (1979). Synthesis of rabbit β-globin in

cultured monkey kidney cells following infection with a SV-40 β-globin recombinant genome. *Nature* **277**, 108–114.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., and Summers, M. D. (1995). "Virus Taxonomy," p. 586. Springer-Verlag, Vienna.

Muzyczka, N. (1992). Use of adeno-associated virus as a general transduction vector for mammalian cells. *Curr. Top. Microbiol. Immunol.* **158**, 97–129.

Myers, M. W., and Carter, B. J. (1980). Assembly of adeno-associated virus. *Virology* **102**, 71–82.

Naeger, L. K., Cater, J., and Pintel, D. J. (1990). The small nonstructural protein (NS2) of the minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner. *J. Virol.* **64**, 6166–6175.

Nakagaki, M., and Kawase, S. (1980). Structural proteins of densonucleosis virus isolated from the silkworm, *Bombyx mori*, infected with the flacherie virus. *J. Invert. Pathol.* **36**, 166–171.

Nuesch, J. P. F., and Tattersall, P. (1993). Nuclear targeting of the parvoviral replicator molecule NS1: Evidence for self-association prior to nuclear transport. *Virology* **196**, 637–651.

Nuesch, J. P. F., Cotmore, S. F., and Tattersall, P. (1992). Expression of functional parvoviral NS1 from recombinant vaccinia virus: Effects of mutations in the nucleotide-binding motif. *Virology* **191**, 406–416.

Nuesch, J. P. F., Cotmore, S. F., and Tattersall, P. (1995). Sequence motifs in the replicator protein of parvovirus MVM essential for nicking and covalent attatchment to the viral origin: Identification of the linking tyrosine. *Virology* **209**, 122–135.

Oppenheim, A., and Peleg, A. (1989). Helpers for efficient encapsidation of SV40 pseudovirions. *Gene* **77**, 79–86.

Oppenheim, A., Peleg, A., Fibach, E., and Rachmilewitz, E. A. (1986). Efficient introduction of plasmid DNA into human hemopoietic cells by encapsidation in simian virus 40 pseudovirions. *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6925–6929.

Paradiso, P. R. (1981). Infectious process of the parvovirus H-1: Correlation of protein content, particle density, and viral infectivity. *J. Virol.* **39**, 800–807.

Park, K.H., Huang, T., Correia, F. F., and Krystal, M. (1991). Rescue of a foreign gene by Sendai virus. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5537–5541.

Pintel, D., Dadachanji, D., Astell, C. R., and Ward D. C. (1983). The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. *Nucleic Acids Res.* **11**, 1019–1038.

Pintel, D. J., Gersappe, A., Haut, D., and Pearson, J. (1995). Determinants that govern alternative splicing of parvovirus pre-mRNA's. *Semin. Virol.* **6**, 283–290.

Podsakoff, G., Wong, K. K., Chatergee, S. (1994). Efficient gene transfer into non-dividing cells by adeno-associated virus-based vectors. *J. Virol.* **68**, 5656–5666.

Porter, D. C., Ansardi, D. C., and Morrow, C. D. (1995). Encapsidation of poliovirus replicons encoding the complete human immunodeficiency virus type 1 *gag gene* by using a complementation system which provides the P1 capsid protein in trans. *J. Virol.* **69**, 1548–1555.

Rhode, S. L. (1976). Replication process of the parvovirus H-1 V. Isolation and characterization of temperature-sensitive H-1 mutants defective in progeny DNA synthesis. *J. Virol.* **17**, 659–667.

Rhode, S. L. (1977). Replication process of the parvovirus H-1. VI. Characterization of a replication terminus of H-1 replicative form DNA. *J. Virol.* **21**, 694–712.

Rhode, S. L. (1978). Defective interfering particles of parvovirus H-1. *J. Virol.* **27**, 347–356.

Rhode, S. L. (1985). Trans-activation of parvovirus p38 promoter by the 76K noncapsid protein. *J. Virol.* **55**, 886–889.

Rhode, S. L. (1989). Both excision and replication of cloned autonomous parvovirus DNA require the NS1 (rep) protein. *J. Virol.* **63**, 4249–4256.

Rhode, S. L., and Klaassen, B. (1982). DNA sequence of the 5′ terminus containing the replication origin of parvovirus replicative form DNA. *J. Virol.* **41**, 990–999.

Richards, R. G., and Armentrout, R. W. (1979). Early events in parvovirus replication: Lack of integration by minute virus of mice into host cell DNA. *J. Virol.* **30**, 397–399.

Richards, R., Linser, P., and Armentrout, R. W. (1977). Kinetics of assembly of a parvovirus, minute parvovirus of mice, in synchronized rat brain cells. *J. Virol.* **22**, 778–793.

Rittner, K., Heilbronn, R., Kleinschmidt, J. A., and Sczakiel, G. (1992). Adeno-associated virus type 2-mediated inhibition of human immunodeficiency virus type 1 (HIV-1) replication: Involvement of $P78_{rep}/P68_{rep}$ and the HIV-1 long terminal repeat. *J. Gen. Virol.* **73**, 2977–2981.

Roemer, K., and Friedmann, T. (1992). Concepts and strategies for human gene therapy. *Eur. J. Biochem.* **208**, 211–255.

Rommelaere, J., and Cornelius, J. J. (1991). Antineoplastic activity of parvoviruses. *J. Virol. Methods* **33**, 233–251.

Rommelaere, J., and Tattersall, P. (1990). Onco-suppression by parvoviruses. *In* "Handbook of the Parvoviruses" (P. Tjissen, ed.), Vol. 2, Ch. 3. CRC Press, Boca Raton, Florida.

Ron, D., and Tal, J. (1985). Co-evolution of cells and virus as a mechanism for the persistence of lymphotropic minute virus of mice in L-cells. *J. Virol.* **55**, 424–430.

Rosenfeld, M. A., Yoshimura, K., Trapnell, B. C., Yoneyama, K., Rosenthal, E. R., Dalemans, W., Fukayama, M., Bargon, J., Stier, L. E., Stratford-Perricaudet, L., Perricaudet, M., Guggino, W. B., Pavirani, A., Lecocq, J.-P., and Crystal, R. G. (1992). In vivo transfer of the human cystic fibrosis transmembrane conductance regulator gene to the airway epithelium. *Cell* **68**, 143–155.

Russell, D. W., Miller, A. D., and Alexander, I. E. (1994). Adeno-associated virus vectors preferentially transduce cells in S phase. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8915–8919.

Russell, S. J., Brandenburger, A., Flemming, C. L., Collins, M. K. L., and Rommelaere, J. (1992). Transformation-dependent expression of interleukin genes delivered by a recombinant parvovirus. *J. Virol.* **66**, 2821–2828.

Salvino, R., Skiadopoulus, M., Faust, E. A., Tam, P., Shade, R. O., and Astell, C. R. (1991). Two spatially distinct elements constitute a bipartite DNA replication origin in the minute virus of mice genome. *J. Virol.* **65**, 1352–1363.

Sambrook, J. (1978). Isolation of defective variants of simian virus 40 whose genomes contain sequences derived from adenovirus-2 DNA. *J. Gen. Virol.* **38**, 313–327.

Samulski, R. J., Berns, K. I., Tan, M., and Muzyczka, N. (1982). Cloning of adeno-associated virus into pBR322: Rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2077–2081.

Samulski, R. J., Chang, L.-S., and Shenk, T. (1989). Helper free stocks of recombinant adeno-associated viruses: Normal integration does not require viral gene expression. *J. Virol.* **63**, 3822–3828.

Samulski, R. J., Zhu, X., Brook, J. D., Housman, D. E., Epstein, N., and Hunter, L. A. (1991). Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J.* **10**, 3941–3950.

Schlehofer, J. R., Ehrbar, M., and zur Hausen, H. (1986). Vaccinia virus, herpes simplex virus, and carcinogens induce DNA amplification in a human cell line and support replication of a herpes dependent parvovirus. *Virology* **152**, 110–117.

Schlesinger, S. (1993). Alphaviruses—vectors for the expression of heterologous genes. *Trends Biotechnol.* **11**, 18–22.

Senapathy, P., Tratschin, J.-D., and Carter, B. J. (1984). Replication of adeno-associated virus DNA: Complementation of naturally occuring *rep*–mutants by wild type genomes and correction of terminal palindrome deletions. *J. Mol. Biol.* **179**, 1–20.

Shull, B. C., Chen, K. C., Lederman, M., Stout, E. R., and Bates, R. C. (1988). Genomic clones of bovine parvovirus: Construction and effect of deletions and terminal sequence inversions on infectivity. *J. Virol.* **62**, 417–426.

Siegl, G. (1984). Biology and pathologenicity of autonomous parvoviruses. *In* "Parvoviruses" (K. Berns, ed.). Plenum Press, New York.

Siegl, G., Bates, R. C., Berns, K. I., Carter, B. J., Kelly, D. C., Kurstak, E., and Tattersall, P. (1985). Characteristics and taxonomy of *Parvoviridae*. *Intervirology* **23**, 61–73.

Smith, A. D., Resnick, D. A., Zhang, A., Geisler, S. C., Arnold, E., and Arnold, G. F. (1994). Use of random systematic mutagenesis to generate viable human rhinovirus 14 chimeras displaying human immunodeficiency virus type 1 V3 loop sequences. *J. Virol.* **68**, 575–579.

Smith, R. L., Geller, A. I., Escudero, K. W., and Wilcox, C. L. (1995). Long term expression in sensory neurons in tissue culture from herpes simplex type 1 promoters in a herpes derived vector. *J. Virol.* **69**, 4593–4599.

Soike, K. F., Iatropoulis, M., and Siegl, G. (1976). Infection of newborn and fetal hamsters induced by inoculation of LuIII parvovirus. *Arch. Virol.* **51**, 235–241.

Spitzer, A. L., Maxwell, F., Corsini, J., and Maxwell, I. H. (1996). Species specificity for transduction of cultured cells by a recombinant LuIII rodent parvovirus genome encapsidated by canine parvovirus or feline panleukopenia virus. *J. Gen. Virol.*, submitted.

Srivastava, C. H., Samulski, R. J., Lu, L., Larsen, S. H., and Srivastava, A. (1989). Construction of a recombinant human parvovirus B19: Adeno-associated virus type 2 (AAV) DNA inverted terminal repeats are functional in an AAV-B19 hybrid virus. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8078–8082.

Strassheim, L. M., Gruenberg, A., Veijalainen, P., Sgro, J.-Y., and Parrish, C. R. (1994). Two dominant neutralizing antigenic determinants of canine parvovirus are found on the threefold spike of the virus capsid. *Virology* **198**, 175–184.

Tal, J., and Attathom, T. (1993). Insecticidal potential of the insect parvovirus GmDNV. *Arch. Ins. Biochem. Physiol.* **22**, 345–356.

Tam, P., and Astell, C. R. (1993). Replication of minute virus of mice minigenomes: Novel replication elements required for MVM DNA replication. *Virology* **193**, 812–824.

Tattersall, P. (1972). Replication of the parvovirus MVM. I. Dependence of virus multiplication and plaque formation on cell growth. *J. Virol.* **10**, 586–590.

Tattersall, P., and Cotmore, S. F. (1990). Reproduction of autonomous parvovirus DNA. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 8. CRC Press, Boca Raton, Florida.

Tattersall, P., and Gardiner, E. M. (1990). Autonomous parvovirus–host cell interactions. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 7. CRC Press, Boca Raton, Florida.

Tattersall, P., Cawte, P. J., Shatkin, A. J., and Ward, D. C. (1976). Three structural polypeptides coded for by the minute virus of mice, a parvovirus. *J. Virol.* **20**, 273–289.

Tijssen, P., and Bergoin, M. (1995). Densonucleosis viruses constitute an increasingly diversified subfamily among parvoviruses. *Semin. Virol.* **6**, 347–355.

Tijssen, P., van den Hurk, J., and Kurstak, E. (1976). Biochemical, biophysical and biological properties of densonucleosis virus. I. Structural proteins. *J. Virol.* **17**, 686–691.

Tijssen, P., Arella, M., and Kawase, S. (1990). Molecular biology of densonucleosis

viruses. *In* "Handbook of Parvoviruses" (P. Tijssen, ed.), Vol. 1, Ch. 14. CRC Press, Boca Raton, Florida.

Tijssen, P., Bergeron, J., Dubuc, R., and Hebert, B. (1995). Minor genetic changes among porcine parvovirus groups are responsible for major distinguishing biological properties. *Semin. Virol.* **6**, 319–328.

Toolan, H. W. (1967). Lack of oncogenic effect of the H-viruses for hamsters. *Nature* **214**, 1036.

Toolan, H. W., and Ledinko, N. (1968). Inhibition by H-1 virus of the incidence of tumors produced by adenovirus 12 in hamsters. *Virology* **35**, 475–478.

Toolan, H., Saunders, E. L., Southam, C. M., Moore, A. E., and Levine, A. G. (1965). H1 virus viremia in the human. *Proc. Soc. Exp. Biol. Med.* **119**, 711–715.

Toolan, H. W., Rhode, S. L., and Girthy, J. F. (1982). Inhibition of 7,12-dimethylbenz(a)anthracene-induced tumors in Syrian hamsters by prior infection with H-1 virus. *Cancer Res.* **43**, 2552–2555.

Tratschin, J.-D., West, M. P., Sandbank, T., and Carter, B. J. (1984a). A human parvovirus, adeno-associated virus, as a eukaryotic vector: Transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Mol. Cell. Biol.* **4**, 2072–2081.

Tratschin, J.-D., Miller, I. L., and Carter, B. J. (1984b). Genetic analysis of adeno-associated virus: Properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. *J. Virol.* **51**, 611–619.

Tratschin, J.-D., Miller, I. L., Smith, M. G., and Carter, B. J. (1985). Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. *Mol. Cell. Biol.* **5**, 3251–3260.

Tratschin, J.-D., Tal, J., and Carter, B. J. (1986). Negative and positive regulation in trans of gene expression from adeno-associated vectors in mammalian cells by a *rep* gene product. *Mol. Cell. Biol.* **6**, 2884–2894.

Truyen, U., Gruenberg, A., Chang, S. F., Obermaier, B., Veijalainen, P., and Parrish, C. R. (1995). Evolution of the feline-subgroup parvoviruses and control of canine host range in vitro. *J. Virol.* **69**, 4702–4710.

Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W., and Parrish, C. L. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science* **251**, 1456–1463.

Tullis, G. E., Labienic-Pintel, L., Clemens, K. E., and Pintel, D. (1988). Generation and characterization of a temperature sensitive mutation in the NS-1 gene of the autonomous parvovirus minute virus of mice. *J. Virol.* **62**, 2736–2744.

Tyson, J. J., Chen, K. C., Lederman, M., and Bates, R. C. (1990). Analysis of the kinetic hairpin transfer model for parvoviral DNA replication. *J. Theor. Biol.* **144**, 155–169.

Urcelay, E., Ward, P., Wiener, S. M., Safer, B., and Kotin, R. (1995). Asymmetric replication in vitro from a human sequence element is dependent on adeno-associated virus Rep protein. *J. Virol.* **69**, 2038–2046.

Van Hill, B., Duponchel, N., Salome, N., Spruyt, N., Cotmore, S. F., Tattersall, P., Cornelius, J. J., and Rommelaere, J. (1989). Limitations to the expression of parvoviral nonstructural proteins may determine the extent of sensitization of EJ-ras-transformed rat cells to minute virus of mice. *Virology* **171**, 89–97.

Vanacker, J.-M., and Rommelaere, J. (1995). Non-structural proteins of parvoviruses: From cellular effects to cytotoxicity. *Semin. Virol.* **6**, 291–297.

Vincent, K. A., Moore, G. K., and Haigwood, N. L. (1990). Replication and packaging of HIV envelope genes in a novel adeno-associated virus vector system. *Vaccine* **90**, 353–359.

Walsh, C. E., Liu, J. M., Ypung, Xiao, X., Young, N. S., Nienhuis, A. W., and Samulski, R. J. (1992). Regulated high level expression of a human gamma-globin gene introduced into erythroid cells by an adeno-associated virus vector. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7257–7261.

Welsh, M. J., Smith, A. E., Zabner, J., Rich, D. P., Graham, S. M., Gregory, R. J., Pratt, B. M., and Moscicki, R. A. (1994). Clinical protocol. Cystic fibrosis gene therapy using an adenovirus vector: in vivo safety and efficacy in nasal epithelium. *Hum. Gene Ther.* **5**, 209–219.

Willwand, K., and Hirt, B. (1991). The minute virus of mice capsid specifically recognizes the 3' hairpin of viral replicative form DNA: Mapping of the binding site by hydroxy radical footprinting. *J. Virol.* **65**, 4629–4635.

Willwand, K., and Hirt, B. (1993). The major capsid protein VP2 of minute virus of mice (MVM) can form particles which bind to the 3'-terminal hairpin of MVM replicative form DNA and package single strand progeny DNA. *J. Virol.* **67**, 5660–5663.

Willwand, K., and Kaaden, O.-R. (1990). Proteins of viral and cellular origin bind to the aleutian disease virus (ADV) DNA 3' terminal hairpin: Presentation of a scheme for encapsidation of ADV DNA. *J. Virol.* **64**, 1598–1605.

Wilson, G. M., Jindal, H. K., Yeung, D. E., Chen, W., and Astell, C. R. (1991). Expression of minute virus of mice nonstructural protein in insect cells: Purification of ATPase and helicase activities. *Virology* **185**, 90–98.

Yacobson, B., Hrynko, T. A., Peak, M. J., and Wincour, E. (1989). Replication of adeno-associated virus in cells irradiated with UV light at 254 nm. *J. Virol.* **63**, 1023–1030.

Yang, O., Chen, F., and Trempe, J. E. (1994). Characterization of cell lines that inducibly express the adeno-associated virus *rep* proteins. *J. Virol.* **68**, 4847–4856.

Yang, Q., Kadam, A., and Trempe, J. P. (1992). Mutational analysis of the adeno-associated *rep* gene. *J. Virol.* **66**, 6058–6069.

Yang, Q., Chen, F., Ross, J., and Trempe, J. P. (1995). Inhibition of cellular and SV40 DNA replication by the adeno-associated virus rep proteins. *Virology* **207**, 246–250.

Yang, T. J. (1987). Parvovirus-induced regression of canine transmissible venereal sarcoma. *Amer. J. Vet. Res.* **48**, 799–800.

Zhao, Q., Schoborg, R. V., and Pintel, D. (1994). Alternative splicing of pre-mRNAs encoding the nonstructural proteins of minute virus of mice is facilitated by sequences within the downstream intron. *J. Virol.* **68**, 2849–2859.

Zhou, S. Z., Cooper, S., Kang, L. Y., Ruggieri, L., Heimfeld, S., Srivastava, A., and Broxmeyer, H. E. (1994). Adeno-associated virus 2-mediated high efficiency transfer into immature and mature subsets of hematopoietic progenitor cells in human umbilical cord blood. *J. Exp. Med.* **179**, 1867–1875.

This Page Intentionally Left Blank

ADVANCES IN VIRUS RESEARCH, VOL. 47

VIRUS-INDUCED IMMUNOPATHOLOGY

Barry T. Rouse

Department of Microbiology
College of Veterinary Medicine
University of Tennessee
Knoxville, Tennessee 37996

I. Introduction

The burgeoning field of viral immunology has been heavily influenced by the lifestyle set by infection of mice with lymphocytic choriomeningitis virus (LCMV). Studies of this virus provided vital clues to the understanding of immunological tolerance (Burnet and Fenner, 1949). T-cell recognition (Doherty and Zinkernagel, 1974), the concept of high-dose paralysis (Hotchin, 1971) [now more popularly termed "immune exhaustion," Moskophidis *et al.* (1993)], and the notion that tissue injury may result from an immune response to virus-infected cells (Buchmeier *et al.,* 1980) or to immune complexes composed of viral protein and antibody (Oldstone and Dixon, 1969). Virus-induced immunopathology likely operates to some degree in many, perhaps even all, animal virus infections. In fact, a degree of immunopathology may be the tariff the body must pay to eliminate infections by most agents, even those which are highly cytopathic. With many noncytolytic viruses it is only because of the immune response against them that any discernable disease occurs. The prime example is set by LCMV, but the mechanism(s) operative in this infection do not repre-

sent the full gamut of cellular and molecular events involved in virus-induced immunopathologies. In this brief review, some principal examples of immunopathogenesis are discussed, accompanied by speculations about management. Indeed, the control of such viral infections requires the use of therapeutic vaccines and immune modulators that suppress the development of lesions, since prophylactic vaccines may not always succeed in preventing infection.

II. Immunopathological Lesions Which Primarily Involve $CD8^+$ T Cells

Infection of mice with LCMV provides the hallmark example of a $CD8^+$ T-cell-mediated immunopathology (Buchmeier *et al.*, 1980). Strains of LCMV differ in tissue tropism and virulence, but all are largely noncytolytic and the virus fails to cause overt tissue damage (Hotchin, 1971). Such damage only occurs after the development of an immune response that primarily involves T cells of the $CD8^+$ subset (Doherty *et al.*, 1990; Kagi *et al.*, 1995). Early clues that lesions in LCM are immunopathological came from the observation that the disease only occurred in immunocompetent animals. Accordingly, tolerized as well as immunosuppressed mice failed to exhibit the characteristic choriomeningitis (Burnet and Fenner, 1949; Buchmeier *et al.*, 1980). Later evidence came from adoptive transfer experiments with T-cell subtypes, the use of monoclonal antibodies to selectively deplete various cell types *in vivo*, and very recently the use of gene knockout and transgenic mice (Kagi *et al.*, 1995; Baenziger *et al.*, 1986; Leist *et al.*, 1987).

The evidence provides a clear-cut conclusion: tissue damage occurring in LCMV infection involves the essential participation of T cells of the $CD8^+$ lineage. However, the exact mechanisms by which $CD8^+$ T cells cause the inflammatory tissue damage remains unresolved. For example, in cerebral infections with LCMV, a significant recruitment and extravasation of immunoinflammatory cells to the sites of viral replication occurs in the meninges, ependymal membranes, and choroid plexus of the brain (Buchmeier *et al.*, 1980). Only a minor proportion of cells in the inflammatory reactions are LCMV antigen-specific $CD8^+$ T cells, yet these are essential for the inflammation to proceed (Doherty *et al.*, 1990). One hypothesis contends that the tissue damage results from direct cytotoxicity by $CD8^+$ cytotoxic T lymphocytes (CTL) to virus-infected cells (Kagi *et al.*, 1995). This idea is supported by the observation that knockout mice unable to generate perforin do not express disease (Kagi *et al.*, 1995). However, in brains

with lesions no evidence of overt cellular pathology is present histologically (Walker *et al.,* 1975). Alternatively, although the primary step may involve recognition by specific $CD8^+$ T cells, the bulk of tissue damage may result from the release of several proinflammatory cytokines from the activated inflammatory cells recruited to the site by factors released from $CD8^+$ T cells (Campbell *et al.,* 1994a).

Cytokines such as IL-1, IL-6, TNFα/β, IFNα, and IFNγ have been advocated as likely proinflammatory participants (Campbell *et al.,* 1994b; Sandberg *et al.,* 1994). Moreover, the actual cellular injury could well involve the generation of nitric oxide from IFNγ-activated macrophages recruited to the site (Campbell *et al.,* 1994a). Contrary to this hypothesis, however, is the recent evidence that lesions and disease susceptibility occur normally in IFNγ knockout mice (Campbell, 1995) and that attempts to ablate nitric oxide production do not ameliorate LCM (Campbell *et al.,* 1995). In LCMV-induced disease, the most intensively studied virus-induced immunopathology, one must conclude that the cellular and especially the molecular events leading to tissue damage remain a poorly understood immunoinflammatory bouillabaisse.

Several other viral infections also appear to induce disease by interacting with $CD8^+$ T cells. At least two human diseases provide candidate examples. These are acute hepatitis caused by hepatitis B virus (HBV) and myocarditis induced by Coxsackie B virus (CBV). A third candidate is certain lesions caused by HIV (Pantaleo *et al.,* 1993). With the former two examples, excellent murine animal models are available, and these have helped considerably to depict a role of $CD8^+$ T cells in their pathogenesis. Particularly informative have been studies on hepatitis B virus infection using the artificial system of transgenic mice expressing various gene constructs of HBV (Chisari and Ferrari, 1995). These investigations reveal that $CD8^+$ T cells are principally involved as mediators of viral clearance as well as immunopathology, and that a variety of cellular and molecular events are at play (Ando *et al.,* 1993). It seems reasonable to assume that similar $CD8^+$ T-cell-mediated events occur in human HBV-induced hepatitis, but this remains to be shown.

Transgenic mice that express HBV envelope antigens in their hepatocytes remain normal. However, they readily develop lesions which resemble human acute viral hepatitis if given adoptive transfers of $CD8^+$ MHC class I restricted HBV-specific cytotoxic T lymphocytes (Chisari and Ferrari, 1995). Sequential analysis of the basic system reveals that the lesions progress in a predictable stepwise fashion. The earliest detectable step involves direct attachment of CTL to HBV

antigen-positive hepatocytes, with the latter being killed by apoptosis. The widely scattered apoptotic hepatocytes in the transgenic mouse livers are reminiscent of acute viral hepatitis lesions in humans (Ando *et al.*, 1994a). Given the nature of apoptosis, death of such hepatocytes is unlikely to release agonists which drive the inflammatory response. However, after the initial apoptotic phase, antigen-nonspecific inflammatory cells such as neutrophils and monocytes are recruited, and these cause far more hepatic cell damage than do CTL and the damage zone extends way beyond the sites of CTL-mediated apoptosis (Ando *et al.*, 1993). Such events are presumed to be mediated by cytokines, particularly IFNγ, probably released by CTL, which can be directly cytotoxic to hepatocytes that express abundant levels of HBV surface antigen (Guidotti *et al.*, 1995).

In some transgenics in which hepatocytes retain high levels of HBV surface antigen, necrosis can be massive, and the mice die of hepatitis (Guidotti *et al.*, 1995). This step can be prevented by prior administration of neutralizing antibody to IFNγ or by depletion of macrophages. Consequently, in the HBV transgene model the immunopathology is initiated and likely orchestrated mainly by $CD8^+$ T cells, but the principal immunopathological effects appear to be mediated nonspecifically by cytokines and recruited cells such as macrophages. This pattern of events is more commonly found in $CD4^+$ T-cell-mediated immunopathology (discussed subsequently). What is of particular interest in the HBV transgene model is that when the transgene is expressed in tissues such as the kidney or brain, no disease ensues in these organs following the subsequent intravenous adoptive transfer of HBV-specific CTL (Ando *et al.*, 1994b). To cause damage in such organs, which compared to the liver have blood vessels that impede the escape of T cells, requires that the CTL be placed directly into the organs. Such observations indicate that the induction of immunopathology to a virus infection requires not only that the agent be present but that it be available for recognition by $CD8^+$ T cells. In the HBV transgene model, as in LCM, the molecular mechanisms of disease expression await further elucidation.

The murine CBV model of immunopathology has received less investigation than LCMV and HBV, and CBV is far more cytolytic than either LCMV or HBV and can alone cause tissue damage (virological pathology) (Woodruff, 1980). The CBV type 3 strain causes myocardial disease, although the extent of this syndrome is subject to numerous variables of virus and host (Chow *et al.*, 1991). Recently, knockout mice were used to follow the pathogenesis of CBV-3-induced myocarditis, which occurs in mice that survive the acute disease (Hanke *et al.*,

1995). A clear role for $CD8^+$ T-cell function was observed. Accordingly, myocarditis was severe in $CD4^+$ knockout mice but only minimal in animals deficient for a $CD8^+$ T-cell response because of β2M knockout (Hanke *et al.*, 1995). Moreover, the severe disease in $CD4^+$ knockout mice was abrogated by *in vivo* depletion of $CD8^+$ T cells with specific monoclonal antibodies. Such data clearly implicate a major role for $CD8^+$ T cells in the immunopathogenesis of CBV-3-induced myocarditis, but the molecular mechanism of tissue damage remains to be established. Some favor the notion that TNFα and IL-1β are involved, since if such cytokines are administered to infected mice the myocarditis is exacerbated (Lane *et al.*, 1992). Furthermore, TNFα is readily demonstrable in inflammatory cells at the site of cardiac lesions (Hanke *et al.*, 1995).

Great interest and alarm recently came from the outbreak of a rapidly progressive influenza-like, often fatal illness in several previously healthy persons in the Four Corners area of the United States (Nichol *et al.*, 1993). The outbreak was associated with a previously unrecognized agent now named Sin Nombre Virus (SNV) (Eliott *et al.*, 1994). The pulmonary histopathology in patients dying of the disease appeared consistent with an acute immunopathological response to virus-infected cells in the lung (Zaki *et al.*, 1995). Prominent among the inflammatory cells were $CD8^+$ T lymphocytes and it seems possible that such cells may be primary mediators of the pathogenesis of the emerging infectious disease. The results of ongoing studies on the pathogenesis of SNV infection should prove intriguing.

