Harren Jhoti
Andrew Leach (Eds.)

Structure-based Drug Discovery

Springer

STRUCTURE-BASED DRUG DISCOVERY

STRUCTURE-BASED DRUG DISCOVERY

Edited by

HARREN JHOTI
Astex Therapeutics, Cambridge, UK

ANDREW R. LEACH
GlaxoSmithkline, Stevenage, UK

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4406-2 (HB)
ISBN-13 978-1-4020-4406-9 (HB)
ISBN-10 1-4020-4407-0 (e-book)
ISBN-13 978-1-4020-4407-6 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

TABLE OF CONTENTS

PREFACE

The last 25 years has seen structure-based drug discovery evolve from an interesting niche activity pursued by a relatively small number of companies to being a fully integrated series of techniques that are part of the core technologies within most large pharmaceutical companies. This increase in popularity has been driven to a large extent by significant technological advances that have allowed the three-dimensional structure of a target protein to be determined in a much shorter time frame. In the 1980's it could take several years to determine the crystal structure of a key drug target; obtaining structures of bound inhibitors could consume several more months. Today, protein crystal structures may be obtained in months rather than years and subsequent protein/inhibitor complexes often only take weeks (if not days) to solve. Another key factor in the uptake of structure-based discovery methods has been the availability of crystal structures for significantly more proteins at the start of a drug discovery program. This increase in the number of protein structures has also helped the development of improved computational chemistry methods for the prediction of the binding modes of compounds and binding energies.

Successful structure-based design thus requires the synthesis of several different techniques, both experimental and theoretical. This book is intended to provide an overview of some of the more recent developments, with a particular focus on structural genomics, biophysical techniques, fragment-based approaches and computational methods. The first two chapters outline the key advances in structural biology that have addressed some of the major bottlenecks, such as protein expression and crystallisation, in the process of solving protein crystal structures. They also include a review of the structural genomics initiatives intended to obtain novel protein structures that are being pursued around the world. The subsequent five chapters describe several aspects of fragment-based discovery as a major new approach to discovering new drug molecules. The essence of this approach involves the use of biophysical techniques, such as X-ray crystallography and NMR, to screen fragments that due to their limited size and complexity typically bind the drug target with significantly lower affinity than drug-sized molecules. The potential advantages of this approach over conventional drug discovery are discussed as well as the technological

advances required to undertake high-throughput X-ray crystallography and NMR experiments on the binding of molecular fragments to proteins. The final two chapters focus on the latest developments in computational techniques that are integral to applying structure-based methods to medicinal chemistry strategies. These include methodologies for improving the success of docking compounds to protein structures and the scoring of these binding modes in order to predict the free energy of binding.

Chapter 1

FIVE YEARS OF INCREASING STRUCTURAL BIOLOGY THROUGHPUT- A RETROSPECTIVE ANALYSIS

Enrique Abola[1], Dennis D. Carlton[1], Peter Kuhn[2] and Raymond C. Stevens [1]
The Scripps Research Institute, Departments of Molecular Biology [1] and Cell Biology [2]. 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

1. INTRODUCTION

1.1 Structural biology and genomics

The completed sequencing and initial characterization of the human genome in 2001 (Lander et al 2001; Venter et al 2001) and that of other organisms such as *Drosophila melanogaster* (Adams et al 2000) and the SARS Corona Virus (Marra et al 2003), have educated us on the vast complexity of the proteome. Full genome characterization efforts highlight how critical it is to understand at a molecular level all of the protein products from multiple organisms. An important issue for addressing the molecular characterization challenge is the need to quickly and economically characterize normal and diseased biological processes in order to understand the basic biology and chemistry of the systems and to facilitate the discovery and development of new therapeutic and diagnostic protocols. In order to fully characterize the proteins at the molecular level, three-dimensional protein structure determination has proven to be invaluable, complementing biological and biochemical information from other types of experiments. Structural information is also the ultimate rational drug design tool, with the potential to save an estimated 50% of the cost of drug discovery (Stevens 2004). However, the best means

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 1–26.

by which to attain structural knowledge is a topic of controversy. The traditional approach was a complex and labor-intensive process in which one protein or complex was studied at a time. The alternative is a high throughput (HT), discovery-oriented approach wherein entire families, pathways or genomes are characterized. Benefits include the economy of scale, the speed of mass production, and a dramatic increase in discovery rates through the systematic collection and analysis of data. Prior to the late 1990's, the technologies and approaches were too slow and unreliable to allow for such larger scale analyses.

In the past, we have reviewed some of the technology developments in miniaturizing and streamlining structure determination pipelines (Stevens 2004; Abola et al 2000). For this chapter, we summarize the input and output of several structural genomics efforts that have validated new technology efforts over the first 5 years of the HT structural biology era. These technologies have been used by various HT pipelines that have contributed to the determination of over 1600 new structures, a high percentage of which were novel folds, and 70% had less than 30% identity to any other protein in the Protein Data Bank (PDB) at the time of release. As an example of the implementation of the HT pipeline, we discuss in some detail the specific approach of the Joint Center for Structural Genomics (JCSG) that we have been involved in.

1.2 Protein structure initiative

In 1999, several initiatives in Japan, Europe, and the United States were created to investigate the feasibility of HT structural biology, structural genomics, and increased throughput structure based drug dicovery (Table 1-1). For this chapter, we will focus on the efforts in the United States since we are most familiar with them, and they have recently completed the first phase of their efforts (PSI-1). However, we would be remiss in not mentioning a number of critical players in the global effort. Japan has prehaps made the largest financial global investment in this area of structural genomics at the Riken Genomics Center through the creation of the NMR Farm, and development of such innovative technologies as cell-free protien expression (Kigawa et al 2004). In Europe, several efforts in the UK (e.g. SPINE) and Germany (e.g. Protein Structure Factory) were both early movers in this area and have contributed significantly to the field.

More recent efforts included MepNet and the Structural Genomics Consortium (Toronto, Oxford, and Sweden). In addition to academic and government-led efforts, a number of structural genomics companies were set up during this period and have also contributed to the rapid growth of the

Table 1-1. History of Structural Genomics.

Feb 1995	LBNL structural genomics expression/crystallization technology development initiated
1995	Proposal of structural genomics projects in Japan
Jan 1997	The workshop on Structural Genomics (Argonne, IL, USA)
Apr 1997	Start of structural genomics pilot project at RIKEN Institute
1997	Initiating study of structural genomics at DOE and NIGMS/NIH in USA
1998	Start of the initial pilot projects in Germany, Canada, and USA
Feb 1999	Formation of the Berlin Protein Structure Factory
Feb 1999	Formation of Syrrx (previously called Agencor)
Jun 1999	Call for grant applications for NIGMS/NIH pilot projects (PSI-1)
Dec 1999	Formation of Structural GenomiX (previously called Protarch)
Dec 1999	Formation of Astex Technology
Apr 2000	First International Structural Genomics Meeting (Hinxton, UK)
Aug 2000	Formation of Affinium Pharmaceuticals (previously called Integrative Proteomics)
Sep 2000	Structural Genomics: From Gene to Structure to Function (Cambridge, UK)
Sep 2000	Start of the NIGMS Protein Structure Initiatives in USA with seven Centers
Nov 2000	1st International Conference on Structural Genomics 2000 (ICSG 2000) (Yokohama, Japan)
Apr 2001	Second International Structural Genomics Meeting (Airlie House, USA) - Start of International Structural Genomics Organization (ISGO)
Jun 2001	Formation of Plexxikon
Sep 2001	Start of the new two centers for NIGMS Protein Structure Initiatives in USA (9 total)
Mar 2002	Start of the European drive for post-genome research, Structural Proteomics in Europe (SPINE)
Apr 2002	Start of the National Project on Protein Structural and Functional Analyses in Japan
Oct 2002	2nd ISGO International Conference on Structural Genomics (ICSG 2002) (Berlin, Germany)
Early 2003	RIKEN– 100th structure solved at Riken deposited in PDB
April 2003	Formation of The Structural Genomics Consortium (SGC)
April 2004	RFA for next generation structural genomics centers in USA
November 2004	3rd ISGO International Conference on Structural Genomics (ICSG 2004) (Washington, D.C. USA)
Feb 2005	PSI 1000th structure milestone achieved
July 2005	Start of PSI-2 with 4 Large Scale and 6 Specialized Centers

field (e.g. Syrrx, Structural GenomiX, Astex Therapeutics, Affinium Pharmaceuticals, Plexxikon).

As one of the initiators of the structural genomics movement in the late 1990's, the National Institute of General Medical Sciences (NIGMS) created the Protein Structure Initiative (PSI), a national program with the long-range goal of making three-dimensional, high-resolution protein structures obtainable from knowledge of their corresponding DNA sequences. Completed in the summer of 2005, the pilot phase (referred here as PSI-1) supported a 5-year effort with 9 pilot centers throughout the U.S. to evaluate "if" HT structural biology pipelines could be established and then incorporated into scaleable production pipelines capable of solving hundreds of protein structures per year. Early in 2005 the NIGMS PSI announced its first major milestone, that the combined output of the nine PSI centers had exceeded 1,000 structures.

In Phase II (referred here as PSI-2), NIGMS is providing additional funding for four large-scale centers that will scale-up their production lines to provide another 3,000 to 5,000 structures (NIH 2005). A critical component of the second phase will be the careful target selection procedures that will be managed by the NIGMS PSI-2 Network. A part of this coordinated target selection management is the focus on biomedically relevant protein structures. In addition to the four production centers, six technology development centers have been created to continue the development of innovative technologies for the more challenging problems including studies on membrane proteins, large protein assemblies, and the more difficult eukaryotic proteins.

2. NEW TECHNOLOGIES IN PSI-1 PIPELINES

2.1 High-throughput structural determination pipeline

Nine centers successfully completed PSI-1 operations in the summer of 2005 (Table 1-2 and Table 1-3) each of which developed HT pipelines using new technologies, most of which were created before the start of PSI-1 and were critically evaluated during PSI-1. The HT structure determination pipelines covered all activities from target selection to analysis and deposition of solved structures in the PDB. Both single crystal X-ray diffraction and solution NMR structural determination approaches were used.

Table 1-2. PSI Centers.

Center	***Home Institution and website***
PSI-1 Pilot Centers	
Joint Center for Structural Genomics	The Scripps Research Institute, http://www.jcsg.org
Midwest Center for Structural Genomics	Argonne National Laboratory, http://www.mcsg.anl.gov
New York Structural GenomiX Research Consortium	Structural GenomiX, Inc., http://www.nysgrc.org
Northeast Structural Genomics Consortium	Rutgers University, http://www.nesg.org
Southeast Collaboratory for Structural Genomics -	University of Georgia, Athens, http://www.secsg.org/
Berkeley Structural Genomics Center	University of California, Berkeley, http://www.strgen.org/
Tuberculosis (TB) Structural Genomics Consortium -	Los Alamos National Laboratory, http://www.doe-mbi.ucla.edu/TB/
Structural Genomics of Pathogenic Protozoa Consortium	University of Washington, http://www.sgpp.org/
Center for Eukaryotic Structural Genomics -	University of Wisconsin, Madison, http://www.uwstructuralgenomics.org/
PSI-2 Large Scale Production Centers	
Joint Center for Structural Genomics	The Scripps Research Institute, http://www.jcsg.org
Midwest Center for Structural Genomics	Argonne National Laboratory, http://www.mcsg.anl.gov
New York Structural GenomiX Research Consortium	Structural GenomiX, Inc. http://www.nysgrc.org
Northeast Structural Genomics Consortium	Rutgers University, http://www.nesg.org
PSI-2 Specialized Centers	
Accelerated Technologies Center for Gene to 3D Structure	deCODE Biostructures and The Scripps Research Institute, http://www.atcg3d.org
Center for Eukaryotic Structural Genomics	University of Wisconsin, Madison, http://www.uwstructuralgenomics.org
Center for High-Throughput Structural Biology	Hauptman-Woodward Medical Research Institute, http://www.chtsb.org
Center for Structures of Membrane Proteins	University of California, San Francisco http://www.csmp.ucsf.edu
Integrated Center for Structure and Function Innovation	Los Alamos National Laboratory http://www.techcenter.mbi.ucla.edu
New York Consortium on Membrane Protein Structure	New York Structural Biology Center http://www.nycomps.org

Table 1-3. Production summary for PSI-1 Structural Genomics Centers based on TargetDB XML distribution file (http://targetdb.pdb.org/target_files/). The table given below was downloaded from http://olenka.med.virginia.edu/mcsg/html/recent_results.html which was last updated on August 2, 2005. Only distinct target sequences are taken into account for each center and in the total count (hence numbers of "distinct" targets reported for centers where sequences are duplicated or missing in XML files may be lower than those reported by the centers; note also that the number of targets in the total count may be less than the sum of targets for the centers due to target overlaps).

Center	All Targets	Cloned	Targets With Crystals	Diffracting Targets	Total Solved (X-ray, NMR)	Median Length
MCSG	15359	5675	838	349	281	319
JCSG	6594	3650	1166	265	226	415
NESGC	12205	5309	162	115	206	193
NYSGRC	2145	1538	388	185	185	454
TB	1756	1547	208	118	107	574
SECSG	14786	14377	223	118	76	214
BSGC	911	812	94	65	60	374
CESG	6582	4476	104	40	52	222
SGPP	19503	10154	175	45	28	200
TOTAL	74899	45189	3257	1277	1206	361

HT pipelines employed a manufacturing style approach in that responsibilities were compartmentalized by function and processes were standardized through the use of quality assurance practices such as standard operating procedures (SOP). Whenever possible, common quality control practices were employed to monitor processes and materials from beginning to end. Data was uploaded to a common database to facilitate target management, process monitoring, and regular reporting. Laboratory information management systems (Zolnai et al 2003; Bertone et al 2001) were used to manage and track experiments. A good example is the java-based SESAME system developed by the CESG group in Wisconsin. As all the projects were run as multi-institutional collaborations, specific pipeline processes were implemented in separate institutions. For example, in the case of the JCSG, steps from protein production to crystallization as well as crystal mounting were carried out at TSRI and GNF, while diffraction screening and data collection were done at the Stanford Synchrotron Radiation Laboratory (SSRL). Most of the pipelines were established as learning platforms wherein experimental results and operational experiences

were applied using a feedback loop to incrementally introduce improvements to the process.

Achievement of the 1000 structure milestone by these pipelines validates the hypothesis that structural genomics pipelines could be constructed and scaled-up. It also demonstrates the feasibility of using HT approaches for protein production, a notion that was not clear at the start of the PSI as it was generally thought that the variability in protein properties would not make them amenable to handling by simplified processes. Much remains to be done, for example expression and purification of eukaryotic proteins, some of which may require folding partners, remains to be developed and is the focus of a number of the PSI-2 specialized centers.

2.2 Pipeline technologies

A central theme in the technology development area has been automation, integration, and miniaturization of processes in the pipeline. These goals have reduced the cost per structure by decreasing time from gene to structure, material usage, and number of personnel needed to accomplish large numbers of tasks. As mentioned above, most centers developed database and software products to manage their pipeline, in addition a number of essential technologies were also developed, most of which are now in general use by the community. In this section we mention a number of notable technologies that have contributed significantly to the effort.

2.2.1 Protein Expression and Fermentation

Several innovative approaches have resulted in a marked increase in productivity and through put in expression particularly in *E. coli*. Studier from the NYSGRC has formulated growth media (Studier 2005) in which expression strains can grow uninduced to relatively high cell densities and then be induced automatically without any intervention by the experimenter. Cell densities attained in these auto-inducing cultures have produced 10-fold more target protein per volume of culture than with the standard IPTG induction protocol. Auto-induction also allows many cultures to be inoculated in parallel and induced simply by growing to saturation, making auto-induction a powerful tool for screening clones for expression and solubility in an automated setting.

Two JCSG-related innovations that greatly increased capacity were a high throughput, 96-tube *E. coli* expression system (Page et al 2004) and a scalable 96-well micro-expression device (Page et al 2004). The GNF fermentor is a production device capable of 96 simultaneous 65 mL fermentations in either native or selenomethionine (SeMet) media. Pelleted cell mass after 6 hours of growth varies from 1-3g/tube for SeMet to

3-5g/tube in native media. This device has already resulted in a more than ten-fold reduction in the culture volume required for protein production when compared to conventional expression in shaker flasks.

Our group at TSRI adapted a low-cost, high-velocity incubating Glas-Col (Glas-Col, LLC, Terre Haute, IN, USA) Vertiga shaker to develop an efficient, HT *E. coli* microliter-scale expression screening protocol which accurately predicts parameters that can be used for scale-up studies (milliliter and liter fermentation) (Page et al 2004). The apparatus shakes cultures in three-dimensions at speeds of up to 1000 rpm, allowing small-scale (~750 μL) cultures grown in 2 mL deep-well 96-well blocks to achieve optical densities (OD_{600}) as high as 10-20. This generates sufficient material for analysis of expression, solubility, binding to affinity purification matrices, and initial crystallization/NMR analysis. Moreover, this screening strategy has also been used to identify clones which express and are soluble under SeMet or $^{15}N/^{13}C$-labeled expression conditions that are necessary for the production of labeled recombinant proteins for direct structural analysis. It also provided an early quality control step in that one 96-well micro-purification step produced enough of each protein for characterization by MALDI, electrophoresis, or size exclusion chromatography.

Protein purification from *E. coli* at the JCSG has also been largely automated using GNFuge, a robot developed at GNF. The fully automated GNFuge harvests, sonicates, centrifuges, and aspirates 96 bacterial cultures in parallel. In addition, it facilitates fully automated affinity purification of tagged proteins from the resulting lysates and for insoluble proteins. An on-column refolding strategy compatible with this automation was recently implemented.

2.2.2 Crystallization

The last five years has seen rapid development and deployment of technologies and systems designed to carry out large-scale crystallization experiments. These include the use of nanoliter volumes (Santarsiero et al 2002), use of microfluidics (Hansen et al 2002), the development of rapid and large-scale crystallization imaging and storage systems (Hosfield et al 2003), and finally integration of these technologies into a complete system (e.g., CrystalMotion, available from MSC, http://www.rigakumsc.com/). The system that we developed along with a team of engineers and scientists at GNF and Syrrx is capable of performing 100,000 sitting drop experiments per-day, imaging one 96-well plate in one minute and storing and managing up to 40,000 plates in a cold room (Hosfield et al 2003). This system has been in operation for almost five years and continues to process reliable and productive experiments.

The majority of the key developments were created just prior to the start of PSI-1, but PSI-1 was critical in the validation of these advancements. Results from JCSG and other groups implementing the nanovolume crystallization technology clearly demonstrated the power of this new approach. Smaller volumes of protein allowed for the exploration of a broader universe of crystallization conditions, leading to significant costs savings, shorter crystallization times, improved crystal quality, and the successful crystallization of targets, previously difficult to achieve with larger volumes (Santarsiero et al 2002; Carter et al 2005). Capitalizing upon increases in intensity and focus of X-rays beams at modern synchrotron facilities, early JCSG-related studies showed that crystals for X-ray data collection could be reproducibly generated in volumes as low as 50nL. Although 100-200nL proved to be more practical in a production setting, an order of magnitude decrease in protein consumption was realized in the TSRI pipeline. Of critical importance is that all of these technologies are now available at "reasonable" cost to the scientific community, with young start-up labs now able to afford crystallization and imaging robotic systems.

2.2.3 X-ray Diffraction Screening and Data Collection

The process of mounting flash-cooled crystals, aligning them with the X-ray beam and evaluating and collecting their diffraction was clearly a major bottleneck for any HT structure determination pipeline. Efforts to automate these processes were in the planning and prototyping stage in 2000 when PSI-1 started (Abola et al 2000). By 2005, the start of PSI-2, the majority of beamlines used in structural genomics efforts have been automated with new robotic and software systems (McPhillips et al 2002; Cohen et al 2002; Snell et al 2004). In addition, new products became available that have been installed for use in-house. A good example is the ACTOR system from Rigaku/MSC which is the first commercially available off-the-shelf system for automatically changing samples for screening or data collection (Muchmore et al 2000).

Beamlines equipped with a crystal mounting robot can now handle hundreds of samples mounted in 96-format cassettes that can be screened in a few hours (e.g. at SRRL it takes about 5 hours to process 3 cassettes) without any human intervention. Automated crystal mounting at the beamline permits a more thorough and systematic approach to the screening process, which in turn translates into a higher structure determination success rate, as the crystal quality cannot be judged solely from their physical appearance. All diffraction data are processed in real-time to evaluate both quality and completeness. Real-time data reduction and analysis allow accurate determination of the amount of data required to solve

any given structure. Data collection is terminated once sufficient data are collected and the sample restored in liquid nitrogen. At the SSRL, all protein crystallography beamlines now have automation systems and are integrated with the Blu-Ice/DCS data collection environment (McPhillips et al 2002).

The PSI-1 also funded the development of the Compact Light Source (CLS, Lyncean Technologies Inc.) through a Small Business Innovation Research (SBIR) program. The CLS is a breakthrough technology that offers the possibility of a "synchrotron beamline" for home laboratory applications. This tunable, tabletop X-ray source can be used in much the same way as a typical X-ray beamline at a large facility; but it is small enough to bring state-of-the-art methods of macromolecular crystallography directly into an experimenter's local laboratory.

2.2.4 NMR

NMR spectroscopy is a well-established technique for protein structure determination, as well as to screen for the folded state of globular proteins (Muchmore et al 2000; Markley et al 2003; Wüthrich 2003). Since NMR spectroscopy has intrinsically low sensitivity, milligram amounts of protein are required for screening and structure determination with conventional equipment. At the beginning of PSI-1 in 2000, about 6 weeks of NMR instrument time per protein structure was considered to be a realistic estimate for ~1 mM protein samples with molecular weights up to 15 kDa. By 2005, the start of PSI-2, NMR had been successfully transformed. It is now being used for HT structural determination efforts as well as for screening protein samples to determine suitability for crystal structural studies. A total of 123 structures were determined by NMR in PSI-1 of which 91 were done at the NESG. Microcoil NMR probes had been developed for use in biomolecular NMR spectroscopy (Olson et al 1995; Peti et al 2004). Specifically, small diameter coils enable up to ten-fold (mass-based) sensitivity gain so that microgram amounts of protein are now sufficient for screening by NMR spectroscopy. At the JCSG, by the 4th year of operations, most samples were being screened for the folded state with the microcoil probe before undergoing crystallization studies and assigned a grade of A, B, C, or D (Table 1-4; Page et al 2005). Using the microcoil probe, such information could be collected with 5 μL of protein and in 5 minutes. At this time, miniaturization is primarily aimed at identifying promising targets for structure determination. This methodology effectively guides efforts to focus on targets with a high probability of success, and either eliminates poor targets or replaces them by improved constructs. Overall, this process increases the efficiency of the entire pipeline and

results in a reduction of the cost per structure. Further developments including optimized miniaturization may, at least for some proteins, lead the way directly to structure determination (Page et al 2005).

Table 1-4. Results of crystallographic studies with 79 mouse homologue proteins that were graded 'A' to 'D' based on 1D ^{1}H NMR screening.

Grade[a]	Proteins[b]	Crystal Hits[c]	> 5.0 Å Diffraction; *No Structure*[d]	< 5.0 Å Diffraction; *No Structure*[e]	Structures Solved[f]
A	24	16 (67%)	0 (0%)	4 (17%)	4 (17%)
B	26	22 (85%)	0 (0%)	1 (4%)	9 (35%)
C	22	18 (82%)	4 (18%)	5 (23%)	2 (9%)
D	7	6 (86%)	2 (29%)	2 (29%)	0 (0%)
Total	79	62	6	12	15

[a]The classification into four grades, 'A' to 'D', by 1D ^{1}H NMR screening is described in the published manuscript by (Page et al., 2005). 'A' and 'B' are proteins that are now routinely forwarded for extensive coarse and fine-screen crystallization trials, while 'C' and 'D' proteins are only subjected to coarse-screen crystallization trials.
[b] Number of proteins in each category.
[c] The number of proteins that crystallized in at least one coarse screen crystallization condition. Two 'A' proteins had been removed from the pipeline for structure determination by NMR.
[d]The number of proteins for which the best crystals diffracted to no higher than 5.0 Å.
[e]The number of proteins for which the best crystals diffracted to better than 5.0 Å, but no structure is as yet available.
[f]The number of proteins for which high resolution crystal structures have been determined.

In addition to the miniaturization efforts for NMR, cryogenic probes were evolving and becoming more robust and useful, offering approximately a 3-fold increase in sensitivity in routine applications in biological NMR spectroscopy (Monleon et al 2002), and potentially an order-of-magnitude reduction in measurement times. The use of this probe has led to the development of G-matrix Fourier Transform (GFT) NMR by the NESG Center which enables researchers to optimally adjust NMR measurement times to sensitivity requirements and allows them to take full advantage of highly sensitive cryogenic probes for HT NMR structure determination.

3. THE JCSG PROTEIN AND CRYSTAL PRODUCTION PIPELINES

3.1 Protein targets

Although target species changed over the 5 years of the project, the goal of proving the feasibility of attacking an entire genome remained the principal focus of the JCSG. *C. elegans* was chosen as the initial target set. Shortly after start of the project, it became clear that the pipeline was not ready to tackle a complete eukaryotic system and hence *T. maritima* became the principal prokaryotic genome of the JCSG. However, within the following year, the mouse genome was providing a eukaryotic source of additional protein targets. By the end of year four, approximately 70% of the total structures solved by the JCSG were proteins from *T. maritima*.