The viral disease whose mechanism of pathogenesis is under the most intensive investigation is, of course, HIV. Unfortunately, suitable murine models of HIV pathogenesis are lacking, and so it remains difficult to assess experimentally the viewpoint that HIV pathogenesis involves $CD8^+$ T cells (Pantaleo *et al.*, 1993; Zinkernagel and Hengartner, 1994). Initially, such cells are considered to play a protective role. However, this defense function is imperfect and virus usually persists in numerous cell types without causing their destruction. Conceivably, destruction of antigen-presenting cells as well as $CD4^+$ T cells which harbor virus by antigen-specific $CD8^+$ T cells may contribute to immunosuppression (Zinkernagel and Hengartner, 1994). The idea that viruses cause immunosuppression by a $CD8^+$ T-cell-mediated immunopathological reaction against infected cells of the immune system is clearly evident in LCMV infection of mice (Mims and Wainwright, 1968; Jacobs and Cole, 1976, Odermatt *et al.*, 1991). Depending on numerous variables affecting both the virus and infected host, a marked immune suppression and resultant enhanced suscepti-

bility to other agents are observed. Indeed, some have strongly advocated that the pathogenesis of immune suppression in HIV infection might best be understood by studying LCMV infection in mice (Zinkernagel and Hengartner, 1994). In LCMV infection, suppression results from an antiviral $CD8^+$ CTL response against infected antigen-presenting cells such as macrophages and perhaps dendritic cells (Jacobs and Cole, 1976). For a more complete discussion on the immunopathogenesis of HIV infection, other reviews are recommended (Pantaleo *et al.*, 1993; Zinkernagel and Hengartner, 1994).

III. Immunopathological Reactions Primarily Involving $CD4^+$ T Lymphocytes

Of the two principal types of $\alpha\beta$TCR T cells, the $CD4^+$ T-cell subset is usually considered to participate in effector activities more by generating an abundance of cytokines than do $CD8^+$ T lymphocytes. Such cytokines express a variety of activities that include recruitment and activation of nonspecific effector cells (Meltzer and Nacy, 1989). $CD4^+$ T cells can mediate direct effects such as cytotoxicity (Kolaitis *et al.*, 1990), but likely a more common function *in vivo* is to conduct an inflammatory reaction (Meltzer and Nacy, 1989). Such responses are usually termed delayed-type hypersensitivity (DTH) reactions. These represent accumulations of numerous cell types. Only a minority of cells are lymphocytes, and these include the antigen-specific $CD4^+$ T cells. The majority of cells are neutrophils and mononuclear cells, and these cells, especially the latter, are assumed to be the principal purveyors of the immunoprotective and tissue-damaging effects (Meltzer and Nacy, 1989). These damaging effects are numerous and include several proteolytic enzymes, reactive radicals of oxygen and nitrogen, and perhaps some cytokines such as TNFα (Laskin and Pendino, 1995).

Inflammatory reactions instigated by $CD4^+$ T cells vary in cellular composition and the nature of the chemical activities generated. In part, this reflects the fact that $CD4^+$ T cells express different profiles of cytokines and perhaps other signaling molecules such as chemokines (Mossman and Coffman, 1989). Usually, $CD4^+$ T cells are divided into two functional subsets: Th1 T cells producing principally IFNγ, IL-2, and TNFβ, and Th2 cells, which mainly produce IL-4, IL-5, and IL-10 (Mosmann and Coffman, 1989; Paul and Seder, 1994). The former cell types largely marshal inflammatory responses dominated by mononuclear cells and neutrophils. Th2-mediated inflammatory responses have more eosinophils and basophils; these are uncommon responses

to viral infections, although they are likely responsible for the alveolitis which occurs in respiratory syncytial virus (RSV) infection (Alwan *et al.,* 1994). The Th1-organized DTH reactions are considered common responses to viruses, and in some situations these are chronic and tissue-damaging. Such reactions occur in response both to cytolytic viruses such as measles (Johnson *et al.,* 1978) and herpes simplex virus (HSV) (Doymaz and Rouse, 1992) and to viruses which are noncytodestructive. Against cytolytic viruses the DTH response may be considered primarily protective unless, as discussed subsequently, the infection causes the exposure of self-derived determinants which in turn perpetuate the response.

A. Virus-Induced Immunopathology Orchestrated Mainly by Type 1 Cytokine-Producing $CD4^+$ T Cells

In the case of persistent viruses that are minimally cytopathic, the $CD4^+$-mediated DTH reaction to them, with its accompanying tissue damage, must be considered as largely immunopathological. Examples of agents which largely follow this script in the mouse include Theiler's murine encephalomyelitis virus (TMEV) (Miller and Karpus, 1994), mouse coronavirus (Fleming *et al.,* 1993), and Semliki Forest virus (SFV) (Mokharian and Swoveland, 1987). These persistent virus infections affect the central nervous system (CNS) of rodents and induce late inflammatory responses that involve the white matter, causing demyelination [reviewed in Fazakerley and Buchmeier (1993)]. Such viruses are of particular interest since the cause of a common demyelinating disease of humans, multiple sclerosis (MS), remains elusive, and involvement of viruses is suspected (Kurtzke, 1993). Despite several candidates, however, no single agent is currently accepted as a cause of MS (Waksman, 1995), but the idea that viruses can trigger MS as well as several other autoimmune diseases remains a viable hypothesis (Theofilopoulos, 1995). The many viruses associated with immune-mediated inflammatory diseases in the CNS of rodents include mouse coronavirus, SFV, and the picornavirus TMEV. In each instance the outcome of infection is markedly affected by virus strain, host genotype, and the dose and route of infection. As these variables apply to TMEV the topic has been comprehensively reviewed by Brahic *et al.* (1991), and for mouse coronavirus by Fazakerly and Buchmeier (1993). In all three viruses, immune-mediated disease appears associated only with persistent noncytolytic infection, but the actual immune mechanisms at play are still in dispute. In the case of TMEV, maybe the best-understood of these infections, overwhelming evidence points to a

crucial role for $CD4^+$ T cells of the Th1 subtype (Miller and Karpus, 1994). Such cells are assumed to organize inflammatory responses that consist largely of monocytic phagocytes which mediate bystander demyelination. Accordingly, monocytes were demonstrated to strip neurons of their myelin lamellae *in vitro* (Dal Canto and Lipton, 1975), but exactly how this is accomplished biochemically is unclear. Candidate mechanisms likely include oxyradical production, nitric oxide, and TNFα released from IFNγ-activated macrophages. Other cytokines such as IL-6 might also play a role, since overexpression of this cytokine as well as some others in the brain causes neuropathologic consequences (Campbell *et al.*, 1993). Evidence also exists for direct destruction of oligodendroglial cells, the cellular source of myelin, by the cytokine TNFβ generated by $CD4^+$ T cells of the Th1 subtype (Miller and Karpus, 1994).

In the TMEV system, compelling evidence indicates that $CD4^+$ T cells of the Th1 subtype act as the essential cell type that organizes the demyelination, at least in the susceptible SJL mouse (Miller and Karpus, 1994). Data indicate that disease severity correlates temporally with the development of MHC class II restricted DTH reactions and Th1-dependent IgG2a antibody production (Peterson *et al.*, 1992). In addition, disease development is suppressed by anti-CD4 or anti-Ia treatment (Gerety *et al.*, 1994). More persuasive evidence for a vital role of $CD4^+$ Th1 cells came from experiments which showed that mice exposed to a modified form of TMEV antigens, which resulted in selective Th1 anergy (tolerance), had delayed and diminished disease (Karpus *et al.*, 1995). In addition, such mice had reduced Th1 cytokine production but elevated levels of Th2 cytokines. The tolerized mice also had minimal DTH reactions, elevated Th2-dependent IgG1 antibody responses and reduced numbers of $CD4^+$ T cells in the CNS. In TMEV, although virus may persist in oligodendroglial cells, most virus appears to be present in brain macrophages (Clatch *et al.*, 1985). Such cells express high levels of MHC class II in inflamed brains, and these cells likely act as the principal activators of the $CD4^+$ virus-specific T cells (Miller *et al.*, 1995). Indeed, macrophages isolated from brains with TMEV demyelination readily stimulate $CD4^+$ TMEV antigen reactive lines *in vitro* (Miller, 1995). Moreover, the disease is potentiated if $CD4^+$ class II restricted Th1-type TMEV-specific clones are adoptively transferred to infected mice (Gerety *et al.*, 1994). In conclusion, current results in the TMEV system clearly point to an immunopathological disease involving principally $CD4^+$ T cells of the Th1 phenotype. These cells, upon recognition of antigens, drive a bystander demyelinating lesion mediated by cytokine-activated mononuclear phagocytes.

Our current understanding of the cellular and molecular pathogene-

sis of mouse coronavirus and SFV lags behind the TMEV system. Several observations are consistent with the notion that the demyelination occurring in the white matter in both diseases results from T-cell-mediated immune responses to persistent noncytolytic virus infection of glial cells [reviewed in Fazakerley and Buchmeier (1993)]. Although one suspects that the mechanisms at play in mouse coronavirus and SFV resemble those occurring in TMEV and primarily involve $CD4^+$ Th1 T-cell-mediated reactions, this idea has yet to be proven.

Whether or not any human demyelinating disease involves mechanisms akin to those discussed above for TMEV in the SJL mouse remains moot. Demyelination occurs occasionally as a sequel to measles infection and vaccination (Johnson *et al.,* 1978), and although an immune-mediated pathogenesis is suspected, viral antigens have not been demonstrated in the demyelinating lesions (Johnson *et al.,* 1984). However, measles virus might induce an autoimmune demyelination by a "hit and run" mechanism such as is suggested to occur in HSV-induced stromal keratitis, described subsequently. Postinfection demyelination also occurs very occasionally after infections with vaccinia, varicella, rubella, mumps, influenza, and EBV (Fazakerley and Buchmeier, 1993), but possible immune mechanisms at play have not been elucidated. There is considerable interest in the fact that HIV causes an inflammatory disease of the CNS (Spencer and Price, 1992). However, this occurs only in individuals who are markedly immunosuppressed and have very few circulating lymphocytes. The virus affects macrophages and these become MHC class II^+ (Kure, 1990). There is no evidence of T-cell involvement, but it may be that the pathogenesis involves overproduction of the cytokine TNFα from macrophages occurring because of the paucity of $CD4^+$ T cells, which normally produce inhibitors of macrophages such as IL-10 (Tyor *et al.,* 1995). However, demyelination certainly appears to have an immune pathogenesis in visna, an HIV-related virus infection of sheep (Narayan and Clements, 1989). Actually, the lesions of visna closely resemble those of TMEV-induced disease. The T cells in the inflammatory lesions appear to mediate their inflammatory effects by elaborating the ovine equivalent of gamma interferon (Narayan, 1989).

B. Virus-Induced Immunopathology Orchestrated Mainly by Type 2 Cytokine-Producing $CD4^+$ T Cells

Of the viruses which commonly cause disease in humans, RSV infection provides the best example of a disease that likely has a $CD4^+$ T-cell-controlled immune-mediated pathogenesis (Graham *et al.,* 1991). The virus persists in the body for a short time only and is minimally

cytopathic (Herman, 1990). When disease occurs, it usually manifests as the virus is being eliminated and involves many symptoms which mimic allergic reactions. Indeed, reports of the presence of IgE and eosinophil breakdown products in nasal secretions were consistent with an allergic pathogenesis (Welliver *et al.*, 1981), but such reports seem not to have been confirmed. Recently, several groups have studied RSV in mouse models and have shown a clear role for T cells as mediators of immunopathology (Alwan *et al.*, 1994; Graham *et al.*, 1991; Connors *et al.*, 1992). Although RSV replicates poorly in most mouse strains, a role for T cells in disease expression is well-established. For instance, lesions are minimal in immunosuppressed mice (Graham *et al.*, 1991), but become severe in immunocompetent virus-infected mice given adoptive transfers of RSV antigen-specific activated T cells (Alwan *et al.*, 1994). The most effective disease-producing cell transfers are $CD4^+$ T cells that express a Th2 cytokine profile (Alwan *et al.*, 1994). Interestingly, in the RSV system the different viral proteins appear to induce T cells of different cytokine-producing phenotypes (Alwan *et al.*, 1993). In fact, whereas the G protein of RSV induces the immunopathologic $CD4^+$ T cells with a type 2 cytokine profile, other proteins such as the F protein induce both $CD4^+$ and $CD8^+$ T cells, but such cells mediate protection rather than pathology (Anderson and Heilman, 1995). Moreover, these protective cells are largely IFNγ-producing and are considered type 1 cells. Exactly how the $CD4^+$ Th2 cells mediate the pathological immune reactions in the mouse lung is not known, but the prominence of eosinophils in the lesions indicates that such cells may participate in the tissue damage (Anderson and Herman, 1995).

With RSV infection in humans, it is well known that past efforts at vaccination led occasionally to augmented disease (Kapikian *et al.*, 1969). Using a mouse model which mimics this situation, Graham *et al.* (1993) have demonstrated that $CD4^+$ T cells of the type 2 cytokine-producing profile appear responsible for vaccine-augmented reactions. In addition, others showed that inhibition of type 2 cytokines with specific anticytokine antibodies eliminated the enhanced pulmonary pathology (Connors *et al.*, 1992). Taken together, the observations on murine RSV infections indicate a pathological role for type 2 cytokine-producing $CD4^+$ T cells. This makes RSV an unusual, possibly unique, viral agent whose pathogenesis mimics a pattern of events found more commonly in parasitic infections (Sher and Coffman, 1992). In the RSV system, it seems likely that shifting the immune response to a type 1 cytokine-producing cell dominance, as occurs by immunization with live virus (Anderson and Heilman, 1995), or perhaps better still by immunization with minimal vaccines that solely induce Th1 responses, would be a beneficial approach to prevent immunopathologic disease

following infection. Accordingly, in RSV infection, which commonly infects and causes repeated disease in children, the aim is to exploit Th1 memory maximally.

C. T-Cell-Mediated Immunopathogenesis in Cytolytic Virus Infections

The majority of viral infections that induce lesions with an immune-mediated component are persistent and minimally cytopathic. However, some highly cytodestructive viruses do induce immunopathic lesions at least in certain locations. Infection of the eye with HSV provides such an example (Doymaz and Rouse, 1992). Another might be rashes caused by viruses such as measles. Interestingly, measles virus rashes usually fail to occur in immunosuppressed patients (Enders, 1962), consistent with the notion that the lesions in immunocompetent individuals involve an immune reaction.

Ocular infection with HSV is one of the most common causes of vision impairment in the United States, with around 300,000 new cases of infection annually (Mader and Stulting, 1992). Lesions caused by HSV are usually confined to the corneal epithelium, but virus always enters the sensory nerve fibers that innervate sites of infection and passes to the trigeminal ganglion where a nonproductive infection (latency) occurs (Roizman and Sears, 1987). Periodically, virus reactivates and travels back to the cornea where replication occurs and an inflammatory reaction results in the underlying stroma. Repeated episodes of such recrudescence in the stroma ultimately result in opacity (Doymaz and Rouse, 1992). Stromal disease likely represents an immunopathological reaction set off by viral infection, since the disease responds to treatment with corticosteroids (Baum, 1995).

Most of our knowledge of the pathogenesis of herpetic stromal keratitis (HSK) has come from studies in animal models, particularly the mouse (reviewed in Doymaz and Rouse, 1992). Susceptible mouse strains routinely develop HSK after primary infection, but spontaneous reactivated disease is rare. The disease in mice is clearly immunopathological and primarily involves $CD4^+$ T cells of the Th1 subset (Niemialtowski and Rouse, 1992; Henricks *et al.*, 1992). Thus, HSK in the mouse represents a DTH reaction in the corneal stroma. Evidence that $CD4^+$ T cells orchestrate the inflammation has come from experiments showing that virus-infected athymic and SCID mice, or mice selectively depleted of $CD4^+$ T cells, fail to express HSK (Newell *et al.*, 1989; Mercadal *et al.*, 1993). Moreover, lymphocytes isolated from the inflamed corneas of mice with HSK are predominantly $CD4^+$ T cells, and these produce mainly type 1 cytokines except during disease re-

mission, when type 2 cytokines become prominent (Niemialtowski and Rouse, 1992; Babu *et al.*, 1995a). With HSK, although several lines of evidence point to an immunopathological reaction organized by $CD4^+$ T cells, the target antigens which drive the response remain unknown.

Interestingly, virus replication and the bulk of viral antigen expression occur in the corneal epithelium, whereas the inflammatory reaction occurs in the stroma (Mitchell *et al.*, 1994). Furthermore, evidence of viral gene and antigen expression is absent by the time the invasion by $CD4^+$ T cells and the recruited nonspecific inflammatory cells occurs (Babu *et al.*, 1995b). HSK can also be induced in SCID mice reconstituted with populations of $CD4^+$ T cells from HSV-naive donors, and recipient animals develop HSK before virus-specific immunity is detectable (Mercadal *et al.*, 1993). Such observations indicate that the HSV infection may cause the expression of some secondary agonist, such as a self-peptide derived from the immune-sequestered cornea, which drives the immune inflammatory response (Avery *et al.*, 1995). In consequence, HSK may represent an autoimmune inflammatory response set off by HSV which more or less acts as a "hit and run" agent.

Further evidence that HSK may represent an autoimmune inflammatory response was provided recently by the Foster group (Avery *et al.*, 1995). In their system the difference in HSK susceptibility between two congenic mouse strains is known to be controlled by an allotypic variation in an immunoglobulin gene (Jayaraman *et al.*, 1993). Mice expressing the IgH^b allele are resistant, whereas congenic animals expressing the IgH^d allele are sensitive. By inducing tolerance to IgH^b-expressing Ig in susceptible mice, HSK fails to develop. The absence of HSK was interpreted to mean that IgH^b-derived peptides provide tolerance to the target autoantigens recognized in the disease (Avery *et al.*, 1995). This concept was additionally supported by data showing that CD4 Th1-type clones specific to the peptide could transfer HSK to athymic recipients, just as could virus-immune T cells. The observations on HSK pathogenesis collectively indicate that HSV infection may be an example of a viral agent that can cause immune inflammatory disease by triggering an autoreactive response. Other examples and possible mechanisms of virus-induced autoimmune inflammatory responses are briefly discussed in the next section.

IV. Immune Inflammatory Responses Involving Antibody

Most examples of immunopathological responses to viruses involve T lymphocytes and the role of antibody in immunopathology is an almost neglected topic. However, there are at least two widely accepted

examples in which humoral mechanisms account for the immunopathogenesis of viral lesions. These are immune complex/complement-dependent lesions and antibody-dependent enhancement of viral infection. The latter phenomenon probably accounts for the pathogenesis of dengue hemorrhagic fever (DHF) (Kurane and Ennis, 1994) and coronavirus-induced infectious peritonitis, a common viral disease in the domestic cat (Trautwein, 1992). In DHF, which only occurs in persons with existing antibody at the time of infection, the syndrome is assumed to be a sequel to the facilitated infection of F_c receptor-bearing cells such as macrophages which take up virus–antibody complexes. The infected cells respond by producing an abundance of proinflammatory cytokines and stimulate $CD4^+$ and $CD8^+$ T cells to do the same (Kurane and Ennis, 1994). The resultant "cytokine storm" and other chemical mediators released are assumed to trigger the plasma leakage and hemorrhage in DHF. One group contends that vascular damage results from the effects of a novel cytokine produced by stimulated $CD4^+$ T cells, termed "cytotoxic factor" (Mukerjee and Chaturvedi, 1995).

Virus-induced immunopathology resulting from the entrapment in tissues of complement-activating immune complexes was first described for LCMV infection (Oldstone and Dixon, 1969). Immune complex disease results only if complexes are generated in excess as can only happen if the virus is not eliminated efficiently by the immune response. This might occur if agents replicate continuously in sites beyond effective access by protective T cells or if some protective component of the immune response is dysfunctional or exhausted because of overstimulation. Immune exhaustion due to overwhelming antigen exposure has been shown most convincingly to occur in transgenic systems, and involves $CD8^+$ T cells (Moskophidis *et al.,* 1993). In human viral disease, immune-complex-induced lesions have been observed, but only in the case of HBV infection has viral antigen been demonstrated to form part of the complexes (Chisari and Ferrari, 1995). However, immune-complex-mediated lesions in the joints and kidney are reasonably common in humans and it is possible that other viral agents may occasionally be involved.

V. Viruses and Autoimmunity

The idea that viruses might trigger autoimmune responses has been popular for some time, but there is little solid evidence to support the notion at least for any human autoimmune disease. The subject has received several recent reviews (Theofilopoulos, 1995; Sercarz *et al.,* 1993; Lehmann *et al.,* 1993; Lanzavecchia, 1995) and so only a few

points will be made. The oldest, simplest, and possibly most likely mechanism is that virus replication in an anatomically sequestered tissue releases autoantigens which activate self-reactive lymphocytes. This could be the principal mechanism at play in keratitis caused by HSV, since the avascular cornea can be considered unavailable to surveillance by the immune system. A similar explanation might apply to a common sequel to TMEV-induced demyelination in SJL mice wherein animals often develop a late response to myelin components such as PLP (Miller *et al.*, 1995).

A more sophisticated derivation of the antigen release concept is the unveiling of cryptic determinants. This idea, made popular recently by the excellent reviews of Sercarz (Sercarz *et al.*, 1993; Lehmann *et al.*, 1993), provides a more compelling explanation for virus-induced autoimmunity occurring in nonsequestered tissues. Examples might include Coxsackie virus-induced myocarditis (Huber and Lodge, 1984), and may explain why animals previously infected in the brain with persistent agents such as SFV become far more susceptible to the subsequent induction of Experimental Allergic Encephalomyelitis (EAE) when exposed to myelin antigens (Mokharian and Swoveland, 1987).

The essence of the cryptic self-hypothesis is that viral infection leads to the expression and altered presentation of determinants that are molecularly sequestered from the immune system and therefore do not induce tolerance (Lanzavecchia, 1995). The molecular unmasking and presentation could have a variety of causes. These include the possibility that viruses or induced cytokines, or even complexes between viruses and antibodies (Simitsek *et al.*, 1995), may modulate the expression or activity of proteases in APC which might result in the generation and presentation of previously cryptic peptides (Elson *et al.*, 1995). Another possibility is that the induction of abundant cytokines from infected or bystander cells might alter the surface expression of host proteins so that they become autoreactive. Some support for this idea comes from the observation of transgenic mouse models constructed to overexpress certain cytokines. Aberrant inflammatory reactions occur frequently (Campbell *et al.*, 1993; Gieger *et al.*, 1994). In one example, IFNγ overexpression in the mouse retina gave rise to retinitis, which interestingly became more intense following HSV infection of the eye (Gieger *et al.*, 1994). Such data are consistent with the notion that viruses may trigger autoimmunity in some instances by causing a cytokine storm. There is, however, no well-accepted example of this effect occurring under natural circumstances.

A further mechanism by which cryptic determinants become unveiled was reported by Salemi *et al.* (1995). It was shown that if $CD4^+$ T

cells are exposed to HIV gp120, $CD4^+$ is taken up more abundantly. In consequence, previously cryptic determinants on CD4 become exposed and these stimulate autoreactive T cells. Conceivably, such cells could account in part for the depletion of activated $CD4^+$ T cells in AIDS.

Another hypothesis used to explain how viruses might break tolerance and induce autoimmunity is the molecular mimicry hypothesis (Oldstone, 1989). This hypothesis has its enthusiasts, but currently a well-documented example of a natural autoimmune viral disease that results from infection by viruses that act as molecular mimics is lacking. The hypothesis states that viruses share determinants with self-tissues, and the effective immune response generated to the viral determinant spills over to the host and an autoreactive response occurs. There are numerous examples of shared peptide sequences between viral and host proteins. For example, several peptides derived from viral sequences were shown to stimulate T-cell clones derived from MS patients (Wucherpfennig and Strominger, 1995). These data were interpreted to support the notion that viruses are involved in the etiology of MS via a molecular mimicry mechanism. However, as discussed before, no single known virus is currently accepted as an initiating factor in MS (Waksman, 1995). For an enthusiastic viewpoint about molecular mimicry as it relates to virus-induced autoimmunity, the article of Wucherpfennig and Strominger (1995) is recommended. It is also worth noting that the molecular mimicry hypothesis has been advocated to explain aspects of the pathogenesis of HIV infection (Silvestris *et al.*, 1995).

Additional hypotheses have been advanced to explain how viruses might trigger autoimmunity. Included among them is the possibility that viral proteins which express superantigen activity might activate normally quiescent autoreactive clones of T cells (Scherer *et al.*, 1993). As with other hypotheses to explain virus-induced autoimmunity, widely accepted actual examples in natural viral diseases are not at hand.

VI. Control of Virus-Induced Immune Inflammatory Disease

The adage that an ounce of prevention is worth a pound of cure is certainly true in the field of viral pathogenesis. Preventing viral infection or manipulating immune processes during the initial phases of infection is far more successful than attempting to counteract pathological events once underway. With virus-induced immunopathologies, we are usually faced with a chronic tissue-damaging response to antigens that are being constantly replenished from a persistent replicat-

ing agent. The therapeutic challenge is either to remove or to neutralize the agonists which drive the inflammation or to redirect the symphony of events occurring so that tissue damage is minimized or ablated. Few viruses are subject to inhibition by drugs and some of them have strategies that hide them from the chemical attack. Herpes simplex virus provides the best example of this scenario: the virus is susceptible to several antiviral drugs during the replication phase but to none during latency.

Most virus-induced immunopathologies are orchestrated by T cells of one type or another. Such T cells usually recognize viral antigens although rarely is the antigen's identity known, particularly in outbred animals. However, one approach worth pursuing is to prevent specific antigen recognition by pathogenic T cells. The experimental induction of immunological tolerance to offending antigens is clearly the most desirable way to control any immune-mediated pathogenesis, but even when the culpable antigens are known, success is hard to achieve and maintain. Moreover, conceptually there are several forms of tolerance (Matzinger, 1994). These include (i) deletion of T cells specific for a particular MHC–peptide combination, (ii) induction of T cells that survive in a form that is hyporesponsive to antigen (anergy), and (iii) T-cell survival in a form that responds strongly to a particular stimulus but in way which differs from the standard response. This latter state, which is often termed "immune deviation," currently represents the most likely practical way to control viral immunopathology.

Immune deviation is an old concept originating from studies on DTH by Geoffery Asherson in the 1960s, which showed that exposure of guinea pigs to antigen by various routes selectively inhibited the DTH response (Asherson and Stone, 1965). Immune deviation is now better understood at a mechanistic level. It has its basis in the fact that subsets of T cells, both $CD4^+$ and $CD8^+$, exist which have different functional activities and that many of the cytokines they produce cross-regulate each subset (Paul and Seder, 1994; Coffman *et al.*, 1991; Croft *et al.*, 1994). Administration of antigen by various nonsystemic routes, e.g., may induce responses dominated by type 2 cytokine-producing cells which serve to down-regulate the induction of the type 1 cytokine producers that normally appear after systemic exposure (Ridgway *et al.*, 1994; Powrie and Coffman, 1993). Other means of achieving immune deviation include the use of analogue peptides for induction (Sette *et al.*, 1994) or the use of reagents which influence the microenvironment of antigen-activated T cells (Bluestone, 1995; Linsley, 1995). Regarding the latter, it is now evident that the cytokine or costimulator microenvironment in tissues during T-cell activation can

profoundly influence the outcome in terms of the functional set of T cells that differentiates (Linsley, 1995). For example, in an environment dominated by IL-4, the $CD4^+$ subset induced from uncommitted precursors is usually of the Th2 phenotype (Paul and Seder, 1994). Such an effect operating during induction of the immune response to TMEV would diminish the induction of lesion-inducing $CD4^+$ Th1 T cells. This result has, in fact, been reported (Karpus *et al.*, 1995). Thus, exposure of mice to viral antigen coupled to syngeneic spleen cells with ethylcarbodiimide (ECDI) abrogates the normal Th1 T-cell response and shifts the response to one dominated by Th2 T cells (Karpus *et al.*, 1995). This procedure not only prevents the demyelinating disease after subsequent TMEV infection but can suppress lesion severity in infected animals, at least if given not later than 2–3 weeks after infection.

Approaches also exist which favor the induction of Th1 cells, a scenario likely to be beneficial in minimizing the pathology associated with RSV infection. Th1-enhancing procedures include administering or promoting the production of the cytokine IL-12 (Manetti *et al.*, 1993) as well as manipulating the costimulator environment with agents which block CD28 stimulation, such as CTLA-Ig (Linsley, 1995). Recently, a surprising observation was reported which achieves a result similar to IL-12 potentiation. Administering the Schiff-base-forming molecule tucaresol along with antigen led to the accentuation of a $CD4^+$ Th1 response (Rhodes *et al.*, 1995). The mechanism of action is unknown but probably involves the bypass of a costimulator pathway which normally activates Th2-like responses (Shearer, 1995). There is some evidence that the costimulator B7-2 on APC is responsible for Th2 activation (Manetti *et al.*, 1993) and that tucaresol may inhibit the B7-2 stimulus in some way (Shearer, 1995).

All of the aforementioned approaches may achieve immune deviation, but they are usually successful only if used during the induction phase of an immune response. Reversing a given pattern of events by immunomodulators once fulminant lesions are present is a challenging problem. Possibly coming closest to this objective is the success being achieved using the oral tolerance approach to suppress certain experimental autoimmune diseases and perhaps even the human diseases multiple sclerosis and rheumatoid arthritis (Chen *et al.*, 1995). Thus, by feeding antigen, clinical disease expression is minimized. This effect works best using a low-dose antigen regimen which seemingly induces a bystander suppressor-type effect mediated by TGFβ and perhaps other cytokines (Friedman and Weiner, 1994). At higher oral tolerizing doses, the mechanism of tolerance induction appears to be T-cell deletion or anergy (Friedman and Weiner, 1994). This situation is probably

less desirable than a bystander suppressor effect, since in natural diseases the specific antigens involved are rarely if ever known with certainty. It remains to be seen if oral tolerance or any immune deviation scheme will successfully control an established virus-induced immunopathological lesion.