3.2 Production strategies

The JCSG adopted a three-tiered shotgun strategy for the crystallization of the *T. maritima* proteome in order to identify and focus the majority of crystallization efforts on those proteins with a demonstrated propensity to crystallize (Lesley et al 2002). This strategy is founded on the hypothesis that proteins which crystallize readily, even under suboptimal conditions, will do so again during focused crystallization attempts. In tier 1, the goal is to identify those targets which have a propensity to crystallize under the conditions tested; the quality of the crystals produced is not significantly important. To maximize throughput, the protein samples are purified with only one round of affinity purification and screened for crystal formation against a limited number of crystallization conditions; it is expected that some of the proteins will not be sufficiently pure or in the optimal state to crystallize. In tier 2, the objective is to obtain diffraction-quality crystals suitable for structure determination. In this stage, the targets that crystallized in tier 1 are reprocessed to contain SeMet, purified extensively and screened against an expanded set of crystallization conditions. Selected difficult targets that did not produce high quality crystals in tier 2 were subjected to further batch processing in tier 3 which used a loosely defined *ad-hoc* batch process referred to as a "*salvage pathway*".

3.3 Cloning and expression

Target constructs were generally produced in multiples of 96-well plates. Upon generation of selected target sequences and primers, the TSRI pipeline

utilized PCR to generate target DNA from appropriate American Type Culture Centre (ATCC) available genomic DNA. Typically, the insert was ligated into a modified Invitrogen pBAD backbone to create a plasmid that specified ampicillin resistance, arabinose inducibility, and that would place a 6 His N-terminal tag on the protein for use in expression quality control testing and purification. Restriction sites for ligation were engineered using Pm1I (N-terminal) and FseI (C-terminal). Variations of the TSRI protocol included a TEV protease cleavage site and T7 promoter. Cell transformation was by heat shock, with competent cell storage as glycerol stocks at -80ºC. A Qiagen BioRobot 3000 provided the necessary automation.

3.4 Purification

Purification starts with cell harvest, sonication, and clarification of the *E. coli* extracts. At JCSG this was accomplished in a single step for up to 96 samples using the GNFuge. Proteolysis and denaturation were minimized by cooling, inclusion of a protease inhibitor cocktail, and the addition of a mild reducing agent. Viscosity reduction for subsequent steps was provided by adding a DNAse. Verification of expression in the clarified extract was provided by SDS PAGE and anti-His western blotting.

Up to 96 gravity fed, immobilized metal chelate columns (IMAC) were run in parallel to provide one-step purification for native proteins entering initial crystallization screening. Elution was via a single step gradient. IMAC was also applied to all targets entering tiers 2 and 3 of the pipeline for structure determination or salvage. At TSRI, an agarose-based cobalt resin provides low non-specific binding and allowed a low salt elution that facilitated a subsequent ion exchange step. For some *E. coli* studies and for targets expressed in insect cells, a second IMAC step was performed after TEV cleavage of the His tag. Post-IMAC quality control included SDS PAGE for purity, a Bradford assay for yield, and MALDI to verify identity by molecular weight.

Anion or cation exchange chromatography (IEX) were used to both purify and concentrate all samples entering tiers 2 or 3 of the pipeline. Sample loading, separation, and peak cutting are automated through the use of various Pharmacia-Amersham automated FPLC systems. The JCSG capacity was as high as 60 targets per day from a single production shift. TSRI found Waters AP-1 columns packed with Poros HQ resins to be most amenable to the 10mL/min gradient conditions required for maximum throughput. Quality control post-IEX typically included SDS PAGE, MALDI of tryptic-digests, and analytical size exclusion chromatography (SEC).

Preparative SEC was optionally employed for samples showing less than 95% purity by analytical SEC or SDS PAGE. Columns and conditions varied between the two pipelines, but TSRI generally employed Superdex 200. Although relatively slow, the separations were automated and thus ran unattended. If multiple species of a single target were readily resolved, each was screened separately in crystallization trials.

3.5 Suitability testing

In year four of the JCSG, NMR screening represented a major improvement to the pipeline that was the result of a collaboration with Professor Kurt Wüthrich (Page et al 2005). Averaging only 5 minutes per measurement, a 1D ^{1}H NMR spectrum was recorded for subsequent evaluation of band broadening. Proteins were then categorized into one of four groupings that reflected their folded state and the likelihood of structure determination by NMR or crystallography. A study of 79 mouse homolog targets showed that despite a nearly equal ability to crystallize, only proteins graded A-B or flagged as potential multimers produced high resolution structures (Peti et al 2005). Most recently the NMR screening process was further refined by implementing use of a 1mm probe to reduce sample requirements to only 5μL of a 0.5-2mM protein solution, as well as by automation of the sample loading and measurement steps.

A second key component of the pipeline was implemented in year 4, stemming from a collaboration with Dr. Virgil Woods at UCSD (Pantazatos et al 2004). Deuterium exchange mass spectrometry (DXMS) measures the solvent exchange rates of amide hydrogen atoms to identify unstructured regions within a protein. Deletion mutants were then generated to reduce the level of disorder, a process which proved effective in the generation of crystals yielding high resolution structures. Further refinement of the technology to increase throughput continues.

3.6 Crystallization

Targets were concentrated and placed into a suitable delivery buffer for crystallization by ultrafiltration over regenerated cellulose membranes having a molecular weight cut-off of ~10,000Da. At TSRI, coarse and fine screening was typically initiated at protein concentrations of 0.5 and 1mM, respectively, with further optimization guided by the solubility results obtained from crystallization trials. Increasing attention was paid to repeated over-concentration during buffer exchange, which could facilitate

aggregation or irreversible precipitation. TSRI utilized a buffer composition of 10mM Tris pH 7.8, 100mM NaCl, and 0.25mM TCEP for proteins entering crystallization trials. Delivery buffer optimization for soluble proteins received attention only as a salvage pathway, but simply reviewing crystallization results for buffers that yielded clear drops after an overnight incubation provided dramatic improvements in tier 3 buffer selection and crystal quality. Further development of this option continues, based upon previously reported successes at the Berkley Center for Structural Genomics (Jancarik et al 2004).

Crystallization screening at JCSG has been ripe with innovation, owing in part to the systematic capture and analysis of production data (Stevens 2000; Page and Stevens 2004). Roughly 480 crystallization conditions were evaluated in over 320,000 individual experiments on 28% of the *T. maritima* proteome (Page et al 2003). Approximately 86% of purified proteins produced crystals. Prioritization of a subset of these proteins for SeMet labeling and more extensive purification resulted in 68 of 69 proteins yielding crystals in tier 2 and the percentage of total crystals that were harvestable nearly doubled (41%). Further review of pipeline processes revealed that over 75% of the commonly used crystallization conditions found in tier 1 were redundant, with a subset of only 108 of the best conditions yielding crystals for all previously successful 465 tier 1 purified proteins. This led to the establishment of a set of core screening conditions that provided an estimate of a protein's compatibility with tier 2 of the crystallomics core pipeline. It should be noted that the more extensively purified SeMet proteins tended to crystallize under different conditions than their less pure tier 1 counterparts and hence, tier 2 screening maintained its more aggressive wide-sampling of crystallization space. In the last year of PSI-1, TSRI was using a 96-well, 200nL sitting drop format that limited consumption to only 20μL of protein for a complete tier 1 screen at two temperatures.

Review of the results from crystallization studies of *T. maritima* proteins has led to a proposed target filtering strategy (Canaves et al 2004). To identify useful criteria for future protein target selection and to determine ways to improve current pipeline protocols to increase crystallization success of active targets, the distribution of various parameters in the proteome and in the subset of crystallized proteins was analyzed for trends in crystallization success. The parameters analyzed were: (a) biophysical properties, including sequence length, isoelectric point, protein hydropathy, and percentage of charged residues, (b) predicted transmembrane helices and signal peptide sequences, (c) predicted bacterial lipoprotein lipid-binding sites (hydrophobicity pockets), (d) predicted coiled-coils, and (e) predicted low-complexity regions that might lead to disorder.

Seven sequence-derived parameters shown to have a direct effect on protein crystallization were selected for these filtering strategies, including protein length, calculated isoelectric point, percent charged residues, gravy index to indicate hydrophobicity, the number of SEG residues to identify low complexity, the number of predicted trans-membrane helices and the number of predicted signal peptides (Table 1-5). The first strategy proposed is based on the absolute maxima and minima at which crystallization has been observed for each parameter, i.e. none of the observed crystals would be lost, but would still result in an increase in the ratio of successfully crystallized proteins and selected targets (37.7%, Table 1-5). The second strategy is based on more stringent cut-offs that tolerate the loss of up to 5% of the crystals per parameter. The goal is to further reduce the pool of potential targets with respect to the first strategy, while further increasing the ratio between successfully crystallized proteins and selected targets (39.5%, Table 1-5). The loss of a small number of outlier crystallized proteins is tolerated because it allows for a higher success rate for new targets, resulting in an overall increase of successfully crystallized targets. Finally, we opted for an even more stringent filtering strategy that uses as limits the area where most of crystallized proteins cluster in the distribution defined by each protein attribute (Maximum Clustering Strategy, MCS). Whereas the number of lost crystallized proteins and solved structures is higher than in the second strategy, the ratio of crystallized and solved proteins to selected targets is even greater (45.1%, Table 1-5), indicating that this is a superior target selection or design strategy.

4. DISCUSSION

4.1 Production results, JCSG and other centers

A comparison of production statistics for all nine of the PSI-1 large-scale centers is given in Table 1-3. Several factors make it difficult to draw inferences from these numbers. For example, it clearly is hard to gauge the relative difficulties of working with the protein sets that each center was targeting, although prokaryotic proteins made up the majority of targets for PSI-1 (e.g. the JCSG focused almost exclusively on protein from *Thermotoga maritima*). These numbers however represent a valid documentation of the success rates of the overall process and provide an estimate of the approximate cost of doing structural studies. A more detailed breakdown of yield by process step for JCSG is shown in Table 1-6. It is interesting to note that the projected success rates for a number of

Table 1-5. Three different target filtering strategies and statistics calculated from primary sequence.

	No Filtering	Target Filtering Schema		
	Proteome Limits	Absolute Limits	95% Crystal Conservation Per Attribute	Maximum Clustering Strategy
Filtering Parameters: Targets not fulfilling these filtering requirements would be discarded as potential targets.	30<=*Length*<=1690 3.7<=*pI*<=12.3 7.4<=*%Charged*<=52.5 -1.56<=*GRAVY*<=1.62 *0*<=*SEG*<=177 *0*<=*TMHMM*<=16 *0*<=*SignalP*<=1	41<*Length*<813 4.3<*pI*<11.2 17<*%Charged*<47 -0.96<*GRAVY*<0.61 *SEG*<62 *TMHMM*<1 *SignalP*<1	65<*Length*<500 4.6<*pI*<10.2 25<*%Charged*<41 -0.70<*GRAVY*<0.10 *SEG*<32 *TMHMM*<1 *SignalP*<1	90<*Length*<480 4.6<*pI*<7.4 25<*%Charged*<40 -0.50<*GRAVY*<0.10 *SEG*<35 *TMHMM*<1 *SignalP*<1
Final number of targets after filtering (Initial target pool = 1877, i.e. *T. maritima* proteome)	1877	1232	875	606
Number of proteins which crystallized eliminated (total = 465)	0	0	-118	-191
Number of protein structures eliminated (total = 86)	0	0	-14	-18
Chances of Crystallization per target	24.7%	37.7%	39.5%	45.1%
Chances of Structure solution per target	4.6%	7.0%	8.2%	11.2%
Number of proteins crystallized with similar JCSG experimental effort	464 (actual)	708 (theoretical)	741 (theoretical)	847 (theoretical)
Number of protein structures produced with similar JCSG experimental effort	86 (actual)	131 (theoretical)	154 (theoretical)	210 (theoretical)
Theoretical gain in number of structures	n/a	+45	+68	+124
Theoretical increase in JCSG pipeline throughput	n/a	52%	79%	145%

The initial pool of potential targets, the *Thermotoga maritima* proteome, contains 1877 ORFs. The analysis shows the predicted effect of different target selection strategies on total pipeline throughput, assuming that the same experimental effort that was devoted to the full shotgun analysis of the *T. maritime* proteome had been used on focused efforts against selected (filtered) sets of targets.

[@] Number of crystals or structures lost with respect to tier1, if the proposed filtering schemas had been applied to the *T. maritima* proteome.

* Chances of crystallization and structural solution are calculated as 100*(Number of tier1 crystals or structures remaining after filtering / final number of targets after filtering).

[#] Theoretical number of structures gained with similar JCSG resources was calculated extrapolating the chances of structure solution per target for a certain filtering schema to a set of initial targets equivalent in size to the *T. maritime* proteome.

process steps which were presented in the original JCSG research proposal came close to what was achieved. Thus, the 3% overall success rate was anticipated in 2000. However, the original plan called for working with about 45,000 targets, clearly this was not achieved.

Table 1-6. Average yields at various stages of the JCSG pipeline as of October 2004 (number of structures as of August 2, 2005 was 226 as listed in Table 1-3). Numbers in parenthesis were the projected success rates as presented in the proposed JCSG plans in 2000.

Step	Total	% Overall	% Stage	Step	Total	% Overall	% Stage
Target Selection	6537	- -	100	Crystallized	985	17 (9.4)	85
Target Activation	5689	100	87	Screened (X-Ray)	384	7	39
Cloned	3131	55	55	Data Sets	205	4	53
Expressed	2811	49	90	Structure	171	3 (4.3)	83
Soluble Protein	1165	20 (24.4)	41				

4.2 Results from TSRI core

In years 3-5, a TSRI based group worked on a smaller subset of targets with a strong reliance on bioinformatics tools to expand the number of initial constructs per target. Of the 1,452 proteins processed by the TSRI pipeline in the final years of PSI-1, 42% represented homologs or orthologs of a parent target that was also being processed by the rest of JCSG. Secondly, a heavy emphasis was placed on extensive crystallization screening (coarse screening), which out of necessity evolved in year 4 to a more focused optimization of crystallization conditions (fine screening) and cryo-protectants. At the end of PSI-1, the TSRI target pipeline recorded a 44% success rate for turning crystallizable proteins into solved structures. A breakdown of the more than 309,000 screening experiments carried out by the TSRI pipeline alone showed: ~95% coarse screens, ~5% fine screens, and ~1% seeding, additive and chemical modification experiments. In contrast, the distribution of protein crystals that ultimately yielded a high resolution structure was 75% fine screen, 18% coarse screen, and 7%

additive screen. The distribution of structure generating crystals between 277 and 293K was nearly 50:50. Clearly opportunities still exist for improving the ability of sparse coarse screening to generate crystals suitable for high resolution X-ray crystallography. A sampling of TSRI production statistics during this period is shown in Table 1-7.

Table 1-7. Production statistics generated from the TSRI technology development crystallomics core using novel target selection filters. Note – this is a subset of the overall JCSG Crystallomics Core Production. The majority of targets were processed as part of the production crystallomics core at GNF.

		Total	Rate	Rate
Fermentation	Total constructs processed	576	100%	
	Total constructs expressing soluble	179	31%	
	Average ferments per construct that expressed soluble at least once	5.1		
	Average ferments per solved structure	5.4		
Purification	Total constructs processed	179	100%	100%
	Total native proteins processed	72	40%	
	Total SeMet proteins processed	154	86%	
	Total constructs purified successfully	164		92%
	Total purification runs	632		
	Average yield per purification (mg)	~6.4		
	Total purified protein generated (g)	>3.4		
	Average purifications per construct entering purification	3.5		
	Average purifications per solved structure	4.1		
Crystallization	Total constructs entering coarse screening	163	100%	
	Total constructs entering fine screening	67		
	Total proteins solved by X-ray diffraction	28	15.6%	
	Total plates	3,699	100.0%	
	Coarse screen plates	3,041	82.2%	
	Fine screen plates	476	12.9%	
	Seeding plates	4	0.1%	
	Additive plates	53	1.4%	
	Reductive methylation plates	12	0.3%	
	Uncharacterized plates	113	3.1%	
	Total experiments (wells)	309,303	100.0%	100.0%
	Native protein (wells)	95,091	30.7%	

Table 1-7. (Continued)

		Total	Rate	Rate
	SeMet protein (wells)	209,510	67.7%	
	Uncharacterized protein (wells)	4,702	1.5%	
	Coarse screen (wells)	289,825		95.2% 93.7%
	Fine screen (wells)	15,490		5.1%
	Seeding (wells)	132		0.0%
	Additive (wells)	2,704		0.9%
	Reductive methylation (wells)	1,152		0.4%
Structure Determination	Total construct structures solved	29		
	Total unique target structures solved	27		

The TSRI technology development crystallomics core target selection list has also been segmented into protein family or technology method approach (Table 1-8). Not surprisingly, those approaches that were particularly successful included the selection of bacterial homologs of eukaryotic targets, utilization of C-terminal truncations and domain isolation to generate multiple target constructs, and the DXMS-guided generation of deletion mutants for targets that had previously crystallized but diffracted poorly. Techniques that performed poorly included *in-silico* bioinformatics-guided rational target design, exploration of yeast homologs of eukaryotic targets, and utilization of physical measurements of disorder (DXMS) to optimize previously non-crystallizable proteins. Excluding the unsuccessful target selection protocols nearly doubles the average yield of structures per protocol from 2.9% to 5.4%.

4.3 Future directions

In PSI-2 the groups that had operated the JCSG have been funded to run a large-scale production center, JCSG-2 (www.jcsg.org), a specialized center, and a separate Road Map Initiative center. The Road Map Initiative center is called the Joint Center for Innovative Membrane Protein Technologies, and is located at TSRI with a focus on developing novel expression and stability systems for integral membrane proteins (JCIMPT; www.jcimpt.org). The PSI-2 specialized center is called the Accelerated Technologies Center for Gene to 3D Structure (ATCG3D; www.atcg3d.org), which is a collaboration centered at deCODE Biostructures and TSRI, with key collaborations at Lyncean Technologies and the University of Chicago. The ATCG3D is now assembling a new integrated pipeline using technologies currently being developed within the collaboration. Its overall

Table 1-8. A breakdown of the strategy of the TSRI technology development crystallomics core pipeline in years 3-5 (2002-2004) showing a heavy emphasis on homologs, orthologs, and rational target design. In general, the utilization of orthologs and generation of multiple constructs through C-terminal truncations proved more reliable than the utilization of more sophisticated bioinformatics techniques. Interestingly, DXMS produced an exceptionally high yield when used to guide the modification of constructs that had previously generated only poorly diffracting crystals.

Target Selection Protocol	Total Constructs	Parent Targets	Constructs Yielding Crystals	Constructs Yielding Diffraction	Unique Parent Structures Per Row	% Parent Structures Solved
Metabolic pathway targets – Enzymes	65	57	7	6	5	8.80%
Optimized bacterial homologs of metabolic pathway targets	79	33	3	2	0	0.00%
Heart mitochondrial proteome bacterialized	167	73	6	6	3	4.10%
Bacterial homologs of Mouse targets	190	175	17	12	7	4.00%
Optimized, bacterial homologs of Mouse targets	190	161	5	5	3	1.90%
Yeast homologs of Mouse targets	190	184	13	11	3	1.60%
Optimized, yeast homologs of Mouse targets	190	184	3	3	0	0.00%
Multiple constructs of viral targets	190	24	5	5	2	8.30%
Multiple bacterial orthologs of poorly diffracting targets	95	40	9	9	3	7.50%
Sequence optimization by DXMS for non-crystallizable targets	65	14	0	0	0	0.00%
Optimization by DXMS for targets that crystallized but diffracted poorly	31	7	2	2	2	28.60%

goal is to significantly reduce the cost of doing eukaryotic protein structures by approximately 10-fold while maintaining the high quality of work carried out by the structural genomics efforts. Four main areas of technology development are now underway:

4.3.1 Integral membrane proteins

Perhaps the proteins most under-represented in terms of three-dimensional protein structure are membrane proteins, particularly eukaryotic membrane proteins. Research at TSRI for the next several years will focus on developing key technologies to improve the success rate for this family of proteins. These research efforts include cell free protein expression, and improved tools for eukaryotic cell expression (insect cells and other mammalian cell lines). Detergent and lipid chemistry efforts are also a key area where combinatorial chemistry methods will be applied, using nanovolumes of proteins to screen for improved stability reagents.

4.3.2 Cloning by whole gene synthesis

Two of the primary problems with eukaryotic protein structure determination are the acquisition of reliable cDNA clones and the expression of protein constructs that will express well and be amenable to the production of diffraction quality crystals or NMR spectra. ATCG3D proposes to eliminate these bottlenecks by synthesizing all genes (Stewart and Burgin 2005) directly from synthetic oligonucleotides and designing the constructs with protein modeling. This approach has been demonstrated successfully in the past and proof of concept experiments have already shown that large genes (>7kb) and even small viral genomes can be produced by Whole Gene Synthesis. Due to the dropping prices for synthetic oligonucleotides and sequencing reactions, the economics of gene synthesis has reached a point where it is easier, more reliable, and often less expensive to synthesize the gene than it is to source, purchase, and validate a cDNA clone. Most importantly, the entire process can be automated to significantly reduce effort and cost.

It has been shown previously by many different structural biology efforts that an increased success rate of structure solutions can be accomplished by processing an extended number of protein constructs that includes homologs, domain boundary variants and mutants of the desired target protein (Derewenda 2004; Cohi et al 2004; Longenecker et al 2001). It has also been previously shown that when a particular construct does not express well, codon optimization is a very powerful alternative strategy. ATCG3D will use molecular modeling and codon based expression optimization to design

multiple constructs, which will be built from synthetic oligonucleotides. This project is expected to lead to significant cost savings and will greatly improve overall success rates. The full system will be a single instrument that runs the computational modeling, process control database and gene synthesis on a compact footprint.

4.3.3 Crystallization using micro-capillary and *in situ* x-ray screening and data collection

Capillary-based microfluidics technology development has radically changed almost every liquid-based instrument from laser printers to DNA/protein/small molecule analysis. During the late 1990's, companies such as Fluidigm demonstrated the feasibility of microfluidics-based protein crystallization. The current cost of microfluidic chips, however, remains prohibitive for most structural genomics efforts, especially in academia. While the microfluidic technology is established, the breakthrough cost reduction and full implementation into a structural proteomics pipeline has yet to be realized. ATCG3D will focus on developing novel microfluidic technology that is inexpensive and can be directly integrated into both upstream and downstream processing steps including purification, imaging, X-ray screening, and data collection. Of particular importance will be the integration of crystallization with direct X-ray screening of protein crystals.

4.3.4 Compact Light Source

The implementation of an in–house, MAD-capable synchrotron light source might at first appear out of reach; however, the prototype development is already in place, and is based on integrating well-established technologies. The CLS is a miniature synchrotron founded on the marriage of two mature technologies—particle accelerator technology and solid-state laser technology. Accelerators and related hardware have been developed over the past 40 years by the Department of Energy for high-energy physics and synchrotron light sources. Over the past 30 years this progress has led to a large number of high-energy synchrotron light sources worldwide with continuing and dramatic improvements in the performance and quality of X-ray beamlines. Lyncean Technologies has miniaturized this technology by reducing the electron beam energy and by replacing conventional undulator magnets with a laser (http://www. lynceantech.com). The miniature synchrotron has an average flux comparable to the most productive beamlines at the large synchrotrons.

ACKNOWLEDGEMENTS

The majority of work described in this chapter resulted from the highly collaborative efforts between researchers at The Scripps Research Institute, Genomics Institute of the Novartis Research Foundation, and Syrrx. Funding for the JCSG was provided through a NIGMS PSI-1 grant GM062411; JCIMPT was funded by a NIH RoadMap Initiative grant GM073197; and the ATCG3D was funded by NIH PSI-2 grant GM074961.

REFERENCES

Abola E., Kuhn P., Earnest T. and Stevens R. C. 2000. Automation of X-ray crystallography. Nat. Struct. Biol. **7**, 973–977.

Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G., Scherer S. E., Li P. W., Hoskins R. A., Galle R. F. et al. 2000. The genome sequence of *Drosophila melanogaster*. Science **287**, 2185–2195.

Bertone P., Kluger Y., Zheng D., Edwards A. M., Arrowsmith C. H., Montelione G. T., and Gerstein M. 2001. SPINE: An integrated tracking database and data mining approach for prioritizing feasible targets in high-throughput structural proteomics. Nucleic Acids Res. **29**, 2884–2898.

Canaves J. M., Page R., Wilson I. A. and Stevens R. C. 2004. Protein biophysical properties that correlate with crystallization success in *Thermotoga maritima*: Maximum clustering strategy for structural genomics. J. Mol. Biol. **344**, 977–991.

Carter D. C., Rhodes P., McRee D. E., Tari L. W., Dougan D. R., Snell G., Abola E. and Stevens R. C. 2005. Reduction in diffuse-convective disturbances in nanovolume protein crystallization experiments. J. Appl. Cryst. **38**, 87–90.

Choi K. H., Groarke J. M., Young D. C., Rossmann M. G., Pevear D. C., Kuhn R. J. and Smith J. L. 2004. Design, expression, and purification of a Flaviviridae polymerase using a high-throughput approach to facilitate crystal structure determination. Protein Sci. **13**, 2685–92.

Cohen A. E., Ellis P. J., Miller M. D., Deacon A. M. and Phizackerley R. P. 2002. An automated system to mount cryocooled protein crystals on a synchrotron beam line, using compact sample cassettes and a small-scale robot. J. Appl. Crystallogr. **35,** 720–726.

Derewenda Z. 2004. Rational protein crystallization by mutational surface engineering structure. **12**, 529–35.