There are other immunomodulatory strategies that might succeed in arresting the advance of an inflammatory lesion and turn the tide to permit repair and recovery to occur. These include interfering with the effector function of lymphocytes and nonspecific inflammatory cells. Strategies include the use of cytokine receptor antagonists, particularly against the proinflammatory cytokines IL-1 and TNFα (Klein and Brailly, 1995). The approach has shown promise in certain model systems, but from a practical viewpoint it is not convenient since a continual administration of the antagonist is required. However, a successful formulation might come from combining proinflammatory cytokine arrest with the simultaneous use of agonists which achieve lymphocyte reeducation. As regards the latter, it is pertinent to note the intriguing observations of the Chisari laboratory showing that immunization of HBV transgenic mice with a DNA vaccine may curtail transgene expression (Chisari, 1995). In other viral immunopathologies, disease remission may be associated with the expression of the cytokine IL-10 (Babu *et al.*, 1995b; Tumpey *et al.*, 1994). It seems likely that the use of DNA vaccines, particularly those encoding regulatory cytokines, may prove useful to manipulate the immune inflammatory state. One report already attests to this possibility (Rogy *et al.*, 1995).

VII. Conclusions

The induction of an immune response which succeeds in eliminating virus-infected cells and extracellular virus is the common outcome of a viral infection. Removing infected cells usually engenders some tissue damage because of a concomitant inflammatory response, but this is not an unreasonable price to pay to control infection by highly cytolytic agents. If the agent is either noncytopathic or minimally so, destroying functionally intact cells may be, however, an undesirable consequence. This is especially true if cellular destruction is massive or occurs in an organ or tissue which is intolerant to damage, such as the cornea of the eye or the brain. Several viruses cause damage to the brain by immunopathological mechanisms, yet the same agents may cause insignificant lesions in other tissues either because function is retained or the damage is repaired rapidly. In Table I, a compilation is presented of

TABLE I

SOME EXAMPLES OF VIRUS-INDUCED IMMUNOPATHOLOGY

Likely principal mechanism involved	Virus	Reference[a]
CDB^+ T-cell-mediated	LCMV	Doherty *et al.* (1980)
	Hepatitis B	Chisari and Ferrari (1995)
	Coxsackie B	Hanke *et al.* (1995)
	HIV	Zinkernagel and Hengartner (1994)
	Sin Nobre Virus	Zaki *et al.* (1995)
$CD4^+$ T-cell-mediated	Theilers virus	Miller and Karpus (1994)
(type 1)	Mouse coronavirus	Fleming *et al.* (1993)
	Semliki Forest Virus	Mokharian and Swoveland (1987)
	Measles	Johnson *et al.* (1978)
	HSV	Doymaz and Rouse (1992)
	Visna	Narayan and Clements (1989)
(type 2)	Respiratory syncytial	Alwan *et al.* (1993)
Antibody-mediated	Dengue	Kurane and Ennis (1994)
	Feline infectious peritonitis	Trautwein (1992)

[a] These are good source references and are not meant to reflect the primary discoveries of the phenomenon.

some viral examples in which at least some of the lesions have an immunopathological pathogenesis.

The most common mechanism of lesion development in virus-induced immunopathology involves T cells. Usually, it seems that when $CD8^+$ T cells act as the controlling cell type, lesions are acute and the outcome is decided quickly. The classic example is provided by LCM in mice. The newest candidate may turn out to be SNV infection in humans. Lesions orchestrated primarily by $CD4^+$ T cells can be either acute or long-lasting. Curiously, in the LCMV example, if $CD8^+$ T cells are removed from the scene, immunopathological responses may still occur and these involve $CD4^+$ T cells (Doherty *et al.*, 1993; Fung-Leung *et al.*, 1991). Such responses are far more chronic and of lower grade than those mediated by $CD8^+$ T lymphocytes. One possible sequel to chronic inflammatory responses to viruses is that autoreactive inflammatory reactions are initiated and an autoimmune disease occurs. Many mechanisms by which viruses trigger autoimmunity have been conceived but all lack concrete examples, at least with respect to human autoimmune disease.

For some viral agents involved in immunopathological lesions, a clear picture of the cellular events and chemical mediators that participate in tissue damage is available. Rarely, however, is the biochemistry of the actual tissue damage fully understood. Conceivably, such knowledge will accrue from the ever-expanding array of *in vivo* models, particularly those which succeed in changing upon demand the expression of some molecular or cellular event. The practical bonus of such knowledge should be the generation of various approaches that will manage lesions and minimize their clinical significance. The challenge to practical viral immunology is to move from the secure territory of viral prophylaxis to the still alien field of lesion immunomodulation.

ACKNOWLEDGMENTS

I appreciate the valuable comments and insights provided by R. Ahmed, S. Miller, F. Chisari, M. Buchmeier, I. Campbell, P. Doherty, D. Griffin, O. Narayan, P. Openshaw, D. Slauson, and H. Zaghouani. I also thank R. Zinkernagel for access to a review still in preparation and to Drs. Campbell, Miller, and Chisari to cite unpublished material. My own research on viral immunopathology is supported by the National Eye Institute (EY 05093). Excellent typographical assistance was performed by Paula Keaton.

REFERENCES

Alwan, W. H., Record, F. M., and Openshaw, P. J. (1993). *J. Immunol.* **150**, 5211–5218.
Alwan, W. H., Koziowska, W. J., and Openshaw, P. J. M. (1994). *J. Exp. Med.* **179**, 81–89.
Anderson, L. J., and Heilman, C. A. (1995). *J. Infect. Dis.* **171**, 1–7.
Ando, K., Moriyama, T., Guidotti, L. G., Worth, S., Schreiber, R. D., Schlicht, H. J., Huang, S. N., and Chisari, F. V. (1993). *J. Exp. Med.* **178**, 1541–1554.
Ando, K., Guidotti, L. G., Wirth, S., Ishikawa, T., Missale, G., Moriyama, T., Schreiber, R. D., Schlicht, H. J., Huang, S., and Chisari, F. V. (1994a). *J. Immunol.* **152**, 3245–3253.
Ando, K., Guidotti, L. G., Cerny, A., Ishikawa, T., and Chisari, F. V. (1994b). *J. Immunol.* **153**, 482–488.
Asherson, G. L., and Stone, S. H. (1965). *Immunology* **9**, 205–217.
Avery, J., Zhae, Z.-S., Rodriquez, A., Bikett, E. K., Soheillan, M., Foster C. S., and Cantor, H. (1995). *Nature* **376**, 431–433.
Babu, J. S., Kanangat, S., and Rouse, B. T. (1995a). *J. Immunol.* **154**, 4822–4829.
Babu, J., Thomas, J., Kanangat, S., Morrison, L. M., Knipe D., and Rouse, B. T. (1995b). *J. Virol.*, in press.
Baenziger, J., Hengartner, H., Zinkernagel, R. M., and Cole, G. A. (1986). *Eur. J. Immunol.* **16**, 387–393.
Baum, J. (1995). *Clin. Infect. Dis.* **21**, 479–488.
Bluestone, J. A. (1995). *Immunity* **2**, 555–559.
Brahic, M., Bureau, J., and McAllister, A. (1991). *Microb. Pathol.* **11**, 77–84.
Buchmeier, M. J., Welsh, R. M., Dutko, F. J., and Oldstone, M. B. A. (1980). *Adv. Immunol.* **36**, 275–331.

Burnet, F. M., and Fenner, J. (1949). "Production of Antibody." Monographs of Walter and Eliza Hall Institute. McMillan, Melbourne.
Campbell, I. L., Abraham, C. R., Masliah, E., Kemper, P., Inglis, J. D., Oldstone, M. A. B., and Mucke, L. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10061–10065.
Campbell, I. L., Samini, A., and Chiang, C.-S. (1994a). *J. Immunol.* **153**, 3622–3625.
Campbell, I. L., Hobbs, M. V., Kemper, P., and Oldstone, M. B. A. (1994b). *J. Immunol.* **152**, 716–723.
Campbell, I. L., Samini, A., and Chiang, C.-S. (1995). *J. Neurochem.* **64**, S32.
Campbell, I. L. (1995). Personal communication.
Chen, Y., Inobe, J.-I., Marks, R., Gonnella, P., Kuchroo, V. K., and Weiner, H. L. (1995). *Nature* **376**, 177–180.
Chisari, F. S. (1995). Personal communication.
Chisari, F. V., and Ferrari, C. (1995). *Annu. Rev. Immunol.* **13**, 29–60.
Chow, L. H., Gauntt, C. J., and McManus, B. M. (1991). *Lab. Invest.* **64**, 55–64.
Clatch, R. J., Melvold, R. W., Miller, S. D., and Lipton, H. L. (1985). *J. Immunol.* **135**, 1408–1414.
Coffman, R. L., Varkilask, S. P., and Chatelain, R. (1991). *Immunol. Revs.* **123**, 323–359.
Connors, M., Kulkarni, A. B., Firestone, C.-Y., Holmes, K. L., Morse III, H. C., Sotnikov, A. V., and Murphy, B. R. (1992). *J. Virol.* **66**, 7444–7451.
Croft, M., Carter, L., Swain, S. L., and Dutton, R. W. (1994). *J. Exp. Med.* **180**, 1715–1728.
Dal Canto, M. C., and Lipton, H. L. (1975). *Lab Invest.* **33**, 626–637.
Doherty, P. C., and Zinkernagel, R. M. (1974). *Transplant. Revs.* **19**, 89–120.
Doherty, P. C., Allan, J. E., Lynch, F., and Ceredig, R. (1990). *Immunol. Today* **11**, 55.
Doherty, P. C., Hou, S., and Southern, P. J. (1993). *J. Neuroimmunol.* **46**, 11–18.
Doymaz, M., and Rouse, B. T. (1992). *Curr. Top. Microbiol. Immunol.* **179**, 121–136.
Eliott, L. H., Ksiazek, T. G., Rollin, P. E., Spiropoulou, C. F., Morzunov, S., Monroe, M., Goldsmith, C. S., Humphrey, C. D., Zaki, S. R., Krebs, J. W., Maupin, G., Gage, K., Childs, J. E., Nichol, S. T., and Peters, C. J. (1994). *Am. J. Trop. Med. Hyg.* **51**, 102–108.
Elson, C. J., Barker, R. N., Thompson, S. J., and Williams, N. A. (1995). *Immunol. Today* **16**, 71–75.
Enders, J. F. (1962). *Am. J. Dis. Child.* **103**, 282–287.
Fazakerley, J. K., and Buchmeier, M. J. (1993). *Adv. Virus Res.* **42**, 249–323.
Fleming, J. O., Wang, F. I., Trousdale, M. D., Hinton, D. R., and Stohlman, S. A. (1993). *Reg. Immunol.* **5**, 37–43.
Friedman, A., and Weiner, H. L. (1994). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6688–6692.
Fung-Leung, W. P., Kundig, T. M., Zinkernagel, R. M., and Mak, T. W. (1991). *J. Exp. Med.* **174**, 1425–1429.
Gerety, S. J., Rundell, M. K., Dal Canto, M. C., and Miller, S. D. (1994). *Eur. J. Immunol.* **23**, 46–55.
Gieger, K., Howes, E. L., and Sarvetnick, N. (1994). *J. Virol.* **68**, 5556–5557.
Graham, B. S., Bunton, L. A., Wright, P. F., and Karzon, D. T. (1991). *J. Clin. Invest.* **88**, 1026–1033.
Graham, B. S., Henderson, G. S., Tang, Y.-W., Lu, X., Neuzil, K. M., and Colley, D. G. (1993). *J. Immunol.* **151**, 2032–2040.
Guidotti, L. G., Matzke, B., Schaller, H., and Chisari, F. V. (1995). *J. Virol.*, in press.
Hanke, A., Huber, S., Stelzner, A., and Whitton, J. L. (1995). *J. Virol.* **69**, 6720–6728.
Heilman, C. A. (1990). *J. Infect. Dis.* **161**, 402–406.
Henricks, R. L., Tumpey, T. M., and Finnegan, A. (1992). *J. Immunol.* **149**, 3023–3030.
Hotchin, J. (1971). *Monogr. Virol.* **3**, 1–211.

Huber, S. A., and Lodge, P. A. (1984). *Am. J. Pathol.* **116**, 21–29.
Jacobs, R. P., and Cole, G. A. (1976) *J. Immunol.* **117**, 1004–1009.
Jayaraman, S., Heiligenhaus, A., Rodriguez, A., Soukiasian, S., Dorf, M. E., and Foster, C. S. (1993). *J. Immunol.* **151**, 5777.
Johnson, K. P., Wolinsky, J. S., and Ginsberg, A. H. (1978). *In* "Handbook of Clinical Neurology" (P. J. Vinken and G. W. Bruyn, eds.), pp. 391–424. Elsevier/North Holland, Amsterdam.
Johnson, R. T., Griffin, D. E., Hirsch, R. L., Wolinsky, J. S., Roedenbeck, S., de Soriano, I. L., and Vaisberg, A. (1984). *New Engl. J. Med.* **310**, 137–141.
Kagi, D., Ladermann, B., Burki, K., Zinkernagel, R. M., and Hengartner, H. (1995). *Immunol. Revs.* **146**, 94–115.
Kapikian, A. Z., Mitchell, R. H., Chanock, R. M., Shredoff, R. A., and Stewart, C. E. (1969). *Am. J. Epidemiol.* **89**, 405–421.
Karpus, W. J., Pope, J. G., Peterson, J. D., Dal Canto, M. C., and Miller, S. D. (1995). *J. Immunol.* **155**, 947–957.
Klein, B., and Brailly, H. (1995). *Immunol. Today* **16**, 216–219.
Kolaitis, G., Doymaz, M., and Rouse, B. T. (1990). *Immunology* **71**, 101–106.
Kurane, I., and Ennis, F. A. (1994). *Semin. Virol.* **5**, 443–448.
Kure, K., Weidenheim, K. M., Lyman, W. D., and Dickson, D. W. (1990). *Acta Neuropathol.* **80**, 393–400.
Kurtzke, J. F. (1993). *Clin. Microbiol. Rev.* **6**, 382–427.
Lane, J. R., Neumann, D. A., Lanford-Walker, A., Herskowitz, A., and Rose, N. R. (1992). *J. Exp. Med.* **175**, 1123–1129.
Lanzavecchia, A. (1995). *J. Exp. Med.* **181**, 1945–1948.
Laskin, D. L., and Pendino, K. J. (1995). *Annu. Rev. Pharmacol.* **35**, 655–677.
Lehmann, P. V., Sercarz, E. E., Forsthuber, T., Dayan, C. M., and Gammon, G. (1993). *Immunol. Today* **14**, 203–207.
Leist, T. P., Cobbold, S. P., Waldman, H., and Zinkernagel, R. M. (1987). *J. Immunol.* **138**, 2278–2281.
Linsley, P. S. (1995). *J. Exp. Med.* **182**, 289–292.
Mader, T. H., and Stulting, R. D. (1992). Infect. Dis. North Am. **6**, 831–849.
Manetti, R., Parronchi, P., Giudizi, M. G., Piccinni, M.-P., Maggi, E., Trinchieri, G., and Romagnani, S. (1993). *J. Exp. Med.* **177**, 1199–1204.
Matzinger, P. (1994). *Annu. Rev. Immunol.* **12**, 991–1045.
Meltzer, M. S., and Nacy, C. A. (1989). *In* "Delayed Hypersensitivity and the Induction of Activated Cytotoxic Macrophages in Fundamental Immunology" (W. E. Paul, ed.), 2nd Ed., pp. 765–780. Raven Press, New York.
Mercadal, C., Bouley, D., DeStephano, D., and Rouse, B. T. (1993). *J. Virol.* **67**, 3404–3408.
Miller, S. D. (1995). Personal communication.
Miller, S. D., and Karpus, W. J. (1994). *Immunol. Today* **15**, 356–361.
Miller, S. D., McCrae, B. L., Vanderlugt, C. L., Nikievich, K. M., Pope, J. C., Pope, L., and Karpus, W. J. (1995). *Immunol. Revs.* **144**, 225–244.
Mims, C. A., and Wainwright, S. (1968). *J. Immunol.* **101**, 717–724.
Mitchell, W. J., Gressens, P., Martin, J. L., and DeSanto, R. (1994). *J. Gen. Virol.* **75**, 1201–1210.
Mokharian, P., and Swoveland, P. (1987). *J. Immunol.* **138**, 3264–3268.
Moskophidis, D., Lechner, F., Pircher, H. P., and Zinkernagel, R. M. (1993). *Nature* **262**, 758–761.
Mosmann, T. R., and Coffman, R. L. (1989). *Annu. Rev. Immunol.* **7**, 145–173.

Mukerjee, R., and Chaturvedi, U. C. (1995). *Clin. Exp. Immunol.* **102**, 496–500.
Narayan, O. (1989). *Curr. Opin. Immunol.* **2**, 399–402.
Narayan, O., and Clements, J. E. (1989). *J. Gen. Virol.* **70**, 1617–1639.
Newell, C. R., Martin, S., and Rouse, B. T. (1989). *Regional Immunol.* **2**, 1–8.
Nichol, S. T., Spiropoulou, C. F., Morzunov, S., Rolin, P. E., Ksiazek, T. G., Feldmann, H., Sanchez, A., Childs, I., Zaki, S., and Peters, C. J. (1993). *Science* **262**, 914–917.
Niemialtowski, M., and Rouse, B. T. (1992). *J. Immunol.* **148**, 1864–1870.
Odermatt, B., Eppler, M., Leist, T. P., Hengartner, H., and Zinkernagel, R. M. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8252–8256.
Oldstone, M. B. A. (1975). *Prog. Med. Virol.* **19**, 84–119.
Oldstone, M. B. A. (1989). *Curr. Top. Microbiol. Immunol.* **145**, 127–135.
Oldstone, M. B. A., and Dixon, F. (1969). *J. Exp. Med.* **129**, 483–505.
Pantaleo, G., Graziosi, C., and Fauci, A. S. (1993). *New Engl. J. Med.* **328**, 327–335.
Paul, W. E., and Seder, R. A. (1994). *Cell* **76**, 241–251.
Peterson, J. D., Waltenbaugh, C., and Miller, S. D. (1992). *J. Immunol.* **75**, 652–658.
Powrie, F., and Coffman, R. L. (1993). *Immunol. Today* **14**, 270–274.
Rhodes, J., Chen, H., Hall, S. R., Beesley, J. E., Jenkins, D. C., Collins, P., and Zheng, B. (1995). *Nature* **377**, 71–75.
Ridgway, W. M., Weiner, H. L., and Fathman, C. G. (1994). *Curr. Opin. Immunol.* **6**, 946–955.
Rogy, M. A., Auffenberg, T., Espat, N. J., Philip, R., Remick, D., Wollenberg, G. K., Copeland, E. M., and Moldawer, L. L. (1995). *J. Exp. Med.* **181**, 2289–2293.
Roizman, B., and Sears, A. E. (1987). *Annu. Rev. Microbiol.* **41**, 543–571.
Salemi, S., Caporossi, A. P., Boffa, L., Longobardi, M. G., and Barnaba, V. (1995). *J. Exp. Med.* **181**, 2253–2257.
Sandberg, K., Eloranta, M. L., and Campbell, I. L. (1994). *J. Virol.* **68**, 7358–7366.
Scherer, M. T., Ignatowicz, L., Winslow, G. M., Kappler, J. W., and Marrack, P. (1993). *Annu. Rev. Cell. Biol.* **9**, 101–128.
Sercarz, E. E., Lehmann, P. V., Ametani, A., Benichou, G., Miller, A., and Moudgil, K. (1993). *Annu. Rev. Immunol.* **11**, 729–766.
Sette, A., Alexander, J., Ruppert, J., Snoke, K., Franco, A., Ishioka, G., and Grey, H. M. (1994). *Annu. Rev. Immunol.* **12**, 413–431.
Shearer, G. M. (1995). *Nature* **377**, 69–70.
Sher, A., and Coffman, R. L. (1992). *Ann. Rev. Immunol.* **10**, 385–409.
Silvestris, F., Williams, Jr., R. C., and Dammacco, F. (1995). *Clin. Immunopathol.* **75**, 197–205.
Simitsek, P. D., Campbell, D. G., Lanzavecchia, A., Fairweather, N., and Watts, C. (1995). *J. Exp. Med.* **181**, 1957–1964.
Spencer, D. C., and Price, R. W. (1992). *Annu. Rev. Microbiol.* **46**, 655–693.
Theofilopoulos, A. N. (1995). *Immunol. Today* **16**, 90–97.
Trautwein, G. (1992). *Vet. Microbiol.* **33**, 19–34.
Tumpey, T. M., Elner, V. M., Chen, S. H., Oakes, J. E., and Lausch, R. N. (1994). *J. Immunol.* **153**, 2258–2265.
Tyor, W. R., Wesselinge, S. L., Griffin, J. W., McArthur, J. C., and Griffin, D. E. (1995). *J. Acquire. Immun. Def. Syndr.* **9**, 379–388.
Waksman, G. H. (1995). *Nature* **377**, 105–106.
Walker, D. H., Murphy, F. A., Whitfield, S. G., and Bauer, S. P. (1975). *Exp. Mol. Pathol.* **23**, 245–265.
Welliver, R. C., Wong, D. T., Sun, M., Middleton, E., Vaughan, R. S., and Ogra, P. L. (1981). *New Engl. J. Med.* **305**, 841–846.

Woodruff, J. F. (1980). *Am. J. Pathol.* **101**, 425–484.
Wucherpfennig, K. W., and Strominger, J. L. (1995). *Cell* **80**, 695–705.
Zaki, S. R., Greer, P. W., Coffield, L. M., Goldsmith, C. S., Nolte, K. B., Foucar, K., Feddersen, R. M., Zumwalt, R. E., Miller, G. L., Khan, A. S., Rollin, P. E., Ksiazek, T. G., Nichol, S. T., Mahy, B. W. J., and Peters, C. J. (1995). *Am. J. Pathol.* **146**, 552–579.
Zinkernagel, R. M., and Hengartner, H. (1994). *Immunol. Today* **15**, 262–268.

ADVANCES IN VIRUS RESEARCH, VOL. 47

GEOGRAPHIC AND MOLECULAR EPIDEMIOLOGY OF PRIMATE T LYMPHOTROPIC RETROVIRUSES: HTLV-I, HTLV-II, STLV-I, STLV-PP, AND PTLV-L

Antoine Gessain and Guy de Thé

Unité d'Epidémiologie des Virus Oncogènes
Département du SIDA et des Rétrovirus
Institut Pasteur
75724 Paris Cedex 15, France

I. Introduction and Historical Background

Since the pioneering work of Ellerman and Bang and of P. Rous in rats at the beginning of this century and of Bittner and L. Gross in mice, animal oncoretroviruses have been instrumental as a model for gaining fundamental knowledge on the epidemiological and molecular determinants of carcinogenesis. However, the situation in humans seemed different, and it was only after about 70 years of intense but

unsuccessful and often very disappointing research that the first human oncoretrovirus was isolated and characterized in 1980 by the R. C. Gallo team (Poiesz *et al.*, 1980, 1981). In the past 15 years, everything moved much more quickly, and several new primate (including human) retroviruses belonging to either the oncovirus or lentiviruses family were discovered.

In this review, we will present the main geographic and molecular epidemiological characteristics of primate T lymphotropic (onco)retroviruses (PTLVs) exhibiting common features, including a characteristic and peculiar microepidemiology with a puzzling distribution throughout the world, similar modes of transmission, mainly through breast feeding and sexual contact, and a high genetic stability over time, which can be used as a molecular means to follow migrations of infected populations in the recent or distant past.

A. HTLV-I

The first human oncoretrovirus was isolated in 1980 in Dr. Gallo's laboratory from cell cultures from an Afro-American patient with a lymphoproliferative disease originally considered to be a cutaneous T cell lymphoma but later characterized as an adult T cell leukemia. This virus was named human T cell leukemia/lymphoma virus (HTLV) (Poiesz *et al.*, 1980, 1981). Quite independently, in Japan, Takatsuki *et al.* had described in 1977 a T cell lymphoproliferation labeled adult T cell leukemia/lymphoma (ATL) (Takatsuki *et al.*, 1977). The epidemiological features suggested a strong environmental factor which prompted experimental researchers to characterize the tumor cells (Miyoshi *et al.*, 1981) and to search for an oncogenic virus that they isolated in 1981 and named adult T cell leukemia/lymphoma virus (ATLV) (Hinuma *et al.*, 1981; Yoshida *et al.*, 1982). Japanese and American scientists rapidly showed that both isolates referred to the same virus, and agreed to name it HTLV-I (Popovic *et al.*, 1982).

In parallel, the causal association between ATL and HTLV-I was established. In 1983 we initiated studies in the French West Indies to investigate the epidemiological and clinical impact of HTLV-I in this area. This led us to discover the association between this virus and an encephaloneuromyelopathy endemic in the Caribbean, originally named tropical spastic paraparesis (TSP; Gessain *et al.*, 1985; Gessain and Gout, 1992). A similar entity was then recognized in Japan and labeled HTLV-I-associated myelopathy (HAM; Osame *et al.*, 1986, 1987). These two conditions are considered to be the same clinicovirological entity, and this myelopathy is now referred to as TSP/HAM.

The question then arose as to whether the same virus could induce, through different pathways, two different diseases, as in the case of the Epstein-Barr virus (de Thé, 1982), or, as in the case of murine leukemia viruses, would specific mutations in certain structural viral genes control tissue tropism and direct pathogenesis (see below). Since HTLV-I [in sharp contrast to HIV-1 and HIV-2 (Boeri *et al.*, 1992; Goodenow *et al.*, 1989; Wain-Hobson, 1992)] exhibits high genomic stability, it was possible to try to answer that question and to see whether the virus was identical in different parts of the world.

B. HTLV-II

In 1982, a second human lymphotropic retrovirus, labeled HTLV-II, was isolated in Dr. Gallo's laboratory (Kalyanaraman *et al.*, 1982) from a cell line derived from the splenic cells of a patient with a lymphoproliferative disease originally considered to be a "T variant of hairy cell leukemia." Highly endemic in disparate tribes of Amerindian populations and intravenous drug users of the Western World, the pathogenicity of this virus remains unclear despite its presence in some rare cases of CD8 lymphoproliferative diseases and neuromyelopathies (Hall *et al.*, 1994a; Fouchard *et al.*, 1995).

C. STLV-I and STLV-II

Isolated in 1982 by Dr. Miyoshi and colleagues (Miyoshi *et al.*, 1983), simian T cell leukemia/lymphoma virus type I (STLV-I), which is highly prevalent in a large variety of Old World monkey species, causes an ATL-like disease, but only in a few species such as African green monkeys, baboons, and gorillas, while no neurological diseases similar to TSP/HAM have been reported to date in such animals. An STLV-II has been isolated from a monkey originating from South America (Chen *et al.*, 1994); however, this observation needs to be confirmed.

D. PTLV-L and STLV-PP

Two new simian retroviruses have been recently isolated from a *Papio hamadryas* from Ethiopia (PTLV-L; Goubau *et al.*, 1994; Van Brussel *et al.*, 1996) and from captive colonies of *Pan paniscus*: STLV-PP, STLV-PP1664 (Liu *et al.*, 1994a; Giri *et al.*, 1994; Vandamme *et al.*, 1996), respectively.

II. Epidemiological Characteristics of the Different PTLVs

A. HTLV-I

HTLV-I is not a ubiquitous virus but is present throughout the world with clusters of high endemicity often located near areas where the virus is nearly absent (Mueller, 1991; Blattner, 1990; Kaplan and Khabbaz, 1993; Blattner and Gallo, 1994; Gessain, 1996). These highly endemic areas are localized in the southwestern part of the Japanese archipelago (mainly the islands of Okinawa, Kyushu, and Shikoku), the Caribbean area and its surounding regions, foci in South America including Colombia, French Guyana, some parts of Brazil, in intertropical Africa (Gabon, Zaire), in the Middle East (the Mashad region in Iran), and in isolated clusters in Melanesia. The origin of this puzzling geographic or, rather, ethnic distribution (with such a microepidemiology) is not well understood, but is probably linked to a founder effect in certain ethnic groups, followed by the persistence of high viral transmission due to favorable environmental and cultural local situations. The apparent progressive decrease of HTLV-I seroprevalence in southern Japan seems to be due to a slow decrease of the viral transmission in the past 20 to 50 years, in turn related to important modifications in the health care system, nutrition, and socioeconomic factors, including the diminution of the length of breast feeding (Oguma *et al.*, 1992).

In every high endemic area, and despite very different socioeconomic and cultural environments, HTLV-I seroprevalence (as defined by specific serum HTLV-I antibodies) increases gradually with age, especially in women between 20 to 40 years of age, this increase possibly being due to either increased sexual exposure with age, to a cohort effect, or to a reactivation of a latent infection in relation to the immunodysregulation occurring with aging. The global population infected by HTLV-I is estimated to be around 15 to 20 million worldwide, with about 2 to 10% developing an HTLV-I associated disease during their lifetime (ATL, TSP/HAM, uveitis) (de Thé and Bomford, 1993).

Three modes of transmission have been demonstrated for HTLV-I. Mother-to-child transmission (Tajima *et al.*, 1990; Takahashi *et al.*, 1991; Katamine *et al.*, 1994) is mainly linked to the persistence of breast feeding after 4 to 6 months of age. This transmission is due to the presence of HTLV-I provirus in milk mononuclear cells, and the seroconversion occurs at 18 to 24 months of age. Around 15 to 20% of breast-fed children of HTLV-I seropositive mothers will be infected and become HTLV-I seropositive carriers. The *in utero* transplacental transmission of HTLV-I seems very low, its reality still being a matter

of debate. Sexual transmission which occurs mainly from male to female is thought to be responsible for the increased seroprevalence in women (Takezaki *et al.,* 1995). Transmission by a contaminated blood product is responsible for seroconversion with the acquisition of HTLV-I infection in around 15 to 60% of the blood recipients (Sandler *et al.,* 1991; Kaplan and Khabbaz, 1993).