Hansen C. L., Skordalakes E., Berger J. M. and Quake S. R. 2002. A robust and scalable microfluidic metering method that allows protein crystal growth by free interface diffusion. PNAS **99**, 16531–16536.

Hosfield D., Palan J., Hilgers M., Scheibe D., McRee D. E. and Stevens R. C. 2003. A fully integrated protein crystallization platform for small-molecule drug discovery. J. Struct. Biol. **142**, 207–217.

Jancarik J., Pufan R., Hong C., Kim S. H. and Kim R. 2004. Optimum solubility (OS) screening: An efficient method to optimize buffer conditions for homogeneity and crystallization of proteins. Acta Crystallogr. **D60**, 1670–1673.

Kigawa T., Yabuki T., Matsuda N., Matsuda T., Nakajima R., Tanaka A. and Yokoyama S. 2004. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. J. Struct. Funct. Genomics **5**, 63–68.

Klock H. E., White A., Koesema E. and Lesley S. A. 2005. Methods and results for semi-automated cloning using integrated robotics. 2005. J. Struct. Funct. Genomics., **6**, 89-94.

Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W. et al. 2001. Initial sequencing and analysis of the human genome. Nature **409**, 860–921.

Lesley S. A., Kuhn P., Godzik A., Deacon A. M., Mathews I., Kreusch A., Spraggon G., Klock H. E., McMullan D., Shin T. 2002. Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline. Proc. Natl. Acad. Sci. **99**, 11664–11669.

Longenecker K. L., Garrard S. M., Sheffield P. J. and Derewenda Z. S. 2001. Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. Acta. Crystallogr. D. Biol. Crystallogr. **57**, 679–88.

Markley J. L., Ulrich E. L., Westler W. M. and Volkman B. F. 2003. Macromolecular structure determination by NMR spectroscopy. Methods Biochem. Anal. **44**, 89–113.

Marra M. A., Jones S. J. M., Astell C. R., Holt R. A., Brooks-Wilson A., Butterfield Y. S. N., Khattra J., Asano J. K., Barber S. A., Chan S. Y. 2003. The genome sequence of the SARS-Associated Coronavirus. Science **300**, 1399–1404.

McPhillips T. M., McPhillips S. E., Chiu H.-J., Cohen A. E., Deacon A. M., Ellis P. J., Garman E., Gonzalez A., Sauter N. K., Phizackerley R. P. et al. 2002. Blu-Ice and the distributed control system: Software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Rad. **9,** 401–406.

Monleon D., Colson K., Moseley H. N. B., Anklin C., Oswald R., Szyperski T. and Montelione G. T. J. 2002. Rapid analysis of protein backbone resonance assignments using cryogenic probes, a distributed Linux-based computing architecture and a set of semi-automated spectral analysis tools. J. Struct. Funct. Genomics **2**, 93–101.

Muchmore S. W., Olson J., Jones R., Pan J., Blum M., Greer J., Merrick S. M., Magdalinos P. and Nienaber V. 2000. Automated crystal mounting and data collection for protein crystallography. Structure **8**, R243-R246.

NIH press release 2005. http://www.nih.gov/news/pr/jul2005/nigms-01.htm.

Olson D. L., Peck T. L., Webb A., Magin R. L. and Sweedler J. V. 1995. High-resolution microcoil ^{1}H-NMR for mass-limited, nanoliter-volume samples. Science **270**, 1967–1970.

Page R., Moy K., Sims E. C., Velasquez J., McManus B., Grittini C., Clayton T. L. and Stevens R. C. 2004. Scalable high-throughput micro-expression device for recombinant proteins. Biotechniques **37**, 364–370.

Page R., Peti W., Wilson I. A., Stevens R. C. and Wuthrich K. 2005. NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. Proc. Natl. Acad. Sci. U. S. A. **102**, 1901–1905.

Page R. and Stevens R. C. 2004. Crystallization data mining in structural genomics: using positive and negative results to optimize protein crystallization screens. Methods **34**, 373.

Page R., Grzechnik S. K., Canaves J. M., Spraggon G., Kreusch A., Kuhn P., Stevens R. C. and Lesley S. A. 2003. Shotgun crystallization strategy for structural genomics: An optimized two-tiered crystallization screen against the thermotoga maritima proteome. Acta Crystallogr. **D59**, 1028–1037.

Pantazatos D., Kim J. S., Klock H. E., Stevens R. C., Wilson I. A., Lesley S. A. and Woods V. L., Jr. 2004. Raid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. Proc. Natl. Acad. Sci. U. S. A. **101**, 751–756.

Peti W., Norcross J., Eldridge G., O'Neil-Johnson M. 2004. Biomolecular NMR using a microcoil NMR probe-new technique for the chemical shift assignment of aromatic side chains in proteins. J. Am. Chem. Soc. **126**, 5873–5878.

Peti W., Page R., Moy K., O'Neil-Johnson M., Wilson I. A., Stevens R. C. and Wuthrich K. 2005. Towards miniaturization of a structural genomics pipeline using micro-expression and microcoil NMR. J. Struct. Funct. Genomics **6**, 259-67.

Rota P. A., Oberste M. S., Monroe S. S., Nix W. A., Campagnoli R., Icenogle J. P., Peñaranda S., Bankamp B., Maher K., Chen M. 2003. Characterization of a Novel Coronavirus associated with severe acute respiratory syndrome. Science **300**, 1394–1399.

Santarsiero B. D., Yegian D. T., Lee C. C., Spraggon G., Gu J., Scheibe D., Uber D.C., Cornell E. W., Nordmeyer R. A., Kolbe W. F. et al. 2002. J. Appl. Cryst. **35**, 278.

Snell G., Cork C., Nordmeyer R., Cornell E., Meigs G., Yegian D., Jaklevic J., Jin J., Stevens R. C. and Earnest T. 2004. Automated sample mounting and alignment system for biological crystallography at a synchrotron source. Structure **12,** 537–545.

Stevens R. C. 2000. High-throughput protein crystallization. Curr. Opin. Struct. Biol. **10**, 558–563.

Stevens R. C. 2004. Long live structural biology. Nat. Struct. Mol. Biol. **11**, 293–295.

Stewart L. and Burgin A. B. 2005. Whole gene synthesis: A Gene-O-Matic future. In *Frontiers in Drug Design and Discovery*. Bentham Science Publishers, Ltd, Co-Editors Atta-ur-Rahman, B. A. Springer, G. W. Caldwell.

Studier F. W. 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr. Purif. **41**, 207–34.

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A. et al. 2001. The sequence of the human genome. Science. **291**, 1304–1351.

Wüthrich K. 2003. NMR studies of structure and function of biological macromolecules. J. Biomol. NMR **27**, 13–39.

Wüthrich K. 1986. NMR of proteins and nucleic acids. New York, Wiley.

Zolnai Z., Lee P. T., Li J., Chapman M. R., Newman C. S., Phillips G. N., Jr., Rayment I., Ulrich E. L., Volkman B. F. and Markley J. L. 2003. Project management system for structural and functional proteomics: Sesame. J. Struct. Funct. Genomics **4,** 11–23.

Chapter 2

PROTEIN PRODUCTION FOR STRUCTURAL GENOMICS - STRATEGIES FOR THE NEXT PHASE

Pär Nordlund[1], Tobias Cornvik[1,2] and Helena Berglund [1]

1) Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden 2) Department of Biochemistry and Biophysics, Stockholm University, 106 91 Stockholm, Sweden

1. INTRODUCTION

Structural genomics (SG) programs aim at establishing efficient pipelines that allow protein structure determination to be attempted on a genome wide scale, at a high pace and at low cost. Structures produced by these programs not only provide fundamental information on molecular mechanisms of well-studied proteins, but can also hint at the functions of previously non-annotated proteins, and can potentially aid in the rational design of drugs. Targets for each of the SG projects have mainly been contained within a genome, focusing on proteins with no or low homology to known structures, but with some emphasis on potential targets for therapeutic interventions. Most current SG projects began by working on prokaryotic proteins, which are significantly easier than eukaryotic proteins to produce in suitable quantity and quality for structural studies. The most urgent technology developments are therefore required for eukaryotic proteins but, of course, such improvements of protein production methodology are likely to significantly improve the success rates of SG projects targeting prokaryotic proteins as well.

The cost of producing proteins in quantities needed for structural studies in most current protein production pipelines is relatively high, partially due to extensive manual interventions required for the scale-up process.

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 27–48.

Therefore, the development and integration of instrumentation in these steps will be essential for cost reduction and efficiency of future large-scale protein production pipelines. Efficient strategies for improving and assessing the quality of proteins produced by SG pipelines will also have great cost benefits by directing only 'high quality' proteins into subsequent SG pipeline steps of crystallisation, crystal optimisation and NMR data collection and interpretation.

Intensive efforts are therefore being made to improve strategies for protein production within SG projects (Braun and LaBaer, 2003). Similarly, in the pharmaceutical sector, strategies for the production of specific proteins with improved success rates are urgently needed. Parallel technologies, related to technologies being established in SG programs, are being developed for the pharmaceutical sector. In the present review, progress towards establishing efficient pipelines for producing proteins for structural genomics is assessed with a particular emphasis on emerging technologies in *E.coli*, including some from our own laboratory that will be particularly useful for producing eukaryotic proteins for SG projects. Important, but less imminent are technologies for eukaryotic expression (Loomis et al. 2005; Geisse and Henke 2005; McCall et al. 2005; Holz et al. 2003) cell free expression (Sawasaki et al. 2002; Shimizu et al. 2001) and refolding of proteins (Trésaugues et al. 2004; Maxwell et al. 2003; Vincentelli et al. 2004), which are not treated in detail in this review.

2. OVERVIEW OF CURRENT STRATEGIES

The pioneering protein production efforts in structural genomics were focused on implementing streamlined and robust protocols for protein production and evaluating their efficiency (Edwards et al. 2000; Heinemann et al. 2000; Dieckman et al. 2002). Still, protein production in structural genomics employs relatively simple expression and purification strategies, outlined in Figure 2-1. Streamlined protein production strategies based on *Ecoli* expression, followed by one or two generic purification steps, have been the common denominator for the successful projects targeting bacterial proteins. Second tier rescue pathways that employ refolding, cell-free expression, or non-*E. coli* cell based expression systems have been used, but have so far only played a minor role in producing structures within SG initiatives (see http://targetdb.pdb.org). A notable exception is the Riken initiatives, which to a significant extent uses cell free expression systems (http://protein.gsc.riken.go.jp/). It is hard to establish reliable estimates for the success rates for different experimental protocols in the current SG

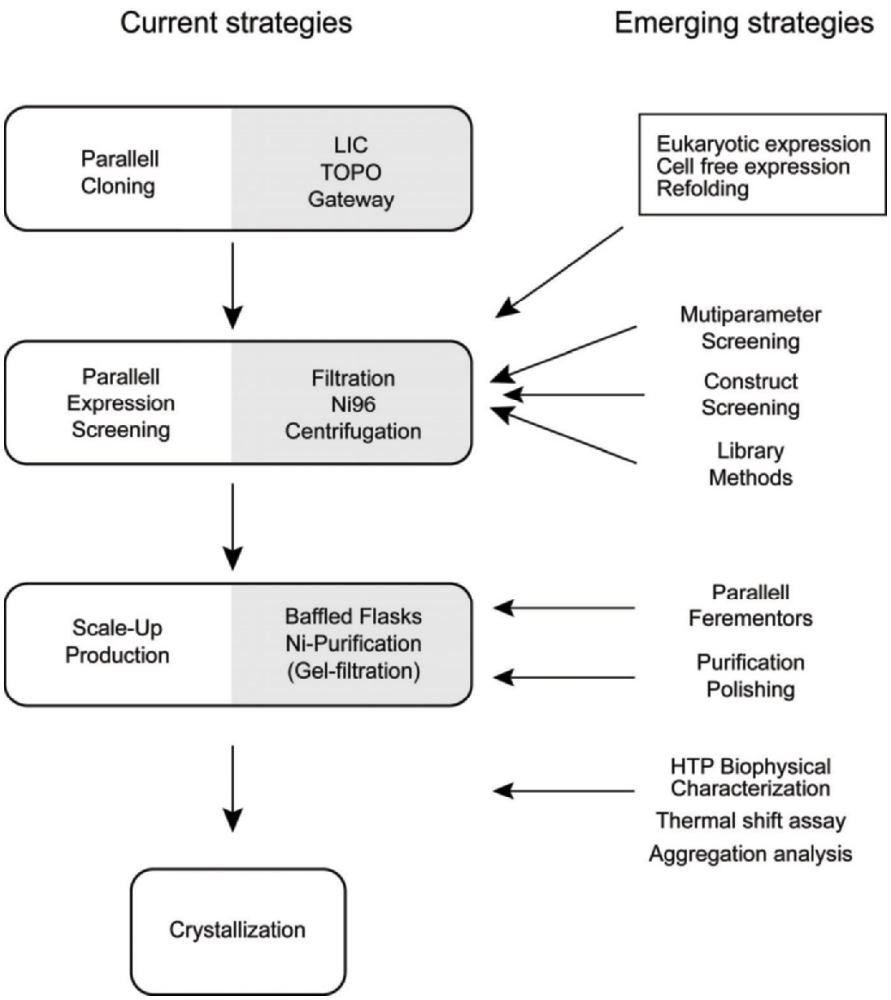


Figure 2-1. Schematic outline of the protein production pipeline in structural genomics.

initiatives, because the available databases do not reveal enough experimental detail and the targets within each SG initiative constitute a relatively small selection of proteins and hence the selection can bias the outcome. A few genome-wide projects have aimed at processing a majority of the proteins from a certain genome using a single well-defined experimental procedure, which helps to minimise statistical bias that could result from the selection process (Lesley et al. 2002; Luan et al. 2004). The most conclusive project is the JCSG project on *T. maritima* where the attrition rate in the protein production step was ~ 60% (# not purified proteins/total # of proteins). The attrition rate in obtaining lead-crystals was

only ~20 % (# not crystallized proteins/# purified proteins), and at present for this ongoing endeavour, the success rates of structure determination from lead-crystals is 20% (# of structures/# proteins with lead-crystals). Thus, for the process as a whole, from cloned gene to structure the success rate is 5 %. For eukaryotic proteins, at present, no reliable statistics from gene to structure are available in the literature for a genome-wide project. However, in an attempt to express 10,167 *C. elegans* genes using a multi-well based small-scale platform, some 15 % allowed affinity purification at levels indicating potential usefulness for structural studies (Luan et al. 2004). The overall attrition rate for crystallisation of eukaryotic proteins is unknown, but likely to be worse than bacterial proteins considering that eukaryotic proteins generally are more complex, having longer loop regions and containing intrinsically disordered regions (Oldfield et al. 2005).

It is anticipated that many more eukaryotic proteins and domains can be expressed in *E. coli* if multi-expression parameter and multi-construct approaches are applied. A significant fraction of proteins are lost upon purification or concentration, indicating that improved methodology for stabilizing proteins during the preparation step will be very valuable. For difficult to express proteins, directed evolution selection strategies of clones from libraries of random mutants or from construct libraries have the potential to improve solubility and stability (Pedelacq et al. 2002; Cornvik et al. 2005). However, the introduction of these strategies into the SG pipeline requires further streamlining and automation of the technologies. The generic scale-up of the protein production process in current SG programs is quite robust but still costly, as it requires many manual interventions.

The potential increased complexity of emerging technologies is likely to increase the managerial challenges. The volume of accumulated data will increase further and sophisticated Laboratory Information Management Systems (LIMS,) will grow in importance.

2.1 Parallel cloning formats

A number of different cloning formats are used in different SG projects although each project preferably employs a single cloning format. Four formats typically used are: 1) streamlined restriction enzyme and ligation based cloning (Christendat et al. 2000; Klock et al. 2005), 2) TOPO based cloning (Chance et al. 2002), 3) recombination based cloning (Gateway) (Hammarstrom et al. 2002), and 4) ligation independent cloning (LIC) (Dieckman et al. 2002). In short, it appears that all of these formats can be implemented in efficient multi-well based format with cloning efficiencies of the available PCR products exceeding 90%. The four formats exhibit

differential advantages and disadvantages. Restriction enzyme and ligation based cloning is the least expensive but requires analysis of restriction profiles of genes to select a suitable restriction sites for cloning. TOPO cloning is unique in that it produces a native sequence without requiring the addition of additional residues to enable cloning. Gateway cloning leaves an additional 8 amino acid linker sequence, which however, does not appear to create particular problems (see below). The Gateway system is particularly useful for comparing different expression vectors. After the gene has been cloned into a Gateway entry vector, it can be sub-cloned into many different expression vectors in a recombinase-based process that reaches 100 % efficiency, does not introduce mutations, and, therefore, does not require additional sequencing. A number of "ORFeomes" are also available in Gateway adapted forms, from which pre-sequenced genes can be directly recombined into expression vectors (Reboul et al. 2003; Rual et al. 2004). LIC is highly efficient and relatively cheap in that no expensive proprietary enzymes or vectors are required for cloning (as is the case for Gateway and TOPO). LIC has emerged as the favored system for recently established SG initiatives.

For all these systems, parallel multi-well procedures for cloning and transformation have been established which depend on pipetting relatively small (μl) volumes using multichannel-pipettes. PCR amplifications from templates are often done efficiently using 'touch down' PCR protocols employing high fidelity polymerases (Eshaghi et al. 2005). The handling of the colony step is potentially time consuming and multi-well formats are often used for speeding up colony handling. Alternatively cloning grills can be used to handle colonies in multiple cloning experiments (Mehlin et al. 2004). The cloning steps can be implemented on commercial liquid handling robots and colony handling could be performed using colony picking automation. The automation of parallel cloning is unlikely to be economical for projects cloning less than a few thousand constructs per year.

2.2 Multi-parameter expression screening platforms in *E.coli*

Recombinant expression in *E.coli* is the method of choice for structural genomics programs since the system is fast and inexpensive and yields protein with a low level of heterogeneity caused by post translational covalent modifications. However, for more challenging proteins, such as human proteins, the processes of identifying optimal expression conditions can be lengthy and generally associated with a large attrition rate. Many parameters can have major effects on the level of soluble expression of different proteins including expression vectors with different promoters,

fusion tags/proteins and copy numbers; host cell types; helper vectors containing chaperones and rare-tRNA's; temperature of growth and induction, culture media, and additives (Sorensen and Mortensen 2005). It is anticipated that a challenging protein may require many parameters to be screened to obtain useful expression. However, in practice when many proteins are processed only a limited number of parameter combinations can be screened and an important goal is to find the parameters and parameter combinations, which are generally more efficient for expressing proteins. For example, low temperature induction appears on average to be better than higher temperature induction for *E. coli* expression of eukaryotic proteins (Vasina and Baneyx 1997; Schein and Noteborn 1988). The long term goal would be to identify conditions which could be selected as sparse matrix parameter screens for expression, in a similar way as sparse matrix screens have been worked out within SG projects for optimization of crystallization experiments (Page et al. 2003). Benchmarking of a large number of expression parameter combinations on a larger set of proteins will be required to define such screens. It is plausible that different classes of proteins might suffer from different types of expression problems and that class-specific expression screens will have to be defined. Current parallel and semi-automated platforms for cloning and expression screening will be very useful for this work. However, the systematic work directed at obtaining optimal sparse matrix expression screen still lies ahead.

2.3 Parallel expression screening platforms

The suitability of a protein for structural studies depends on the ability to express the protein in a soluble form in *E. coli* cytoplasm rather than in inclusion bodies. This 'solubility' is a useful initial measure of the quality of an expressed protein, and often correlates quite well with activity. Soluble bacterial fractions can be separated from inclusion bodies for analytical screening in multi-well plate formats using several techniques, including centrifugation, filtration, magnetic bead affinity purification and "big bead" affinity purification. All four methods can be adapted for automation. The two last methods, allowing binding of proteins to beads in unclarified extracts, provide samples that can be directly analyzed on gels. However, the current cost for using these medias on larger number of proteins and conditions is relatively high. The pre-clarification of bacterial lysates using centrifugation or filtration on multi-well plates allow the direct monitoring of the soluble fraction on SDS-PAGE. However, this is often not sufficient for identifying protein expressing at low or medium level but are therefore most often combined with a subsequent affinity purification step using standard affinity resins dispensed in multi-well plates. Soluble proteins are

normally analyzed in high throughput studies by either dot-blots or SDS-PAGE. SDS-PAGE after affinity purification provide more information, for example, whether the protein has been affected by partial proteolysis. However, SDS-PAGE are time consuming and expensive for screening larger numbers of samples. Microfluidic devices can potentially substitute SDS-PAGE but are still costly and suffer from lower resolution. Dot-blots allow more cost efficient screening of solubility and is emerging as the method of choice when larger numbers of expression parameters are to be screened. In a SG setting, dot blots could give sufficient information about solubility to decide if a scale-up experiment should be performed. We have made extensive use of the sub-micron filtration method followed by Dot-blots for screening soluble expression, the Filtration Dot-blot (FiDo) method (Knaust and Nordlund 2001). Our experience is that the FiDo screen constitutes an efficient and readily automated method for the detection of soluble expression. When we use INDIA- His Probe (Pierce Biotechnology), which to some extent mimics a Ni-resin, for dot-blot detection, we notice that these blots correlate very well with blots obtained after a Ni-purification. Also, the sub-micron filtration correlates well with centrifugation, identifying 95% of the proteins which had been identified as soluble by centrifugation.

A number of laboratories have implemented such parallel expression screening procedures on various liquid handling platforms (Qiagen 8000, Freedom Tecan, Hamilton Star or MWG) (Vincentelli et al. 2005; Nguyen et al. 2004). These platforms are likely to be very useful for benchmarking of expression parameters for sparse matrix definitions. It is also likely that SG initiatives targeting eukaryotic or other difficult protein classes will have to screen relatively large numbers of expression conditions and that these automated platforms will be a core unit of such initiatives.

3. PARAMETERS - CURRENT VIEW

3.1 Vectors

Most structural genomics projects use pET-based vectors with 6xHis-tags on either the N- or C-terminus (Christendat et al. 2000; Vincentelli 2005; Kim et al. 2004). The use of vectors with linker regions allowing proteolytic cleavage to remove the 6xHis-tag is common, thereby generating the potential for obtaining additional protein constructs, which will lead to improved success rates for crystallization. The small size of the 6xHis-tag makes it very popular as a purification tag, but could sometimes affect the solubility (Woestenenk et al. 2004). To facilitate the production of soluble

protein, larger N-terminal fusion proteins have been examined for the effect on soluble protein expression. Until recently, fusion partners had gained their reputation based upon their performance together with a low number of target proteins, but now several comparative studies have been performed using larger numbers of proteins (Braun and LaBaer 2003; Hammarstrom et al. 2002; Dyson et al. 2004; Shih et al. 2002). The proteins Glutathione-s-Transferase (GST), Thioredoxin, Maltose binding protein (MBP), NusA, GB1, and Z have been exploited as N-terminal fusion partners. It is also beneficial if the solubilising fusion partner can be used as a purification handle, as is the case with GST, MBP, the IgG binding domains GB1 and the Z-domain. In addition solubilising fusions proteins can be used in conjunction with a 6xHis-tag.

Although the behavior of a fusion-target protein pair is largely target specific and difficult to predict, some general trends can be established based on recent studies of larger assembles of genes. Small, stable, highly soluble, fast-folding fusion-proteins seem generally to be able to increase the final yield of fusion target protein pair. Larger fusion partners are more efficient at promoting solubility for larger target-fusion protein pairs (Hammarstrom 2006, Braun and LaBaer, 2003) It is not possible to point out one outstanding fusion protein that is successful for all types of target proteins because different studies have come to somewhat varying conclusions regarding the usefulness of specific fusions partners. The varying results are probably due to different target types and other variations that were present in the vector construct used, as well as in the experimental set-up. It is however clear that significant overlap is seen between different tags and that optimization of other vector elements is also important for successful protein production. Moreover, for crystallization purposes it is highly desirable that fusion proteins are removed using a linker specific protease. An outstanding issue is to what extent the target is useful for structural studies after the fusion tag has been removed. From a recent study it was revealed that some 50 % of proteins, which express in a soluble form as MBP fusions could not be purified (Jeon et al. 2005). This demonstrates that MBP can solubilise proteins which are not per-se sufficiently soluble for structural studies. It is not yet known whether the purified proteins from the MBP fusion construct were similarly useful for structural studies as protein expressed with 6xHis-tags.

3.2 Strains

Numerous different strains have been developed to cope with the potential problems during heterologous expression in *E. coli*. Strains containing plasmids coding rare tRNAs e.g. Rosetta (Novagen) and BL21

CodonPlus (Stratagen) and strains co-expressing chaperones have been successfully used to obtain soluble protein expression of previously insoluble targets. There are also strains that have been developed to provide an appropriate oxidising environment for disulfide formation such as the Origami strain (Novagen). A major challenge ahead is the benchmarking of these strains with other expression parameter combinations. For a recent review of strains and conditions see Sorensen and Mortensen (2005).

3.3 Multi-construct approaches

Multi-construct expression strategies are routinely exploited in the pharmaceutical industry, and many academic groups, to improve success rates for obtaining soluble expression in *E. coli* of eukaryotic proteins. Typically 3-10 different expression constructs of each target are generated by varying the choice for the initial and final amino acids of the target protein. Quite often this strategy results in more than one soluble variant, thereby also improving the probability of crystallization. Domain borders are predicted by bioinformatics approaches, primarily using homology modeling on related structures and secondary structure predictions. An additional experimental method for defining protein domains employs limited proteolysis of the full-length protein followed by mass spectrometry of the resulting fragments. This proteolytic domain mapping is however a relatively labor-intensive method, which also requires expression and purification of a larger construct in a soluble form. A more rapid method utilizing deuterium exchange followed MS has been used successfully to improve crystallization rates for 21 *T. maritime* proteins (Pantazatos et al. 2004). Multi-construct methods are likely to be a very important strategy for applications on eukaryotic proteins; however, the extent to which they improve success rates, as well as optimal strategies for selecting domain boarders, still need to be defined.

4. LIBRARY-SELECTION TECHNOLOGIES FOR GENERATING SOLUBLE PROTEINS

The identification of soluble proteins from libraries of random mutants or of random-length expression constructs is expected to be a viable route to producing soluble cytosolic and extracellular targets as well as for membrane proteins (Waldo 2003). Randomly introduced mutations might promote a productive folding path of proteins or modify problematic regions of proteins such as hydrophobic patches or regions sensitive to proteolysis. Similarly, the optimization of well-defined domain borders for expression

constructs could minimize flexible regions to yield more stable and compact domains. Both of these strategies do also potentially allow selection of expression variants, which have effects on transcription, translation or general folding events, as well as the toxicity of the protein. Therefore the library strategies have the potential to improve the quality of the physical behavior of a target protein, as well as the quantity of soluble protein.

The implementation of efficient library-selection strategies has turned out to be challenging. Nevertheless, a number of different strategies for the selection of improved proteins from combinatorial gene libraries has been implemented and tested on a limited number of proteins. Robust strategies for the generation of combinatorial DNA libraries containing random mutations or randomly fragmented DNA have been developed, originally for other applications (Farinas et al. 2001). Generation of random mutations is relatively straightforward using error prone PCR (Smith 1985) but more appropriate methods accessing a larger sequence space are also available (Farinas et al. 2001). Selected mutations can in subsequent steps be recombined using DNA-shuffling (Stemmer 1994). Methods have also been established to generate libraries of random length DNA fragments that can then be screened for the expression of different length protein constructs. Constructs where both the 5' and 3' end of the ORF (both N- and C-termini of the protein) is varied can be generated using random priming (Kawasaki and Inagaki 2001) or fragmentation by sonication (Nakayama and Ohara 2003). Alternatively, deletions at either the N- or C-termini can be generated using exonuclease -based protocols.

While robust protocols for library generation are available, the establishment of technologies allowing the selection or screening of clones expressing more soluble proteins from large combinatorial expression libraries has been challenging. To date, the most widely used method to monitor expression levels at the colony level is the fusion of green fluorescent protein (GFP) to the C-terminus of the target protein (Waldo et al. 2001). Using cycles of error prone PCR followed by GFP-based selection and subsequent recombination of mutations using DNA-shuffling, several *M.tuberculosis* proteins have been evolved (Cabantous et al. 2005). Kawasaki and Inagaki (2001) have used GFP to select constructs from libraries which express soluble fusion proteins. Different length ORFs were generated with a random PCR approach and screened for soluble protein expression using GFP as a C-terminal reporter fusion. However, only 12 soluble constructs were obtained when screening more then 100,000 colonies.

Resistance selection, based on C-terminal antibiotic degrading enzymes is an interesting possibility and potentially selection can be made from very

large libraries. Chloramphenicol acetyl transferase (CAT) has been shown to be a useful selection tool for folded proteins but has not yet been systematically applied for evolving proteins for structural studies (Sieber et al. 2001; Maxwell et al. 1999). Complementation strategies based on split proteins has also been developed. The S-protein split assay consists of 15 terminal amino acids, which compliment a truncated RNAase (Kelemen et al. 1999). Similarly the split protein Lac Z, with a 100 amino acid complementary fragment, is potentially useful for selecting soluble proteins (Wigley et al. 2001). The recent development of a split GFP appears to be a significant improvement on the original GFP method as a better correlation to solubility is seen (Cabantous et al. 2005). Also, the relatively short tag of 10 amino acids can potentially be left on for crystallization and the protein is ready for scale-up production.

Display methods such as phage display or ribosome display have been shown to have some potential for selecting proteins with improved properties. Phage display has been used to find proteins more resistant to protease degradation (Finucane et al. 1999). Ribosome display has been used to select proteins which have less exposed hydrophobic amino acids (Matsuura and Pluckthun 2003). A method for directly monitoring solubility on colonies was developed by Peabody and co-workers who noted that protein diffusion through an agarose layer from self-lysed colonies, as judged by immunoblotting, correlated to the solubility of the protein (Peabody and Al-Bitar 2001).

Alternative direct monitoring procedures are multi-well based. We have selected constructs from an N-terminally deleted construct library of an Ephrin dependent tyrosine kinase receptor using the FiDo screen (see above). Some 5-10 % of these constructs were soluble suggesting a relatively high success rate with this method (Figure 2-2).

The most recent addition to the repertoire of solubility screening methods is the Colony Filtration blot (CoFi-blot), where soluble proteins are separated from inclusion bodies through a filtration step at the colony level (Cornvik et al. 2005). The released proteins are captured on a nitrocellulose membrane and detected using standard immunochemicals in a process conceptually related to the FiDo screen previously developed in our laboratory. The CoFi blot method has several advantages over methods utilising C-terminal fusions of a reporter protein or protein fragment. Clones identified as positives with the CoFi blot can be directly subjected to scale-up production without additional recloning, if a small tag is sufficient for detection. Furthermore, there is no risk for artificial solubilisation or precipitation by the fusion protein. The CoFi blot was shown to correlate well with traditional solubility analysis using centrifugation (Cornvik et al. 2005).

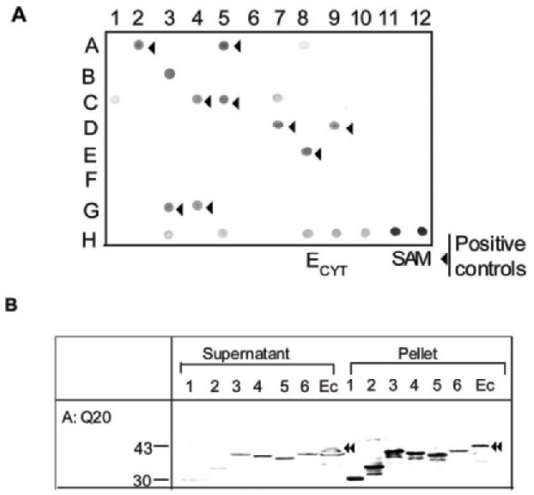


Figure 2-2. FiDo screen of a deletion mutagenesis library of the Ephrin dependent tyrosine kinase receptor. (Knaust & Nordlund, unpublished) (A) A dot blot of the soluble fraction after the filtration step (FiDO screen). Arrows indicate clones expressing at higher level as compared to the wild type. (B) Clones identified as soluble in the FiDo screen were grown in liquid culture and insoluble material was removed by a high-speed centrifugation. The soluble fraction was analyzed on a Western blot.

The CoFi-blot is now a standard method in our lab and has, for example, been used to screen for soluble clones from construct libraries. To benchmark the methodology, N-terminal deletion libraries were generated for 19-proteins which did not express in a traditional expression screen using either an N-terminal Flag or 6xHis-tag vector. The libraries were screened using the CoFi-blot and positive clones were analyzed using multi well affinity purification methodologies. For 11 targets several constructs were found expressing at high levels and were directly suitable for scale-up and structural studies. A vast majority of the sequenced deletion mutants start close to predicted domain borders and, surprisingly, many clones that correspond to the full-length or near full-length protein were identified (Figure 2-3).

5. SCALE-UP FERMENTATION AND PURIFICATION

The cost and efficiency of the scale-up protein production process is a key factor for the success of a HTP protein production pipeline. Furthermore

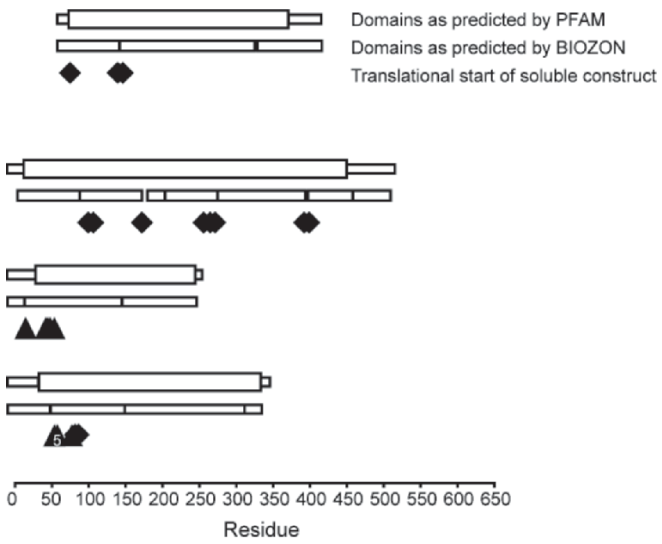


Figure 2-3. Random deletion libraries were screened using the newly developed colony filtration blot. The translational start of the soluble constructs are shown as squares or triangles. When compared to the predicted domain composition, a large fraction of the constructs start close to predicted domain borders.

the established scale-up process has to be reproducible, as well as correlate well with small-scale screening experiments. Baffled shake flasks are currently the main solutions for scale-up fermentation in structural genomics programs, even though they require extensive manual interventions. PET-bottles have been suggested as an alternative, as they are provided as cheap and sterile flasks ready for fermentation (Millard et al. 2003). A parallel scale-up system has been developed by Leslie et al based on cell production in parallel bubble fermentors (reviewed elsewhere in this issue). State of the art fermentors with more excessive fermentation control provide the potential of producing high-density cell cultures, as well as a high degree of automation. However, such fermentors have not yet been extensively applied in structural genomics programs. One major reason for this is that a well suited fermentation system has been lacking on the market. Belach AB, Solna, has recently developed a parallel fermentation system that was tested in our laboratory, directly designed with structural genomics programs in mind. The fermentor Greta requires minimal manual intervention as it has online washing, autoclaving, induction, culture attenuation etc. It allows facile parallel processing, pH control and feed batch processes to obtain high cell densities - typically in the range of OD_{600} 70-90 at harvest (Figure 2-4). Due to a tight control of temperature, oxygenation and pH, the cultures are likely to yield more reproducible data. In a preliminary study we see that

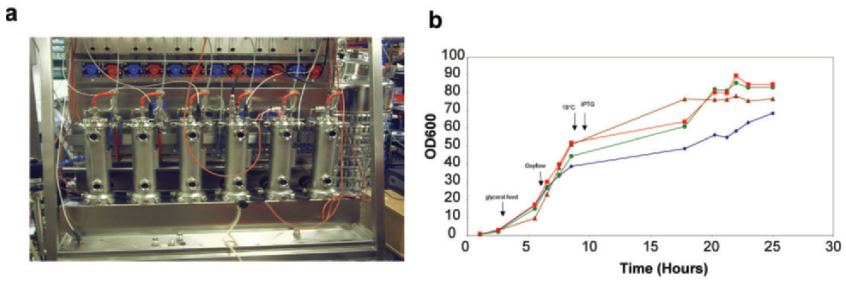


Figure 2-4. (A) Photo of the parallel fermentor Greta. (B) Optical density traces from 4 fermentations of 4 different proteins. Optical densities at 600 nm between 75-100 can be routinely archived.

protein yield per gram of cells is similar in shake flasks and high OD_{600} cultures in Greta.

Structural genomics initiatives are dependant on affinity purification as the main purification tool and the 6xHis-tag is almost exclusively used for this purpose. Step elusions are normally applied to obtain relatively concentrated samples, which can be directly attributed to size exclusion chromatography (SEC), if required. Simple parallel affinity set ups using peristaltic pumps or gravity flow columns are often used. Automation and streamlining of protein purification have been implemented at JCSG, based on in house developed robotics (Lesley et al. 2002). GE HealthCare (previously Amersham Biosciences) have established an automated multi-column purification system based on the Äkta 3d set up and more recently the parallel Äkta Xpress system. These systems are very robust and are already being used in numerous structural genomics laboratories worldwide. In a typical implementation on a four unit Äkta Xpress system, 16 proteins can be purified overnight through a metal-chelate affinity step followed by a SEC. Further polishing of purity is often needed for eukaryotic proteins and might involve extensive column scouting, which can be conveniently established for HTP work on the Äkta Explorer platform (Zingmark and Nordlund, unpublished).

The correlation between small-scale purification screening and scale-up experiments is an important issue. The emerging view is that the correlation between small scale and large scale experiments is good enough to support small-scale multi-parameter screening platforms being used to predict whether a certain expression condition will produce soluble protein when being scaled-up. In a specific study on *P. furiosus* proteins expressed in *E. coli* the correlation of predicting solubility of some 60 proteins was in the range of 80 % (Jenney et al. 2005).

6. HIGH-THROUGHPUT BIOPHYSICAL CHARACTERIZATION

Even when expression of soluble proteins has been established, the quality of the produced proteins can vary dramatically. It is clear that the possibility to purify and concentrate a protein sample does not equate with proper protein folding or sample homogeneity. Biophysical characteristics such as protein integrity, "foldedness", solubility, and stability can be used for characterisation of a protein as well as for screening conditions to optimise the state of a sample. It is not known to what extent these parameters correlate to crystallisability, but the emerging view is that biophysical properties of the protein correlate better to crystallisability than for example protein activity, as partially unfolded and aggregated proteins can show activity but usually do not allow crystallization. So far only a small set of biophysical methods are available in the HTP setting but as instruments reach new limits of detection the range of methods available for HTP work will increase. From a crystallisation point of view, methodologies to identify stabilizing/solubilizing buffers or additives to be incorporated in subsequent crystallization trials will be of great value.

For investigation of the chemical integrity and homogeneity of the proteins, SDS-PAGE and MS analysis are already available for HTP quality control using small amounts of sample. Limited proteolysis combined with SDS gels or MS analysis can also give an indication of flexible regions that might disturb the crystallization process. It has been shown that proteins stable to limited proteolysis are significantly more likely to produce well diffracting crystals (Gao et al. 2005). The coupling of Deuterium exchange with MS methods have been used to identify unstructured regions (Pantazatos et al. 2004). This technology has potential to directly identify constructs which are more likely to crystallise. Whether a protein is folded or not can be addressed directly using NMR spectroscopy. Compared to other characterisation methods, NMR spectroscopy cannot be considered high throughput, but the information content is high. Attempts have been made to correlate ^{15}N HSQC data with protein crystallisation but the methods have been suggested to be complementary rather than correlated (Savchenko et al. 2003). However, the quality of 1d ^{1}H NMR data has been shown to correlate, not to crystallisability, but to the diffracting quality of the formed crystals (Page et al. 2005).

It is anticipated that a poorly folded protein, with a higher content of molten globular or unstructured regions is also more likely to have poor solubility and stability. The selection of samples with low aggregation tendencies and high stability might, in fact, also co-select for "foldedness" of the sample. Dynamic light scattering (DLS) is a well established technology

for measuring size distributions within a sample (mono/poly-dispersity). It has been shown to correlate to some extent to crystallisation, as the likelihood of successful crystallisation is lower if the DLS size distribution is polydisperse (Zulauf and Darcy 1992). However, the method has not yet been extensively used for large buffer or ligand screens, since it is only recently that DLS instruments that allow larger screens using reasonable amount of proteins have been available. DLS, as well as static light scattering (SLS) instruments with increased sensitivity, have been customized to accommodate plate readers, which has increased the throughput. Using this instrument it was recently shown that buffers could be selected using DLS which yield more homogenous proteins in SEC. (http://www.wyatt.com/literature/platereaderbuffer.pdf).

Direct analytical gel-filtration has the potential as a medium throughput tool for monitoring aggregation/oligomerisation. Unfortunately, commercially available parallel instrumentation for performing such experiments in a HTP mode does not yet appear to be available. Solubility screens based on direct monitoring of precipitations by ocular inspection, turbidity measurements at λ=340 nm (Vincentelli et al. 2004) or by protein yield after filtration (Bondos and Bicknell 2003) can easily be performed in a multi-well format and are a potential means for pre-screening the intrinsic solubility of the sample in different buffers. However, smaller aggregates, which might be deleterious to crystallization, are not detected with these methods. The output from the different technologies for monitoring solubility and aggregation tendencies is probably to some extent overlapping and the evaluation of which method and/or combinations of methods that would be the most useful for predicting crystallizability is still to be established.

The measurement of the thermal stability of a protein at different conditions is likely to help in predicting how well the protein will tolerate the protein production and structure determination process. Furthermore, by identification of additive binding using a thermal stability shift assay, additives can be identified which potentially improve a protein's stability and homogeneity, providing valuable information for the crystallization experiment. Traditionally, thermal stability is studied using calorimetric technologies or melting curves based on CD measurements or UV spectrophotometry. In the recently established thermofluor-type experiments (Pantoliano et al. 2001), a hydrophobic fluoroprobe binds to the molten globule of the melting protein, inducing fluorescence (the probe is quenched in polar solvents). This method now provides a true HTP and a small-scale technology for rapid measurements of thermal melting points as well as ligand-induced stabilisation of proteins in a multi-well format using commercially available real-time PCR instruments (Lo et al. 2004). Only relatively small quantities of proteins are needed, ~1-15 μg per well. The

method is generic in the sense that no prior knowledge of the protein is required in order to screen for stabilising conditions or ligands. However, for proteins having hydrophobic binding sites for substrates or other exposed hydrophobic surfaces it will not work well as the flouroprobe will be quenched at ambient temperatures.

We have used the thermoflour method for identifying stabilising or destabilising additives for ten *E. coli* proteins (Figure 2-5). The additives identified were used in 45,000 subsequent crystallisation experiments. The use of a stabilising additive in the crystallisation screens was shown to improve the proteins probability to yield crystals to 70 %.

In our studies we have not seen any direct correlation between the melting temperature of a protein and the probability of obtaining crystals, which has also been reported for the *T Maritima* project (Canaves et al. 2004). Preliminary data indicates that the shape of the melting curves might give information about the condition or 'foldedness' of the protein, when it appears to correlate to crystallisability (data not shown). The thermofluor method can also be used to rapidly identify stabilizing buffers which might be useful in protein preparation. However, the extent of correlation between the thermal stability and solubility/aggregation tendencies still remains to be elucidated.

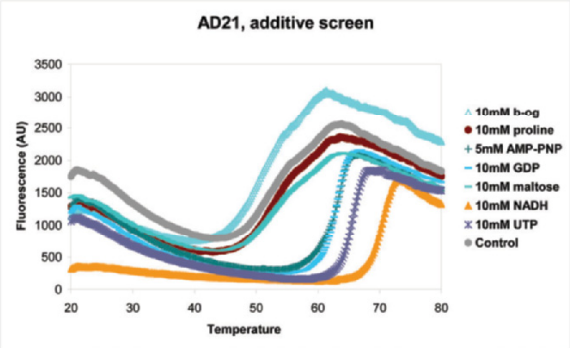


Figure 2-5. Melting curves obtained from an additive screen in a thermoflour experiment. Several stabilizing additives can be identified, seen as a change in the melting temperature as compared to the control.

7. CONCLUSIONS AND FUTURE PERSPECTIVES

A number of SG centres have been able to establish relatively efficient platforms for the production and structural characterisation of prokaryotic

proteins. Some of these centres are now focusing on eukaryotic proteins. It is clear that efficient SG methodologies on eukaryotic proteins will require extension of current protein production strategies. It is very likely that *E. coli* expression will constitute the main platform for some time ahead. Multi-construct approaches combined with sparse matrix expression screens are also likely to be core technologies in these efforts. Automated domain identification using solubility selection schemes have potential to impact the pipeline. Methods for improving protein quality will be critical to improve the current attrition rates. Some biophysical methods are available in HTP formats but more will be required. The prioritisation for biophysical characterisation measurements and their role in decision-making in the pipeline still needs to be developed.

REFERENCES

Bondos S. E. and Bicknell A. 2003. Detection and prevention of protein aggregation before, during, and after purification. Anal. Biochem. **316**, 223–231.

Braun P. and LaBaer J. 2003. High throughput protein production for functional proteomics. Trends Biotechnol. **21**, 383–388.

Cabantous S., Pedelacq J. D., Mark B. L., Naranjo C., Terwilliger T. C. and Waldo G. S. 2005. Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from mycobacterium tuberculosis. J. Struct. Funct. Genomics **6**, 113–119.

Cabantous S., Terwilliger T. C. and Waldo G. S. 2005. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat. Biotechnol. **23**, 102–107.

Canaves J. M., Page R., Wilson I. A. and Stevens R. C. 2004. Protein biophysical properties that correlate with crystallization success in thermotoga maritima: Maximum clustering strategy for structural genomics. J. Mol. Biol. **344**, 977–991.

Chance M. R., Bresnick A. R., Burley S. K., Jiang J.-S., Lima C. D., Sali A., Almo1 S. C., Bonanno J. B., Buglino J. A., Boulton S. et al. 2002. Structural genomics: A pipeline for providing structures for the biologist. Protein Sci. **11**, 723–738.

Christendat D., Yee A., Dharamsi A., Kluger Y., Savchenko A., Cort J. R., Booth V., Mackereth C. D., Saridakis V., Ekiel I. et al. 2000. Structural proteomics of an archaeon. Nat. Struct. Biol. **7**, 903–909.

Cornvik T., Dahlroth S. L., Magnusdottir A., Herman M. D., Knaust R., Ekberg M. and Nordlund P. 2005. Colony filtration blot: A new screening method for soluble protein expression in Escherichia coli. Nat. Methods **2**, 507–509.

Dieckman L., Gu M. Y., Stols L., Donnelly M. I. and Collart F. R. 2002. High throughput methods for gene cloning and expression. Protein Expr. Purif. **25**, 1–7.

Dyson M. R., Shadbolt S. P., Vincent K. J., Perera R. L. and McCafferty J. 2004. Production of soluble mammalian proteins in Escherichia coli: Identification of protein features that correlate with successful expression. BMC Biotechnol. 4.

Edwards A. M., Arrowsmith C. H., Christendat D., Dharamsi A., Friesen J. D., Greenblatt J. F. and Vedadi M. 2000. Protein production: Feeding the crystallographers and NMR spectroscopists. Nat. Struct. Biol. **7**, 970–972.

Eshaghi S., Hedrén M., Ignatushchenko M., Nasser A., Hammarberg T., Thornell A. and Nordlund P. 2005. An efficient strategy for high-throughput expression screening of recombinant integral membrane proteins. Protein Sci. **14,** 676–683.

Farinas E. T., Bulter T. and Arnold F. H. 2001. Directed enzyme evolution. Curr. Opin. Biotechnol. **12,** 545–551.

Finucane M. D., Tuna M., Lees J. H. and Woolfson D. N. 1999. Core-directed protein design. I. An experimental method for selecting stable proteins from combinatorial libraries. Biochemistry **38,** 11604–11612.

Gao X., Bain K., Bonanno J. B., Buchanan M., Henderson D., Lorimer D., Marsh C., Reynes J. A., Sauder J. M., Schwinn K. et al. 2005. High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. J. Struct. Funct. Genomics **6**, 129–134.

Geisse S. and Henke M. 2005. Large-scale transient transfection of mammalian cells: A newly emerging attractive option for recombinant protein production. J. Struct. Funct. Genomics **6**, 165–170.

Hammarström M. 2006. Protein production and purification in structural genomics *Doctoral thesis, KTH.*

Hammarstrom M., Hellgren N., Van, den, Berg S., Berglund H. and Hard T. 2002. Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. Protein Sci. **11**, 313–321.

Heinemann U., Frevert J., Hofmann K.-P., Illing G., Maurer C., Oschkinat H. and Saenger W. 2000. An integrated approach to structural genomics. Prog. Biophys. Mol. Biol. **73,** 347–362.

Holz C., Prinz B., Bolotina N., Sievert V., Büssow K., Simon B., Stahl U. and Lang C. 2003. Establishing the yeast Saccharomyces cerevisiae as a system for expression of human proteins on a proteome-scale. J. Struct. Funct. Genomics **4**, 97–108.

Jenney F. E., Jr., Brereton P. S., Izumi M., Poole F. L., Shah C., Sugar F. J., Lee H. S. and Adams M. W. 2005. High-throughput production of Pyrococcus furiosus proteins: Considerations for metalloproteins. J. Synchrotron Rad. **12**, 8–12.

Jeon W. B., Aceti D. J., Bingman C. A., Vojtik F. C., Olson A. C., Ellefson J. M., McCombs J. E., Sreenath H. K., Blommel P. G., Seder K. D. et al. 2005. High-throughput purification and quality assurance of arabidopsis thaliana proteins for eukaryotic structural genomics. J. Struct. Funct. Genomics **6**, 143–147.

Kawasaki M. and Inagaki F. 2001. Random PCR-Based screening for soluble domains using green fluorescent protein. Biochem. Biophys. Res. Comm. **280,** 842–844.

Kelemen B. R., Klink T. A., Behlke M. A., Eubanks S. R., Leland P. A. and Raines R. T. 1999. Hypersensitive substrate for ribonucleases. Nucleic Acids Res. **27**, 3696–3701.

Kim Y., Dementieva I., Zhou M., Wu R., Lezondra L., Quartey P., Joachimiak G., Korolev O., Li H. and Joachimiak A. 2004. Automation of protein purification for structural genomics. J. Struct. Funct. Genomics **5**, 111–118.

Kimber M. S., Vallee F., Houston S., Necakov A., Skarina T., Evdokimova E., Beasley S., Christendat D., Savchenko A., Arrowsmith C. H. et al. 2003. Data mining crystallization databases: Knowledge-based approaches to optimize protein crystal screens. Proteins **51,** 562–568.