B. HTLV-II

HTLV-II is known to be highly endemic among some disparate native New World Amerindian tribes, including the Navajo and Pueblo in New Mexico (Hjelle *et al.,* 1993), the Seminole in Florida (Levine *et al.,* 1993), the Guayami in Panama (Lairmore *et al.,* 1990; Pardi *et al.,* 1993b), the Kayapo and Kraho in Brazil (Maloney *et al.,* 1992; Black *et al.,* 1994; Ishak *et al.,* 1995), the Wayu (Ijichi *et al.,* 1993; Duenas-Barajas *et al.,* 1993) and Orinoco (Fujiyama *et al.,* 1993) in Colombia, and the Tobas and Matacos in north Argentina (Ferrer *et al.,* 1993; Biglione *et al.,* 1993). In these populations, HTLV-II seroprevalence varies greatly, but can reach 20% of the general adult population and up to 50% in women of 50 years or older, as described in some Kraho groups (Maloney *et al.,* 1992; Ishak *et al.,* 1995).

In the Western World HTLV-II takes an endemo-epidemic course in intravenous drug users in the United States (Lee *et al.,* 1989, 1993), in some South and Central American countries, and, to a lesser extent, in Europe (Italy and Spain) (De Rossi *et al.,* 1991; Zella *et al.,* 1993). While breast feeding appears predominant in the developing world (Heinene *et al.,* 1992; Tuppin *et al.,* 1996), transmission occurs mainly through sharing contaminated needles (Lee *et al.,* 1989; Hall *et al.,* 1992b, 1994a) and blood transfusion in the occidental world. Furthermore, some studies are consistent with the hypothesis that heterosexual transmission may play a significant role in the spread of HTLV-II in both environments (Tuppin *et al.,* 1996).

C. STLV-I

Seroepidemiological studies have detected the presence of STLV-I antibodies in a large variety of Old World monkey species and apes including chimpanzees (*Pan troglodytes*), gorillas (*Gorilla gorilla*), grivet monkeys (*Cercopithecus aethiops aethiops*), baboons (*Papio anubis*), cynomolgus or crab-eating macaques (*Macaca fascicularis*), pig-tailed macaques (*Macaca nemestrina*), stump-tailed macaques (*Macaca arctoides*), rhesus macaques (*Macaca mulatta*), bonnet ma-

caques (*Macaca radiata*), lion-tailed macaques (*Macaca silenus*), toque monkeys (*Macaca sinica*), and Celebes macaques (*Macaca nigrescens, M. nigra, M. hecki, M. tonkeana, M. maura, M. ochreata, and M. brunnescens*) (Miyoshi *et al.*, 1983; Hayami *et al.*, 1984; Ishida *et al.*, 1983; Ishikawa *et al.*, 1987; Ibrahim *et al.*, 1995). In contrast, New World monkey species and prosimians lack antibodies to STLV-I (Kaplan *et al.*, 1993; Ibrahim *et al.*, 1995; and unpublished data). In the few Old World monkey large colonies studied, STLV-I seroprevalence increases with age and is more elevated in the female (Ibrahim *et al.*, 1995), and, as observed in humans, the transmission of STLV-I seems to occur mainly from mother to offspring through breast feeding and from male to female by sexual contact. These observations support the view of the great similarities, if not identities, of simian and human oncoretrovirus characteristics.

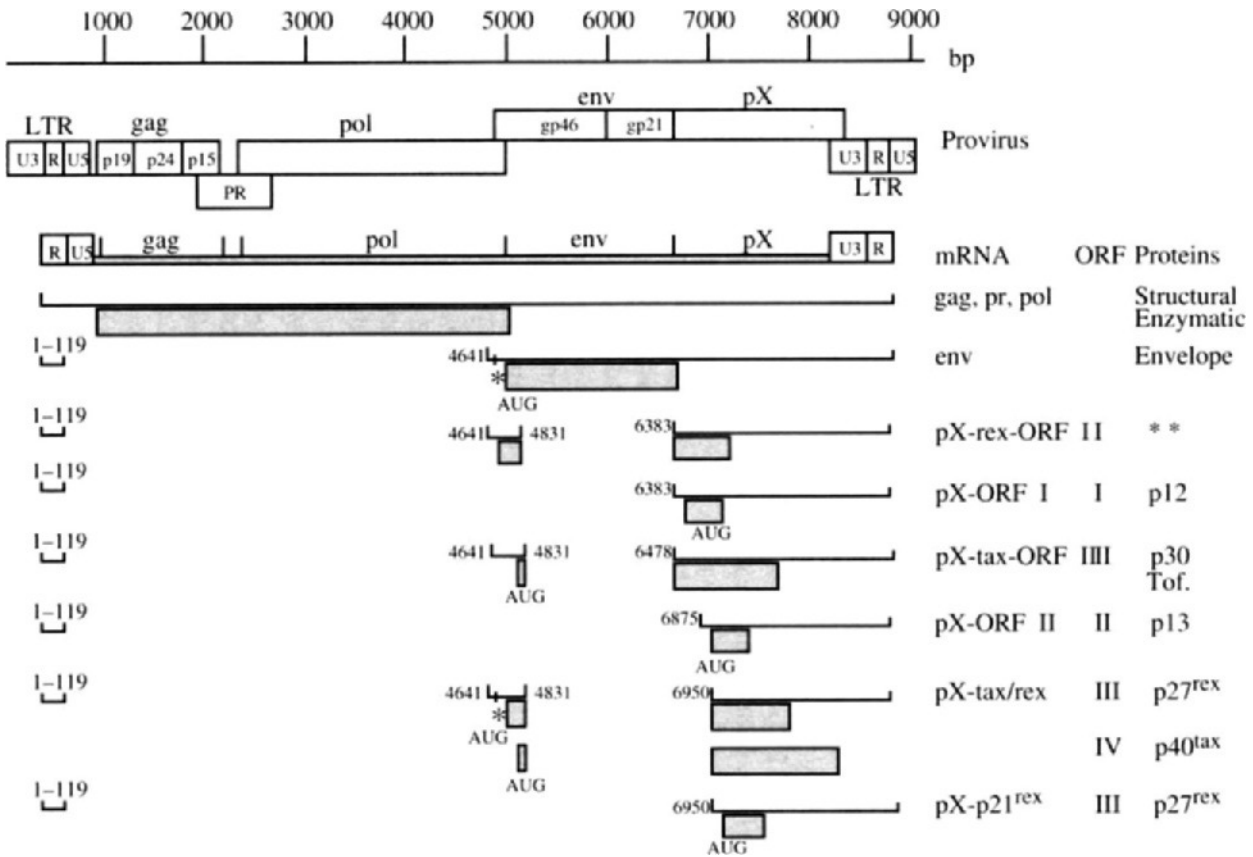
D. PTLV-L and STLV-PP

No data are available on the transmission of PTLV-L, isolated from a *Papio hamadryas* kept in captivity in Leuven, Belgium (Goubau *et al.*, 1994). Concerning STLV-PP, a familial study in the United States on a captive colony from which the virus was isolated suggested transmission from mother to offspring and from male to female (Giri *et al.*, 1994).

III. Genetic Structure of HTLV-I/II

HTLV-I/II viruses with simian T lymphotropic retroviruses thus form the group of primates T lymphotropic (PTLV) exogenous (onco)-retroviruses sharing the same overall genetic organization (Cann *et al.*, 1990). In addition to *gag, pol,* and *env* coding regions and to the LTRs containing major functional sequences, they possess inserted between the *env* gene and the 3′ LTR (see Fig. 1) a "pX" region which encodes regulatory proteins named, for HTLV-I/II, *tax* (p40/I, p37/II, a viral transcriptional transactivator), *rex* (p27/I, p26/II, a posttranscriptional regulator of viral expression), and a polypetide (p21I/II) of yet unknown function (Cann *et al.*, 1990; Koralnik, 1996). The existence of four new alternatively spliced mRNAs, encoded by the ORF I and ORF

FIG. 1. Genomic structure of HTLV-I provirus and mRNAs with corresponding encoded proteins.


1000
2000
3000
4000
5000
6000
7000
8000
9000
bp
LTR
gag
pol
env
pX
gp46
gp21
U3
R
U5
p19
p24
p15
PR
Provirus
mRNA
ORF
Proteins
gag, pr, pol
Structural
Enzymatic
env
Envelope
AUG
pX-rex-ORF II
pX-ORF I
pX-tax-ORF III
pX-ORF II
pX-tax/rex
pX-p21rex
p12
p30
Tof.
p13
p27rex
p40tax
1–119
4641
4831
6383
6478
6875
6950

II of the pX region for HTLV-I has been discovered (Berneman *et al.*, 1992; Ciminale *et al.*, 1992; Koralnik, 1996; Koralnik *et al.*, 1992) (see Fig. 1). Yet their products are ill-defined, with unknown functions except for the P12 (Franchini *et al.*, 1993; Franchini, 1995). Studies are ongoing for similar proteins in other members of the PTLVs. Preliminary data seem to indicate that the HTLV-II pX region also encodes for some new proteins (Ciminale *et al.*, 1994, 1996), but that P12 is not present in some STLV-I (Ibrahim *et al.*, 1995).

The first complete sequence of HTLV-I, from a Japanese ATL patient (labeled ATK), was published by Yoshida's team in 1983 (Seiki *et al.*, 1983) and has since been considered to be the prototype HTLV-I sequence. However, one year earlier, the same group had sequenced the LTR from another Japanese ATL named ATM (Seiki *et al.*, 1982). This was followed by the sequencing of the LTR of American ATL isolates (Josephs *et al.*, 1984). Sequences from TSP/HAM isolates were published as early as 1988 (Imamura *et al.*, 1988; Tsujimoto *et al.*, 1988). In these early data, great homology in sequence with the original ATK prototype was noted, with minor nucleotide divergence ranging from 0 to 3%, according to the genomic region considered and the origin of the isolate. Ratner *et al.* (1985, 1991) published the sequence of the first HTLV-I molecular variant from an ATL patient originating from Zaire, which exhibited an overall 3.3% divergence compared to the ATK prototype.

Thus, to date, only eight complete nucleotide sequences of HTLV-I have been published, four originating from ATL patients, namely ATK (Seiki *et al.*, 1983), HS35 (Malik *et al.*, 1988), EL (Ratner *et al.*, 1991), and YS (Chou *et al.*, 1995), two originating from TSP/HAM, namely TSP1 (Evangelista *et al.*, 1990) and BOI (Bazarbachi *et al.*, 1996), and two from healthy seropositive individuals, namely CH (Ratner *et al.*, 1991) and MEL5 (Gessain *et al.*, 1993a)

Concerning HTLV-II, the first sequence (MO), considered the prototype, was published in 1985 (Shimotono *et al.*, 1985), and, since then, only three other complete sequences have been reported: G12 from a healthy Guaymi Amerindian from Panama (Pardi *et al.*, 1993b), NRA from an American patient with a concomitant CD8 oligoclonal lymphoproliferation and a B hairy cell leukemia (Lee *et al.*, 1993), and from an Italian IVDU (Salemi *et al.*, 1996).

Despite the fact that partial sequences of an STLV-I genome (Ptm3 from Indonesia) have been available since 1984 (Watanabe *et al.*, 1985, 1986), only one complete nucleotide sequence of an STLV-I has just been reported by our group (Ibrahim *et al.*, 1995). For PTLV-L and STLV-PP, only partial sequences of the pX are available to date (Giri *et al.*, 1994; Goubau *et al.*, 1994; Liu *et al.*, 1994a).

IV. Methods Used in Molecular Epidemiology of PTLVs

Most of the work performed on the molecular epidemiology of PTLVs has focused on the analysis of fragments of the *env* gene, especially in the region coding for the TM transmembrane gp 21 protein (Ehrlich *et al.,* 1992; Gessain *et al.,* 1991, 1992b, 1994a; Mahieux *et al.,* 1994), the noncoding LTR region (Bangham *et al.,* 1988; Daenke *et al.,* 1990; Kinoshita *et al.,* 1991; Komurian *et al.,* 1991, 1992; Saksena *et al.,* 1992; Miura *et al.,* 1994; Ureta Vidal *et al.,* 1994a, 1994b; Vandamme *et al.,* 1994; Switzer *et al.,* 1995a,b), and a small fragment of the *pol* gene (Dube *et al.,* 1993, 1994, 1995; Poiesz *et al.,* 1993). Fewer data concern the *gag,* most of the *pol,* or pX regions (Gray *et al.,* 1987; Kwok *et al.,* 1988; Komurian *et al.,* 1991; Nerurkar *et al.,* 1993a, 1994b; Song *et al.,* 1995; Vandamme *et al.,* 1995). HTLV-I, HTLV-II, and STLV-I sequences established before 1988 were obtained using classic molecular cloning in phages, followed by subcloning in pBR322 vectors (Seiki *et al.,* 1983; Malik *et al.,* 1988). The use of the polymerase chain reaction (PCR) since 1988 has led to rapid progress in molecular epidemiology due to the simplicity, efficiency, and rapidity of this method. Furthermore, PCR represents a powerful tool for studying directly the *ex vivo* genomic variability with the posssibility of detecting a rare event in the genomic DNA of an infected individual, and this, without cell culture, eliminates the possibility of an *in vitro* viral selection, as demonstrated in the HIV/SIV system (Meyerhans *et al.,* 1989). PCR is also efficient for searching for the presence of different viral strains within an individual at a given time (intrastrain variability, quasispecies), allowing the sequencing of multiple clones from a single amplified product (Daenke *et al.,* 1990; Ehrlich *et al.,* 1992; Gessain *et al.,* 1992b; Nerurkar *et al.,* 1993b). In the HIV system, the existence of quasispecies has been established using this technique (Goodenow *et al.,* 1989; Meyerhans *et al.,* 1989).

The direct sequencing of amplified DNA without cloning has enabled the dominant HTLV-I species from each sample to be rapidly sequenced, thus representing a procedure well adapted to molecular epidemiology, especially when the aim is to search for geographic variations of *ex vivo* HTLV-I or HTLV-II sequences (Komurian *et al.,* 1991; Major *et al.,* 1993).

Another technique well adapted to molecular epidemiology because of its simplicity and rapidity is the study of the restriction fragment length polymorphism (RFLP) of amplified PCR fragments. This technique has been applied in our laboratory to the study of the HTLV-I LTR and has demonstrated the existence of at least four major geo-

graphic HTLV-I subtypes (Komurian *et al.*, 1992; Ureta Vidal *et al.*, 1994a,b) (Table I). Similarly, it led to the discovery of two major subtypes of HTLV-II (Hall *et al.*, 1992a; Switzer *et al.*, 1995a).

Several different techniques can be applied to perform comparative sequence analysis, including phylogenetic studies. The most frequently used methods are the neighbor-joining (NJ) method, the DNA maximum parsimony, the maximum likelihood technique, and the UPGMA methods (Vandamme *et al.*, 1994, 1995; Song *et al.*, 1994). The new, modified NJ method (Nei *et al.*, 1995; Saitou and Nei, 1987) and the weighted parsimony method are incisive methods for analyzing closely related sequences (around 2 to 15% nucleotide sequence differences) as encountered with HTLVs. The use of a bootstrap analysis is important to test the robustness of the obtained trees (Vandamme *et al.*, 1994, 1995; Ibrahim *et al.*, 1995).

TABLE I

RESULTS OF RFLP ANALYSIS OF 180 HTLV-I PROVIRAL DNA SAMPLES FROM SEVERAL GEOGRAPHIC ORIGINS SHOWING 12 RESTRICTION PROFILES CLASSIFIED INTO FOUR MAJOR SUBTYPES[a,b,c]

HTLV-I LTR RFLPs					Previous Subtypes[a]	Number of Specimens	Proposed Subtypes
Apa I	Nde I	Mae III	Dra I	Sac I			
+	–	+	–	+		17	African a
–	–	+	–	+	I	6	African b
+	+	+	–	+		1	African c
+	–	+	+	+		1	African d
+	+	–	–	+	II	52	Cosmopolitan a
+	+	–	+	–	(II)	13	Cosmopolitan b
+	+	–	–	–		5	Cosmopolitan c
+	+	–	+	+		2	Cosmopolitan d
–	+	–	–	+		4	Cosmopolitan e
–	+	–	+	–		1	Cosmopolitan f
+	+	+	+	–	III	77	Japanese
+	–	–	–	+		1	Melanesian
						180	

[a] Gray shading indicates features consistent in a given subtype. Adapted from Ureta Vidal *et al.* (1994b).

[b] + indicates the presence of a restriction enzyme site and – its absence.

[c] According to the initial classification of Komurian *et al.* (1992).

V. HTLV-I

A. Do Tissue-Specific Genomic Mutations Exist for HTLV-I?

The question as to whether specific diseases were linked to specific sequences of HTLV-I arose when it was recognized that HTLV-I was associated with at least two life-threatening diseases of totally different characteristics, namely ATL (Takatsuki *et al.,* 1977) and TSP/HAM (Gessain *et al.,* 1985; Osame *et al.,* 1986, 1987; Gout *et al.,* 1990; Gessain and Gout, 1992). In the experimental murine leukemia viruses system (Moloney, Friend, and BRE strains) hematological or neurological degenerative disorders are specifically associated with viral mutations in the *env* and/or LTR sequences (Lenz *et al.,* 1984; Rassart *et al.,* 1986; Li *et al.,* 1987; Szurek *et al.,* 1988). Early British (Daenke *et al.,* 1990) and Japanese (Kinoshita *et al.,* 1991) studies indicated that there were no obvious differences in the HTLV-I *env* and LTR sequences in isolates from the peripheral blood mononuclear cells of either ATL or TSP/HAM patients. In addition to these early studies involving only patients from one country (Jamaica or Japan), we compared proviral DNA from four leukemic and five TSP/HAM patients from Japan, Ivory Coast, and the French West Indies (Komurian *et al.,* 1991). The sequencing of 1918 bp covering the U3/R region of the LTR, most of SU gp46 *env,* and pX regions of these nine *ex vivo* lymphocyte DNAs showed 1.7% variation in *env,* 2.7% in *tax,* and 6.2% in the LTR. While no mutation could be specifically linked to either hematological or neurological processes, a suggestion was made that geographic subtypes of HTLV-I might exist (Komurian *et al.,* 1991). Several other studies from different groups similarly failed to detect any specific sequences related to the outcome of a specific disease (Schulz *et al.,* 1991; Gessain *et al.,* 1991, 1992b; Paine *et al.,* 1991; Ratner *et al.,* 1991; Major *et al.,* 1993; Ureta Vidal *et al.,* 1994a,b). The recent claim that a specific *tax* mutation was associated with TSP/HAM (Renjifo *et al.,* 1995) remains to be confirmed and may be due to a bias in the sampling of the specimens studied (most of them came from a Colombian area with a high TSP/HAM prevalence). In fact, this mutation seems to be present only in the HTLV-I strains of the Cosmopolitan geographic subtype but not related to any specific clinical outcome. In our data, nearly all the ATL and healthy carriers infected with the Cosmopolitan HTLV-I strain also possess this "so-called *tax*-specific TSP/HAM mutation" (Mahieux *et al.,* 1995a).

Such negative results concerning tissue tropic sequences of HTLV-I do not preclude the possibility of yet undetected differences in HTLV-I

virions homing to either hematological or neurological tissues. Thus, comparison of HTLV-I strains present in the cerebrospinal fluid (CSF) lymphocytes or possibly in not yet definitely discovered HTLV-I infected cells from the central nervous system would be instrumental for the search of a possible neurological variant. This would predict the existence of an HTLV-I quasispecies in patients, and, indeed, evidence for minor intrastrain variants has been obtained from patients (see below). Recently the Bangham group (Niewiesk *et al.*, 1994, 1995) has shown that the *tax* gene is more variable within and between healthy carriers than in patients with TSP/HAM. They suggest that the *tax* sequence heterogeneity, rather than the presence of a particular sequence, distinguishes healthy seropositive carriers from TSP/HAM patients. The functional relevance of such a finding, which would indicate that the selection for amino acid conservation in *tax* is stronger for patients with TSP/HAM, needs further studies.

In contrast, Saito *et al.* (1995) recently repeated the existence of frequent mutations in the pX region of HTLV-I in TSP/HAM compared to ATL or healthy seropositives but found no differences between PBNCs and CSF in a single patient. This also contrasts with the findings of Kira *et al.* (1994), who described sequence heterogeneity of HTLV-I from viral DNA in the central nervous system of TSP/HAM patients.

Ongoing studies using transgenic mice transfected with LTR or *tax* constructs from either ATL or TSP/HAM are necessary to investigate the possible functional relevance of the subtle nucleotide differences observed in certain viral strains originating from ATL, TSP/HAM, or arthopathies (Gonzalez-Dunia *et al.*, 1992; Gessain *et al.*, 1996a).

Even the study of an HTLV-I full-length sequence obtained from a patient a few weeks after the development of TSP/HAM following a blood transfusion did not reveal any specific mutations in favor of a neurotropic variant (Bazarbachi *et al.*, 1995).

B. Geographic Molecular HTLV-I Subtypes

Malik and co-workers (1988), by comparing their ATL isolate HS35 from Jamaica with the available sequences, first suggested the possible existence of a closer genetic homology between HTLV-I isolates from the same geographic area than among isolates from the same diseases but obtained from different geographic areas. This theory was not shared by Gray *et al.* (1990) based on data from the *env* gene of only three isolates. Because of these inconsistencies and because few sequence data are available on isolates from Africa, the West Indies, and South America, we and several other groups began, in 1988–1990, studies on a large variety of specimens from different geographic areas

(De *et al.*, 1991; Dekaban *et al.*, 1992; Fukasawa *et al.*, 1987; Gessain *et al.*, 1991, 1992a,b, 1993a; Komurian *et al.*, 1991, 1992; Kwok *et al.*, 1988; Paine *et al.*, 1991; Ratner *et al.*, 1991; Schulz *et al.*, 1991; Shirabe *et al.*, 1990; Sherman *et al.*, 1992, 1993, Major *et al.*, 1993; Mahieux *et al.*, 1994; Miura *et al.*, 1994; Ureta Vidal *et al.*, 1994a,b; Vandamme *et al.*, 1994; Yanagihara, 1994). The HTLV-I molecular clusters emerged from the accumulated data which demonstrate, without ambiguity, that the nucleotide changes observed in some regions of the HTLV-I genome are specific for the geographic origin of the patients (Table II), but not for the associated pathologies. The initial studies suggested the presence of a distinct molecular cluster in patients from Ivory Coast and of a more distant one in those from Zaire (Ratner *et al.*, 1991; Boeri *et al.*, 1992; Gessain *et al.*, 1992a,b). Furthermore, the HTLV-I sequences from West African countries (Mauritania and Guinea-Bissau)

TABLE II

GEOGRAPHIC GENOTYPES OF HTLV-I

		% Nucleotide divergence[a]	
	Location	LTR (U3/R)[b]	env (gp21)[c]
Cosmopolitan	West and South Africa Americas Caribbean Iran Japan	0–1	0–1
Japanese	Japan India	1–2	0.5–1.5
West African	Ivory Coast Mali Burkina Faso Caribbean	1.5–2	1–2
Central African[d]	Zaire Gabon Cameroon Central African Republic	3.5–4.5	2–4
Australo-Melanesian	Papua New Guinea Solomon Islands Australia	6–7	6.5–7.5

[a] As compared to Cosmopolitan strains.

[b] The region compared corresponds to 413 base pairs in the U3:R region (22–434).

[c] The region compared corresponds to 522 base pairs in the gp21 env region (6047–6567).

[d] A new molecular phylogenetic HTLV-I subtype (D) has been reported recently in some inhabitants of Central Africa, especially in Pygmies (Mahieux *et al.*, 1995b).

and some from Ivory Coast and the Central African Republic were virtually identical to those observed in the French West Indies, Haiti, French Guyana, and Peru, strongly suggesting that HTLV-I was introduced into the New World through viral carriers during the slave trade (Gessain *et al.*, 1992b). Another study of 10 new HTLV-I strains from the Americas confirms a common origin with a relatively recent introduction of the virus in the Americas (Gessain *et al.*, 1994a). These HTLV-I strains, very conserved among themselves and widespread throughout the world, were defined as the Cosmopolitan subtype (Gessain *et al.*, 1991). An interesting recent phylogenetic report describes in detail the possible origin of such HTLV-I strains found in Amerindians from British Columbia (Canada).

Recent data obtained from our laboratory (Mahieux *et al.*, 1995b and unpublished results) as well as from other groups (Vandamme *et al.*, 1994; Liu *et al.*, 1994b; Gasmi *et al.*, 1994; Dube *et al.*, 1994; Yamashita *et al.*, 1995; Moynet *et al.*, 1995) indicate a diversity of HTLV-I molecular variants in Africa, again exhibiting a clear clustering according to the geographic origin of the specimens, the most distant variant being present in central African countries (Zaire, Gabon, Cameroon, and Central African Republic). Furthermore, we have recently discovered, in Central Africa, molecular variants (particularly present in Pygmies) constituting a new phylogenetic molecular subtype of HTLV-I strains (Mahieux *et al.*, 1995b).

The genomic stability of HTLV-I isolates being very high (Table II), the most variable region found was the LTR, mostly the U3 region (up to 5.2% nucleotide variability with the exception of the very divergent Melanesian strains; see the following section), and the most conserved regions being the sequences encoding for the *tax* and *rex* regulatory proteins. Thus, nucleotidic divergence for the *tax* region ranged from 0 to 2.6% (mean 0.9%) and for the *rex* region from 0 to 1.4% (mean 0.4%). On the amino acid level the variability was about 1% for both these proteins. Recent data indicate that the region located between the *env* gene and the splice acceptor for *tax* in the pX region exhibits greater nucleotide diversity (Ratner *et al.*, 1991; Gessain *et al.*, 1993a). This region of 900 bp, initially considered a noncoding region (NCR), has recently been shown, as mentioned already, to encode in part for new proteins (Berneman *et al.*, 1992; Ciminale *et al.*, 1992; Koralnik *et al.*, 1992).

Besides the sequencing of critical regions of the HTLV-I genome, a procedure which could be carried out on only a few samples, we used the RFLP technique applied to the U3-R/LTR (the most variable region) to test a larger series of specimens. In an early series (Komurian *et al.*, 1992) involving 30 proviral HTLV-I from Japan, the Carribean,

and West and Central Africa, three subtypes emerged: an African, a Cosmopolitan, and a Japanese subtype. More recently, we enlarged the study to 180 specimens covering ATL, TSP/HAM patients, and asymptomatic carriers from different parts of the world (Table I). As seen in Figs. 2 and 3, the results confirmed the existence of geographic subtypes, adding a fourth one from the South Pacific (Ureta Vidal *et al.*, 1994a,b). Furthermore, we validated the RFLP LTR approach by establishing that subtypes thus obtained correlated perfectly with mutations observed by sequencing LTR and other parts of the proviral genomes (Figs. 2 and 3).

Assuming a similar evolutionary rate for the different HTLV-I strains in the different geographic locations, such geographic clusters of HTLV-I variants (Figs. 2 and 4) should reflect in part the genetic drift of *in vivo* evolution during thousands of years of specific viral strains in remote and isolated human populations.