Klock H., White A., Koesema E. and Lesley S. 2005. Methods and results for semi-automated cloning using integrated robotics. J. Struct. Funct. Genomics **6**, 89–94.

Knaust R. C. and Nordlund P. 2001. Screening for soluble expression of recombinant proteins in a 96-well format. Anal. Biochem. **297**, 79–85.

Lesley S. A., Kuhn P., Godzik A., Deacon A. M., Mathews I., Kreusch A., Spraggon G., Klock H. E., McMullan D., Shin T. et al. 2002. Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline. Proc. Natl. Acad. Sci. U. S. A. **99**, 11664–11669.

Lo M. C., Aulabaugh A., Jin G., Cowling R., Bard J., Malamas M. and Ellestad G. 2004. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal. Biochem. **332**, 153–159.

Loomis K. H., Yaeger K. W., Batenjany M. M., Mehler M. M., Grabski A. C., Wong S. C. and Novy R. E. 2005. Insectdirect system: Rapid, high-level protein expression and purification from insect cells. J. Struct. Funct. Genomics **6**, 189–194.

Luan C. H., Qiu S., Finley J. B., Carson M., Gray R. J., Huang W., Johnson D., Tsao J., Reboul J., Vaglio P. et al. 2004. High-throughput expression of C-elegans proteins. Genome Res. **14**, 2102–2110.

Matsuura T. and Pluckthun A. 2003. Selection based on the folding properties of proteins with ribosome display. FEBS Lett. **539**, 24–28.

Maxwell K. L., Bona D., Liu C., Arrowsmith C. H. and Edwards A. M. 2003. Refolding out of guanidine hydrochloride is an effective approach for high-throughput structural studies of small proteins. Protein Sci. **12**, 2073–2080.

Maxwell K. L., Mittermaier A. K., Forman-Kay J. D. and Davidson A. R. 1999. A simple in vivo assay for increased protein solubility. Protein Sci. **8**, 1908–1911.

McCall E. J., Danielsson A., Hardern I. M., Dartsch C., Hicks R., Wahlberg J. M. and Abbott W. M. 2005. Improvements to the throughput of recombinant protein expression in the baculovirus/insect cell system. Protein Expr. Purif. **42**, 29–36.

Mehlin C., Boni E., Andreyka J. and Terry R. 2004. Cloning grills: High throughput cloning for structural genomics. J. Struct. Funct. Genomics **5**, 59–61.

Millard C. S., Stols L., Quartey P., Kim Y., Dementieva I. and Donnelly M. I. 2003. A less laborious approach to the high-throughput production of recombinant proteins in Escherichia coli using 2-liter plastic bottles. Protein Expr. Purif. **29**, 311–320.

Nakayama M. and Ohara O. 2003. A system using convertible vectors for screening soluble recombinant proteins produced in Escherichia coli from randomly fragmented cDNAs. Biochem. Biophys. Res. Comm. **312**, 825–830.

Nguyen H., Martinez B., Oganesyan N. and Kim R. 2004. An automated small-scale protein expression and purification screening provides beneficial information for protein production. J. Struct. Funct. Genomics **5**, 23–27.

Oldfield C. J., Yugong E. L. U., Cheng A. K., Dunker J. U. L. and Markley J. L. 2005. Addressing the intrinsic disorder bottleneck in structural proteomics. Proteins: Structure, Function, and Bioinformatics **59**, 444–453.

Page R., Grzechnik S. K., Canaves J. M., Spraggon G., Kreusch A., Kuhn P., Stevens R. C. and Lesley S. A. 2003. Shotgun crystallization strategy for structural genomics: An optimized two-tiered crystallization screen against the Thermotoga maritima proteome. Acta Crystallogr. D **59**, 1028–1037.

Page R., Peti W., Wilson I. A., Stevens R. C. and Wuthrich K. 2005. NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. Proc. Natl. Acad. Sci. U. S. A. **102**, 1901–1905.

Pantazatos D., Kim J. S., Klock H. E., Stevens R. C., Wilson I. A., Lesley S. A. and Woods V. L., Jr. 2004. Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. Proc. Natl. Acad. Sci. U. S. A. **101**, 751–756.

Pantoliano M. W., Petrella E. C., Kwasnoski J. D., Lobanov V. S., Myslik J., Graf E., Carver T., Asel E., Springer B. A., Lane P. et al. 2001. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J. Biomol. Screen. **6**, 429–440.

Peabody D. S. and Al-Bitar L. 2001. Isolation of viral coat protein mutants with altered assembly and aggregation properties. Nucleic Acids Res. 29.

Pedelacq J. D., Piltch E., Liong E. C., Berendzen J., Kim C. Y., Rho B. S., Park M. S., Terwilliger T. C. and Waldo G. S. 2002. Engineering soluble proteins for structural genomics. Nat. Biotechnol. **20**, 927–932.

Reboul J., Vaglio P., Rual J. F., Lamesch P., Martinez M., Armstrong C. M., Li S., Jacotot L., Bertin N., Janky R. et al. 2003. C. Elegans ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. Nature Genetics **34**, 35–41.

Rual J. F., Hirozane-Kishikawa T., Hao T., Bertin N., Li S., Dricot A., Li N., Rosenberg J., Lamesch P., Vidalain P. O. et al. 2004. Human ORFeome Version 1.1: A platform for reverse proteomics. Genome Res. **14**, 2128–2135.

Savchenko A., Yee A., Khachatryan A., Skarina T., Evdokimova E., Pavlova M., Semesi A., Northey J., Beasley S., Lan N. et al. 2003. Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches. Proteins: Structure, Function, and Genetics **50**, 392–399.

Sawasaki T., Ogasawara T., Morishita R. and Endo Y. 2002. A cell-free protein synthesis system for high-throughput proteomics. Proc. Natl. Acad. Sci. U. S. A. **99**, 14652–14657.

Schein C. H. and Noteborn M. H. M. 1988. Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperature. **6**, 291–294.

Shih Y. P., Kung W. M., Chen J. C., Yeh C. H., Wang A. H. and Wang T. F. 2002. High-throughput screening of soluble recombinant proteins. Protein Sci. **11**, 1714–1719.

Shimizu Y., Inoue A., Tomari Y., Suzuki T., Yokogawa T., Nishikawa K. and Ueda T. 2001. Cell-free translation reconstituted with purified components. **19**, 751–755.

Sieber V., Martinez C. A. and Arnold F. H. 2001. Libraries of hybrid proteins from distantly related sequences. Nat. Biotechnol. **19**, 456–460.

Smith G. P. 1985. Filamentous fusion phage - novel expression vectors that display cloned antigens on the virion surface. Science **228**, 1315–1317.

Sorensen H. and Mortensen K. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microb. Cell Fact. **4**, 1.

Stemmer W. P. C. 1994. Rapid evolution of a protein in-vitro by DNA shuffling. Nature **370**, 389–391.

Tresaugues L., Collinet B., Minard P., Henckes G., Aufrere R., Blondeau K., Liger D., Zhou C. Z., Janin J., Van, Tilbeurgh H. et al. 2004. Refolding strategies from inclusion bodies in a structural genomics project. J. Struct. Funct. Genomics **5**, 195–204.

Vasina J. A. and Baneyx F. 1997. Expression of aggregation-prone recombinant proteins at low temperatures: A comparative study of the Escherichia coli csp A and tac promoter systems. Protein Expr. Purif. **9**, 211–218.

Vincentelli R., Canaan S., Campanacci V., Valencia C., Maurin D., Frassinetti F., Scappucini-Calvo L., Bourne Y., Cambillau C. and Bignon C. 2004. High-throughput automated refolding screening of inclusion bodies. Protein Sci. **13**, 2782–2792.

Vincentelli R., Canaan S., Offant J., Cambillau C. and Bignon C. 2005. Automated expression and solubility screening of His-tagged proteins in 96-well format. Anal. Biochem. **346**, 77–84.

Waldo G. S. 2003. Genetic screens and directed evolution for protein solubility. Curr. Opin. Chem. Biol. **7**, 33–38.

Waldo G. S., Standish B. M., Berendzen J. and Terwilliger T. C. 1999. Rapid protein-folding assay using green fluorescent protein. Nat. Biotechnol. **17,** 691–695.

Wigley W. C., Stidham R. D., Smith N. M., Hunt J. F. and Thomas P. J. 2001. Protein solubility and folding monitored in vivo by structural complementation of a genetic marker protein. Nat. Biotechnol. **19**, 131–136.

Woestenenk E. A., Hammarstrm M., van, den, Berg S., HÃ¤rd T. and Berglund H. 2004. His tag effect on solubility of human proteins produced in Escherichia coli: A comparison between four expression vectors. J. Struct. Funct. Genomics **5**, 217–229.

Zulauf M. and Darcy A. 1992. Light-scattering of proteins as a criterion for crystallization. J. Cryst. Growth **122**, 102–106.

Chapter 3

INTRODUCTION TO FRAGMENT SCREENING

Andrew R. Leach and Michael M. Hann
GlaxoSmithKline Research and Development, Gunnels Wood Road, Stevenage Hertfordshire SG1 2NY, United Kingdom

1. INTRODUCTION

To achieve within one molecule all the complex properties required to ensure the desired target specificity and potency, bioavailability, duration of action and lack of toxicity required of a new therapeutic drug is exceedingly difficult. A consequence of this is that many drug discovery projects result in failure even after many years of work. Moreover, new challenges (e.g. requirements for improved drug quality and new diseases and targets) continue to be introduced that further test our current understanding. The concept of fragment screening has arisen from the realisation that finding a good starting point for the evolution of a new molecule is a critical issue. Although it would be preferable to start as close as possible to the desired end point, the sheer number and diversity of molecular entities (Ertl 2003) that might be considered suggests that evolution by selection will invariably be the way in which new drugs are discovered in the foreseeable future. Fragment screening aims to identify molecules which are substantially smaller than "drug like" molecules yet which still have some activity at the required target protein and correspond to more attractive starting points for a drug discovery effort.

2. DRUG LIKENESS

The development of technologies for the synthesis and screening of large numbers of compounds has provided some unique opportunities - and

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 49–72.

challenges - in drug discovery. It became clear early in the application of these methods that success would not result simply from increasing the throughput (both with regard to the numbers of molecules synthesized and the numbers of molecules screened). A seminal contribution was that of Lipinski and colleagues, who examined a series of clinically tested drug molecules to try to determine whether they possessed any distinguishing properties (Lipinski et al 1997; Lipinski 2001). These studies resulted in the "Lipinski rule of five" which is a set of simple filters designed to predict whether or not a molecule is likely to have absorption problems due to poor solubility and/or poor permeability. The rule of five states that poor oral absorption and/or distribution are more likely when:

1. The molecular weight is greater than 500
2. The logP (the logarithm of the n-octanol/water partition coefficient, calculated using the ClogP program) is greater than 5
3. There are more than 5 hydrogen bond donors (defined as the sum of OH and NH groups)
4. There are more than 10 hydrogen bond acceptors (defined as the number of N and O atoms).

Compounds that are substrates for biological transporters are specifically excluded from this analysis. The rule of five is usually implemented by identifying compounds that exceed two or more of the above parameters; Lipinski and colleagues found that less than 10% of their data set of clinical drug candidates had any two parameters of the above parameters outside the specified range.

Lipinski's work acted as a catalyst for others to investigate the properties of "drug-like" molecules. Most of these studies involved comparisons of drug and non-drug molecules in order to identify the key characteristics of drugs in terms of their chemical and physicochemical properties. For example, in one study (Oprea et al 2001) 70% of the "drug-like" compounds had between zero and two hydrogen bond donors, between two and nine hydrogen bond acceptors, between two and eight rotatable bonds and between one and four rings. Veber and colleagues proposed (Veber et al 2002) that the number of rotatable bonds (10 or fewer) and the polar surface area (a value of 140$Å^2$ or less) were important properties to achieve oral bioavailability in the rat. The polar surface area is a widely used descriptor for drug-like properties (Clark and Pickett 2000) and is defined as that part of the molecular surface arising from oxygen or nitrogen atoms or from hydrogen atoms attached to nitrogen or oxygen atoms.

The simplicity of the Lipinski rule and the ease and speed with which it can be calculated was another important reason for its widespread adoption Moreover, it can be applied not only to "real" compounds (e.g. from those in

one's screening collection or which are available for purchase from an external vendor) but also to "virtual" compounds (i.e. molecules that have not yet been made). As such it can be used as an "in silico" filter. Other types of in silico filters can also be identified and are also widely used in drug discovery (Rishton 2003, Walters and Murcko 2003). Some filters are used to identify compounds that contain reactive or otherwise undesirable functionality, such as Michael acceptors, alkyl halides or aldehydes. These filters can be implemented by defining substructures or substructure counts which flag molecules for removal. Other filters use calculated physicochemical properties to classify molecules; an example is the use of a calculated octanol/water partition coefficient in the Lipinski rule of 5. A third type of filter uses a mathematical model to score or classify molecules according to their degree of "drug-likeness". Many forms of mathematical model can be used; typically these are based on calculated molecular descriptors and other properties (Leach and Gillet 2004) in conjunction with some form of model-building methodology such as multiple linear regression, neural networks or genetic algorithms.

3. THE HISTORICAL BASIS OF LEAD-LIKENESS AND FRAGMENT SCREENING

Following their introduction the concepts of drug likeness rapidly gained acceptance and were widely incorporated into many areas of drug discovery. A key refinement, introduced in 1999 by Teague, Davis and Oprea of AstraZeneca (Teague et al 1999), was the notion of *lead likeness*. The AstraZeneca group first identified lead/drug pairs from the literature. Such pairs comprise the drug molecule together with the initial lead molecule from which it was derived. Various properties were then calculated for these pairs of molecules to determine how the properties of the lead molecules differed (if at all) from the corresponding drugs. The outcome was that many properties did show a statistically significant change. In particular, most properties increased in value or number, suggesting that *optimised drugs are more complex than their initial leads*. Four of the properties that follow this pattern are the molecular weight, the logP, the numbers of hydrogen bond donors and the numbers of hydrogen bond acceptors. The same group expanded their data set in a subsequent publication (Oprea et al 2001) and reported the values shown in Table 3-1; in all cases these increase from lead to drug.

The GlaxoSmithKline group of Hann, Harper and Leach published a similar study shortly after the first AstraZeneca paper (Hann et al 2001) in which a larger data set of lead/drug pairs derived from the compendium published by Sneader (Sneader 1996) (see Table 3-2).

Table 3-1. Increases in the median values of molecular properties from leads to drugs (Oprea et al 2001).

Property	Increase
Molecular weight	69 Da
Hydrogen bond acceptors	1
Rotatable bonds	2
Number of rings	1
ClogP	0.43
Hydrogen bond donors	0

Table 3-2. Average property values for Leads and Drugs (Hann et al 2001).

Property	Average value for Leads	Average value for Drugs	Increment
MW	272.0	314.0	42.0
H-bond donors	0.85	0.8	-0.05
H-bond acceptors	2.2	2.5	0.3
ClogP	1.9	2.4	0.5
Number of heavy atoms	19.0	22.0	3.0

Although these studies considered different data sets the common conclusion was that the initial leads for successful drug discovery programmess have statistically different properties to the final drugs. There are several possible explanations for these observations. Leads are often less potent than is required in the final drug, and increases in potency are often most easily achieved by adding functionality, which in turn increases the molecular weight and properties such as the numbers of donors and acceptors. LogP also frequently increases during lead optimisation. This may be partly due to the introduction of specific hydrophobic interactions (giving increased potency) and/or due to a need to increase the apparent concentration of the drug in the lipophilic environment of a membrane bound target. One caveat to note, at least in the case of the Sneader data set, is that many of the leads were small hormones such as biogenic amines. These starting points are often of such low complexity (i.e low molecular weight) that adding mass is inevitable.

A theoretical analysis of lead-likeness was proposed by Hann and colleagues (published in the same paper described above (Hann et al 2001)). This analysis was based on a simple model designed to predict how the chances of finding a hit varied with the complexity of the molecule. The ligand and its binding site are represented as 1-dimensional bitstrings of interaction points. The number of such interaction points in the ligand is considered a measure of its complexity. These interaction points represent the various molecular properties of the ligand that might influence binding, such as shape, electrostatics and other properties such lipophilicity. In the

model each of the ligand interaction points must exactly match a corresponding point in the binding site for a binding event to be permitted. Thus each positive element in the ligand must match a negative element in the binding site and vice versa. Figure 3-1 shows examples of successful and unsuccessful matches. Given this model it is then possible to calculate the probability that a ligand of size L will match a binding site of size B. A typical result is shown in Figure 3-2 for a binding site of size 12. Specifically, we show the probability that a ligand of size 2, 3, 4… can match the binding site in one, two, three… ways. Also calculated is the overall probability that a ligand can match at all, which is the sum of these individual results. As can be seen, the probability that the ligand can match at all shows a smooth and rapid decay to zero as its complexity increases. Of particular interest is the probability that the ligand has just one (i.e. unique) match. This probability passes through a maximum; in this specific case this occurs at a ligand complexity of 3.

The second part of the model considers the probability of measuring the binding of a ligand as a function of its complexity. Here, the complexity is considered a crude indicator of the strength of the interaction and thus of the potency. As the number of pairwise interactions increases then the probability of measuring experimentally the interaction also grows. This is represented in Figure 3-3 as a hyperbolic curve. This indicates that when the number of interactions is below a critical number then the binding cannot be measured because it is too weak (too few interactions). A rapid increase in the probability then occurs, levelling off at a value of 1, consistent with the notion that once the potency exceeds some threshold it will always be possible to measure the interaction. The probability of a "useful event" is defined as the product of these two probability distributions. This probability reflects the true balance of the probability of having a useful matching interaction **and** being able to measure it.

```
Receptor:           1 2 3 4 5 6 7 8 9
features:           - - + - + - - + -

Successful match    + + -
Successful match                  + + -
Unsuccessful match      + + -
```

Figure 3-1. The simple interaction model requires an exact match between ligand and receptor for a successful interaction, as shown for a receptor with 9 interaction points.

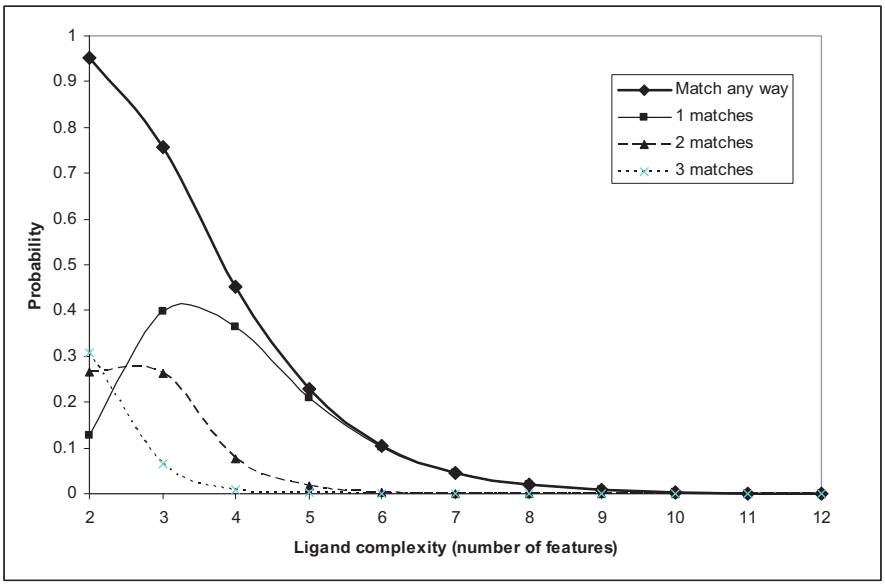


Figure 3-2. Graph showing the probability of finding one, two, three… any matches for varying ligand complexity (for a receptor with 12 interaction sites).

Of course, this is a very simple model. The difficulty is to determine where for real systems the maximum in the combined curve lies. This is a challenging problem. Nevertheless, an understanding of the reasons underlying the distributions can help focus synthesis and acquisition efforts towards particular types of compounds. The two probabilities have competing distributions insofar as the probability of finding a match goes from high to low while the probability of measuring the interaction varies from low to high. This competition leads to a combined probability having a bell shaped form. At low complexity, the probability of a useful event is zero; even though there is a high probability that the properties match they are too few to contribute to an observable binding in an assay. At high complexity the probability of there being a complete match is very small, though if such a match does occur it will be easily measured. In the intermediate region there is the highest probability of a useful event being found, due to the realistic probabilities of both having a match and being able to measure it.

Another aspect of the problem of molecular complexity and the value in using less complex ligands (hereafter referred to as molecular fragments, or just fragments) concerns the effectiveness with which chemical space can be explored. A number of groups have tried to estimate the number of potential "drug like" molecules (i.e. whose structure would fall within what is

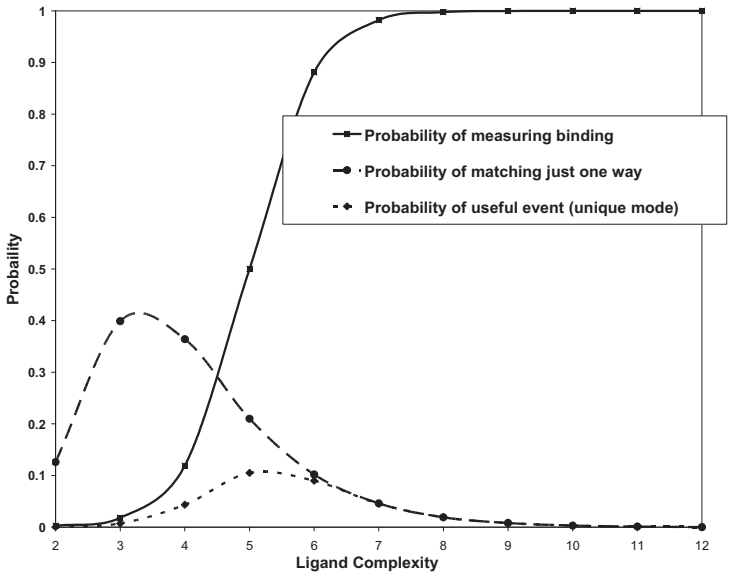


Figure 3-3. Figure showing the result of multiplying the probability of a single match by the probability of being able to measure the binding event. The resulting probability of a "useful event" passes through a maximum.

generally considered "drug like" in terms of their elements, the ways in which the elements are bonded together, and properties such as molecular weight and the ratio of heteroatoms to carbon atoms). These estimates vary widely, but all suggest that drug-like chemical space is very large – far larger than the number of molecules made to date and possibly larger than the number of atoms in the known universe. Consider Figure 3-4 which shows a target protein with two binding sites. Suppose we have a set of five fragments. To identify a molecule that binds to both binding sites may require the synthesis of the full 25-member combinatorial library. By contrast, if we can identify the fragments that bind at the individual sites and then combine them, we need only make one full molecule. Such an approach enables the chemical space to be explored in an additive, rather than multiplicative, manner. This can provide some very dramatic savings; if a target contains S subsites that are combined using L linkers then a physical fragment library containing M members can act as a surrogate for a virtual library of the following size:

$$V \sim M^S L^{S-1}$$

A 1000-member fragment library with 20 linkers can act as a probe for a much larger 20 million-member library for a two-site target.

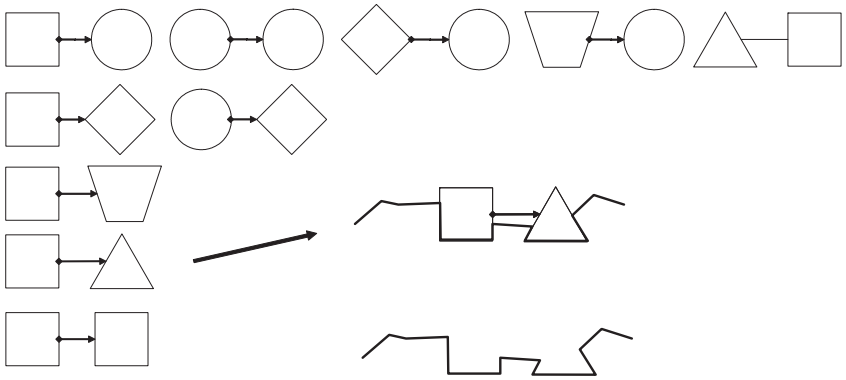


Figure 3-4. For a set of five fragments there are 25 different pairwise combinations. By contrast, if the fragments are screened individually then the desired combination can be obtained directly.