1. Existence of Distantly Related HTLV-I in Melanesia

Since 1987, a series of seroepidemiological studies have shown in remote populations of Melanesia the presence of a high level of HTLV-I ELISA seroreactivity, not positively confirmed by Western blot (WB). Such seroindeterminate sera exhibit antibodies mainly to "gag proteins," but lack any *env* reactivity (Yanagihara *et al.*, 1991a,b, 1994). This suggests the existence of either a variant of HTLV-I or of an immunologically related virus. Investigating HTLV-I DNA specimens obtained by Dr. Gajdusek's team (Yanagihara *et al.*, 1991b) from Papua New Guinea (PNG) and the Solomon Islands, Gessain *et al.* (1991) focused on a region of 522 bp (encompassing most of the TM gp21 and a small portion of the SU gp46 *env* gene) in the DNA extracted from either *ex vivo* uncultured or cultured PBMCs of six Melanesian Papua New Guinean and Solomon Islanders. Unexpectedly, the isolates from healthy Melanesian carriers and from a Solomon Islander with TSP/HAM diverged by 7 to 8% from the ATK HTLV-I. Furthermore, the HTLV-I variants from Papua New Guinea differed by nearly 3–4% from the Solomon Islands strain, indicating the existence of an HTLV-I intervariability (quasispecies) within this geographic region. By contrast, HTLV-I strains from two residents of Bellona Island, a Polynesian outlier within the Solomon Islands, were closely related to the Cosmopolitan group (>97% homology in the gp21), suggesting a more recent introduction of HTLV-I in this area (Figs. 4 and 5). It is worthwhile to note that most of the nucleotide changes identified in these Melanesian HTLV-I proviruses corresponded primarily to single base substitutions within a given codon resulting in no amino acid changes in 85% of

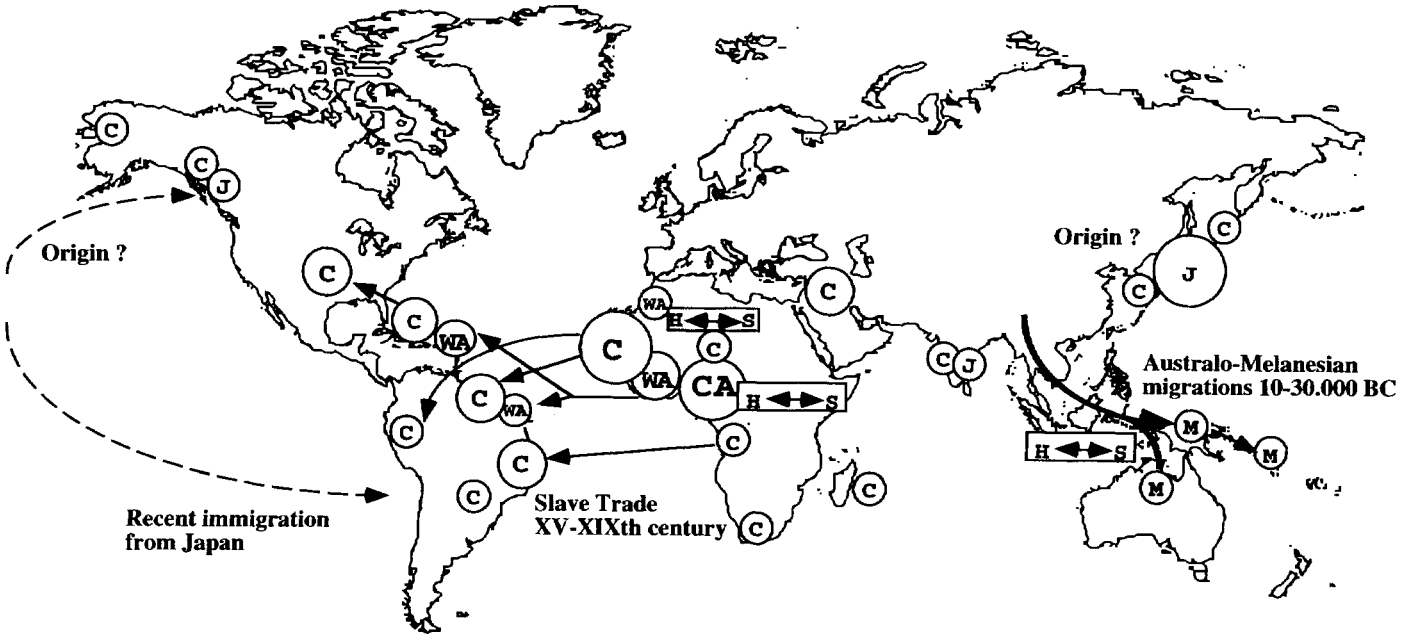
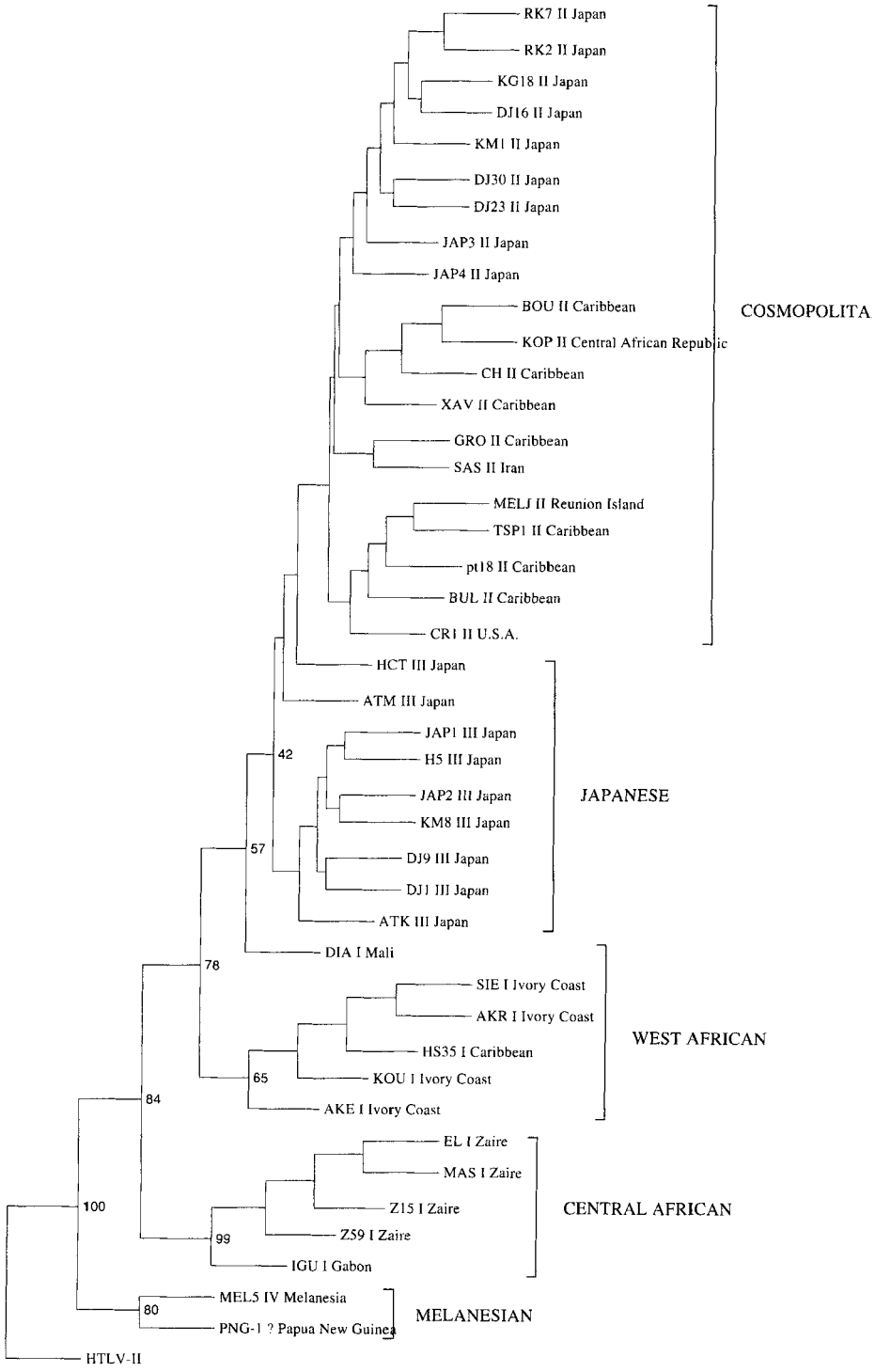


FIG. 2. Geographic localization of the different HTLV-I molecular subtypes based on RFLP and/or sequence analysis of the LTR of 270 samples originating from most of the HTLV-I endemic areas. The size of the circles represents the number of specimens studied for a given subtype: C (Cosmopolitan subtype), J (Japanese subtype), WA (West African subtype), CA (Central African subtype), and M (Melanesian subtype). Three main suspected interspecies transmissions of STLV-I from simian to human (double arrows) led to recognizable phylogenetic classes that persist in modern species.


RK7 II Japan
RK2 II Japan
KG18 II Japan
DJ16 II Japan
KM1 II Japan
DJ30 II Japan
DJ23 II Japan
JAP3 II Japan
JAP4 II Japan
BOU II Caribbean
KOP II Central African Republic
CH II Caribbean
XAV II Caribbean
GRO II Caribbean
SAS II Iran
MELJ II Reunion Island
TSP1 II Caribbean
pt18 II Caribbean
BUL II Caribbean
CR1 II U.S.A.
COSMOPOLITA
HCT III Japan
ATM III Japan
JAP1 III Japan
H5 III Japan
JAP2 III Japan
KM8 III Japan
DJ9 III Japan
DJ1 III Japan
ATK III Japan
JAPANESE
42
57
DIA I Mali
78
SIE I Ivory Coast
AKR I Ivory Coast
HS35 I Caribbean
KOU I Ivory Coast
AKE I Ivory Coast
WEST AFRICAN
65
84
EL I Zaire
MAS I Zaire
Z15 I Zaire
Z59 I Zaire
IGU I Gabon
CENTRAL AFRICAN
100
99
MEL5 IV Melanesia
PNG-1 ? Papua New Guinea
MELANESIAN
80
HTLV-II

cases. These data suggest that a high genetic pressure maintains the integrity of the HTLV-I transmembrane *env* protein structure. The complete nucleotide sequence of one of these HTLV-I variants, Mel 5 isolated from a Solomon Islander, showed an overall nucleotide divergence from the prototype ATK of 8.5%: the LTR with 9.5% of overall divergence, the U5 (12.5%), and the U3 (10%) were more variable than the R region (6.5%) (Gessain *et al.*, 1993a). The corresponding degree of variability at the amino acid level of the structural genes ranged from 3% (p24) to 11% (p19), being unexpectedly higher (8.5–29%) for both the regulatory (*tax* and *rex*) and the new pX encoded proteins. In these variants, the conservation of the *env* neutralizing epitope (88–98 AA) within the SU gp46 (external envelope glycoprotein) suggests that a vaccine prepared against the Cosmopolitan prototype of HTLV-I might also protect against these HTLV-I variants (Gessain *et al.*, 1993a; Benson *et al.*, 1994).

In parallel, Sherman *et al.* (1992) and Saksena *et al.* (1992) have sequenced parts of PCR amplified regions of LTR, *pol*, and pX genes of one of these HTLV-I variants isolated from a Papua New Guinean (mel1/PNG1). Similar to our results, they showed a marked heterogeneity with an overall divergence of 7–8% from the ATK prototype. Based on dendrograms of the available sequences of the LTR and *env* genes, these authors suggested that the HTLV-I PNG1 diverged from the other PTLVs, probably after the emergence of the progenitors of STLV-I from Asia, but before that of the ancestors from Africa thousands of years ago. Nerurkar *et al.* (1993b) also studied the interfamilial and intrafamilial genomic diversity of such variant strains.

Furthermore, HTLV-I was recently isolated from Australian Aboriginals (Bastian *et al.*, 1993). Sequence analysis of the *env* and segments of the *pol* and pX regions of an ATL isolate (MSHR-1) confirmed the presence in these isolated populations of an HTLV-I variant exhibiting a sequence divergence from the ATK prototype of about 5 to 7%, depending on the genes studied, but being closely related to Melanasian isolates (Fig. 4). Oligonucleotide primers specific for such Australo-

FIG. 3. Phylogenetic tree obtained by the maximum parsimony method after sequence alignment and bootstrapping. The compared sequences correspond to a 315-bp fragment of the LTR region encompassing most of the U3 and R regions. The number indicated at some nodes represents their frequency of occurrence out of 100 trees and therefore a measure of the robustness of the proposed tree. The length of the branches is not proportional to nucleotide subtitutions. The topology of the tree demonstrates the existence of five molecular clusters of HTLV-I genotypes related to the geographic origin of the specimens studied and not the disease presented by the patients (ATL, TSP/HAM). Adapted from Ureta Vidal *et al.* (1994b).

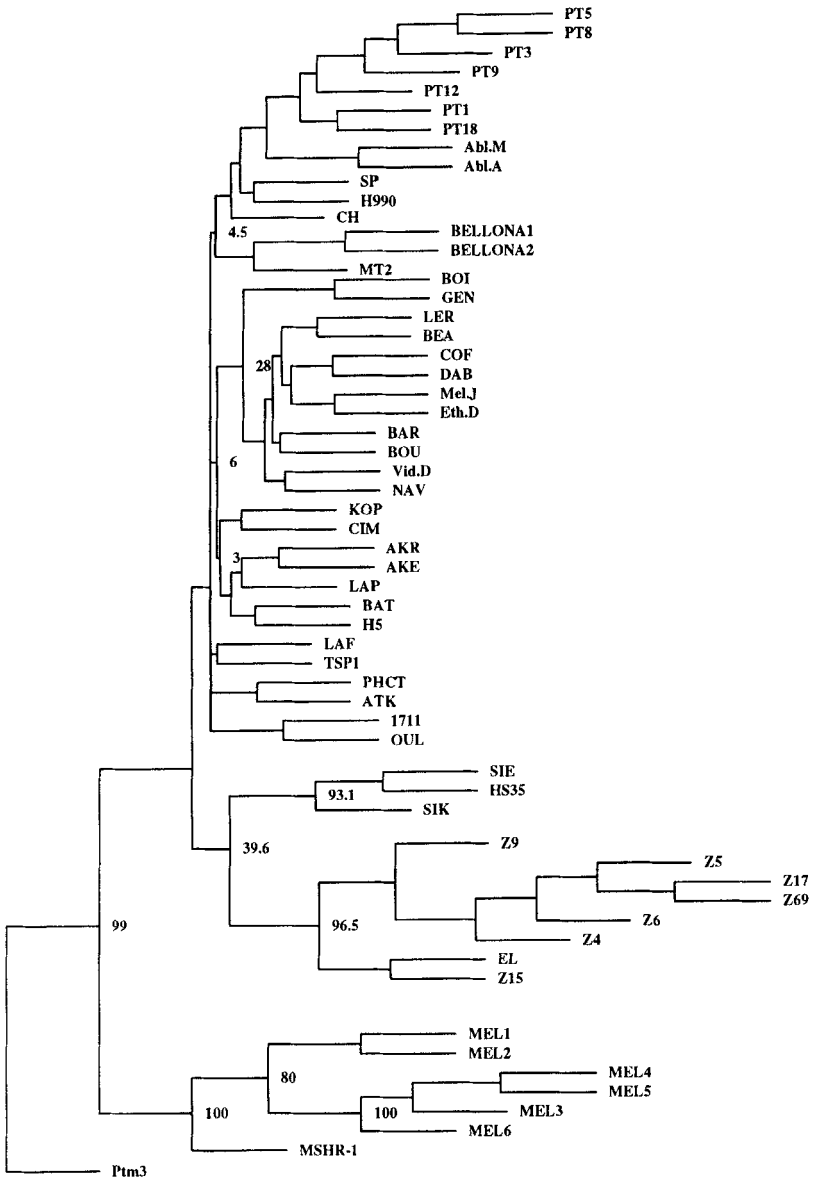
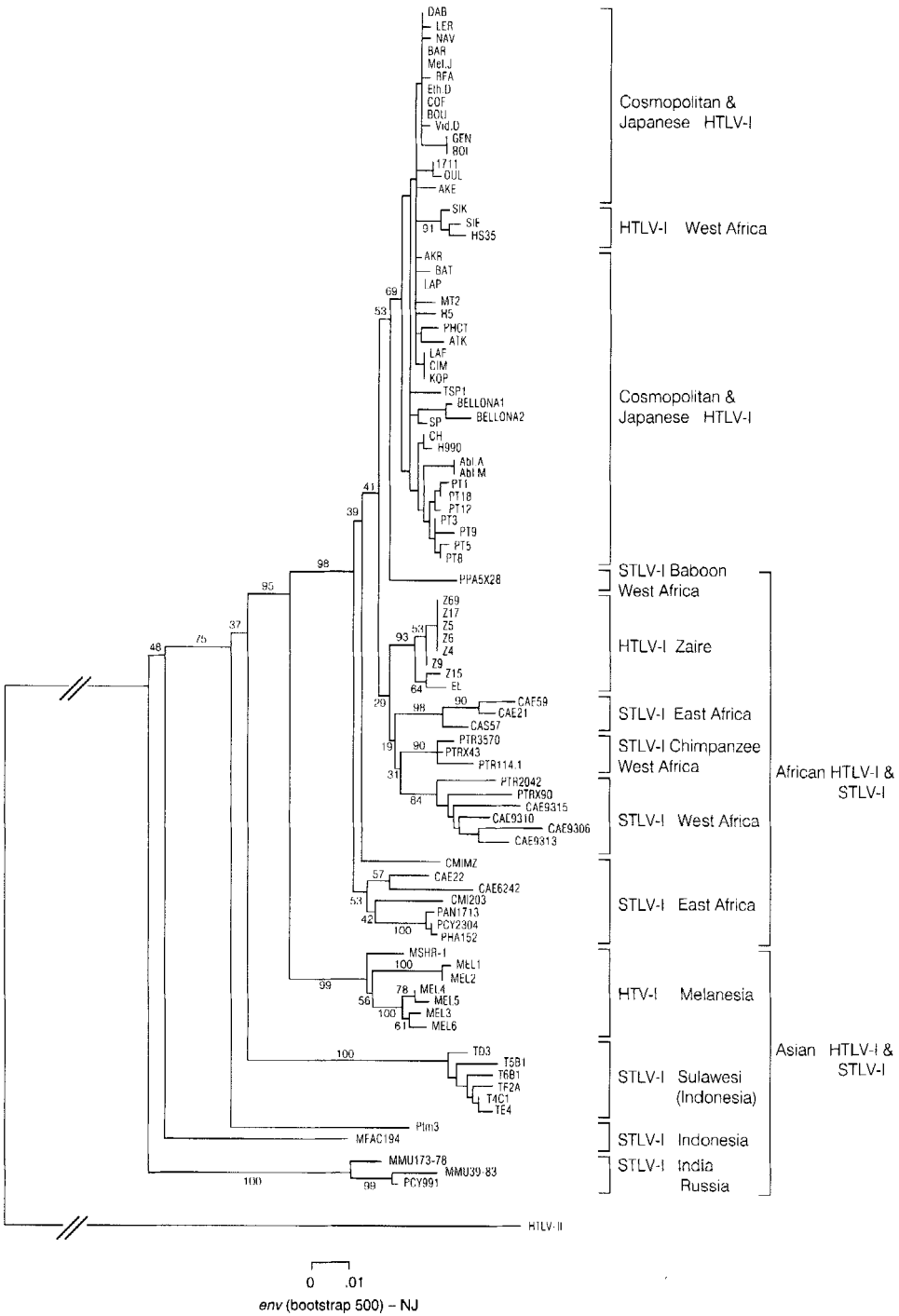


FIG. 4. Phylogenetic tree obtained by the maximum parsimony method for 62 different HTLV-I strains. The compared sequences correspond to a 522-bp region of the *env* gene encompassing most of the gp21. The largest cluster, from PT5 to OUL, represents the widespread Cosmopolitan group. The cluster with SIE, SIK, and HS 35 represents samples from Ivory Coast and Jamaica. The cluster with Z4, Z5, Z6, Z9, Z15, Z17, Z69, and EL corresponds to isolates from Zaire. The cluster containing Mel 1 to Mel 6 and MSHR-1 corresponds to the distant molecular variants from Papua New Guinea, the Solomon Islands, and Australia. The Ptm3 isolate is the prototype STLV-I from Asia. Adapted from Mahieux *et al.* (1994).


DAB
LER
NAV
BAR
Mel.J
RFA
Eth.D
COF
BOU
Vid.D
GEN
BOI
1711
OUL
AKE
SIK
SIE
HS35
AKR
BAT
LAP
MT2
H5
PHCT
ATK
LAF
CIM
KOP
TSP1
BELLONA1
SP
BELLONA2
CH
H990
Abt.A
Abt.M
PT1
PT10
PT12
PT3
PT9
PT5
PT8
PPA5X28
Z69
Z17
Z5
Z6
Z4
Z9
Z15
EL
CAE59
CAE21
CAS57
PTR3570
PTRX43
PTR114.1
PTR2042
PTRX90
CAE9315
CAE9310
CAE9306
CAE9313
CMIMZ
CAE22
CAE6242
CMI203
PAN1713
PCY2304
PHA152
MSHR-1
MEL1
MEL2
MEL4
MEL5
MEL3
MEL6
TD3
T5B1
T6B1
TF2A
T4C1
TE4
Ptm3
MFAC194
MMU173-78
MMU39-83
PCY991
HTLV-II
0
.01
env (bootstrap 500) – NJ
Cosmopolitan & Japanese HTLV-I
HTLV-I West Africa
Cosmopolitan & Japanese HTLV-I
STLV-I Baboon West Africa
HTLV-I Zaire
STLV-I East Africa
STLV-I Chimpanzee West Africa
STLV-I West Africa
STLV-I East Africa
African HTLV-I & STLV-I
HTV-I Melanesia
STLV-I Sulawesi (Indonesia)
STLV-I Indonesia
STLV-I India Russia
Asian HTLV-I & STLV-I

Melanesian HTLV-I strains have been generated, allowing rapid HTLV-I genotyping (Nerurkar *et al.*, 1994a).

It is critical to note here that all these Australo-Melanesian HTLV-I molecular variants were isolated from individuals with complete HTLV-I seroreactivites by WB and not from subjects exhibiting only "gag encoded proteins" on their WB. The question of the existence of more distantly related HTLVs in the latter individuals, also frequently found in Central Africa, remains an open question, but all attempts to isolate in culture or to detect by PCR retroviral markers in these individuals with such undeterminate WB profiles (from Australo-Melanesia or Central Africa) have been totally negative up to now (Nerurkar *et al.*, 1992; Lal *et al.*, 1992; Yanaghiara, 1994; Gessain *et al.*, 1995a; Mahieux *et al.*, in preparation).

2. *Evidence for Two Distinct HTLV-I Ancestral Lineages in Japan with a Particular Geographic Distribution*

For unknown reasons, southeastern Japan is one of the highest HTLV-I endemic areas in the world (Mueller, 1991; Blattner and Gallo, 1994; Tajima *et al.*, 1994). Based on historical data, several hypotheses have been proposed invoking population migrations and the possible introduction of the virus through commercial activities in the past few centuries. The possible high prevalence of HTLV-I in the Ainu population, considered to be one of the two oldest native Japanese populations, argues against the latter hypothesis and suggests that HTLV-I was present in Aboriginal Japanese in prehistoric times, some 2300 years ago.

By studying the amplified LTRs of Japanese HTLV-I samples using an RFLP technique, we suggested in 1992, in a preliminary report, that two genomic subtypes of HTLV-I exist in this region, one being specifically Japanese (then labeled "subtype III") and the other corresponding to the Cosmopolitan (previously named "subtype II"). Since our initial work was based on only 15 samples from Kagoshima on Kyushu Island, we subsequently studied by sequence analysis and/or the LTR/RFLP technique 67 new HTLV-I samples from Japanese living in different geographic areas (Hokkaido, Honshu, Kyushu, or the Ryukyu Islands) of Japan (59 cases) or Americans of Japanese ancestry

FIG. 5. Phylogenetic tree using a neighbor-joining approach and performed with the MEGA program on a fragment of 522 bp encompassing most of gp21 and the carboxy-terminus of the gp46. All the available published STLV-I and HTLV-I sequences were used for these studies and the HTLV-II MO isolate was used as the outgroup to root the trees. The numbers indicated at some nodes represent their frequency of occurrence out of 500 trees. Adapted from Ibrahim *et al.* (1995).

living in Hawaii (8 cases) (Ureta Vidal *et al.*, 1994a). The results, together with the data from the 19 previously published LTR sequences, clearly demonstrated the presence of two genomic subtypes of HTLV-I in Japan. The first, which we proposed naming Japanese subtype (previously labeled "subtype III"), was more frequent (67/86 = 78%) than the Cosmopolitan subtype (previously named "subtype II") (19/86 = 22%). In parallel, a fragment of 413 bp of the U3/R region (nucleotides 22 to 434) was cloned and sequenced from 10 of these new Japanese samples. The alignment and comparison of these sequences clearly confirmed the existence of two distinct molecular subtypes of HTLV-I in Japan, diverging in this LTR region by about 1.6%. Furthermore, we observed that the geographic distribution of the two subtypes, among the 80 subjects studied whose place of residence in Japan was known, showed that while the Japanese subtype was present in all parts of Japan, the Cosmopolitan subtype clustered in the southern islands of the archipelago (i.e., Kyushu and the Ryukyu Islands) as well as in immigrants from those areas who had lived in Hawaii for decades.

Historically at least three distinct populations constitute the roots of the Japanese population of today: the Wajin, the Ainu, and the Ryukyuan. The Wajin, considered to be descendants of postneolithic migrants (300 BC–600 AD) from mainland China and South Korea, presently form the largest part of the population. The two other groups, the Ainu and the Ryukyuan, considered to be descendants of the Aboriginal native populations, live principally in the north (Hokkaido) (the Ainu) and in the south (Okinawa) (the Ryukyuan). These two latter groups are considered to have been present in Japan since the prehistoric Jomon period some 2300 years ago. In more recent times (seventeenth to nineteenth centuries), Kyushu Island was regularly open for international trade, while the main central and northern Japanese islands remained tightly closed to any contact with foreigners.

Two main questions merit discussion: What could be the origin of the gradient of seroprevalence of HTLV-I in Japan (0–1% in the north versus 3–20% in the south) and of the significant differences in the geographic distribution of the Japanese versus the Cosmopolitan HTLV-I molecular subtypes?

As noted above, it is generally agreed that one of the oldest population groups in Japan, found today in the south, is the Ryukyuan, and thus one could hypothesize that this group came to Japan with a high HTLV-I infection rate. This is supported by the fact that the population on Okinawa, many of Ryukuan origin, has the highest prevalence within Japan. In contrast, the Wajin, of more recent origin, who cur-

rently represent the major population group on the three main islands, have a very low HTLV-I prevalence, perhaps as a result of a low level of intermingling with the infected population they encountered when they arrived on the islands. The data concerning the claimed high prevalence of HTLV-I in the Ainu population need to be confirmed. Thus, the existence of a gradient of HTLV-I infection in Japan supports the hypothesis of an ancient origin of the virus. To test such a hypothesis, one should search for HTLV-I endemic areas or infected ethnic groups in other Asian or American regions where descendants of the original carrier group still exist. As yet, no such endemic area or infected groups have been found, possibly because of the nearly complete disappearance of such an infected Aboriginal group or the disappearance of the virus within the descendants of such populations.

Not only is the Cosmopolitan the predominant subtype in Okinawa, but there also appears to be a gradient in the proportion of the Cosmopolitan vs the Japanese subtype as one moves to the north. In Kyushyu Island, the Japanese subtype is clearly more prevalent (76%) than the Cosmopolitan type. We suspect that such a gradient is related to a migration of the Cosmopolitan-infected population group from the south toward the north. Of interest is the recent demonstration that an HTLV-I isolate from an Ainu clustered with the Cosmopolitan type on a phylogenetic tree (Miura *et al.,* 1994). However, these data need to be confirmed on other Ainu samples from Hokkaido and Sakhalin Islands where HTLV-I-infected individuals have recently been found (Gurtsevitch *et al.,* 1995; Gessain *et al.,* 1996b). In contrast to our classification in five subtypes, Miura *et al.* (1994), who studied nine Japanese strains, propose classifying the HTLV-I isolates in three main genotypes: a Cosmopolitan including all Japanese, Caribbean, and West African strains, a Central African, and a Melanesian. Their Cosmopolitan group is divided into three lineages or subtypes: Subtype A corresponds to some Japanese isolates (including the Ainu sample), to most of the Caribbean and South American isolates, and to one Indian isolate (this subtype A corresponds to our "Cosmopolitan subtype"). The subtype B would contain most of the Japanese isolates and another Indian strain (this corresponds to our "Japanese subtype"). Their third subgroup (within their Cosmopolitan subtype) corresponds to West African and some Caribbean isolates (this corresponds to our West African subtype). In this study, Miura *et al.* (1994) suggest that the presence of the Cosmopolitan subtype both in Japan and in the Caribbean could be linked to the movement of people in the Paleolithic period from Asia to the Americas. However, their genetic studies in-

volved only nine samples from Japan, and to confirm or invalidate this interesting hypothesis, a much larger study involving samples from various Amerindian groups is needed. In this context, it is of interest to note that recent sequence data from some East Indian HTLV-I isolates indicate the presence of both subtypes (Cosmopolitan and Japanese) in this region (Nerurkar *et al.,* 1993a; Hashimoto *et al.,* 1993).

Since the Okinawan population, or at least a part of it, is considered to be the oldest Japanese population, why should the Cosmopolitan subtype be the most prevalent there? In the present phylogenetic analysis, it was difficult to determine whether the Japanese or the Cosmopolitan subtype is the most ancient, because the nucleotide divergence between the two groups is very low as compared to all other HTLV-I isolates, including variants from Africa and Melanesia. Since the Cosmopolitan subtype is found throughout the world and has a very high genomic stability, it is believed to have been dispersed in the relatively recent past through the migration of infected individuals during the past few centuries. Thus, it may have arrived in Okinawa during the latter time period by a route which remains to be determined. There are also alternative possibilities. Both subtypes could have been present in Japan since ancient times, introduced by the original settlers in the pre-agricultural Jomon period by two waves of populations infected, respectively, by one or the other of these two subtypes. In such a case, the uneven distribution of the two HTLV-I subtypes now found in Japan could reflect either a difference in the mode or efficiency of transmission or a particular migration pattern within Japan other than that discussed above. Furthermore, the pressure of various environmental and/or genetic cofactors could result in an increase in the prevalence rate of one of the two viruses in a given population over a long period of time. These questions will, it is hoped, be answered by further molecular studies of the HTLVs in Japan and other Asian countries, and by linking genetic data of human ethnic subgroups with viral genetic profiles (Tajima *et al.,* 1994).

3. HTLV-I in the Middle East

Initial reports of sporadic cases of ATL or TSP/HAM in immigrants from the area of Mashad (Iran) living in the United States, Israel, or France (Gabarre *et al.,* 1993) have shown that the northeastern part of Iran is a major HTLV-I endemic area (Meytes *et al.,* 1990). A recent study has shown a 2% seroprevalence in the blood donors of the town of Mashad, with evidence of intrafamilial clustering of HTLV-I infection (Farid *et al.,* in preparation). From a molecular point of view, the

HTLV-I strains present in Mashad belong to the Cosmopolitan group (Gabarre *et al.*, 1993; Ureta Vidal *et al.*, 1994b; Nerurkar *et al.*, 1995; Voevodin *et al.*, 1996). However, specific mutations in the *pol* and *env* genes seem to occur in this Cosmopolitan subtype (Voevodin *et al.*, 1995). Furthermore, few ATL and TSP/HAM cases infected with an HTLV-I Cosmopolitan strain similar to the molecular strain present in Mashad have been reported in Iraq (Gessain *et al.*, unpublished results) or in Kuwait (Voevodin *et al.*, 1995). The phylogenetic study of more HTLV-I strains from the Middle East should be instrumental in obtaining a better understanding of the relationships which occurred in the distant past between the Far East and the eastern and northern part of Africa.