Another illustration of the enhanced sampling obtained with fragment screening approaches is provided by Figure 3-5 which shows the number of carboxylic acids (of all types) registered in the GlaxoSmithKline (GSK) registry system and plotted as a function of molecular weight (top curve; historical data). The lower curve shows the incremental change for each 25

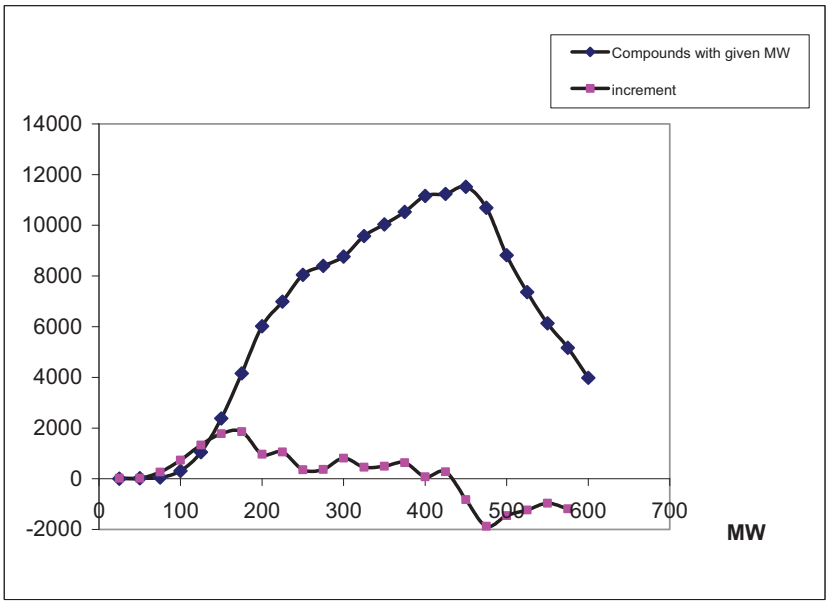


Figure 3-5. The number of carboxylic acids with a given molecular weight in the GSK compound collection, together with the change in 25Da increments.

dalton increase. As can be seen, the number of carboxylic acids in a particular molecular weight band initially increases rapidly, in an approximately exponential manner. However, around 150 Daltons this exponential behaviour ceases; indeed the curve goes through a maximum at around a molecular weight of 400. The set therefore significantly under-samples the virtual space of carboxylic acids in a manner that is progressively worse as the molecular weight increases. Thus in the lower molecular weight region (e.g. <350 daltons, typical of many fragment sets) the set of acids provides a more effective sampling than is the case at higher Molecular Weight (say 425) because there is a smaller difference between the numbers of available compounds and the number theoretically possible. This is schematically illustrated in Figure 3-6 which includes an extrapolation of the initial rate of increase in Figure 3-5. Such an exponential increase was also observed in studies to enumerate all possible molecules containing up to 11 non-hydrogen atoms (Fink et al 2005).

One of the challenges in fragment-based lead discovery is that the activities of the molecules identified will often be lower than for larger, more drug-like molecules. This may result in a promising fragment of relatively weak potency being overlooked in favour of more complex molecules that have higher initial potency but which are ultimately less developable. It has therefore proved useful to have ways to compare molecules that takes account of their size. A particularly useful concept in this context is the *maximal affinity* of a ligand, first introduced by Andrews and colleagues (Andrews et al 1984). The maximal affinity equals the maximum free energy of interaction

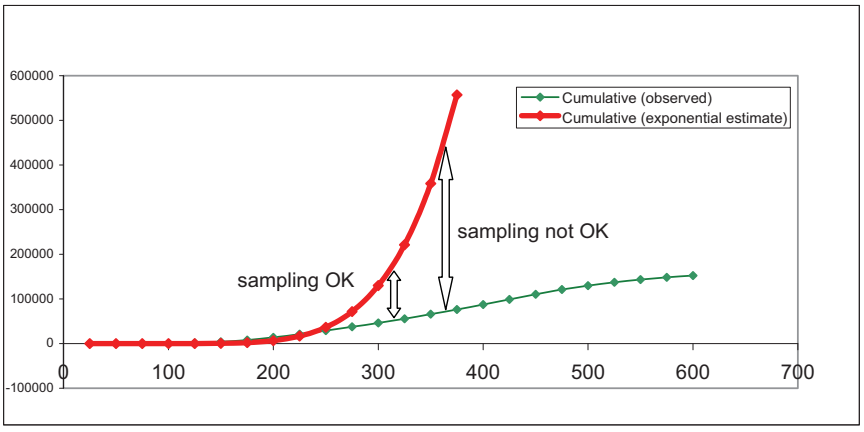


Figure 3-6. Graph to illustrate that the set of available carboxylic acids is a reasonable representation of the total number possible for low molecular weight, but that due to the exponential increase in the number of structures this sampling becomes inappropriate at higher molecular weight values.

with a biological macromolecule. Andrews and colleagues extracted a set of ligands with known binding affinities from the literature. The functional groups present in each ligand were identified and a multiple linear regression analysis performed in order to determine the contributions of each functional group to the binding affinity (together with an entropy term related to the freezing of rotatable degrees of freedom). Summing the corresponding contributions for any ligand gives its maximal binding energy (i.e. the maximal interaction that would be expected if all functional groups make their optimal contribution).

In a more recent study, Kuntz and colleagues analysed a data set of 160 ligands (Kuntz et al 1999). When the binding affinity per non-hydrogen atom was plotted as a function of the number of atoms then the graph shown in Figure 3-7 results. The initial slope of this graph has a value of approximately 1.5kcal/mol and this was therefore proposed as the maximal free energy contribution per non hydrogen atom. As can be seen from the graph, once a molecule contains approximately 15 non hydrogen atoms the free energy tends to increase little with molecular mass. A number of reasons for this shift were proposed and were primarily considered to arise, not from any fundamental thermodynamic reason, but due to the properties of very tight binding ligands (such as very long dissociation times). Moreover, many

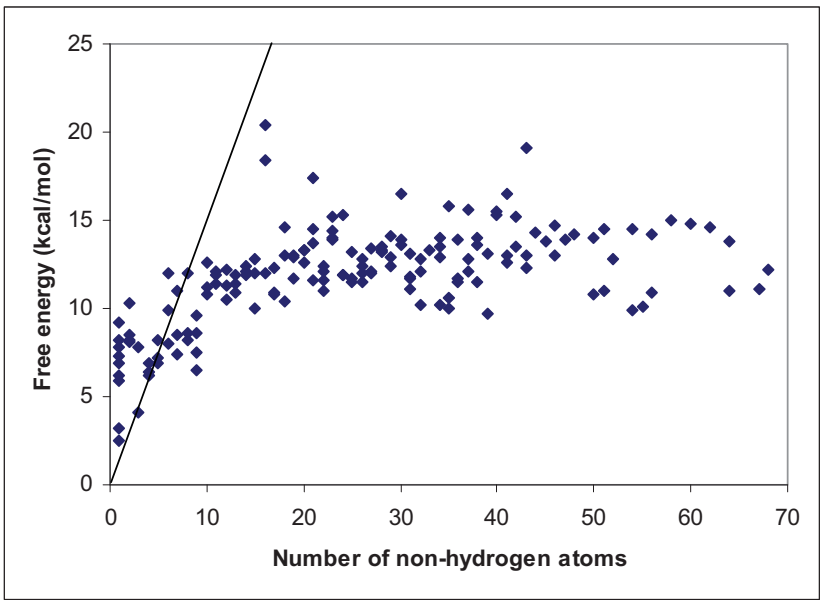


Figure 3-7. Free energy of binding plotted against number of non-hydrogen atoms (Kuntz et al 1999). The initial line has a slope of 1.5kcal/mol-atom.

assays are configured so that affinities greater than nanomolar cannot be effectively measured.

Researchers at Pfizer have termed the experimental binding affinity per atom the "ligand efficiency" and have proposed that it is a useful parameter to use when prioritising the hits from screening experiments (Hopkins et al 2004). Moreover, a lower limit on the ligand efficiency can be estimated by assuming that the goal is to achieve a binding constant of 10nM in a molecule with molecular weight of 500. An analysis of the Pfizer screening collection revealed that the mean molecular mass for a non hydrogen atom in their "drug like" compounds is 13.3 and so a molecule with a molecular weight of 500 and a binding constant of 10nM would have 38 non-hydrogen atoms and a ligand efficiency of 0.29kcal/mol per non-H atom. This is significantly less than the maximal value of 1.5kcal/mol per atom. The Pfizer proposal was that the hits with the highest ligand efficiencies should be selected for optimisation, all other factors being equal. A straightforward extension of these ideas enables other properties to be taken into account. For example to achieve compounds with an acceptable logP (not too high) whilst retaining potency one could use the difference between the log potency and the logP as the term for comparison. Burrows and colleagues at Astra Zenecca have proposed that when this term is greater than 2 log units then it is likely that the compound will be a good lead compound (Burrows 2004). Further expansion of these ideas was provided by Abad-Zapatero and Metz who in addition to normalising the binding affinity by molecular weight also defined a surface-binding efficiency index in which the polar surface area was used as the normalising factor (Abad-Zapatero and Metz 2005).

When considering potency it should always be remembered that there is a logarithmic relationship between the free energy and the equilibrium constant. Thus when a molecule with low nanomolar potency is split into two fragments the individual potencies of the derived fragments are unlikely to have a potency better than tens if not hundreds of micromolar. The converse, of course, is that if one can successfully link two weakly potent fragments then this may afford a low nanomolar compound. Moreover, the expected affinity of the joined molecule should be larger than the sum of the affinities of the two individual fragments (Page and Jencks 1971; Murray and Vedonk 2002). The reason for this is that a fragment loses significant rigid body translational and rotational entropy when it forms the intermolecular complex. This unfavourable entropic term is only weakly dependent on molecular weight. Thus, whereas two unfavourable terms are present when two fragments bind this is replaced by just one unfavourable term for the combined molecule.

Further theoretical insights into the properties required for small molecular fragments to be successful leads were provided by Rejto and Verkhiver, whose starting premise was that the primary molecular

recognition event between a protein binding site could be accomplished using a core fragment, or "molecular anchor" (Rejto and Verkhivker 1996). Such a fragment could then be grown into a complete ligand. The system used for their initial investigations was the interaction between the FK506 binding protein (FKBP-12) and the FK506 inhibitor (Figure 3-8). The FK506 inhibitor contains a pipecolinyl moiety that anchors the ligand into the binding site. The pipecolinyl moiety itself is not a particularly strong binding ligand. It was therefore proposed that binding affinity alone might not be the sole factor in determining whether a fragment would be a good molecular anchor.

When several fragments derived from the structure of FK506 were taken and docked into the binding site on FKBP-12 it was observed that for the

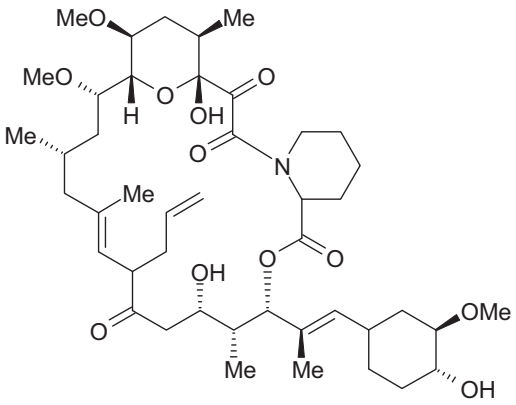


Figure 3-8. FK506.

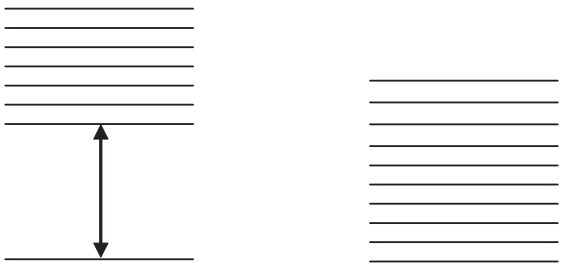


Figure 3-9. The preferred binding energy spectrum for a molecular anchor has a significant gap between the lowest energy mode and the next lowest (left) rather than many binding modes of similar energy (right).

pipecolinyl fragment just a single binding mode was derived, in contrast to other fragments for which multiple binding modes of comparable (predicted) binding affinity could be found. This suggested that the binding energy landscape for a successful molecular anchor would have one binding mode much lower in energy than the next lowest level; unsuccessful fragments were predicted to have an energy spectrum with many different binding modes of approximately the same energy (Figure 3-9). This stability gap was the unique feature of successful fragments compared to random ones.

4. CONSTRUCTING A FRAGMENT SET FOR SCREENING

The concepts of "lead-likeness" rapidly gained currency within the biotechnology and pharmaceutical industries. How though can some of the ideas described above be put into practice? The two key issues to be considered are which molecules to include in any "lead-like" or fragment screening set, and which technique (or techniques) should be used to screen the set against biological targets. Clearly the two problems are closely related; all experimental assay techniques impose some kind of constraint on the properties of the molecules involved. Further details on the various types of assay methodology used in fragment screening can be found elsewhere in this book and so will not be considered here in any detail. Rather, we will concentrate on the design and construction of the screening sets themselves from a computational design perspective. We will nevertheless recognise that the assay technique or techniques being used will impose restrictions on the properties of the molecules being screened. For example, lead-like starting points are likely to have less potency and will not normally be found in high-throughput screening where concentrations of the order 10uM are typically used. The obvious solution is to screen compounds at higher concentrations but this then introduces other problems related to compound solubility, purity and interference with readout (e.g. by fluorescence quenching) that need to be considered.

There are many ways to design screening sets. A distinction is often made between the large, "diverse" sets such as those used in high throughput screening and the smaller, "focussed" sets that are selected with a specific target or group of related targets in mind. The methods that can be used to construct diverse and focussed screening sets have been discussed and reviewed extensively in the literature and so will not be considered in detail here except where there are specific issues arising from the use of fragments. One factor worth considering at this point is the balance between diversity and focus. The knowledge plot shown in Figure 3-10 (Hann et al 2004) has

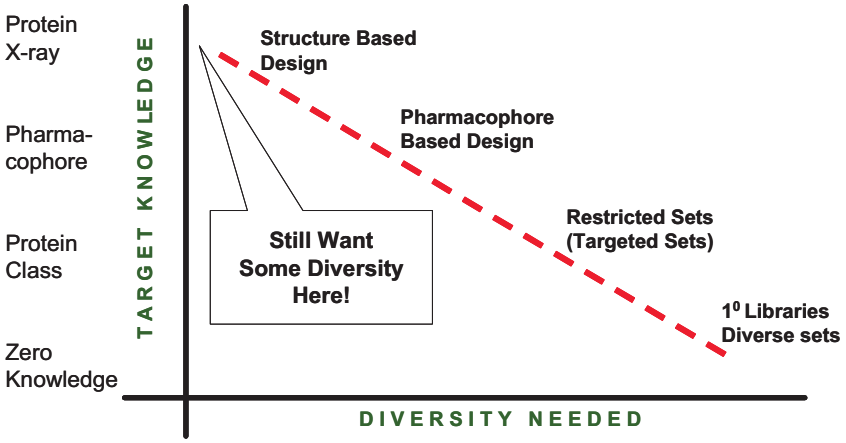


Figure 3-10. Knowledge plot illustrating that the degree of diversity required is generally inversely proportional to the amount of knowledge available.

proved particularly useful in this regard. This diagram indicates that the need for diversity is inversely proportional to the knowledge that is available on the biological target.

A number of computational techniques have proved useful when deciding which fragments to include in a fragment screening set. Many of the filters described above (for example those to remove reactive compounds) are applicable and are often used as initial filters. With regard to the physicochemical properties of the ligands then a more restrictive set of parameters than those used to characterise drug-likeness is generally considered more appropriate, consistent with the conclusions of the retrospective analyses indicated above. One of the reasons for the widespread adoption of Lipinski's "rule of five" was the ease with which it could be implemented as an *in silico* filter. Within the area of fragment screening an analogous rule has also been suggested (Congreve et al 2003). Analysis of the hits obtained by screening collections of fragments using X-ray crystallography against a variety of targets such as kinases and proteases suggested that a "rule of three" might be applicable. This rule suggests that the molecular weight be limited to less than 300; that the number of hydrogen bond donors should be three or fewer; the number of hydrogen bond acceptors should also be three or fewer; and the calculated octanol/water partition coefficient (using ClogP) should be less than or equal to 3. Three or fewer rotatable bonds and a polar surface area of 60Å^2 or less were also proposed as useful criteria.

The relatively low capacity of many of the screening techniques used in fragment-based discovery (compared to high-throughput screening) means

that it is usually necessary to further refine the initial set of compounds that meet such basic filters; some form of subset-selection algorithm is required. A particularly popular approach is to identify fragments related to those that commonly occur in drug-like molecules. Certain fragments (often called *privileged structures* (Müller 2003)) appear frequently in drug molecules, often conferring activity against a number of biological targets. Whilst some of the most common of these privileged structures have been identified manually a number of computational methods have also been developed to automatically identify pharmaceutically-relevant fragments from collections of drug-like molecules. The fragments are then sorted according to frequency and after removal of trivial examples such as simple alkyl groups the highest scoring fragments are identified for potential inclusion in the screening set.

Bemis and Murcko defined an hierarchial approach in which a molecule is converted into its graph representation and then broken down into ring systems, linker atoms and side chains (Bemis and Murcko 1996; Bemis and Murcko 1999). The ring systems and linkers together define a *framework*, as illustrated in Figure 3-11. Bemis and Murcko identified the top scoring frameworks in the Comprehensive Medicinal Chemistry database, as shown in Figure 3-12. They found that just 32 frameworks accounted for 50% of the 5120 drug molecules in the entire set. An alternative approach that is widely used is the RECAP algorithm (Lewell et al 1998) (RECAP stands for Retrosynthetic Combinatorial Analysis Procedure). In RECAP the fragmentation is performed by cleaving bonds that can be easily formed in a reaction sequence, such as amides and ethers.

It is important to recognize that the fragments generated by procedures such as RECAP or that of Bemis and Murcko will not necessarily directly correspond to "real" molecules that can be purchased or synthesised; more usually it is necessary to identify an available molecule (or molecules) that contains the desired fragment for inclusion in the set.

By way of practical illustration, four examples of actual fragment sets are now described in outline. The first is an evolving set used at GlaxoSmithKline for high-concentration screening. From a large set of available in-house and external compounds a series of 2D substructure and property filters were first applied to identify potential candidates for inclusion (MW <400, rotatable bonds < 6, heavy atoms < 22, Donors < 3, acceptors < 8, ClogP < 2.2). The selection criteria also required there to be a synthetic handle present in order to facilitate the rapid synthesis of further analogues. A diversity measure based on 3D pharmacophore keys was then used to select a subset of compounds from the initial filtered selection (Leach et al 2000; Hann et al 2005). Examples of the types of generic structures selected as a result of this process are shown in Figure 3-13. Scientists at AstraZeneca have used a broadly similar approach to select a set

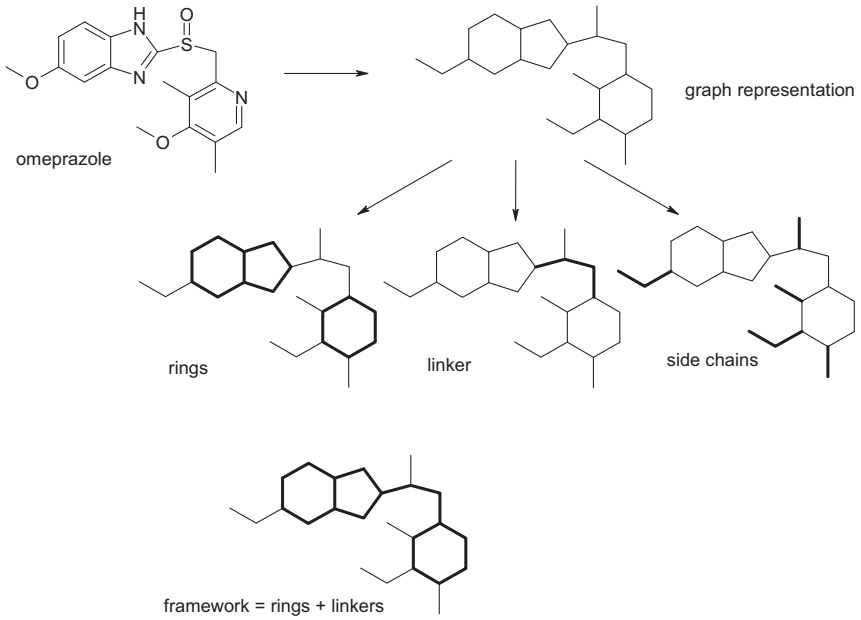


Figure 3-11. The creation of rings, linkers, side chains and frameworks from the molecular graph.

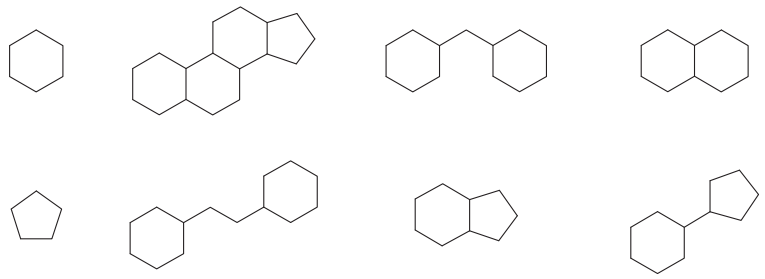


Figure 3-12. Top-scoring frameworks in drug molecules as identified by Bemis and Murcko.

of 2000 compounds designed to have a roughly equal proportion of acidic, basic and neutral compounds (with a small number of zwitterionic molecules) and with a pre-defined physicochemical property distribution (Burrows 2004). Vernalis have described four generations of a low molecular weight fragment library for use in NMR-based screening (Baurin et al 2004). Again, *in silico* filters and property calculations were

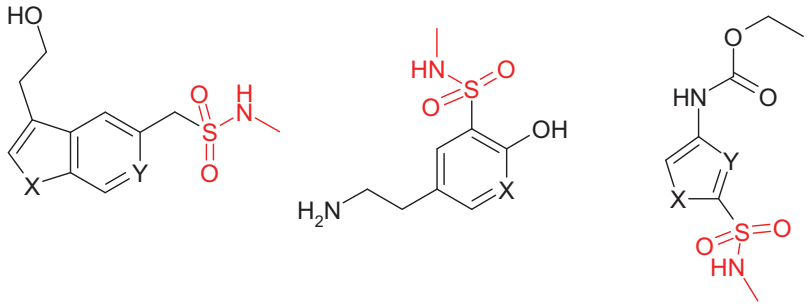


Figure 3-13. Typical fragments in GSK fragment set. X and Y represent heteroatoms and the sulphonamide groups correspond to positions where rapid synthesis expansion can occur should a hit be identified.

developed to help automate the selection process. Both "general purpose" sets together with those directed towards particular protein targets (kinases) were constructed. Pharmacophore-based descriptors were also used (analogous to those employed by the GSK group) as a measure of diversity and complexity. Finally, scientists at Astex have described the construction of screening sets for use in X-ray crystallographic fragment screening (Hartshorn et al 2004). Sets directed against a specific target or groups of related targets have been constructed together with a more general-purpose set. The starting point for the latter was a fragmentation analysis of drug molecules which identified a small set of commonly-found, simple organic ring systems. These ring systems were then combined with a set of desirable side chains. Three sources of side chains were used: those observed frequently in drug molecules, lipophilic/secondary side chains (intended to pick up hydrophobic interactions in a protein binding site), and a set of nitrogen substituents. Each of the relevant side chains was combined with the ring systems to give a virtual library (of size 4513); the structures in this virtual library were then compared against databases of compounds available from external sources giving a final set of 327 compounds.

As the examples above illustrate, there are several approaches that can be used to construct fragment screening sets, though certain steps are common to most of these. Differences arise from the varying emphasis placed on the various properties and diversity of the structures in the set and also on the screening technology or technologies that are used to perform the screening. Physical properties are often important, with solubility being particularly critical as the compounds will often be screened at a much higher concentration than is usual in typical high-throughput screens.

5. CONVERTING FRAGMENT HITS INTO LEADS

Having successfully identified one or more fragment "hits" and confirmed their activity it is then necessary to optimise the structure to give a sustainable lead series. It is frequently the case that the activity of the initial hits will be weak (sometimes very weak) and so a key requirement is to increase the potency. However, other properties should also be taken into account, such as selectivity against other targets and ADMET properties. Four scenarios have been identified (Rees et al 2004): fragment evolution; fragment linking; fragment self-assembly and fragment optimisation.

Fragment evolution is perhaps the most straightforward procedure, involving the addition of functionality that binds to adjacent regions of the active site and thereby improves the potency of the initial hit. Where structural information is available on the binding mode of the initial fragment hit (e.g. from X-ray crystallography) then the full range of structure-based design approaches can be used. When such structural information is not available then the screening of appropriate analogues of the original hit would be performed in order to try and establish a structure-activity relationship. The related compounds may already be available in-house or may be obtained from 3rd-party compound vendors; in other cases specific synthesis may be required.

An early example of the fragment evolution approach is the design and synthesis of DNA gyrase inhibitors (Boehm et al 2000). In this particular example a set of low molecular weight "needle" inhibitors was docked into the active site. 3D pharmacophore searching techniques were also employed. From these calculations a set of compounds was identified for testing; from a total of 3000 compounds tested 150 hits were obtained. Several hit validation techniques were employed to confirm which of these bound to the desired active site. The subsequent optimisation process was facilitated by having available X-ray structures of the compounds bound to the enzyme; during the optimisation the activity of the compounds increased by approximately four orders of magnitude (Figure 3-14).

Fragment linking is illustrated in Figure 3-15; this involves joining two fragments that bind at adjacent sites. There are two key challenges with this approach; it requires fragments that bind to two different sites and it requires a procedure by which these fragments can be linked together. Even in those cases where it is possible to find fragments binding to more than one site, the linking step can also be difficult to achieve. Having access to structural information on the binding modes is not a prerequisite but can obviously be of significant benefit.