C. HTLV-I Intrastrain and Temporal Variations in the Same Individual

Although there is 2–8% variation between geographic subtypes, the HTLV-I intrastrain variability appears much lower (<0.5%). This is in sharp contrast to the HIV and SIV systems (Boeri *et al.*, 1992; Meyerhans *et al.*, 1989; Wain-Hobson, 1992), in which the *in vivo* diversity of the genomic strain present in one individual represents one of the main problems in developing an anti-HIV vaccine. For HTLV-I, a few studies have tried to cope with this problem (Daenke *et al.*, 1990; Ehrlich *et al.*, 1992; Gessain *et al.*, 1992a) by studying multiple clones of the same PCR product of *ex vivo* PBMC proviral DNAs. Daenke *et al.* (1990) first demonstrated the existence of intrastrain variability by studying 12 HTLV-I seropositive West Indian patients (mostly with TSP/HAM), and observing HTLV-I sequence variants (in the *env* gene) in 5 of these individuals: in 4 patients, two variants were identified, and in a fifth patient, three variants. However, in all these cases, a dominant variant emerged. In a second study, Ehrlich *et al.* (1992) studied multiple clones of a 235-bp portion of the TM p21 protein in fresh or cultured cells of 19 HTLV-I seropositive individuals mostly from the United States. By studying two or more clones in 16 individuals, they demonstrated the presence of multiple HTLV-I genotypes in patients with TSP/HAM or ATL. In a third report, Gessain *et al.* (1992a) investigated 71 clones of 522 bp (encompassing most of the TM gp21 *env* gene) of 28 *ex vivo* DNA samples from ATL or TSP/HAM patients and healthy carriers from various geographic areas including the West Indies, West and Central Africa, and South America. Of possible relevance, a quasispecies distribution of viral sequences was observed

in nine of the nine African proviral DNAs, but in only 2 of the 14 samples from other regions of the world. In most of the cases, it was possible to identify a prevalent molecular clone within a viral quasispecies population. The nucleotide changes observed in the microvariants in the individual host were exclusively single base substitutions, half of which led to amino acid changes. In the same report (Gessain *et al.,* 1992a), the genetic variability of HTLV-I *in vivo* was determined over time in the same individuals by studying a well-documented HTLV-I transmission involving three individuals: a blood donor who transmitted HTLV-I-infected cells to a cardiac transplant patient, who a few months later developed TSP/HAM, and transmitted the virus to his wife by sexual contact (Gout *et al.,* 1990). The proviral sequences revealed a total identity in the gp21 *env* region, including specific mutations for these three individuals. This demonstrated the lack of any genetic drift during the *in vivo* passage of the same virus in three different hosts (two asymptomatic and one with TSP/HAM) over a period of 5 years.

As noted, Niewiesk *et al.* (1994) have shown that tax is more variable within and between healthy carriers than in patients with TSP/HAM. This was demonstrated by sequencing 20 clones of the full-length proviral *tax* gene in four TSP/HAM patients and in four healthy HTLV-I seropositive individuals of Afro-American origin living in Great Britain. Few studies have been performed to characterize the intrafamilial genetic variability of HTLV-I strains. The interfamilial sequence variation between HTLV-I strains from the Solomon Islands and those from Papua New Guinea was 3.4–4.2%, whereas the genetic heterogeneity among virus strains from the three Solomon Island families was 0.2 to 0.9% (Nerurkar *et al.,* 1993b; Yanagihara, 1994). In a study performed in a Zairian area with a high prevalence of TSP/HAM, Liu *et al.* (1994a) showed that identical sequences were found within the two families studied except for one woman infected with two variants, one the familial strain, the other a mutated one with a single substitution in the 755 sequenced nucleotides of the LTR region. Furthermore, in a family of Afro-Caribbean origin living in England (the father having ATL, the mother TSP/HAM, and three of the five children being HTLV-I seropositive healthy), Major *et al.* (1993) found that all five infected family members had identical *tax* nucleotide sequences as determined by direct sequencing of PCR products. However, by studying clones of this gene, they found more sequence variation (point mutations) within the TSP/HAM patient than in the ATL patient, and there was no conservation of mutations between the two individuals.

VI. STLV-I: Close Phylogenetic Relationship with HTLV-I

The first molecular characterization of an STLV-I sequence was performed on an Asian pig-tail monkey (Ptm3), and demonstrated an overall nucleotide homology of 90% in the *env*, pX, and LTR regions with the HTLV-I ATK prototype (Watanabe *et al.*, 1985, 1986). Furthermore, Watanabe *et al.* (1986), by sequencing the LTR of two STLV-I (from an African green monkey and from a chimpanzee), showed that the nucleotide sequence of HTLV-I (ATK) was found to be included within the divergency among STLV-I but showed closer homology (95%) to the African subtype of STLV-I than to the Asian subtype of STLV-I (90%). They proposed the existence of a distinct group of primate retroviruses that may be collectively named primate T cell leukemia virus type I (PTLV-I). Since then, numerous STLV-I have been isolated and their sequences partially characterized (Watanabe *et al.*, 1986; Schätzl *et al.*, 1992; Saksena *et al.*, 1993, 1994; Koralnik *et al.*, 1994; Song *et al.*, 1994; Vandamme *et al.*, 1994, 1995; Voevodin *et al.*, 1996).

In two reports (Koralnik *et al.*, 1994; Saksena *et al.*, 1994), phylogenetic analyses of fragments of the *env* and *pol* genes of both HTLV-I and STLV-I from Central and West Africa have provided evidence that interspecies transmission of HTLV-I and STLV-I between primates, including humans, has occurred in the distant and recent past in these areas. This has led to recognizable phylogenetic clades that persist in modern species. Vandamme *et al.* (1994), by comparing full-length LTR of various HTLV-I and STLV-I, suggest that an ancient African origin of PTLV-I is more compatible than an Asian origin. Song *et al.* (1994) suggest also that there is evidence for independent virus evolution in Asia.

We have seen that the Australo-Melanesian HTLV-I strain differs from the other subtypes by around 7% of nucleotides and occupies a unique phylogenetic position between Asian STLV-I strains and all the other, less divergent HTLV-I subtypes (Cosmopolitan, Japanese, African). One can hypothesize that, similar to what might have occurred in Africa, interspecies virus transmission may have taken place in Asia and that STLV-I strains genetically related to the Australo-Melanesian strains of HTLV-I may exist in populations of nonhuman primates presently inhabiting islands that served as the migratory pathways or routes for the early settlers of Papua New Guinea and Australia 30,000 to 40,000 years ago (Yanagihara, 1994). However, despite the wide geographic spread and large diversity of STLV-I-infected monkeys in Asia, only three partial STLV-I sequences are available from Southern Asia and Indonesia, their exact geographic origins not being defined: two sequences

from Indonesia, Ptm3 from a pig-tail monkey (*Macaca nemestrina*) (Watanabe *et al.*, 1985) and C194 from a crab-eating macaque (*Macaca fascicularis*) (Koralnik *et al.*, 1994), and one from southern Asia, MM39-83, from a rhesus macaque (*Macaca mulata*) (Koralnik *et al.*, 1994).

We recently isolated a new STLV-I originating from a troop of Celebes macaques (*Macaca tonkeana*) and performed the first complete nucleotide sequencing of an STLV-I on one of these isolates (STLV-I TE4) (Ibrahim *et al.*, 1995). *Macaca tonkeana* lived originally on Sulawesi Island, which is one of the nearest natural monkey habitats to Papua New Guinea, an area devoid of monkeys, but is also found on one of the two suspected migratory pathways taken by the early Australoid migrants. Phylogenetic trees performed on the LTR and *env* (gp46, gp21) regions demonstrated that this new STLV-I occupies a unique position, being, by most analysis, more closely related to the Australo-Melanesian HTLV-I topotype than to any other Asian STLV-I (Fig. 5). These data raise new hypotheses on the possible interspecies viral transmission between monkeys carrying STLV-I and early Australoid settlers, ancestors of the present-day Australo-Melanesian inhabitants, during their migration from the Southeast Asian land mass to the Greater Australian continent.

VII. HTLV-II

A. Existence of Two Molecular Subtypes

The first complete sequence of HTLV-II (MO) was published in 1985 by Shimotohno *et al.* (1985). This virus was originally cloned from the splenic cells of a patient from the United States with an atypical T hairy cell leukemia (Kalyanaraman *et al.*, 1982). The comparison of its sequence with that of HTLV-I ATK demonstrated an overall homology at the nucleotide level of 60% in the coding regions. However, some regions exhibited greater homology, 85% for the p24 *gag* and 78% for the *tax* region, than other regions such as the p19 *gag* (59%), the SU gp46 envelope (63%), or the *rex* region (61%).

In 1990, Lairmore *et al.* published sequences of small fragments of the *gag* (107 bp) and *pol* (112 bp) genes of two different HTLV-II from Guaymi Indians living in Panama. The comparison of these small sequences exhibiting a high homology with HTLV-II MO (94% and 98%, respectively) suggested that, in line with the HTLV-I situation, low genetic variability between different HEV-II isolates exists. However, Hall *et al.*, in 1992a, demonstrated the existence of two molecular sub-

types of HTLV-II that they tentatively designated as HTLV-IIA and HTLV-IIB by comparing the restriction map analysis of several isolates from intravenous drug users in the United States. The study of the nucleotide sequence of the *env* gene region encoding the TM gp21 glycoprotein confirmed the existence of these two subtypes, which diverge by about 4% in this genomic region (Hall *et al.,* 1992a). Furthermore, in this study there was a marked conservation of the sequences within the two subtypes, with a maximum of 0.4% difference within a subtype (Fig. 6). This work has been confirmed by the same group with analysis of the sequences of the entire *env* gene, a portion of the *pol*, and the LTR, the LTR exhibiting the greatest divergence (5.8%) within subtypes, as in the case of HTLV-I (Hall *et al.,* 1992b; Takahashi *et al.,* 1993). Several other groups have confirmed the existence of these two subtypes (Hjelle *et al.,* 1993; Lee *et al.,* 1993; Zella *et al.,* 1993; Ijichi *et al.,* 1993; Pardi *et al.,* 1993a,b; Igarashi *et al.,* 1993; Lal *et al.,* 1994; Ferrer *et al.,* 1993; Calabro *et al.,* 1993; Switzer *et al.,* 1995a,b; Dube *et al.,* 1993, 1994, 1995; Gessain *et al.,* 1994b, 1995b; Mauclere *et al.,* 1995; Salemi *et al.,* 1995; Ishak *et al.,* 1995; Eiraku *et al.,* 1995; Tuppin *et al.,* 1996; Biglione *et al.,* in preparation). In the early studies it was clear that the subtype A (prototype MO) was present mainly in intravenous drug abusers in the United States and only rarely encountered in Amerindians such as the Pueblo Indians, whereas subtype B (prototypes NRA and G12) was found mainly in most of the Amerindian groups tested, scattered throughout North, Central, and South America, including some remote groups in Brazil, Argentina, and Colombia. Subtype B was thus referred to as the Paleoindian strain. The situation is currently changing, since HTLV-II subtype A has been found to be the natural host of several Amerindian tribes, including the Kayapo of Brazil (Ishak *et al.,* 1995). Furthermore, both subtypes have been found to be present in intravenous drug abusers from the United States, Italy, and Spain (Salemi *et al.,* 1995). Moreover, as we will see, several HTLV-II subtype B strains very similar to the Paleoindian strains have been found in remote populations of Central Africa, including Pygmies (Gessain *et al.,* 1995; Tuppin *et al.,* 1996).

B. An Ancient Presence of HTLV-II in Africa

As we have seen, epidemiological and phylogenetic analyses indicate that HTLV-I and STLV-I have been present in the Old World (Africa and Asia) for several millennia, whereas HTLV-II, which is highly endemic in certain native Amerindian tribes, has been considered to be a New World virus, brought from Asia to the Americas some 10,000 to

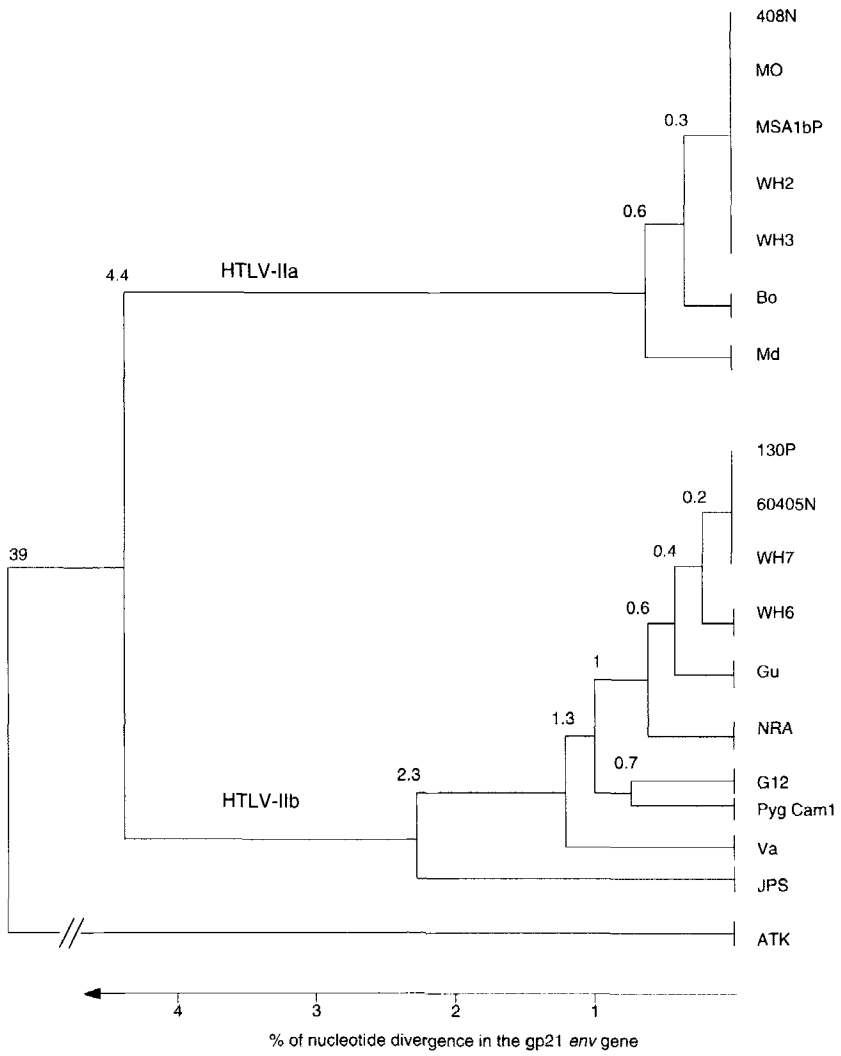


FIG. 6. Dendrogram comparing the nucleotide sequences of 589 base pairs of the HTLV-II *env* region (bp 6052–6640) coding for gp21. The branch lengths are not proportional to nucleotide substitutions. This dendrogram clearly shows the existence of two molecular clusters of HTLV-II genotypes, with a new African variant (JPS) included within the subtype B cluster. This analysis was performed on 17 different HTLV-II isolates, including a new Gabonese variant strain (JPS) and 16 other available published sequences comprising HTLV-II subtype A (MO), subtype B (NRA), G12 prototypes, and HTLV-II isolates representative of the different geographic origins: Amerindians (408N, 60405N, 130P, MSA1bP), intravenous drug abusers from the United States (WH2, WH3, WH6, WH7) or Italy (Gu, Va, Bo, Md), and the recently described HTLV-II b (Pyg Cam1) from a Pygmy living in a remote area of Cameroon (Gessain *et al.*, 1995). The HTLV-I ATK isolate was used as outgroup to root the phylogenetic trees. Adapted from Tuppin *et al.* (1996).

40,000 years ago by the migration over the Bering land bridge of HTLV-II-infected populations (Maloney *et al.*, 1992; Hall *et al.*, 1994; Neel *et al.*, 1995a). Serological and molecular evidence of sporadic cases of HTLV-II infection in Western and Central Africa (Mauclère *et al.*, 1993, 1995a; Igarashi *et al.*, 1993; Bonis *et al.*, 1994), however, raises the possibility that HTLV-II or a related retrovirus has also been present in Africa for a long time (Gessain *et al.*, 1996a). This hypothesis was first suggested after the detection of HTLV-II-like serological markers in Pygmies in remote areas of Zaire and Cameroon. Pygmies are considered to belong to the oldest populations in Central Africa, having lived in these areas for at least 20,000 years (Goubau *et al.*, 1992, 1993, 1996; Froment *et al.*, 1993).

We have reported (Gessain *et al.*, 1995b) the isolation, molecular characterization, and phylogenetic analysis of an HTLV-II from a healthy, 59-year-old Pygmy woman of the Bakola population, a group living in a remote area of the equatorial forest of southwest Cameroon in which HTLV-II-like serological markers were found 10 years ago in four individuals (Froment *et al.*, 1994). The culture of PBMCs from this woman resulted in the establishment of a continuous CD8+ cell line containing an HTLV-II provirus that produces HTLV-II antigens and retroviral particles. Molecular analysis of this new HTLV-II, PYGCAM-1, which is present in both uncultured and long-term cultured cells, demonstrated close homology (Fig. 6) with the HTLV-II subtype B prototypes (G12 and NRA) originally isolated from a Guaymi Amerindian in Panama and from a patient from the United States with T hairy cell leukemia, respectively (Lee *et al.*, 1993; Pardi *et al.*, 1993b).

We have proposed that this new isolate, a genuine African HTLV-II, has been present in this remote population for a very long time for a number of reasons. First, serum from the father of this woman as well as from three unrelated Bakola Pygmies living in the same area had shown similar seroreactivity (p24 >> P19, rgp 21, and K55), indicating the presence of HTLV-II 10 years previously. Isolation of an HTLV-II subtype B from the woman suggests that the four individuals were infected by the same virus. Furthermore, the presence of HTLV-II antibodies in the father of the index case supports the hypothesis that HTLV-II has been endemic in this population for at least two generations. Close contact between the Bakola Pygmy group and neighboring populations has been rare until recently. Second, a retrospective seroepidemiological survey of sera collected between 1967 and 1970 from 214 Cameroonian Pygmies showed HTLV-II-like seroreactivity in five (Goubau *et al.*, 1993). HTLV-II appears to be virtually absent in other populations in Cameroon: no antibodies to HTLV-II were detected

in more than 4000 samples of sera from adult individuals of different ethnic groups and from various areas of Cameroon, although HTLV-I was present in 0.5–2% (Mauclère *et al.*, 1995b and in preparation). Third, while most of the Pygmy groups tested were seronegative for HTLV-II antibodies, the only other African population described to date in which HTLV-II is endemic is the Bambuti Pygmies in the Ituri region of northeast Zaire, where 14 of 102 individuals were found to be HTLV-II-like seropositive in 1970 and 4 of 12 in 1991 (Goubau *et al.*, 1992, 1993). We failed to detect any HTLV-II-infected individuals in representative samples of Biaka Pygmies from the Lobaye region of Central African Republic and of Twa Pygmies from the Lake Tomba area of Zaire or in the Bantu population of the surrounding areas (Gessain *et al.*, 1993b). It is interesting to note that the Bambuti and Bakola Pygmies, both of which are infected with HTLV-II, are located at the western and eastern ends, respectively, of the present area of Pygmy habitation in Central Africa, and are of different paleoanthropological origins, with a divergence, based on linguistic and cultural features, of 10,000 years. Taken together, these results strongly suggest that some isolated Pygmy populations (considered to be the earliest inhabitants of Central Africa) have for a very long time been an African reservoir for HTLV-II infection. Genetic studies suggest that about 20,000 years of isolation would have been required for Pygmies to become genetically distinct.

Since any close contact between these central African Pygmies and Amerindian populations is very unlikely, African HTLV-II may have diverged from a common ancestor long before HTLV-II was introduced into the Americas by perhaps two waves of migration over the Bering land bridge, each carrying one of the two molecular subtypes of HTLV-II.

We were surprised to find a genuine African HTLV-II so close genetically to the Paleoindian subtype B, but genetic variability within HTLV-II subtypes seems to be extremely low. Thus, a nucleotide divergence of only 0–0.4% (over 1000 base pairs of the *env* gene) was observed among HTLV-IIs of different Indian groups in North, Central, and South America, who have probably had no contact for several thousands of years.

The possibility that HTLV-II subtype B is a genuine African HTLV-II is reinforced by our finding of a slightly variant (2% in the *env* gene) subtype B in a Gabonese Bahoumbou family which has harbored this variant virus for at least three generations (Gessain *et al.*, 1994b; Tuppin *et al.*, 1996). An HTLV-II B provirus has also been detected in two individuals in another area of Gabon (Delaporte *et al.*, personal communication, 1994) and in Zaire (Dube *et al.*, 1994). Only two HTLV-II subtype A African isolates have been described to date, one in Ghana

(Igarashi *et al.,* 1993) and the other in Cameroon (Mauclère *et al.,* 1995), both in prostitutes working in large towns, suggesting more an infection imported through commercial sexual activities than a long-standing situation. Alternatively, we have to bear in mind that these two HTLV-II subtype A strains found in prostitutes from Central/West Africa are closely related to but slightly different from those found in intravenous drug abusers from the United States. Thus, they may also represent a different molecular strain of HTLV-II subtype A of African origin (Mauclère *et al.,* 1995)

Preliminary data from Mongolia show the presence of a typical HTLV-II subtype A similar to other subtype A strains present in the Americas (Hall *et al.,* 1994b and personnal communication). This again supports the hypothesis that genetic drift is much lower for HTLV-II than for HTLV-I, for which (assuming a similar evolutionary rate in the different geographic locations) an estimated 0.2–1% nucleotide divergence (in the *env* and *pol* genes) during 1000 years of evolution has been proposed (Poiesz *et al.,* 1993; Yanagihara *et al.,* 1994).

Our data thus support the hypothesis that HTLV-II subtype B has been present for a long time (20,000 to 30,000 years) in Central Africa. In the context of recent evidence for interspecies transmission in Central and West Africa of PTLV-I (HTLV-I/STLV-I) strains, we would like to suggest that some STLV-II (closely related to HTLV-II subtype B) do exist or might have existed in Africa. The recent finding of new PTLVs in several pygmy chimpanzees of Zairian origin (PTLV-PP1664 and STLV-PP) and in wild-caught baboons in Eritrea (PTLV-L) thus support the hypothesis of an African origin for some PTLVs.

VIII. STLV-PP AND PTLV-L

A new virus called PTLV-L was isolated from a wild-born baboon (*Papio hamadryas*) from Eritrea (Goubau *et al.,* 1994). In a cDNA library, a 1802-bp-long fragment was identified that extends from the *env* region, including the complete transmembrane protein gene, to part of the *tax/rex* gene. Homologies at the nucleotide sequence level of PTLV-L—prototype simian T lymphotropic virus PH969—with HTLV-I and HTLV-II, respectively, were 62% and 64% overall, 65% and 70% in the *env* region, and 80% and 80% in the partial *tax/rex* sequence. Phylogenetic analysis based on the gene encoding the transmembrane protein indicates that PTLV-L represents a PTLV type with a long independent evolution, longer than any strain within the PTLV-I or PTLV-II groups.

Another new virus has been isolated independently by two different groups from captive colonies of pygmy chimps (*Pan paniscus*) originating from Zaire but living in primate centers in the United States and Belgium (STLV-PP: Giri *et al.*, 1994; Diglio *et al.*, 1995; STLV-PP1664: Liu *et al.*, 1994a; Vandamme *et al.*, 1996). The "Belgium" virus was designated STLV PP1664, and phylogenetic analysis performed on fragments of the *tax* region indicates that although this virus is more closely related to HTLV-II than to the other PTLV types, it is clearly separated from HTLV-II, indicating a long independent evolution (Liu

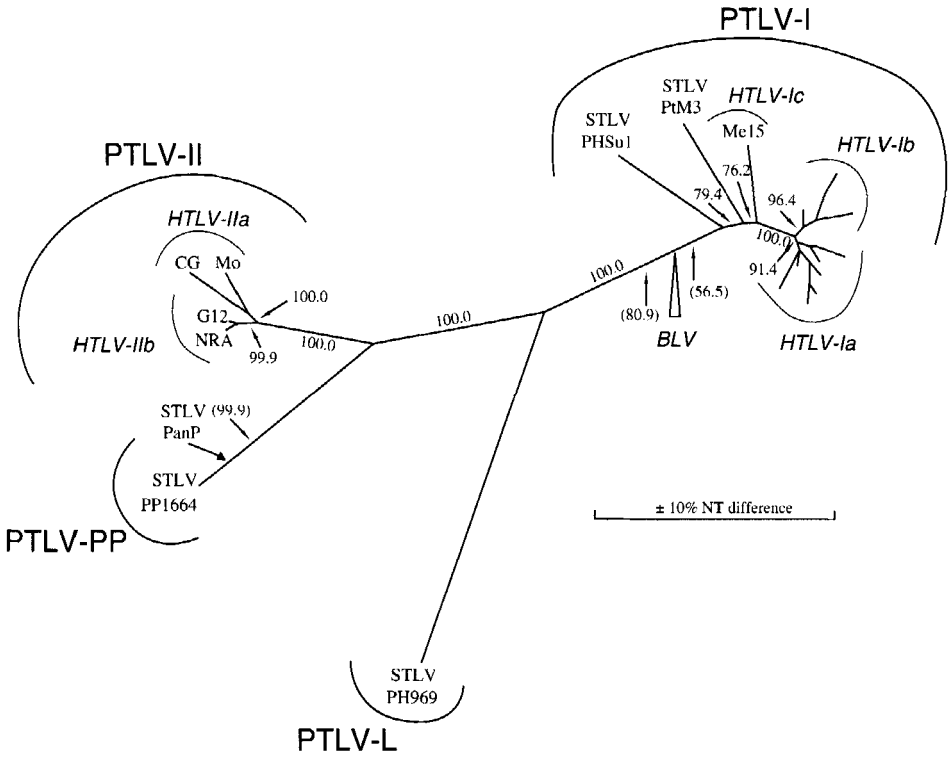


FIG. 7. Unrooted phylogenetic analysis of a 710-nucleotide *tax/rex* fragment using a neighbor-joining approach comparing the different prototypes of HTLV-I, HTLV-II, STLV-I, STLV-II, and PTLV-PP and PTLV-L. The distances were calculated using the Kimura two-parameter model, scoring transitions twice more likely than transversions. The bootstrap statistical analysis was applied using 1000 bootstrap samples. The values on the branches represent the percentage of trees for which the sequences at one end of the branch are a monophyletic group. Several phylogenetic groups with high bootstrap values can be seen clearly. Adapted from Vandamme *et al.* (1996).

et al., 1994a) (Fig. 7). Concerning STLV-PP, sequence data from viral DNA amplified with the SK43-SK44 primers (*tax* region) indicated that STLV-PP is slightly more closely related to HTLV-II (87%) than to HTLV-I ATK (78%). Furthermore, STLV-PP appears to be equally distant from STLV Ptm3 and the new PTLV-L PH 969. It is interesting to note that the animals infected by these new PTLV (PTLV-L and STLV-PP and STLV-PP1664) exhibit an HTLV-II like serology on Western blot with at least a P24 >> P19 and an rgp21 with or without the reactivity against the specific gp46 peptide (K55) of HTLV-II (Giri *et al.,* 1994; Goubau *et al.,* 1994). Further sequencing of other genes of such PTLV is ongoing and will be crucial for a better understanding of the phylogenetic relationship between all the PTLV. However, these data indicate that the number of PTLV types should be considered open and suggest that the variety of indigenous viruses in the PILV group is to date the largest in the African continent (Fig. 7) (Vandamme *et al.,* 1996).

IX. Conclusions and Perspective: HTLV-I/II as Privileged Markers of Migration of Ancient Human Populations

The high stability of the HTLV-I/II genomes, their integration, and their lifelong latency make the subtle mutations observed unique molecular markers for tracing the geographic path of evolution of such viruses, paralleling the migration of the carriers (Gessain *et al.,* 1992b, 1995b; Yanagihara, 1994).

Based on the comparison of available historical, paleoanthropological, and sequence data for various HTLV-I from Africa, the Americas, the Far East, and the Pacific region, one can estimate the *in vivo* genetic drift of HTLV-I to be around 0.02 to 0.1% per century (1% per 1000–5000 years) for the gp21 *env* gene. An estimation of the substitution rate for STLV-I is about 1% per 20,000–122,000 years (Song *et al.,* 1994). However, accumulation of new data on HTLV-I sequences from several parts of the world is necessary to improve our knowledge concerning the estimation rate and patterns of HTLV-I sequence evolution (Picard *et al.,* 1995). All these data are based on the undemonstrated premises that the rate of *in vivo* genetic drift is similar in the different geographic HTLV-I endemic areas and that it remains constant over time. Hypothetical routes of HTLV-I dissemination throughout the world have been suggested by some authors (Saksena *et al.,* 1992; Yanagihara, 1994), with a possible origin of HTLV-I in the Indo-Malay region followed by a dissemination of the virus to Africa through migrations of populations via Madagascar and both East and South

Africa. However, the lack of knowledge of the frenetic structure of HTLV-I from the Indian Ocean (especially Madagascar and the Seychelles Islands), East, South, and North Africa, and most of the Asian continent, including China, India, Siberia, and Southeast Asia, makes any definite interpretation concerning the origin and the phylogenetic relationship of these primate T lymphotropic retroviruses much too premature (Miura *et al.,* 1994; Tajima *et al.,* 1994; Picard *et al.,* 1996).

The present world repartition of the HTLV-I/STLV-I genotypes is the result of at least three independent events: interspecies transmission from simian to human, evolution in some remote populations (as for the Melanesians), and migration of infected populations (as during the slave trade) (Fig. 2).

In regard to HTLV-II, its presence in isolated Amerindian populations disseminated throughout the Americas and its apparent absence elsewhere had led some authors to the speculation that this virus was exclusively a "New World virus" endemic in native Amerindians over centuries or millennia. The virus is supposed to have been introduced to the Americas through the successive waves of migrations which occurred from Siberia through the Bering strait area millennia ago (Dube *et al.,* 1993; Noel *et al.,* 1994). However, the recent discovery of HTLV-II infection in Central and West African populations, especially in remote areas of Zaire (Goubau *et al.,* 1992) and Cameroon (Gessain *et al.,* 1995b), raises new questions concerning the origin, evolution, and global dissemination of this human retrovirus.

ACKNOWLEDGMENTS

We thank Fredj Tekaia and Abel Ureta Vidal for the preparation of the dendrogram and phylogenetic trees and Mrs. Liliane Lozano and Monique Van Beveren for the preparation of the manuscript. Financial support from the World Laboratory (subproject MCD-2), the Comité de Paris of the Ligue Nationale contre le Cancer (contract RS44), l'Agence Nationale de Recherches sur le SIDA and Virus Cancer Prévention is acknowledged. We thank Dr. A.-M. Vandamme for Fig. 7.

REFERENCES

Bangham, C. R., Daenke, S., Phillips, R. E., Cruickshank, J. K., and Bell, J. I. (1988). Enzymatic ampification of exogenous and endogenous retrovirual sequences from DNA of patients with tropical spastic paraparesis. *EMBO J.* **7**, 4179–4184.

Bastian, I., Gardner, J., Webb, D., and Gardner, I. (1993). Isolation of a human T-lymphotropic virus type I strain from Australian Aboriginals. *J. Virol.* **67**, 843–851.

Bazarbachi, A., Huang, M., Gessain, A., Saal, F., Saib, A., Peries, J., de Thé, H., and Galibert, F. (1995). Human T-cell leukemia virus type I in post-transfusional spastic

paraparesis: Complete proviral sequence from uncultured blood cells. *Int. J. Cancer* **63**, 494–499.

Benson, J., Tschachler, E., Gessain, A., Yanagihara, R., Gallo, R. C., and Franchini, G. (1994). Cross-neutralizing antibodies against Cosmopolitan and Melanesian strains of human T cell leukemiallymphotropic virus type I in sera from inhabitants of Africa and the Solomon islands. *AIDS Res. Hum. Retroviruses* **10**, 91–96.

Berneman, Z., Gartenhaus, R., Reitz, M., Blattner, W., Manns, A., Lkehara, O., Gallo, R. C., and Klotman, M. (1992). Expression of alternatively spliced human T lymphotropic virus type 1 (HTLV-I) pX mRNA in infected cell lines and in primary uncultured cells from patients with adult T cell leukemia/lymphoma and healthy carriers. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3005–3009.

Biglione, M., Gessain, A., Quiruelas, S., Fay, O., Taborda, M. A., Femandez, E., Lupo, S., Panzita, A., and de Thé, G. (1993). Endemic HTLV-II infection among Tobas and Matacos Amerindians from North Argentina. *J. Acquir. Immune Defic. Syndr.* **6**, 631–633.

Black, F. L., Biggar, R. J., Neel, J. V., Maloney, E. M., and Waters, D. J. (1994). Endemic transmission of HTLV type II among Kayapo Indians of Brazil. *AIDS Res. Hum. Retroviruses* **10**, 1165–1171.

Blattner, W. A. (1990). Epidemiology of HTLV-I and associated diseases. *In* "Human Retrovirology: ATLL," pp. 251–256. Raven Press, New York.

Blattner, W. A., and Gallo, R. C. (1994). Epidemiology of HTLV-I and HTLV-II infection. *In* "Adult T-cell Leukaemia" (K. Takatsuki, ed.), pp. 45–90. Oxford University Press, Oxford.

Boeri, E., Giri, A., Lillo, F., Ferrari, G., Vamier, O. E., Ferro, A., Sabbatani, S., Saxinger, W. C., and Franchini, G. (1992). In vivo genetic variability of the human immunodeficiency virus type 2 V3 region. *J. Virol.* **66**, 4546–4550.

Boeri, E., Gessain, A., Garin, B., Kazadi, K., de Thé, G., and Franchini, G. (1993). Qualitative changes in the HTLV-I *env* gene sequence in the spastic versus nonspastic tropical paraparesis are not correlated with disease specificity. *AIDS Res. Hum. Retroviruses* **9**, 1, 1–5.

Bonis, J., Verdier, M., Dumas, M., and Denis, F. (1994). Low human T cell leukemia virus type II seroprevalence in Africa. *J. Infect. Dis.* **169**, 225–227.

Calabro, M. L., Luparello, M., Grottola, A., Del Mistro, A., Fiore, J. R., Angarano, G., and Chieco-Bianchi, L. (1993). Detection of human T lymphotropic virus type II/b in human immunodeficiency virus type 1-coinfected persons in southeastern Italy. *J. Infect. Dis.* **158**, 1273–1277.

Cann, A. J., and Chen, I. S. Y. (1990). Human T-cell leukemia virus Types I and II. *In* "Virology" (B. N. Fields, D. R. Knipe, *et al.*, eds.), pp. 1501–1527. Raven Press, New York.

Chen, Y.-M. A., Jang, Y.-J., Kanki, P. J., Yu, Q.-C., Wang, J.-J., Montali, R. J., Samuel, K. P., and Papas, T. S. (1994). Isolation and characterization of simian T-cell leukemia virus type II from New World monkeys. *J. Virol.* **68**, 1149–1157.

Chou, K. S., Okayama, A., Tachibana, N., Lee, T. H., and Essex, M. (1995). Nucleotide sequence analysis of a full length human T+-cell leukemia virus type I from adult T-cell leukemia cells: A prematurely terminated pX open reading frame II. *Int. J. Cancer* **60**, 701–706.

Ciminale, V., Pavlakis, G., Derse, D., Cunningham, C., and Felber, B. (1992). Complex splicing in the HTLV family of retroviruses: Novel mRNAs and proteins produced by HTLV-I. *J. Virol.* **66**, 1737–1745.

Ciminale, V., D'Agostino, D. M., Franchini, G., Felber, B. K., and Chieco-Bianchi, L. (1994). Expression and characterization of proteins encoded in the proximal X region of HTLV-II. *AIDS Res. Hum. Retroviruses* **10**, 461.

Ciminale, V., D'Agostino, D. M., Zotti, L., and Chicco-Bianchi, L. (1996). Coding potential

of the X region of human T-cell leukemia/lymphotropic virus type II. *J. Acquir. Immune Defic. Syndr.*, in press.

Daenke, S., Nightingale, S., Cruickshank, J. K., and Bangham, R. M. (1990). Sequence variants of human T-cell lymphotropic virus type I from patients with tropical spastic paraparesis and adult T-cell leukemia do not distinguish neurological from leukemic isolates. *J. Virol.* **64**, 1278–1282.

De, B. K., Lairmore, M. D., Griffis, K., Williams, L. J., Vilinger, F., Quinn, T. C., Brown, C., Nzilambi, Sugimoto, M., Araki, S., and Folks, T. M. (1991). Comparative analysis of nucleotide sequences of the partial envelope gene (5′ domain) among human T lymphotropic virus type I (HTLV-I) isolates. *Virology* **182**, 413–419.

Dekaban, G. A., King, E. E., Waters, D., and Rice, G. P. A. (1992). Nucleotide sequence analysis of an HTLV-I isolate from a Chilean patient with HAM/TSP. *AIDS Res. Hum. Retroviruses* **8**, 1201–1207.

De Rossi, A., Mammano, F., De Mistro, A., and Chieco-Bianchi, L. (1991). Serological and molecular evidence of infection by human T-cell lymphotropic virus type II in Italian drug addicts by use of synthetic peptides and polymerase chain reaction. *Eur. J. Cancer* **27**, 835–838.

de Thé, G. (1982). Epidemiology of Epstein-Barr virus and associated diseases in man. *In* "The Herpes Viruses" (B. Roizman, ed.), pp. 25–103. Plenum Press, New York.

de Thé, G., and Bomford, R. (1993). An HTLV-I vaccine: Why, how, for whom? *AIDS Res. Hum. Retroviruses* **9**, 381–386.

Digilio, L., Giri, A., Fullen, J., Gallo, R. C., and Franchini, G. (1995). Molecular and biological characterization of STLV pan-p, a novel infectious retrovirus. *J. Acquir. Immune Defic. Syndr.* **10**, 262. [Abstract]

Dube, D. K., Sherman, M. P., Saksena, N. K., Bryz-Gornia, V., Mendelson, J., Love, J., Arnold, C. B., Spicer, T., Dube, S., Glaser, J. B., Williams, A. E., Nishimura, M., Jacobsen, S., Ferrer, J. F., del Pino, N., Quiruelas, S., and Poiesz, B. J. (1993). Genetic heterogeneity in human T-cell leukemia/lymphoma virus type II. *J. Virol.* **67**, 1175–1184.

Dube, D. K., Dube, S., Erensoy, S., Jones, B., Bryz-Gornia, V., Spicer, T., Love, J., Saksena, N., Lechat, M. F., Shrager, D. I., Dosik, H., Glaser, J., Levis, W., Blattner, W., Montagna, R., Blumberg, B., and Poiesz, B. J. (1994). Serological and nucleic acid analyses for HIV and HTLV infection on archival human plasma samples from Zaire. *Virology* **202**, 379–389.

Dube, S., Spicer, T., Bryz-Gornia, V., Jones, B., Dean, T., Love, J., Ferrer, J., Esteban, N., Harrington, W. Jr., Glasser, J., Dube, D. K., Williams, A., Dosik, H., Siegal, F., and Poiesz, B. J. (1995). A rapid and sensitive method of identification of HTLV-II subtypes. *J. Med. Virol.* **45**, 1–9.

Duenas-Barajas, E., Bernal, J. E., Vaught, D. R., Nerurkar, V. R., Sarmiento, P., Yanagihara, R., and Gajdusek, D. C. (1993). Human retroviruses in Amerindians of Colombia: High prevalence of human T cell lymphotropic virus type II infection among the Tunebo Indians. *Am. J. Trop. Med. Hyg.* **49**, 657–663.

Ehrlich, G. D., Andrews, J., Sherman, M. P., Greenberg, S. J., and Poiesz, J. (1992). DNA sequence analysis of the gene encoding the HTLV-I P21E transmembrane protein reveals inter- and intraisolate genetic heterogeneity. *Virology* **186**, 619–627.

Eiraku, N., Monken, C., Kubo, T., Zhu, S. W., Rios, M., Bianco, C., Hjelle, B., Nagashima, K., and Hall, W. W. (1995). Nucleotide sequence and restriction fragment length polymorphism analysis of the long terminal repeat of human T cell leukemia virus type II. *AIDS Res. Hum. Retroviruses* **11**, 625–636.

Evangelista, A, Maroushek, S., Minnigan, H., Larson, A., Retzel, E., Haase, A, Gonzales-

Dunia, D., McFarlin, D., Mingioli, E., Jacobson, S., Osame, M., and Sonoda, S. (1990). Nucleotide sequence analysis of a provirus derived from an individual with tropical spastic paraparesis. *Microbiol. Pathogenesis* **8**, 259–278.

Ferrer, J. F., Del Pino, N., Esteban, E., Sherman, M. P., Dube, S., Dube, D. K., Basombrio, M. A., Pimentel, E., Segovia, A., Quirulas, S., and Poiesz, B. J. (1993). High rate of infection with the human T-cell leukemia retrovirus type II in four Indian populations of Argentina. *Virology* **197**, 576–584.

Fouchard, N., Flageul, B., Bagot, M., Avril, M. F., Hermine, O., Sigaux, F., Merle-Beral, H., Troussard, X., Delfraissy, J. F., de Thé, G., and Gessain, A. (1995). Lack of evidence of HTLV-I/II infection in T CD8 malignant or reactive lymphoproliferative disorders in France: A serological and/or molecular study of 169 cases. *Leukemia* **9**, 2087–2092.

Franchini, G. (1995). Molecular mechanisms of human T-cell leukemia/lymphtotropic virus type I infection. *Blood* **86**, 3619–3639.

Franchini, G., Mulloy, J. C., Koralnil, I. J., Lomonico, A., Sparkowski, J. J., Andresson, T., Goldstein, D. J., and Schlegel, R. (1993). The human T-cell leukemia/lymphotropic virus type I $p12^{I}$ protein cooperates with the E5 oncoprotein of bovine papillomavirus in cell transformation and binds the 16kD subunit of the vacuolar H^+ ATPase. *J. Virol.* **67**, 7701–7704.

Froment, A., Delaporte, E., Dazza, M.-C., and Larouze, B. (1993). HTLV-II among Pygmies from Cameroon. *AIDS Res. Hum. Retroviruses* **9**, 707.

Fujiyama, C., Fujiyoshi, T., Miura, T., Yashiki, S., Matsumoto, D., Zaninovic, V., Blanco, O., Harrington, Jr. W., Byrnes, J. J., Hayami, M., Tajima, K, and Sonoda, S. (1993). A new endemic focus of human T lymphotropic virus type II carriers among Orinoco natives in Colombia. *J. Infect. Dis.* **168**, 1075–1077.

Fukasawa, M., Tsujimoto, H., Ishikawa, K. I., Miura, T., Ivanoff, B., Cooper, R. W., Frost, E., Delaporte, E., Mingle, J. A. A., Grant, F. C., and Hayami, M. (1987). Human T-cell leukemia virus type 1 isolates from Gabon and Ghana: Comparative analysis of proviral genomes. *Virology* **161**, 315–320.

Gabarre, J., Gessain, A., Raphael, M., Merle-Beral, H., Dubourg, O., Fourcade, C., Gandjbakhch, I., Jault, F., Delcourt, A., and Binet, J. L. (1993). Adult T-cell leukemia/lymphoma revealed by a surgically cured cardiac valve lymphomatous involvement in an Iranian woman: Clinical, immunopathological and viromolecular studies. *Leukemia* **7**, 1904–1909.

Gasmi, M., Farouqi, B., d'Incan, M., and Desgranges, C. (1994). Long terminal repeat sequence analysis of HTLV type I molecular variants identified in four North African patients. *AIDS Res. Hum. Retroviruses* **10**, 1313–1315.

Gessain, A. (1996a). Epidemiology of HTLV-I and associated diseases. *In* "Human T cell Lymphotropic Virus Type 1" (P. Hollsberg, ed.). John Wiley & Sons, England (in press).

Gessain, A. (1996b). Virological aspects of tropical spastic paraparesis/HTLV-I associated myelopathy and HTLV-I infection. *J. Neurovirol.* **2**, in press.

Gessain, A., and Gout, O. (1992). Chronic myelopathy associated with human T-lymphotropic viras type I (HTLV-I). *Ann. Intern. Med.* **117**, 933–946.

Gessain, A., Barin, F., Vernant, J. C., Gout, O., Maurs, L., Calender, A., and de Thé, G. (1985). Antibodies to human T-lymphotropic virus type-l in patients with tropical spastic paraparesis. *Lancet* **2**, 407–410.

Gessain, A., Yanahigara, R., Franchini, G., Garruto, R. M., Jenkins, C. L., Ajdukiewicz, A. B., Gallo, R. C., and Gajdusek, D. C. (1991). Highly divergent molecular variants of human T-lymphotropic virus type I from isolated populations in Papua New Guinea and the Solomon Islands. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 7684–7698.

Gessain, A., Boeri, E., Kazadi, K., Garin, B., Salaun, J. J., Gallo, R. C., de Thé, G., and

Franchini, G. (1992a). Variant rétroviral HTLV-I au Zaire chez un patient ayant une neuromyélopathie chronique, séquence nucléotidique du gène d'enveloppe. *Compte Rendu de l'Acadeémie des Sciences* **314**, 159–164.

Gessain, A., Gallo, R. C., and Franchini, G. (1992b). The low degree of HTLV-I genetic drift in vivo as a mean to follow viral transmission and movement of ancient human population. *J. Virol.* **66**, 2288–2295.

Gessain, A., Boeri, E., Yanahigara, R., Gallo, R. C., and Franchini, G. (1993a). Complete nucleotide sequence of a highly divergent HTLV-I variant from Melanesia. Genetic and phylogenetic relationship to HTLV-I strains from other geographical regions. *J. Virol.* **67**, 1015–1023.

Gessain, A., Hervé, V., Jeannel, D., Garin, B., Mathiot, C., and de Thé, G. (1993b). HTLV-1 but not HTLV-2 found in Pygmies from Central African Republic. *J. Acquir. Immune Defic. Syndr.* **6**, 1373–1375.

Gessain, A., Koralnik, I. J., Fullen, J., Boeri, E., Mora, C., Blank, A., Salazar-Grueso, E. F., Kaplan, J., Saxinger, W. C., Davidson, M., Lairmore, M. D., Levine, P., and Franchini, G. (1994a). Phylogenetic study of ten new HTLV-I strains from the Americas. *AIDS Res. Hum. Retroviruses* **10**, 103–106.

Gessain, A., Tuppin, P., Kazanji, M., Cosnefroy, J. Y., Georges-Courbor, M. C., Georges, A., and de Thé, G. (1994b). A distinct molecular variant of HTLV-IIb in Gabon, Central Africa. *AIDS Res. Hum. Retroviruses* **10**, 753–755.

Gessain, A., Mahieux, R., and de Thé, G. (1995a). HTLV-I "indeterminate" Western blot patterns observed in sera from tropical regions: The situation revisited. *J. AIDS Hum. Retrovirol.* **9**, 316–318.

Gessain, A., Mauclère, P., Froment, A., Biglione, M., Le Hesran, J. Y., Tekaia, F., Millan, J., and de Thé, G. (1995b). Isolation and molecular characterization of a human T lymphotropic virus type II, subtype B, from a healthy Pygmy living in a remote area of Cameroon: An ancient origin for HTLV-II in Africa. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4041–4045.

Gessain, A., Malet, C., Robert-Lamplin, J., Armelle, L., David, P., Chichlo, B., Sousova, O., Stepina, V., Gurtsevitch, V., Tortevoye, P., Hubert, A., and de Thé, G. (1996a). Serological evidence of HTLV-I but not HTLV-II infection in ethnic groups of northern and eastern Siberia. *J. Acquir. Immune Defic. Syndr.* **11**, 413–414.

Gessain, A., and de Thé, G. (1996b). What is the situation of human T cell lymphotropic virus type II (HTLV-II) in Africa? Origin and dissemination of genomic subtypes. *J. Acquir. Immune Defic. Syndr.*, in press.

Giri, A., Markharn, P., Digilio, L., Hurteau, G., Gallo, R.C., and Franchini, G. (1994). Isolation of a novel simian T-cell lymphotropic virus from *Pan paniscus* that is distantly related to the human T-cell leukemia/lymphotropic virus types I and II. *J. Virol.* **68**, 8392–8395.

Gonzalez-Dunia, D., Grimber, G., Briand, P., Brahic, M., and Ozden, S. (1992). Tissue expression pattern directed in transgenic mice by the LTR of an HTLV-I provirus isolated from a case of Tropical spastic paraparesis. *Virology* **187**, 705–710.

Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J., and Wain Hobson, S. (1989). HIV isolates are rapidly evolving quasispecies: Evidence for viral mixtures and preferred nucleotide substitution. *J. AIDS* **2**, 344–352.

Goubau, P., Desmyter, J., Ghesquiere, J., and Kasereka, B. (1992). HTLV-II among Pygmies. *Nature* **359**, 201.

Goubau, P., Liu, H.-F., de Lange, G. G., Vandamme, A.-M., and Desmyter, J. (1993). HTLV-II seroprevalence in Pygmies across Africa since 1970. *AIDS Res. Hum. Retroviruses* **9**, 709–713.

Goubau, P., van Brussel, M., Vandamme, A.-M., Liu, H-F., and Desmyter, J. (1994). A primate T-lymphotropic virus, PTLV-I, different from human T-lymphotropic viruses types I and II, in a wild-caught baboon (*Papio hamadryas*). *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2848–2852.

Goubau, P., Vandamme, A.-M., and Desmyter, J. (1996). Questions on the evolution of primate T-lymphotropic viruses raised by molecular and epidemiological studies of divergent strains. *J. Acquir. Immune Defic. Syndr.*, in press.

Gout, O., Baulac, M., Gessain, A., Semah, F., Saal, F., Périès, J., Cabrol, C., Foucault-Fretz, C., Laplane, D., Sigaux, F., and de Thé, G. (1990). Rapid development of myelopathy after HTLV-I infection acquired by transfusion during cardiac transplantation. *New Engl. J. Med.* **322**, 383–388.

Gray, G. S., Bartman, T., and White, M. (1987). Nucleotide sequence of the core (*gag*) gene from HTLV-I isolate MT2. *Nucleic Acids Res.* **17**, 7998.

Gray, G. S., White, M., Bartman, T., and Mann, D. (1990). Envelope gene sequence of HTLV-I isolate MT-2 and its comparison with other HTLV-I isolates. *Virology* **177**, 391–395.

Gurtsevitch, V., Senyuta, N., Shih, J., Stepine, V., Pavlish, O., Syrtsev, A., Susova, O., Yakovleva, L, Scherbak, L., and Hayami, M. (1995). HTLV-I infection among Nivkhi people in Sakhalin. *Int. J. Cancer* **60**, 432–433.

Hall, W. H., Takahashi, H., Zhu, S. W., and Ijichi, S. (1992b). Human T cell leukemia virus, type II (HTLV-II). *In* "Focus on HIV," pp. 209–226. Brocket Hall, United Kingdom.

Hall, W. W., Kubo, T., Tjichi, S., Takahashi, H., and Zhu, S. W. (1994a). Human T cell leukemia/lymphoma virus, type II (HTLV-II): Emergence of an important newly recognized pathogen. *Semin. Virol.* **5**, 165–178.

Hall, W. H., Takahashi, H., Liu, C., Kaplan, M. H., Scheewing, O., Ijichi, S., Nagashima, K., and Gallo, R. C. (1992a). Multiple isolates and characteristics of human T-cell leukemia virus type II. *J. Virol.* **66**, 2456–2463.

Hall, W. W., Zhu, S. W., Horal, P., Furuta, Y., Zagaany, G., and Vahlne, A. (1994b). HTLV-II infection in Mongolia (abstract 2). *In* "Program and Abstracts from the 6th International Conference on Human Retrovirology, Absecon, New Jersey." *AIDS Res. Hum. Retroviruses* **10**, 443. [Abstract]

Hashimoto, K., Lalkaka, J., Fujisawa, J.-I., Singhal, B. S., Machigashira, K., Kubota, R., Suehara, M., Osame, M., and Yoshida, M. (1993). Limited sequence divergence of HTLV-I of Indian HAM/TSP patients from a prototype Japanese isolate. *AIDS Res. Hum. Retroviruses* **9**, 495–497.

Hayami, M., Komuro, A., Nozawa, K., Shotake, T., Ishikawa, K., Yamamoto, K., Ishida, T., Honjo, S., and Hinuma, Y. (1984). Prevalence of antibody to adult T-cell leukemia virus-associated antigens (ATLA) in Japanese monkeys and other non-human primates. *Int. J. Cancer* **33**, 179–183.

Heneine, W., Woods, T., Green, D., Fukuda, K., Giusti, R., Castillo, L., Armien, B., Gracia, F., Kaplan, J. E. (1992). Detection of HTLV-II in breast milk of HTLV-II infected mothers. *Lancet* **340**, 1157–1158.

Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K. I., Shirakawa, S., and Miyoshi, I. (1981). Adult T-cell leukemia: Antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6476–6480.

Hjelle, B., Zhu, S. W., Takahashi, H., Ijichi, S., and Hall, W. W. (1993). Endemic human T cell leukemia virus type II infection in Southwestern US Indians involves two prototype variants of virus. *J. Infect. Dis.* **168**, 737–740.

Ibrahim, F., de Thé, G., and Gessain, A. (1995). Isolation and characterization of a new Simian T cell leukemia virus type I from naturally infected Celebes Macaques (*Macaca tonkeana*): Complete nucleotide sequence and phylogenetic relationship with the Australo-Melanesian HTLV-I. *J. Virol.* **69**, 6980–6993.

Igarashi, T., Yamashita, M., Miura, T., Osei-Kwasi, M., Aysi, N. K., Shiraki, H., Kurimura, T., and Hayami, M. (1993). Isolation and genomic analysis of human T lymphotropic virus type II from Ghana. *AIDS Res. Hum. Retroviruses* **9**, 1039–1042.

Ijichi, S., Zaninovic, V., Leon-S., F. E., Katahira, Y., Sonoda, S., Miura, T., Hayami, M., and Hall, W. W. (1993). Identification of human T cell leukemia virus type IIb infection in the Wayu, an aboriginal population of Colombia. *Jpn. J. Cancer Res.* **84**, 1215–1218.

Imamura, J., Tsujimoto, A., Ohta, Y., Hirose, S., Shimotohno, K., Miwa, M., and Miyoshi, I. (1988). DNA blotting analysis of human retroviruses in cerebro-spinal fluid of spastic paraparesis patients: The viruses are identical to human T-cell leukemia virus type-1 (HTLV-I). *Int. J. Cancer* **42**, 222–224.

Ishak, R., Azevedo, V. N., Harrington Jr., W. J., Eiraku, N., Ishak, M. O. G., Guerreiro, J. F., Santos, S. B., Kubo, T., Monken, C., and Hall, W. W. (1995). Identification of human T-cell lymphotropic virus type IIa (HTLV-IIa) infection in the Kayapo, an indigenous population of Brazil. *AIDS Res. Hum. Retroviruses* **11**, 813–821.

Ishida, T., Yamamoto, K., Kaneko, R., Tokita, E., and Hinuma, Y. (1983). Seroepidemiological study of antibodies to adult T-cell leukemia virus-associated antigen (ATLA) in free-ranging Japanese monkeys (*Macaca fuscata*). *Microbiol. Immunol.* **27**, 297–301.

Ishikawa, K., Fukasawa, M., Tsujimoto, H., Else, J. G., Isahakia, M., Ubhi, N. K., Ishida, T., Takenaka, O., Kawamoto, Y., Shotake, T., Ohsawa, H., Ivanoff, B., Cooper, R.W., Frost, E., Grant, F. C., Spriatna, Y., Sutarman, Y., Abe, K., Yamamoto, K., and Hayami, M. (1987). Serological survey and virus isolation of simian T-cell leukemia/T-lymphotropic virus type I (STLV-I) in non-human primates in their native countries. *Int. J. Cancer* **40**, 233–239.

Josephs, S. F., Wong-Staal, F., Manzari, V., Gallo, R. C., Sodroski, J. G., Trus, M. D, Perkins, D., Patarca, R., and Haseltine, W. A. (1984). Long terminal repeat structure of an American isolate of type I human T cell leukemia virus. *Virology* **139**, 340–345.

Kalyanaraman, V. S., Sarngadharan, M. G., Robert-Ciuroff, M., Miyoshi, I., Blayney, D., Golde, D., and Gallo, R. C. (1982). A new subtype of human T-cell leukemia virus (HTLV-II) associated with a T-cell variant of hairy cell leukemia. *Science* **218**, 571–573.

Kaplan, J. E., and Khabbaz, R. F. (1993). The epidemiology of human T-lymphotropic virus types I and II. *Rev. Med. Virol.* **3**, 137–148.

Kaplan J. E., Holland, M. U., Green, D. B., Gracia, F., and Reeves, W. C. (1993). Failure to detect human T-lymphotropic virus antibody in wild-caught new world primates. *Am. J. Trop. Med. Hyg.* **49**, 236–238.

Katamine, S., Moriuchi, R., Yamamoto, T., Terada, K., Eguchi, K., Tsuji, Y. *et al.* (1994). HTLV-I proviral DNA in umbilical cord blood of babies born to carrier mothers. *Lancet* **343**, 1326–1327.

Kinoshita, T., Tsujimoto, A., and Shimotohno, K. (1991). Sequence variations in LTR and *env* regions of HTLV-I do not discriminate between the virus from patients with HTLV-I-associated myelopathy and adult T-cell leukemia. *Int. J. Cancer* **47**, 491–495.

Kira, J., Koyanagi, Y., Yamada, T., *et al.* (1994). Sequence heterogeneity of HTLV-1 proviral DNA in the central nervous system of patients with HTLV-1-associated myelopathy. *Ann. Neurol.* **36**, 149–156.

Komurian, F., Pelloquin, F., and de Thé, G. (1991). In vivo genomic variability of human T cell leukemia virus type I depends more upon geography than upon pathologies. *J. Virol.* **65**, 3770–3778.

Komurian-Pradel, F., Pelloquin, F., Sonoda, S., and de Thé, G. (1992). Geographical subtypes demonstrated by RFLP following PCR in the LTR region of HTLV-I. *AIDS Res. Hum. Retroviruses* **8**, 429–434.

Komuro, A., Watanabe, T., Miyoshi, I., Hayami, M., Tsujimoto, H., Seiki, M., and Yoshida, M. (1984). Detection and characterization of simian retroviruses homologous to human T-cell leukemia virus type I. *Virology* **138**, 373–378.

Koralnik, I. (1996). Genomic structure of HTLV-I. *In* "Human T cell Lymphotropic Virus Type 1" (P. Höllsberg, ed.), in press. John Wiley & Sons, England.

Koralnik, I., Boeri, E., Gessain, A., Klotman, M., Lo Monico, A., Berneman, Z., and Franchini, G. (1992). Protein isoforms encoded by the pX region of human T cell leukemia/lymphotropic virus type I. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8813–8817.

Koralnik, I. J., Fullen, J., and Franchini, G. (1993). The $p12^{I}$, $p13^{II}$ and $p30^{II}$ proteins encoded by human T-cell leukemia/lymphotropic virus type I open reading frames I and II are localized in three different cellular compartments. *J. Virol.* **67**, 2360–2366.