An early example of the fragment linking approach was the identification of a 49-nM inhibitor of the KF506-binding protein (FKBP) using the

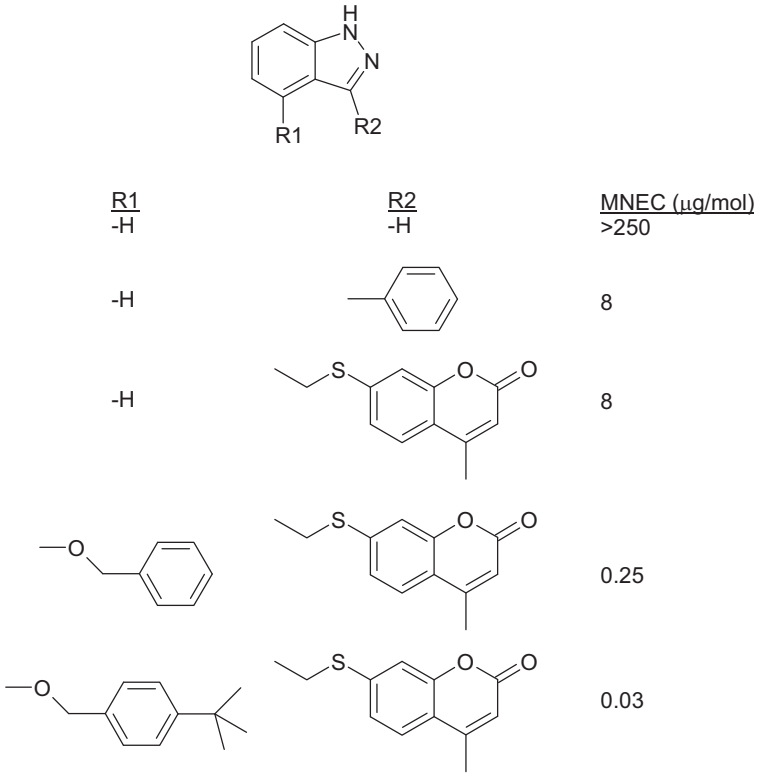


Figure 3-14. Evolution of indazole inhibitors of DNA Gyrase. MNEC is the maximal noneffective concentration, a measure of activity.

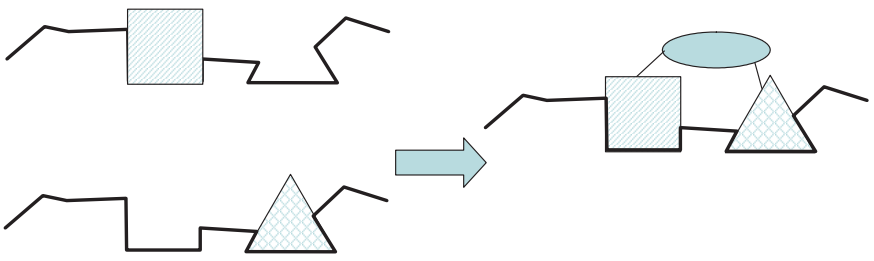


Figure 3-15. Fragment linking schematic.

SAR-by-NMR method (Shuker et al 1996). First, compounds that bound weakly to FKBP were identified. These included a trimethyoxyphenyl pipecolinic acid derivative (K_d=2.0uM). A second round of screening was then performed using the same library but in the presence of saturating

amounts of this pipecolinic acid fragment. This afforded a benzanilide derivative that bound with an affinity of 0.8mM. Subsequent screening of close analogues enabled the SAR to be expanded, and a model for the binding of these fragments to be developed. Four compounds that were designed to link the two sites were subsequently synthesized and found to have nanomolar activities (see Figure 3-16).

Fragment self-assembly involves the use of reactive fragments that link together to form an active inhibitor in the presence of the protein target. The essence of the approach is that the protein serves to select those combinations of reagents that act as inhibitors. Perhaps the first example of

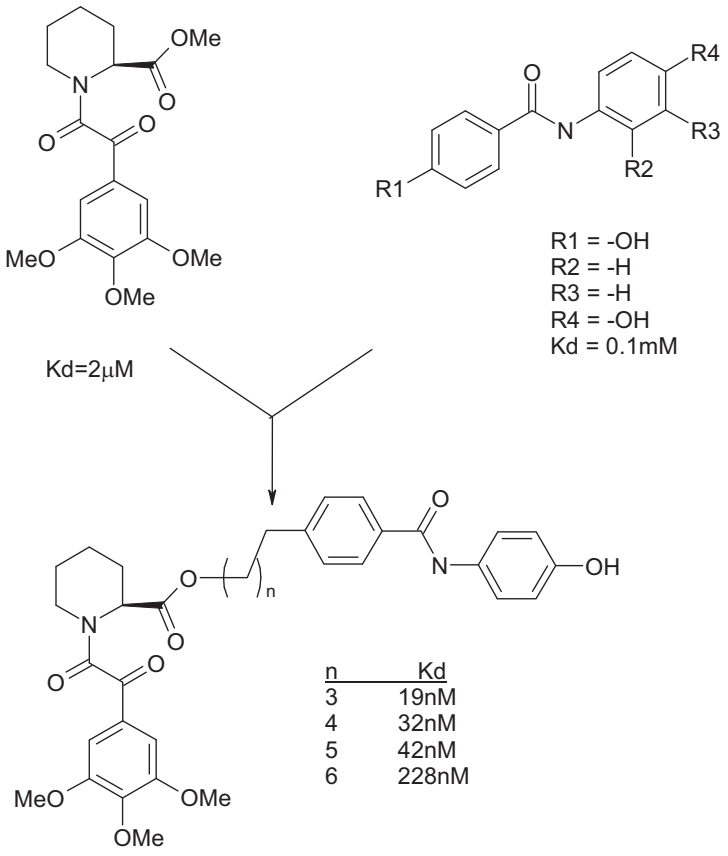


Figure 3-16. FK506 fragment linking example (Shuker et al 1996).

this approach was the reaction between four amines and three aldehydes to give imines (which were subsequently reduced to amines) (Huc and Lehn 1997). Twelve possible amines result from this reaction. When performed in the presence of carbonic anhydrase the proportion of one amine was increased, presumed to correspond to the most active inhibitor (Figure 3-17).

Fragment optimisation involves the optimisation or modification of a part of the molecule, often to enhance properties other than the inherent potency of the original molecule or to address some other problem. An example of this approach is the incorporation of alternative S1-binding fragments into a series of trans-lactam thrombin inhibitors (Hann et al 2001). In this works novel proflavin displacement assay was used to identify candidate fragments. Proflavin had been shown by X-ray crystallography to bind into the S1 pocket of thrombin (Conti et al 1998). This provided the basis for a simple absorbance-based assay to probe for S1 binders. Among the hits from this assay was 2-aminoimidazole whose binding mode in this region of the enzyme was subsequently confirmed using X-ray crystallographic analysis. It was subsequently incorporated into the trans-lactam series of inhibitors. The complexity of the trans-lactam system made it desirable to have a mechanism to prioritise potential S1 substituents in advance of committing chemistry resource (Figure 3-18).

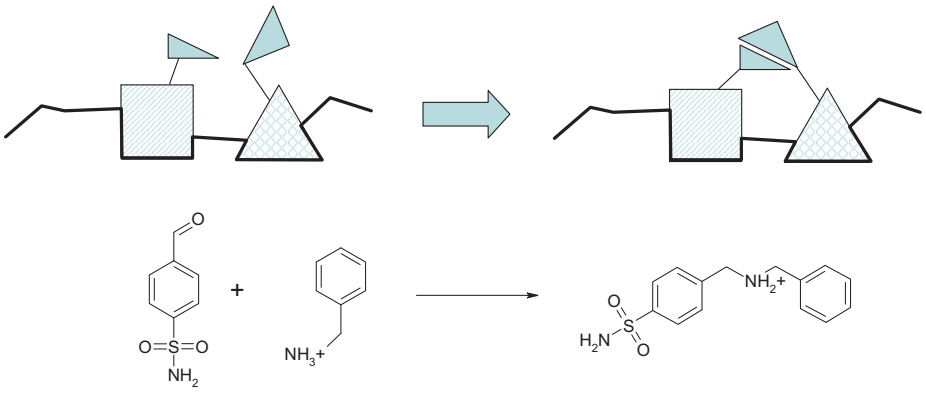


Figure 3-17. The general principle of the fragment self-assembly approach together with the amine whose yield was increased when the reaction was performed in the presence of carbonic anhydrase.

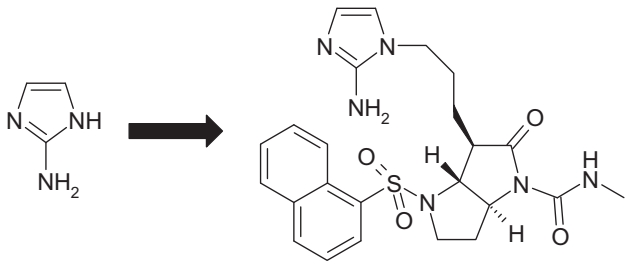


Figure 3-18. 2-amino imidazole was identified as a novel thrombin S1 binding group. This moiety was subsequently incorporated into the trans-lactam series as shown; an illustration of fragment optimisation.

6. SUMMARY

In this chapter we have provided an overview of the theoretical background to fragment methods and lead-likeness. Fragment-based approaches are still in an early stage of development and are just one of several techniques that can be used to identify novel lead compounds for drug development. There are in particular some practical challenges associated with fragment screening that relate to the generally lower level of potency that such compounds possess. Nevertheless, the approach also offers some significant advantages by providing less complex molecules which may have better potential for drug optimisation and by enabling chemical space to be more effectively explored. The next few years will undoubtedly see a maturing of the area and improvements in our understanding of how the concepts can be applied more widely to drug discovery.

REFERENCES

Abad-Zapatero C. and Metz J. T. 2005. Ligand efficiency indices as guidposts for drug discovery. Drug Discovery Today **10**, 464–469.

Andrews P. R., Craik D. J. and Martin J. L. 1984. Functional group contributions to drug-receptor interactions. J. Med. Chem. **27**, 1648–1657.

Baurin N., Aboul-Ela F., Barril X., Davis B., Drysdale M., Dymock B., Finch H., Fromont C., Richardson C., Simmonite H. et al. 2004. Design and characterization of libraries of molecular fragments for use in NMR screening against protein targets. J. Chem. Inf. Comput. Sci. **44**, 2157–2166.

Bemis G. W. and Murcko M. A. 1996. The properties of known drugs. 1. Molecular frameworks. J. Med. Chem. **39**, 2887–2893.

Bemis G. W. and Murcko M. A. 1999. The properties of known drugs. 2. Side chains. J. Med. Chem. **42**, 5095–5099.

Boehm H.-J., Boehringer M., Bur D., Gmuended H., Huber W., Klaus W., Kostrewa D., Kuehne H., Luebbers T., Muenier-Keller N. et al. 2000. Novel inhibitors of DNA Gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimisation. a promising alternative to random screening. J. Med. Chem. **43**, 2664–2674.

Bohacek R. S., Martin C. and Guida W. C. 1996. The art and practice of structure-based drug design: A molecular modeling perspective. Med. Res. Rev. 16, 3–50.

Burrows J. 2004. High concentration screening: integrated lead generation. Society for Medicines Research conference, December 2nd. See, http://webcasts.prous.com/SMR_DEC_2004/program.asp.

Clark D. E. and Pickett S. D. 2000. Computational methods for the prediction of 'drug-likeness'. Drug Discovery Today **5**, 49–58.

Congreve M., Carr R., Murray C. and Jhoti H. 2003. A 'Rule of Three' for fragment-based lead discovery. Drug Discovery Today **8**, 876–877.

Conti E., Rivetti C., Wonacott A. and Brick P. 1998. X-ray and spectrophotometric studies of the binding of proflavin to the S1 specificity pocket of human α-thrombin. FEBS Lett. **425**, 229–233.

Erlanson D. A., McDowell R. S. and O'Brien T. 2004. Fragment-based drug discovery. J. Med. Chem. **47**, 3463–3482.

Ertl P. 2003. Cheminformatics analysis of organic substituents: Identification of the most common substituents, calculation of substituent properties, and automatic identification of drug-like bioisosteric groups. J. Chem. Inf. Comput. Sci. **43**, 374–380.

Fattori D. 2004. Molecular recognition: The fragment approach in lead generation. Drug Discovery Today **9**, 229–238.

Fink T., Bruggesser H. and Reymond J.-L. 2005. Virtual exploration of the small-molecule chemical universe below 160 Daltons. Angew. Chem. Int. Ed. **44**, 1504–1508.

Hann M. M., Leach A. R. and Harper G. 2001. Molecular complexity and its impact on the probability of finding leads for drug discovery. J. Chem. Inf. Comput. Sci. **41**, 856–864.

Hann M. M., Leach A. R. and Green D. V. S. 2004. Computational chemistry, molecular complexity and screening set design in 'Chemoinformatics in Drug Discovery', Oprea T editor Wiley-VCH. 43–57.

Hartshorn M. J., Murray C. W., Cleasby A., Frederickson M., Tickle I. J. and Jhoti H. 2004. Fragment-based lead discovery using X-ray crystallography. J. Med. Chem. **48**, 403–413.

Hopkins A. L., Groom C. R. and Alex A. 2004. Ligand efficiency: A useful metric for lead selection. Drug Discovery Today **9**, 430–431.

Huc I., and Lehn J.-M. 1997. Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly. Proceedings of the National Academy of Sciences USA **94**, 2106–2110

Kuntz I. D., Chen K., Sharp K. A. and Kollman P. A. 1999. The maximal affinity of ligands. Proc. Natl. Acad. Sci. U. S. A. **96**, 9997–10002.

Leach A. R., Green D. V. S., Hann M. M., Judd D. B. and Good A. C. 2000. Where are the GaPs? A rational approach to monomer acquisition and selection. J. Chem. Inf. Comput. Sci. 40**,** 1262–1269.

Leach A. R. and Gillet V. J. 2004. An introduction to cheminformatics. Kluwer.

Lewell X.-Q., Judd D. B., Watson S. P. and Hann M. M. 1998. RECAP-retrosynthetic combinatorial analysis procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. J. Chem. Inf. Comput. Sci. **38**, 511–522.

Lipinski C. A. 2001. Drug-like properties and the causes of poor solubility and poor permeability. J. Pharmacol. Toxicol. Methods **44**, 235–249.

Lipinski C. A., Lombardo F., Dominy B. W. and Feeney P. J. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **23**, 3–25.

Müller G. 2003. Medicinal chemistry of target family-directed masterkeys. Drug Discovery Today **8**, 681–691.

Murray C. W. and Verdonk M. L. 2002. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. J. Comput. Aided Mol. Des. **16**, 741–753.

Oprea T. I., Davis A. M., Teague S. J. and Leeson P. D. 2001. Is there a difference between leads and drugs? A historical perspective. J. Chem. Inf. Comput. Sci. **41**, 1308–1315.

Page M. I. and Jencks W. P. 1971. Proc. Natl. Acad. Sci. U. S. A. **68**, 1678–1683.

Rees D. C., Congreve M., Murray C. W. and Carr R. 2004. Fragment-based lead discovery. Nature Rev. Drug Disc. **3**, 660–672.

Rejto P. A. and Verkhivker G. M. 1996. Unraveling principles of lead discovery: From unfrustrated energy landscapes to novel molecular anchors. Proc. Natl. Acad. Sci. U. S. A. **93**, 8945–8950.

Rishton G. M. 2003. Nonleadlikeness and leadlikeness in biochemical screening. Drug Discovery Today **8**, 86–96.

Shuker S. B., Hajduk P. J., Meadows R. P. and Fesik S. W. 1996. Discoverying high-affinity ligands for proteins: SAR by NMR. Science **274**, 1531–1534.

Sneader W. 1996. Drug prototypes and their exploitation. John Wiley and Sons, Ltd.

Teague S. J., Davis A. M., Leeson P. D. and Oprea T. 1999. The design of leadlike combinatorial libraries. Angew. Chem. Int. Ed. **38**, 3743–3748.

Veber D. F., Johnson S. R., Cheng H.-Y., Smith B. R., Ward K. W. and Kopple K. D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. **45**, 2615–2623.

Vieth M., Siegel M. G., Higgs R. E., Watson I. A., Robertson D. H., Savin K. A., Durst G. L. and Hipskind P. A. 2004. Characteristic physical properties and structural fragments of marketed oral drugs. J. Med. Chem. **47**, 224–232.

Walters W. P. and Murcko M. A. 2000. Library filtering systems and prediction of druglike properties. In Methods and principles in medicinal chemistry, vol. 10 (Virtual Screening for Bioactive, Molecules) 15–32.

Chapter 4

FRAGMENT-BASED NMR SCREENING IN LEAD DISCOVERY

Christopher A. Lepre and Jonathan M. Moore
Vertex Pharmaceuticals Incorporated, 130 Waverly Street, Cambridge, MA 02139

1. INTRODUCTION

Traditional lead discovery has historically been driven by screening very large libraries of compounds to identify molecules with enzymatic or biological activity. However, traditional methods may not provide drug leads with suitable potency, novelty, molecular diversity or physicochemical properties. As a result, fragment-based screening has recently gained acceptance as an alternative approach for generating high quality lead molecules in pharmaceutical discovery. Why has fragment-based screening become so popular? A fragment-based approach can effectively represent the chemical diversity of a large, fully enumerated library without requiring the purchase or synthesis of enormous numbers of compounds. In addition, smaller scaffolds can provide better starting points for medicinal chemistry, as they can be elaborated into larger, more potent compounds, without pushing the limits of physicochemical properties such as molecular weight, polar surface area and clogP (which are known to correlate with oral bioavailability) (Lipinski et al., 1997; Teague et al., 1999).

From a practical perspective, fragment-based screening may be carried out via any physical method that is capable of detecting binding of a small molecule to a macromolecular target. The most popular techniques employ NMR spectroscopy, X-ray crystallography, and mass spectroscopy. NMR screening in particular has evolved into a proven method for lead generation, and in this chapter we will describe theoretical aspects and practical applications of the most commonly used NMR-based screening approaches.

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 73–98.

X-ray crystallographic-based methods will be examined elsewhere in this volume.

All fragment-based design strategies, whether they use biophysical or biochemical methods to detect binding of small molecules to a drug target, share common elements: i) a purified drug target, ii) a means of detecting binding, and iii) a strategy for use of binding information to generate drug leads. When the field of fragment-based screening was born, literature descriptions of NMR-based screening work focused on the method of detection, rather than the lead generation strategy used. For example, the SAR by NMR approach (Shuker et al., 1996), as initially proposed, used ^{15}N-^{1}H heteronuclear NMR to detect binding, and a fragment linking strategy to identify and optimize leads. Similarly, the SHAPES strategy, as originally described (Fejzo et al., 1999), used ligand-directed rather than protein-directed methods of detection, and a combination or fragment fusion strategy to generate more potent binders. As the number of studies expanded, it became evident that experimental approaches and ligand design strategies could be combined in a wide variety of ways to best address each target and drug design problem. For this reason, it is best to consider the physical methods used in NMR-based screening separately from the strategies, as we have done in this review.

Although the techniques described in this chapter were sometimes initially proposed as standalone technologies, the examples we provide clearly show that NMR screening is best deployed as one component of an integrated platform of biophysical, biochemical, computational and chemical approaches such as X-ray crystallography, enzymology, virtual screening and combinatorial chemistry. In this context, significant synergies exist that can accelerate the identification of novel, drug-like lead classes of compounds for synthesis and optimization.

While the early literature focused primarily on proof-of-concept studies with model systems (Fejzo, 2002; Shuker et al., 1996) and development of the experimental techniques required to detect ligand binding by NMR, most recent work has described applications of NMR fragment-based screening and the insight into inhibitor design that these methods provide to numerous real-life drug discovery programs. Because there are many excellent, comprehensive reviews available on the subject (Pellecchia et al., 2002b; Peng et al., 2004; Peng et al., 2001; Stockman and Dalvit, 2002; van Dongen et al., 2002b; Wyss et al., 2002), we will not attempt to review the entire field of NMR screening, but rather focus on some key examples from the recent literature. We will illustrate the common ligand design strategies by providing examples of how NMR methods have been applied to generate and optimize new chemical classes of drug leads for therapeutically relevant drug targets.

2. EXPERIMENTAL METHODS – DETECTION OF BINDING BY NMR

2.1 Protein-directed methods

It is well established that NMR is a sensitive method for detecting binding of small molecules to proteins or other macromolecular targets. Binding may be detected by observing either target or ligand NMR resonances. There are advantages and disadvantages associated with either experimental method, and the optimum strategy will vary depending on such factors as the nature of the drug design problem at hand, the molecular weight of the target, and the feasibility of expressing reasonable quantities of isotopically labelled protein.

Detection of ligand binding by observation of receptor resonances is accomplished by using heteronuclear correlation spectroscopy to obtain an initial "fingerprint" of receptor amide or methyl protons in the unliganded state, followed by additional experiments with different mixtures of compounds. By observing and comparing the chemical shifts of ^{1}H-^{15}N (or ^{1}H-^{13}C) resonances with and without compounds present, it is possible to identify which mixtures contain compounds that bind to the target. Additional spectra are subsequently collected to deconvolute the mixtures and determine which compounds are binding. If sequence-specific resonance assignments exist for the observed peaks, it is possible to localize the binding site(s) of the small molecule within the three-dimensional structure of the target. These experiments alone are often mistakenly referred to as "SAR by NMR" (Shuker et al., 1996). In fact, the term "SAR by NMR" refers to a process for ligand assembly that incorporates heteronuclear 2-dimensional (2D) NMR-based screening of molecular fragments followed by design of linkers to connect fragments bound at adjacent subsites on the target protein, such that larger, higher affinity compounds are obtained. Many applications and modifications of SAR by NMR have been reported (Hajduk et al., 2000a; Hajduk et al., 1999b; Ross et al., 2000; Ross and Senn, 2001), and have been covered in a number of excellent reviews (Hajduk et al., 1999c; Lepre et al., 2004; Moore, 1999; Pellecchia et al., 2002b; Peng, 2004; Roberts, 2000; Stockman and Dalvit, 2002; van Dongen et al., 2002b; Wyss et al., 2002).

Ligand binding is manifested as shifting of peak positions in the HSQC spectrum for the apo- versus liganded spectra, an example of which is shown in Figure 4-1. The SAR by NMR process, shown schematically in Figure 4-2, consists of several steps. A molecule is identified that binds to one subsite. This scaffold is optimized for the first subsite. Next, a compound that binds

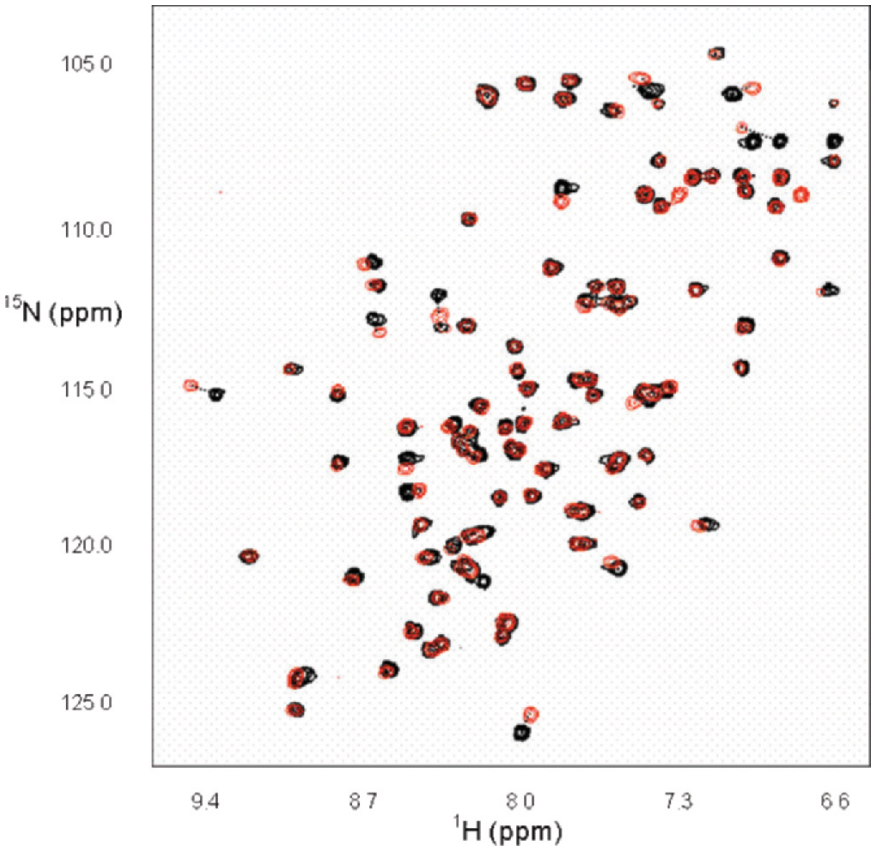


Figure 4-1. Example of SAR by NMR chemical shift perturbation. A superposition of 15N-HSQC spectra for FKBP in the absence (red contours) and presence (black contours) of compound 3. Both spectra were acquired in the presence of saturating amounts (2.0 mM) of compound 2. Significant chemical shift changes are observed for annotated residues. Compounds 2 and 3 are small molecules that bind to adjacent subsites within the larger FK506 binding site, as described in Shuker et al. (Shuker et al., 1996). Figure and legend were reproduced with permission from (Shuker et al., 1996). Copyright 1996 Science. Figure kindly provided by Dr. Phil Hajduk, Abbott Laboratories.

to an adjacent subsite is identified, and then optimized for maximum affinity to that subsite. Finally, based on the structure, a linker is designed to connect the scaffolds.