Koralnik, I. J., Boeri, E., Saxinger, W. C., Lo Monico, A., Fullen, J., Gessain, A., Guo, H.-G., Gallo, R. C., Markham, P., Kalyanaraman, V., Hirsch, V., Allan, J., Murphy, K., Alford, P., Slattery, J. P., O'Brien, S. J., and Franchini, G. (1994). Phylogenetic associations of human and simian T-cell leukemia/lymphotropic virus type I strains: Evidence for interspecies transmission. *J. Virol.* **68**, 2693–2707.

Kwok, S., Kellogg, D., Ehrlich, G., Poiesz, B., Bhagavati, S., and Sninskyn J. J. (1988). Characterization of a sequence of human T-cell leukemia virus type 1 from a patient with chronic progressive myelopathy. *J. Infect. Dis.* **158**, 1193–1197.

Lairmore, M. D., Jacobson, S., Gracia, F., De, B. K., Castillo, L., Larreategui, M., Roberts, B. D., Levine, P. H., Blattner, W. A., and Kaplan, J. E. (1990). Isolation of human T-cell lymphotropic virus type 2 from Guaymi Indians in Panama. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8840–8844.

Lal, R., Rudolph D. L., Nerurkar, V., and Yanagihara, R. (1992). Humoral responses to the immunodominant *gag* and *env* epitopes of human T-lymphotropic virus type I among Melanesians. *Viral Immunol.* **4**, 265–272.

Lal, R. B., Gongora-Biachi, R. A., Pardi, D., Switzer, W. M., Goldman, I., and Lal, A. A. (1993). Evidence for mother-to-child transmission of human T lymphotropic virus type II. *J. Infect. Dis.* **168**, 586–591.

Lal, R. B., Owen, S. M., Rudolph, D., and Levine, P. H. (1994). Sequence variation within the immunodominant epitope-coding region from the external glycoprotein of human T lymphotropic virus type II in isolates from Seminole Indians. *J. Infect. Dis.* **169**, 407–411.

Lee, H., Swanson, P., Shorty, V. S., Zack, J. A., Rosenblatt, J. D., and Chen, I. S. (1989). High rate of HTLV-II infection in seropositive IV drug abusers from New Orleans. *Science* **244**, 471–475.

Lee, H., Idler, K. B., Swanson, P., Aparicio, J. J., Chin, K. K., Lax, J. P., Nguyen, M., Mann, T., Leckie, G., Zanetti, A., Marinucci, G., Chen, I. S. Y., and Rosenblatt, J. D. (1993). Complete nucleotide sequence of HTLV-II isolate NRA: Comparison of envelope sequence variation of HTLV-II isolates from U.S. blood donors and U.S. and Italian IV drug users. *Virology* **196**, 57–69.

Lenz, J., Celander, D., Crowther, R. L., Patarca, R., Perkins, D. W., and Haseltine, W. A. (1984). Determination of the leukaemogenicity of a murine retrovirus by sequences within the long terminal repeat. *Nature* **308**, 467–470.

Levine, P. H., Jacobson, S., Elliot, R., Cavallero, A., Colclough, G., Dorry, C., Stephson, C., Knigge, R. M, Drummond, J., Nishimura, M, Taylor, M. E., Wiktor, S., and Shaw, G. M. (1993). HTLV-II infection in Florida Indians. *AIDS Res. Hum. Retroviruses* **9**, 123–127.

Li, Y., Golemis, E., Hartley, J. W., and Hopkins, N. (1987). Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides. *J. Virol.* **61**, 693–700.

Liu, H.-F., Vandamme, A.-M., Van Brussel, M., Desmyter, J., and Goubau, P. (1994a). New retroviruses in human and simian T-lymphotropic viruses. *Lancet* **344**, 265–266.

Liu, H-F., Vandamme, A.-M., Kazadi, K., Carton, H., Desmyter, J., and Goubau, P. (1994b). Familial transmission and minimal sequence variability of human T-lymphotropic virus type I (HTLV-I) in Zaire. *AIDS Res. Hum. Retroviruses* **10**, 1135–1142.

Mahieux, R., Gessain, A., Dandelot, J., Vitrac, D., Hubert, A., Truffert, A. Cnudde, F., Musenger, C., and de Thé, G. (1994). Seroepidemiology, viral isolation and molecular characterization of HTLV-I from the Reunion island, Indian Ocean. *AIDS Res Hum. Retroviruses* **10**, 745–753.

Mahieux, R., de Thé, G., and Gessain, A. (1995a) The *tax* mutation at nucleotide 7959 of human T-cell leukemia virus type 1 (HTLV-1) is not associated with tropical spastic paraparesis/HTLV-1-associated with myelopathy but is linked to the Cosmopolitan molecular genotype. *J. Virol.* **69**, 5925–5927.

Mahieux, R., Ibrahim, F., Mauclère, P., *et al.* (1995b). Isolation and molecular characterization of HTLV-I variants from 8 Pygmies living in remote areas of Cameroon and Central African Republic: Phylogenetic comparison with 52 new isolates from central, west and south Africa. *J. Acquir. Immune Defic. Syndr.* **10**. [Abstract]

Major, M. E., Nightingale, S., and Desselberger, U. (1993). Complete sequence conservation of the human T cell leukaemia virus type 1 *tax* gene within a family cluster showing different pathologies. *J. Gen. Virol.* **74**, 2531–2537.

Malik, K. T., Even, J., and Karpas, A. (1988). Molecular cloning and complete nucleotide sequence of an adult T cell leukaemia virus/human T cell leukaemia virus type I (ATLV/HTLV-I) isolate of Caribbean origin: Relationship to other members of the ATLV/HTLV-I subgroup. *J. Gen. Virol.* **69**, 1695–1710.

Maloney, E. M., Biggar, R. J., Nell, J. V., Taylor, M. E., Hahn, B. H., Shaw, G. M., and Blattner, W. A. (1992). Endemic human T-cell lymphotropic virus II infection among isolated Brazilian Amerindians. *J. Infect. Dis.* **166**, 100–107.

Manns, A., and Blattner, W. A. (1990). The epidemiology of the human T-cell lymphotropic virus type I and type II: Etiologic role in human disease. *Transfusion* **31**, 67–75.

Mauclère, P., Gessain, A., Garcia-Calleja, J. M., Le Hesran, J. Y., Salla, R., Rehle, T., and de Thé, G. (1993). HTLV-II in African prostitutes from Cameroon. *AIDS* **10**, 1394–1395.

Mauclère, P., Mahieux, R., Garcia-Calleja, J. M., Salla, R., Tekaia, F., Millan, J., de Thé, G., and Gessain, A. (1995a). A new HTLV-II subtype a isolate in an HIV-1 infected prostitute from Cameroon, Central Africa. *AIDS Res. Hum. Retroviruses* **11**, 889–993.

Mauclère, P., le Hesran, J. Y., Mahieux, R., Salla, R., Millan, J., de Thé, G., and Gessain, A. (1995b). HTLV like gag indeterminate western blot pattern in Cameroon revealed major differences from typical HTLV-I infection in epidemiological determinants including age, sex and geoclimatical repartition. *J. Acquir. Immune Defic. Syndr.* **10**. [Abstract]

Meyerhans, A., Cheynier, R., Albert, J. Martina, S., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B., and Wain Hobson, S. (1989). Temporal fluctuation in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58**, 901–910.

Meytes, D, Schochat, B., Lee, H., Nadel, G., Sidi, Y., Cerney, M., Swanson, P., Shaklai, M., Kilim, Y., Elgat, M., Chin, E., Danon, Y., and Rosenblatt, J. D. (1990). Serological and molecular survey for HTLV-I infection in a high-risk Middle Eastern group. *Lancet* **336**, 1533–1535.

Miura, T., Fukunaga, T., Igarashi, T., Yamashita, M., Ido, E., Funahashi, S.-I., Ishida, T., Washio, K., Ueda, S., Hashimoto, K.-I, Yoshida, M., Osame, M., Singhal, B. S., Zaninovic, V., Cartier, L., Sonoda, S., Tajima, K., Ina, Y., Gojobori, T., and Hayami, M. (1994). Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1124–1127.

Miyoshi, I., Kubonishi, I., Yoshimoto, S., Akagi, T., Ohtsuki, Y., Shiraishi, Y., Nagata, K., and Hinuma, Y. (1981). Type C virus particles in a cord T-cell line derived by cocultivating normal human cord leukocytes and human leukaemic T cells. *Nature* **294**, 770–771.

Miyoshi, I., Fujishita, M., Taguchi, H., Matsubayashi, K., Miwa, N., and Y. Tanioka. (1983). Natural infection in non-human primates with adult T-cell leukemia virus or a closely related agent. *Int. J. Cancer* **33**, 333–336.

Moynet, D., Cosnefroy, J. Y., Bedjabaga, I., Roelants, G., Georges-Courbot, M. C., and Guillemain, B. (1995). Identification of new genetic subtypes of human T-cell leukemia virus type I in Gabon from encoding sequence of the surface envelope glycoprotein. *AIDS Res. Hum. Retroviruses* **11**, 1407–1411.

Mueller, N. (1991). The epidemiology of HTLV-I infection. *Cancer Causes Control* **2**, 37–52.

Neel, J. V., Biggar, R. J., and Sukarnik, R. I. (1994). Virologic and genetic studies relate Amerind origins to the indigenous people of the Mongolia/Manchuria/southeastern Siberia region. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10737–10741.

Nei, M., Takezaki, N., and Sitnikova, T. (1995). Assessing molecular phylogenies. *Science* **267**, 253–255.

Nerurkar, V. R., Miller, M. A., Leon-Monzon, M. E., *et al.* (1992). Failure to isolate human T-cell lymphotropic virus type I and to detect variant-specifte genomic sequences by polymerase chain reaction in Melanesians with indeterminate Western immunoblot. *J. Gen Virol.* **73**, 1805–1810.

Nerurkar, V. R., Babu, P. G., Song, K.-J., Melland, R. R., Gnanarnuthu, C., Saraswathi, N. K., Chandy, M., Godec, M. S., John, T. J., and Yanagihara, R. (1993a). Sequence analysis of human T cell lymphotropic virus type I strains from southern India: Gene amplification and direct sequencing from whole blood blotted onto filter paper. *J. Gen. Virol.* **74**, 2799–2805.

Nerurkar, V. R., Song, K.-J., Saitou, N., Melland, R. R., and Yanagihara, R. (1993b). Interfamilial and intrafamilial genomic diversity and molecular phylogeny of human T-cell lymphotropic virus type I from Papua New Guinea and the Solomon Islands. *Virology* **196**, 506–513.

Nerurkar, V. R., Song, K-J., Bastian, I.B., Garin, B., Franchini, G., and Yanagihara, R. (1994a). Genotyping of human T cell lymphotropic virus type I using Australo-Melanesian topotype-specific oligonucleotide primer-based polymerase chain reaction: Insights into viral evolution and dissemination. *J. Infect. Dis.* **170**, 1353–1360.

Nerurkar, V. R., Song, K.-J., Melland, R. R., and Yanagihara, R. (1994b). Genetic and phylogenetic analyses of human T-cell lymphotropic virus type I variants from Melanesians with and without spastic myelopathy. *Mol. Neurobiol.* **8**, 155–173.

Nerurkar, V. R., Achiron, A., Song, K. J., Melland, R. R., Pinhas-Hamiel, O., Melamed, E., Shohat, B., and Yanagihara, R. (1995). Human T-cell lymphotropic virus type I in Iranian-born Mashhadi Jews: Genetic and phylogenetic evidence for common source of infection. *J. Med. Virol.* **45**, 361–366.

Niewiesk, S., Daenke, S., Parker, C. E., Taylor, G., Weber, J., Nightingale, S., and Bangham, C. R. M. (1994). The transactivator gene of human T-cell leukemia virus type I is more variable within and between healthy carriers than patients with tropical spastic paraparesis. *J. Virol.* **68**, 6778–6781.

Niewiesk, S., Daenke, S., Parker, C. E., Taylor, G. Weber, J., Nightingale, S., and Bangham, C. R. M. (1995). Naturally occurring variants of human T-cell leukemia virus type I *tax* protein impair its recognition by cytotoxic T lymphocytes and the transactivation function of *tax*. *J. Virol.* **69**, 2649–2653.

Oguma, S., Imamura, Y., Kusumoto, Y., Nishimura, Y., Yamaguchi, K., Takatsuki, K., *et al.* (1992). Accelerated declining tendency of human T-cell leukemia virus type I carrier rates among younger blood donors in Kumamoto, Japan. *Cancer Res.* **52**, 2620–2623.

Osame, M., Usuki, K., Izumo, S., Izichi, N., Arnitani, H., Igata, A., Matsumoto, M., and Tara, M. (1986). HTLV-I associated myelopathy, a new clinical entity. *Lancet* i, 1031–1032.

Osame, M., Matsumoto, M., Usuki, K., Izumo, S., Ijichi, N., Amitani, H., Tara, M., and Igata, A. (1987). Chronic progressive myelopathy associated with elevated antibodies to Human T lymphotropic virus type I and adult T-cell leukemia-like cells. *Ann. Neurol.* **21**, 117–122.

Paine, E., Garcia, J., Philpott, T. C., Shaw, G., and Ratner, L. (1991). Limited sequence variation in human T-lymphotropic virus type 1 isolates from North American and African Patients. *Virology* **182**, 111–123.

Pardi, D., Kaplan, J. E., Coligan, J. E., Folks, T. M., and Lal, R. B. (1993a). Identification and characterization of an extended tax protein in human T-cell lymphotropic virus type II subtype b isolates. *J. Virol.* **67**, 7663–7667.

Pardi, D., Switzer, W. M., Hadlock, K. G., Kaplan, J. E., Lal, R. B., and Folks, T. M. (1993b). Complete nucleotide sequence of an Amerindian human T-cell lymphotropic virus type II (HTLV-II) isolate: Identification of a variant HTLV-II subtype b from a Guaymi Indian. *J. Virol.* **67**, 4659–4664.

Picard, F. J., Coulthart, M. B., Oger, J., King, E. E., Kim, S., Arp, J., Rice, G. P. A., and Dekaban, G. A. (1995). Human T-lymphotropic virus type-I (HTLV-I) in coastal natives of British Columbia: Phylogenetic affinities and possible origins. *J. Virol.* **69**, 7248–7256.

Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D., and Gallo, R. C. (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7415–7419.

Poiesz, B. J., Ruscetti, F. W., Reitz, M. S., Kalyanaraman, V. S., and Gallo, R. C. (1981). Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sézary T-cell leukaemia. *Nature* **294**, 268–271.

Poiesz, B., Sherman, M., Saksena, N., Dube, D., Dube, S., Gavalchin, J., Fan, N., Lane, M., and Paul, B. (1993). The biology and epidemiology of the human T-cell lymphoma/leukemia viruses. *In* "Frontiers of Infectious Diseases" (H. C. Neu, J. A. Levy, and R. A. Weiss, eds.), pp. 189–205. Churchill-Livingston, London.

Popovic, M., Reitz Jr., M. S., Sarngadharan, M. G., Robert-Guroff, M., Kalyanaraman, V. S., Nakao, Y., Miyoshi, I., Minowada, J., Yoshida, M., Ito, Y., and Gallo, R. C. (1982). The virus of Japanese adult T-cell leukaemia is a member of the human T-cell leukaemia virus group. *Nature* **300**, 63–66.

Rassart, E., Nelbach, L., and Jolicocur, P. (1986). Cas-Br-E murine leukemia virus: Sequencing of the palalytogenic region of its genome and derivation of specific probes to study its origin and the structure of its recombinant genomes in leukemic isolates. *J. Virol.* **60**, 910–919.

Ratner, L., Josephs, S., Starcich, B., Hahn, B., Shaw, G M., Gallo, RC., and Wong-Staal, F. (1985). Nucleotide sequence analysis of a variant human T cell leukemia virus (HTLV-Ib) provirus with a deletion in pX-I. *J. Virol.* **54**, 781–790.

Ratner, L., Philpott, T., and Trowbridge, D. B. (1991). Nucleotide sequence analysis of

isolates of human T-lymphotropic virus type I of diverse geographical origins. *AIDS Res. Hum. Retroviruses* **7**, 923–941.

Renjifo, B., Borrero, I., and Essex, M. (1995). *Tax* mutation associated with tropical spastic paraparesis/human T-cell leukemia virus type I-associated myelopathy. *J. Virol.* **69**, 2611–2616.

Saito, M., Furukawa, Y., Kubota, R., *et al.* (1995). Frequent mutation in pX region of HLV-1 is observed with HAM/TSP patients, but it is not specifically associated with the central nervous system lesions. *J. Neurovirol.* 1, 286–294.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evolution* **4**, 406–425.

Saksena, N. K., Sherman, M. P., Yanagihara, R., Dubs, D. K., and Poiesz, B. J. (1992). LTR sequence and phylogenetic analyses of a newly discovered variant of HTLV-I isolated from the Hagahai of Papua New Guinea. *Virology* **189**.

Saksena, N. K., Hervé, V., Sherman, M. P., Durand, J. P., Mathiot, C., Müller, M., Love, J. L., LeGuuenno, J. L., Barré-Sinoussi, F., Dube, D. K., and Poiesz, B. J. (1993). Sequence and phylogenetic analyses of a new STLV-I from a naturally infected tantalus monkey from Central Africa. *Virology* **192**, 312–320.

Saksena, N. K., Hervé, V., Durand, J. P., LeGuenno, B., Diop, O. M., Digoutte, J. P., Mathiot, C., Muller, M. C., Love, J. L., Dube, S., Sherrnan, M. P., Benz, P. M., Erensoy, S., Galat-Luong, A., Galat, G., Paul, B., Dube, D. K., Barré-Sinoussi, F., and Poiesz, B. J. (1994). Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukemia viruses (STLV-I) from various naturally infected monkey species from Central and Western Africa. *Virology* **198**, 297–310.

Salemi, M., Cattaneo, E., Casoli, C., and Bertazzoni, U. (1995). Identification of IIa and IIb molecular subtypes of human T-cell lymphotropic virus type II among Italian injecting drug users. *J. Acquir. Immune Defic. Syndr.* **8**, 516–520.

Salemi, M., Vandamme, A.-M., Guano, F., Gradozzi, C., Cattaneo, E., Casoli, C., and Bertazzoni, U. (1996). Complete nucleotide sequence of the Italian human T-cell lymphotropic virus type II isolate Gu and phylogenetic identification of a possible origin of South European epidemics. *J. Gen. Virol.* **77**, 1193–1201.

Sandler, S. G., Fang, C. T., and Williams, A. E. (1991). Human T-cell lymphotropic virus type I and II in transfusion medicine. *Transfusion Med. Rev.* **5**, 93–107.

Schätzl, H., Tschikobava, M., Rose, D., Voevodin, A., Nitschko, H., Singers E., Busch, U., von der Helm, K., and Lapin, B. The Sukhumi primate monkey model for viral lymphomogenesis: High incidence of lymphomas with presence of STLV-I and EBV-like virus. *Leukemia* **7**, 86–92.

Schulz, T. F., Calabro, M. L., Hoad, J. G., Carrington, C. V. F., Matutes, E., Catovsky, D., and Weiss, R. (1991). HTLV-I envelope sequences from Brazil, the Caribbean, and Romania: Clustering of sequences according to geographic origin and variability in an antibody epitope. *Virology* **184**, 483–491.

Seiki, M., Hattori, S., and Yoshida, M. (1982). Human T cell leukemia virus: Molecular cloning of the provirus DNA and the unique terminal struture. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6899–6902.

Seiki, M., Hattori, S., Hirayama, Y., and Yoshida, M. (1983). Human adult T-cell leukemia virus: Complete nucleotide sequence of the provirus genome integrated in leukemia cell DNA. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3618–3622.

Sherman, M. P., Saksena, N. K., Dube, D. K., Yanagihara, R., and Poiesz, B. J. (1992). Evolutionary insights on the origin of human T-cell lymphoma/leukemia virus type I (HTLV-I) derived from sequence analysis of a new HTLV-I variant from Papua New Guinea. *J. Virol.*, 2556–2563.

Sherman, M. P., Dube, S., Spicer, T. P., Kane, T. D., Love, J. L., Saksena, N. K., Iannone, R.,

Gibbs, C. J., Yanagihara Jr., R., Dube, D. K., and Poiesz, B. J. (1993). Sequence analysis of an immunogenic and neutralizing domain of the human T-cell lymphoma/leukemia virus type I gp46 surface membrane protein among various primate T-cell lymphoma/leukemia virus isolates including those from a patient with both HTLV-I-associated myelopathy and adult T-cell leukemia. *Cancer Res.* **53**, 6067–6073.

Shimotohno, K., Takashi, Y., Shimizu, N., Gojobori, T., Golde, D. W., Chen, I. S. Y., Miwa, M., and Sugimura, T. (1985). Complete nucleotide sequence of an infectious clone of human T cell leukemia virus type II: An open reading frame for the protease gene. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3101–3105.

Shirabe, S., Nakamura, T., Tsujihata, M., Nagataki, S., Seiki, M., and Yoshida, M. (1990). Retrovirus from human T-cell leukemia virus type I associated melopathy is the same strain as a prototype human T-cell leukemia virus type I. *Arch. Neurol.* **47**, 1258–1260.

Song, K.-J., Nerurkar, V. R., Saitou, N., Lazo, A., Blakeslee, J. R., Miyoshi, I., and Yanagihara, R. (1994). Genetic analysis and molecular phylogeny of simian T-cell lymphotropic virus type I: Evidence for independent virus evolution in Asia and Africa. *Virology* **199**, 56–66.

Song, K.-J., Nerurkar, V. R., Pereira-Cortez, A. J., Yamamoto, M., Taguchi, H., Miyoshi, I., and Yanagihara, R. (1995). Sequence and phylogenetic analyses of human T cell lymphotropic virus type I from a Brazilian woman with adult T cell leukemia: Comparison with virus strains from South America and the Caribbean basin. *Am. J. Trop. Med. Hyg.* **52**, 101–108.

Switzer, W. M., Pieniazek, D., Swanson, P., Samdal, H. H., Soriano, V., Khabbaz, R. F., Kaplan, J. E., Lal, R. B., and Heneine, W. (1995a). Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. *J. Virol.* **69**, 621–632.

Switzer, W. M., Owen, S. M., Pieniazek, D. A., Newrkar, V. R., Duenas-Barajas, E., Heneine, W., and Lal, R. B. (1995b). Molecular analysis of human T-cell lymphotropic virus type II from Wayuu Indians of Colombia demonstrates two subtypes of HTLV-IIb.

Szurek, P. F., Yuen, P. H., Jerzy, R., and Wong, P. K. Y. (1988). Identification of point mutations in the envelope gene of Moloney murine leukemia virus TB temperature-sensitive paralytogenic mutant tsl: Molecular determinants for neurovirulence. *Virology* **62**, 357–360.

Tajima K., Ito S. I., and the Tsushima ATL Study Group (1990). Prospective studies of HTLV-I and associated diseases in Japan. *In* "Human Retrovirology: HTLV (W. A. Blattner, ed.), pp. 267–279. Raven Press, New York.

Tajima, K., Inoue, M., Takezaki, T., Ito, M., and Ito, S. I. (1994). Ethnoepidemiology of ATL in Japan with special reference to the Mongoloid dispersal. *In* "Adult T-cell Leukaemia" (K. Takatsuki, ed.), pp. 91–112. Oxford University Press, Oxford.

Takahashi K., Takezaki T., Oki T., Kawakami K., Yashiki S., Fujiyoshi T., *et al.* (1991). Inhibitory effect of maternal antibody on mother-to-child transmission of human T-lymphotropic virus type I. *Int. J. Cancer* **49**, 673–677.

Takahashi, H., Zhu, S. W., Ijichi, S., Vahlne, A., Suzuki, H., and Hall, W. W. (1993). Nucleotide sequence analysis of human T cell leukemia virus, type II (HTLV-II) isolates. *AIDS Res. Hum. Retroviruses* **9**, 721–732.

Takatsuki, K., Uchiayama, T., Sagawa, K., and Yodoi, J. (1977). Adult T-cell leukemia in Japan. *In* "Topics in Hematology" (S. Seno, F. Takaku, and S. Irino eds.), pp. 73–77. Excerpta Medica, Amsterdam.

Takezaki, T., Tajima, K., Komoda, H., and Imai, J. (1995). Incidence of human T lymphotropic virus type I seroconversion after age 40 among Japanese residents in an area where the virus is endemic. *J. Infec. Dis.* **171**, 559–565.

Tsujimoto, A., Teruuchi, T., Imamura, J., Shimothono, K., Miyoshi, I., and Masanao, M. (1988). Nucleotide sequence analysis of a provirus derived from HTLV-1 associated myelopathy (HAM). *Mol. Biol. Med.* **50**, 481–492.

Tuppin, P., Gessain, A., Kazanji, M., Mahieux, R., Cosnefroy, J. Y., Tekaia, F., Georges-Courbot, M. C., Georges, A., and de Thé, G. (1996). Evidence in Gabon, Central Africa, for an intrafamilial clustering with mother to child and sexual transmission of a new molecular variant of human T lymphotropic virus type-II subtype B. *J. Med. Virol.*, in press.

Ureta Vidal, A., Gessain, A, Yoshida, M, Mahieux, R., Nishioka, K., Tekaia, F., Rosen, L., and de Thé, G. (1994a). Molecular epidemiology of HTLV type I in Japan: Evidence for two distinct ancestral lineages with a particular geographical distribution. *AIDS Res. Hum. Retroviruses* **10**, 1557–1566.

Ureta Vidal, A., Gessain, A, Yoshida, M, Tekaia, F., Garin, B., Guillemain, B., Schulz, T., Farid, R., and de Thé, G. (1994b). Phylogenetic classification of human T cell leukaemia/lymphoma virus type I genotypes in five major molecular and geograhical subtypes. *J. Gen. Virol.* **75**, 3655–3666.

Van Brussel, M., Goubau, P., Rousseau, R., Desmyter, J., and Vandamme, A.-M. (1996). The genomic structure of a new simian T-lymphotropic virus, STLV-PH969, differs from that of human T-lymphotropic virus types I and II. *J. Gen. Virol.* **77**, 347–358.

Vandamme, A.-M., Liu, H.-F., Goubau, P., and Desmyter, J. (1994). Primate T-lymphotropic virus type I LTR sequence variation and its phylogenetic analysis: Compatibility with an African origin of PTLV-I. *Virology* **202**, 212–223.

Vandamme, A.-M., Liu, H.-F., van Brussel, M., de Meurichy, W., Desmyter, J., and Goubau, P. (1996). The presence of a divergent T-lymphotropic virus in a wild-caught pygmy chimpanzee (*Pan paniscus*) supports an African origin for the human T-lymphotropic/simian T-lymphotropic group of viruses. *J. Gen. Virol.* **77**, 1089–1099.

Voevodin, A., Al-Mufti, S., Farah, S., Khan, R., and Miura, T. (1995). Molecular characterization of human T-lymphotropic virus, type 1 (HTLV-1) found in Kuwait: Close similarity with HTLV-1 isolates originating from Mashhad, Iran. *AIDS Res. Hum. Retroviruses* **11**, 1255–1259.

Voevodin, A., Miura, T., Samilchuk, E., and Schätzl, H. (1996). Phylogenetic characterization of simian T lymphotropic virus type I (STLV-I) from the Ethiopian sacred baboon (*Papio hamadryas*). *AIDS Res. Hum. Retroviruses* **12**, 255.

Wain-Hobson, S. (1992). Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo. *Curr. Top. Microbiol. Immunol.* **176**, 181–193.

Watanabe, T., Seiki, M., Tsujimoto, H., Miyoshi, I., Hayarni, M., and Yoshida, M. (1985). Sequence homology of the simian retrovirus genome with human T-cell leukemia virus type I. *Virology* **144**, 59–65.

Watanabe, T., Seiki, M., Hirayama, Y., and Yoshida, M. (1986). Human T-cell leukemia virus type I is a member of the African subtype of simian viruses (STLV). *Virology* **148**, 385–388.

Yamashita, M., Takehisa, J., Miura, T., Ido, E., Becker, W. B., Robson, B. A., Becker, M. L. B., and Hayami, M. (1995). Presence of the Widespread subtype of HTLV-I in South Africa. *AIDS Res. Hum. Retroviruses* **11**, 645–647.

Yanagihara, R., Nerurkar, V. R., and Ajdukiewicz, A. B. (1991a). Comparison between strains of human T lymphotropic virus type I isolated from inhabitants of the Solomon Islands and Papua New Guinea. *J. Infect. Dis.* **164**, 443–449.

Yanagihara, R., Nerurkar, V. R., Garruto, R. M., Miller, M. A., Leon-Monzon, M. E., Jenkins, C. L., Sanders, R. C., Liberski, P. P., Alpers, M. P., and Gajdusek, D. C. (1991b). Characterization of a variant of human T lymphotropic virus type I isolated

from a healthy member of a remote, recently contacted group in Papua New Guinea. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1446–1450.

Yanagihara, R. (1994). Geographic-specific genotypes or topotypes of human T cell lymphotropic virus type I as markers for early and recent migrations of human populations. *Virus Res.* **43**, 147–186.

Yoshida, M., Miyoshi, J., and Hinuma, Y. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2031–2035.

Zella, D., Mori, L., Sala, M., *et al.* (1990). HTLV-II infection in Italian drug abusers. *Lancet* **335**, 575–576.

Zella, D., Cavicchini, A., Salemi, M., Casoli, C., Lori, F., Achilli, G., Cattaneo, E., Landini, V. and Bertazzoni, B. (1993). Molecular characterization of two isolates of human T-cell leukemia virus type I from Italian drug abusers and comparison of genome structure with other isolates. *J. Gen. Virol.* **74**, 437–444.

INDEX

C

I

J

L

ISBN 0-12-039847-8