Detecting binding via protein resonances provides two immediate advantages over detecting binding via ligand resonances. Since the protein is assigned and the binding sites are localized, it is possible to immediately

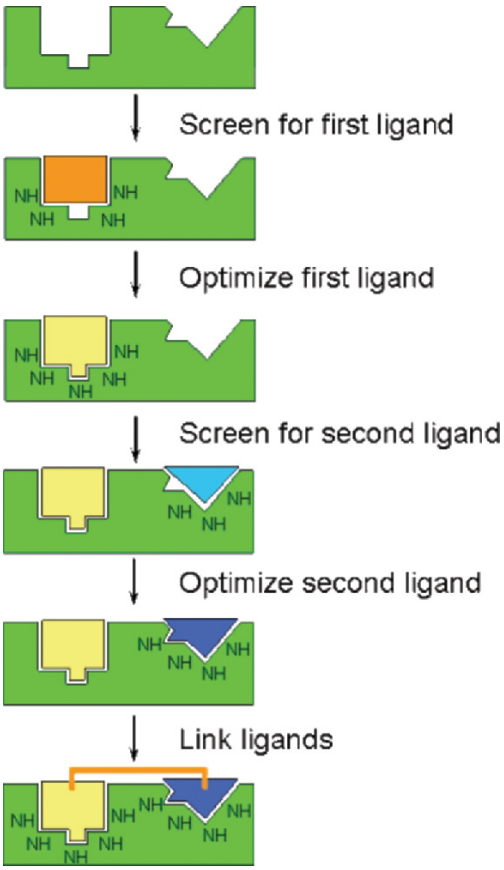


Figure 4-2. Outline of the SAR by NMR method. Figure reproduced with permission from (Shuker et al., 1996). Copyright 1996 Science. Figure kindly provided by Dr. Phil Hajduk, Abbott Laboratories.

distinguish binding to desirable vs. undesirable sites. Another advantage of receptor-based methods is that it is possible to detect binding of both low and high affinity hits. For example, high affinity compounds, defined here as compounds that bind with dissociation constant, $K_{D,}$ below 0.1 micromolar, will display two distinct sets of peaks in 2D heteronuclear correlation spectra corresponding to the free and bound states of the target. Alternatively, binding of lower affinity compounds, e.g. with K_D > 1 μM, will be manifested by changes in chemical shift, broadening, and/or disappearance of a subset of amide peaks, depending on whether the ligand is in intermediate or fast exchange with the target. For a detailed treatment of

chemical exchange and the influence of exchange rates on NMR parameters such as chemical shift and transverse and longitudinal relaxation rates, the reader is referred to some of the more physical reviews in the literature (Lepre et al., 2004; Peng et al., 2001).

Although receptor-based methods may be considered more information-rich than ligand-detected methods, limitations exist that must also be considered. Many important targets are not accessible for receptor-based screening simply because their molecular mass is beyond the range (>30 kDa) for which sequence-specific resonance assignments are practical, and even for targets below 30 kDa, significant challenges remain. For example, current resonance assignment strategies require milligram quantities of soluble, monodisperse, uniformly isotope labelled (^{13}C, ^{15}N, ^{2}H) protein. For targets above 20 kDa, additional strategies such as aminoacid specific labelling, and stereoselective ^{13}C and ^{2}H methyl labelling are often invoked. The requirement for isotopic labelling in turn implies that a suitable expression host must be available. Unfortunately, it is often the case that mammalian proteins are difficult to express in a bacterial expression host such as *E. coli*, where overexpression of these proteins results in toxicity to the host cell. In these cases, and for targets that do not express to high levels, isotope labelling can be very costly or impossible. Beyond the cost and difficulty of isotope labelling, there are additional requirements for sample stability. Samples needed to collect a complete data set for resonance assignment must be stable at room temperature for a week or more. (Note that this differs from the data collection times required for the actual screening, which is much more rapid). In addition to difficulties with sample preparation, the process of resonance assignment can be lengthy, requiring weeks or months, after data collection, for a large monomeric protein of < 30 kDa. The time required for assignment can be severely limiting and frustrating, in that other approaches, such as X-ray crystallography, can provide structural data for modeling and chemistry with much faster turnaround. However, when the target is amenable, and particularly when complementary structural information can be generated for bound ligands, protein-directed methods can be a very powerful and elegant means for obtaining and optimizing leads.

2.2 Ligand-directed methods

Because of the practical limitations outlined above, ligand-based screening methods have become more widely used in industrial settings than protein-directed methods. These methods can be applied to any target without regard to molecular weight. In fact, the most popular methods for ligand-based screening become more sensitive as molecular mass of the

target increases, and they have even been successfully applied to receptors bound to membranes, liposomes or solid supports (Claasen et al., 2005; Klein et al., 1999; Meinecke and Meyer, 2001). Using these methods, it is feasible to screen libraries of 1000 compounds with less than a milligram of unlabeled protein. Also, since isotope labelling is not required, expression is not limited to bacterial hosts, which substantially increases the pool of potential targets that can be examined.

Ligand-based detection of binding is generally accomplished via one of two mechanisms. The first mechanism exploits the differences in mobility of the ligand in the free state versus bound to the receptor. NMR parameters are very sensitive to these differences. NMR screening libraries typically consist of small molecules with masses < 500 Da. Compounds of this size exhibit small longitudinal (R_1) and transverse (R_2) relaxation rates, weak or vanishing 2D NOESY cross peaks, and large translational diffusion coefficients (D_t). Bound compounds, however, will assume the NMR parameters of the significantly larger receptor (e.g. with MW > 30 kDa). These properties include large R_2, large selective R_1, positive 2D NOESY cross peaks, and smaller diffusion coefficients. Under suitable conditions, even small fractions of ligand populations bound to the receptor will produce easily observable differences in the ligand NMR spectra, and provide a simple means to detect binding. For the interested reader, more physical descriptions of the exchange processes and how they affect NMR parameters may be found elsewhere (Lepre et al., 2004; Peng, 2004).

A second mechanism to detect binding by NMR involves detecting ^{1}H magnetization that is transferred from the receptor to the bound ligand. In this mechanism, molecules bound to the receptor "pick up" magnetization from the target protein, while ligands that do not bind are unaffected. Because experiments based on this mechanism are very sensitive, easy to implement, and simple to interpret, this mechanism of detection is highly popular and has been widely adopted. For this reason we will discuss two such methods, the Saturation Transfer Difference (STD) (Mayer and Meyer, 1999), and WATERLOGSY (Dalvit et al., 2001; Dalvit et al., 2000) methods, in more detail below.

2.3 Saturation transfer difference (STD) methods

Saturation transfer difference, as the name indicates, is an NMR difference experiment. The experiment is shown schematically in Figures 4-3 and 4-4. In the basic experiment, a frequency selective pulse train is applied that excites a subset of receptor resonances. Care must be taken that the selective saturation occurs at frequencies for which no small molecule ligand resonances are present, for example, the upfield shifted methyl groups of

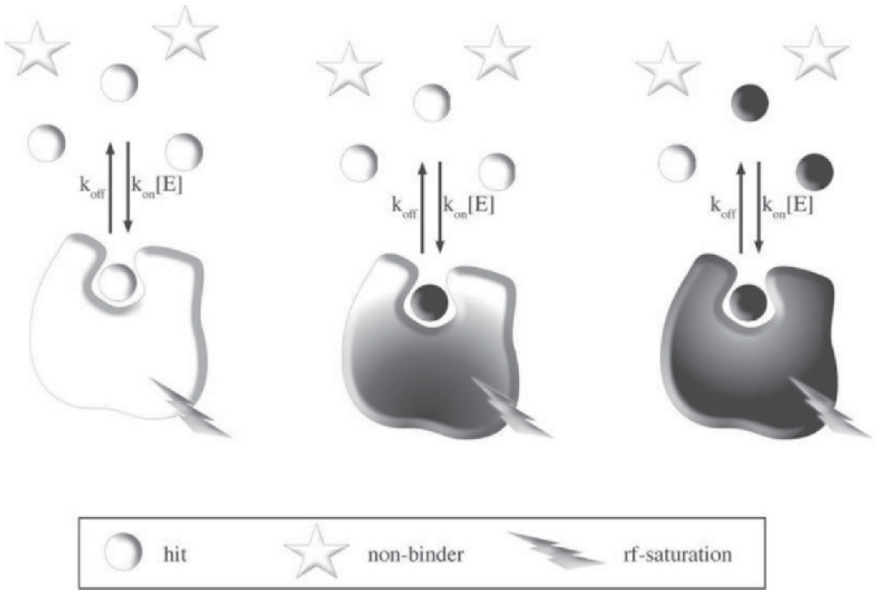


Figure 4-3. Detection of binding using the Saturation Transfer Difference (STD) experiment. Frequency selective irradiation (lightning bolt) causes selective ^{1}H saturation (shading) of the target receptor. Irradiation is applied for a sustained interval during which saturation spreads throughout the entire receptor via ^{1}H-^{1}H cross-relaxation (spin-diffusion). Saturation is transferred to binding compounds (circles) during their residence time in the receptor binding site. The number of ligands having experienced saturation transfer increases as more ligand exchanges on and off the receptor during the sustained saturation period. Non-binding compounds (stars) are unaffected. Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

protein Val, Leu, and Ile side chains. The saturation, which arises locally from the selectively excited protons, then spreads throughout the network of receptor protons via intramolecular ^{1}H-^{1}H cross relaxation pathways. Saturation transfer is highly efficient in a large molecular weight receptor, as these receptors have large rotational correlation times, which enhance spin-diffusion and thus transfer of magnetization within the receptor. As saturation spreads to the ligand binding site (Figure 4-3, right), saturation is transferred from the receptor to the bound small molecule via intermolecular cross relaxation. After the ligand dissociates, the retained magnetization persists due to the small longitudinal relaxation rate R_1 (or alternatively, long relaxation time T_1) of the free ligand protons. Simultaneously, more unsaturated ligand binds and dissociates as the target is continuously saturated, accumulating a population of saturated ligand molecules. With prolonged saturation times, it is possible for a low concentration of target to produce a much higher concentration of saturated ligand, in effect

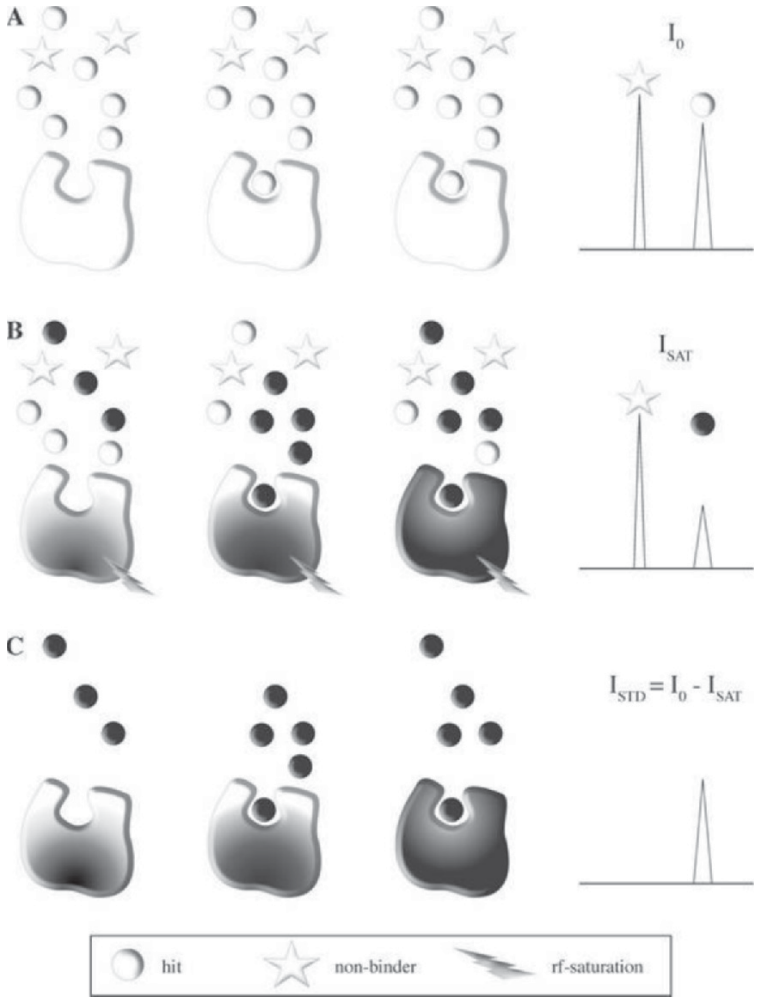


Figure 4-4. Schematic diagram depicting difference spectroscopy in the STD experiment. Circles and stars indicate binding and non-binding compounds, respectively. STD involves two experiments: an off-resonance and on-resonance experiment. Top panel A: off-resonance (reference) applies rf-irradiation off-resonance from both receptor and compound protons. Detection produces spectra with intensity I^0. Middle panel B: in on-resonance experiment, the rf-irradiation selectively saturates receptor and any binding compounds (indicated by dark shading). This manifests as the decreased signal intensity I^{SAT}. Bottom panel C: the STD response is the spectral difference $I^{STD} = I^0 - I^{SAT}$, which yields only resonances of the receptor and binding compounds. Receptor resonances are usually invisible due to either low concentration or relaxation filtering. The STD sensitivity depends on the number of ligands receiving saturation from the receptor and can be described in terms of the average number of saturated ligands produced per receptor molecule. Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

amplifying the binding signal. After the on-resonance experiment is collected, a reference experiment is then collected that applies the saturating RF pulse train "off-resonance", in a spectral region where no NMR resonances are excited. The NMR pulse sequence used encodes the two different experiments such that data is collected in an interleaved manner and subtracted by the NMR spectrometer software. After data is processed, the resulting spectrum will show signals arising only from ligands that bind the receptor. A pictorial description of how this is achieved is shown schematically in Figure 4-4.

It is straightforward to identify which compound(s) in a mixture bind the receptor by comparison with reference spectra for the small molecule compound mixtures, and no further deconvolution is necessary. An example of a processed STD spectrum and a reference spectrum is shown in Figure 4-5. A noteworthy advantage of the STD method over other earlier ligand-directed methods is that it is not necessary to subtract or otherwise account for free-state contributions of ligand signals. This permits the use of large ligand excesses, which provide further increases in sensitivity. As an approximation, this holds as long as the dissociation constant, $K_D < L_T$, where L_T is the total concentration of ligand. Generally, the STD method can be used to detect binding for ligands within a K_D range of $10^{-8} < K_D < 10^{-3}$ M (Mayer and Meyer, 1999). For a more detailed discussion regarding experimental implementation of these experiments, the reader is referred to some recent reviews (Lepre et al., 2004; Peng, 2001).

2.4 WaterLOGSY

Another experimental method for transfer of ^{1}H magnetization from the receptor to the bound ligand is the waterLOGSY (*Water-L*igand *O*bserved via *G*radient *S*pectroscop*Y*) (Dalvit et al., 2001; Dalvit et al., 2000) experiment. In waterLOGSY, selective excitation of the receptor/ligand complex is carried out as in the STD experiment. However, instead of direct excitation of protein resonances followed by spin diffusion, in the waterLOGSY experiment target resonances are perturbed indirectly by excitation or inversion of bulk water magnetization. Bulk water, once excited, can then transfer magnetization to the receptor via several pathways, which are illustrated in Figure 4-6. One pathway involves direct ^{1}H-^{1}H cross-relaxation between ligands and tightly bound water molecules at the binding site. Another pathway occurs when inverted bulk water undergoes chemical exchange with OH and NH groups at the receptor binding site. The inverted NH and OH groups then directly transfer inversion to bound ligand protons. Finally, a third mechanism involves inversion of NH and OH groups at distal

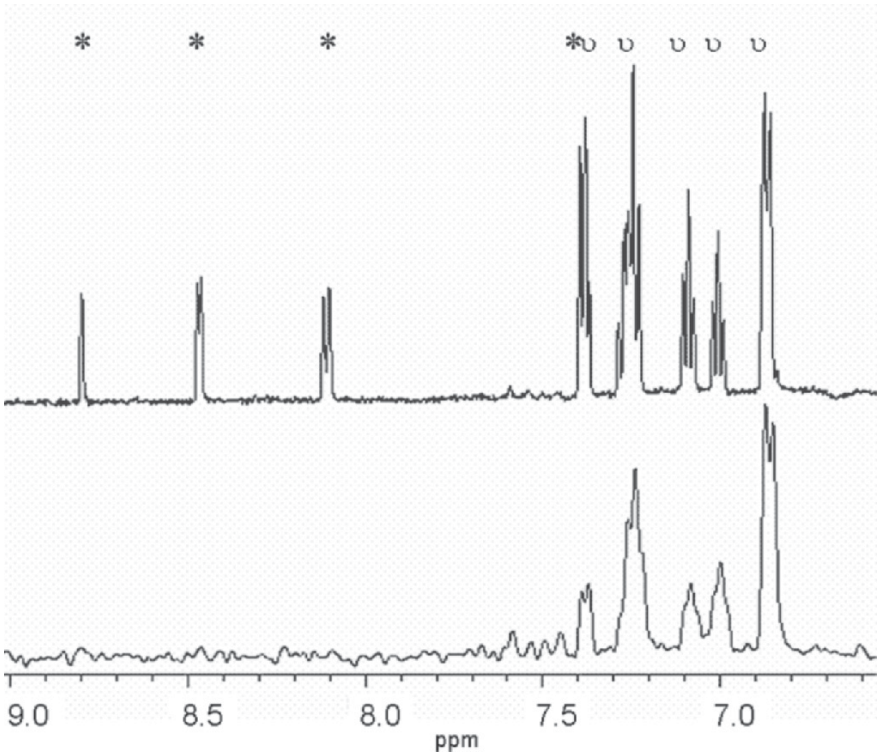


Figure 4-5. Example of the STD experiment with p38 kinase domain (42 kDa) Top panel A: ^{1}H NMR spectra of two compounds in the absence of receptor using 1D version of a standard NOESY. Resonances from nicotinic acid and 2-phenoxy benzoic acid are marked with asterisks and diamonds, respectively. Bottom panel B: result of the STD experiment in the presence of receptor; the resonances of the binding compound (2-phenoxybenzoic acid) are present. Receptor protons are invisible due to relaxation filtering. Sample conditions were as follows: 1mM compounds, 35 mM receptor dissolved in D^2O buffer (25 mM d-Tris, 10% d-glycerol, 20 mM d-DTT, pD = 8.4). Both spectra were acquired at 11.74 Tesla. The 1D spectrum (top) was collected using a standard NOESY pulse sequence with 16K data points, 128 transients and a relaxation delay of 3 s. The STD spectrum (bottom) was recorded at 278K, 2K data points with 256 transients acquired for both on- and off-resonance spectra. A 3s train of 50ms Gaussian pulses separated by 1ms was used for selective receptor saturation. The proton carrier was placed at 0.74 and –20 ppm for on- and off-resonance saturation, respectively. Excitation-sculpting was used to eliminate residual H2O and provide the aforementioned relaxation filtering. Reproduced with permission from Peng et al. (2001) (Peng et al., 2001). Copyright 2001 Elsevier.

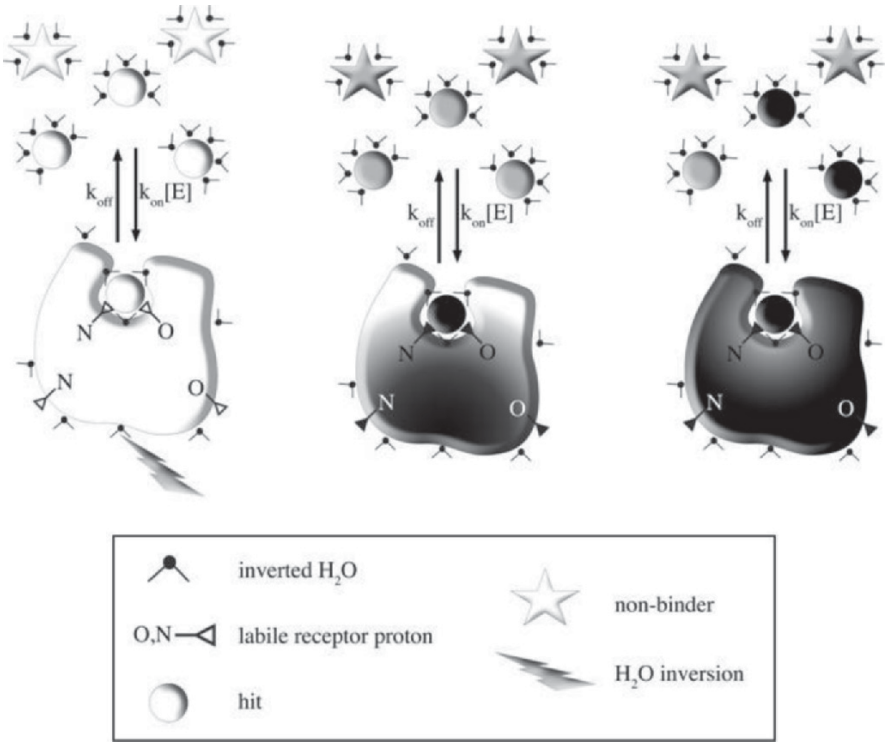


Figure 4-6. Magnetization transfer mechanisms underlying waterLOGSY (Dalvit et al., 2001; Dalvit et al., 2000). Magnetization transfer from bulk water to ligand occurs via labile receptor protons within and remote from the ligand-binding site as well as from long-lived water molecules within the binding pocket. Dark gray and light gray shading indicate magnetization transfer from inverted water to ligand protons in the slow tumbling (i.e. receptor-ligand complex) and fast tumbling (i.e. free ligand) limits, respectively. Only the hits experience both types of magnetization transfer. The pool of free ligands having experienced inversion-transfer from bulk water builds up as ligand continues to exchange on and off the receptor. Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

sites by chemical exchange with bulk water, followed by spin diffusion through the protein, analogous to the STD experiment.

As in the STD experiment, once a ligand has "picked up" magnetization from the receptor, it then dissociates where it is easily observed. Similarly, using an excess of ligand can amplify the effect significantly making the waterLOGSY a highly sensitive experiment. However, distinguishing compounds that bind from ones that do not is slightly different in waterLOGSY versus STD. Because of the differential mechanisms of cross-relaxation between water spins mediated through the receptor versus free in solution, ligands that bind will exhibit peaks of opposite sign in the NMR

spectrum from ligands that do not bind the receptor. This makes identification of binding compounds in a mixture fairly straightforward, as shown in Figure 4-7. However, in the waterLOGSY experiment, unlike the STD experiment, magnetization can be transferred directly to the free ligands as well if they contain exchangeable protons. This complicates the interpretation of spectral data, but can be addressed by collecting appropriate reference data for the free ligands.

An attractive feature of the waterLOGSY approach occurs for targets that exhibit decreased or poor spatial proton density, such as nucleic acid targets. Because water can interact directly with groups at the binding site, and is not

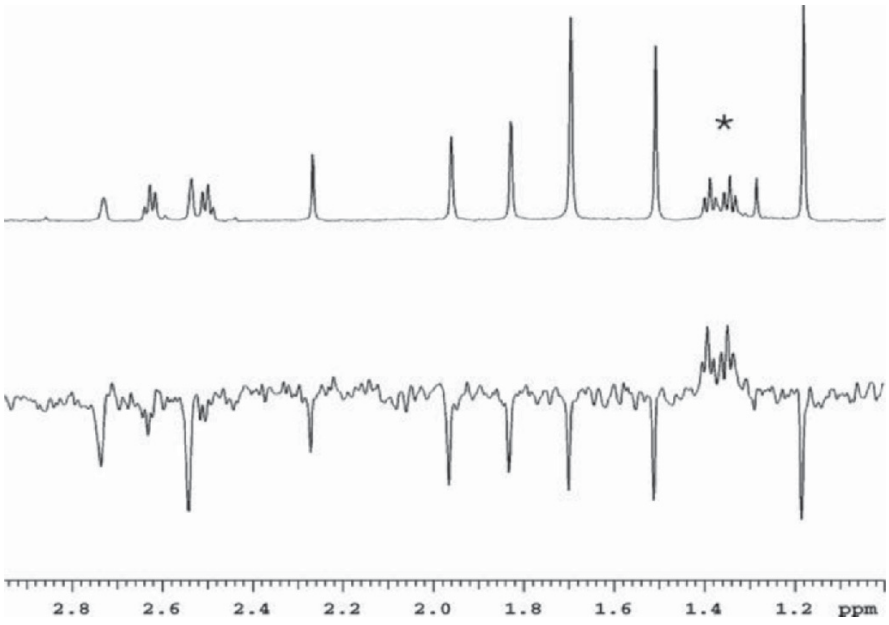


Figure 4-7. Example of the WaterLOGSY experiment. One-dimensional reference (upper) and WaterLOGSY (lower) spectra recorded for a 10-compound mixture in the presence of 10 μM cdk2. The WaterLOGSY and the reference spectra were recorded at 300^{o} K with 256 and 128 scans, respectively. The H_2O solvent suppression in both experiments was achieved with the H_2O excitation sculpting sequence (Hwang and Shaka, 1995). The WaterLOGSY was recorded with a 38 ms long 180^{o} H_2O selective Gaussian pulse. The relaxation and mixing times were 2.6 s and 2.0 s respectively. Positive and negative signals in the lower spectrum identify ckd2 binding and non-binding compounds, respectively. The asterisk indicates the methyl group resonances of the cdk2 ligand ethyl alpha-(ethoxycarbonyl)-3-indoleacrylate. Figure and legend are reprinted with permission from J. Biomol. NMR 18: 65-68 (2000). The figure was kindly provided by Dr. Claudio Dalvit.

dependent on transfer of magnetization from remote target sites to the binding site, the waterLOGSY experiment will provide significantly improved sensitivity over the STD experiment for this class of receptors (Johnson et al., 2003; Lepre et al., 2002).

3. APPLICATIONS OF NMR FRAGMENT-BASED SCREENING

By design, the primary hits generated by fragment-based screening are typically simple, low molecular weight compounds, and they consequently tend to have low binding affinities. In order to convert primary screening results into viable leads for a medicinal chemistry effort, hits must be validated (i.e. it must be confirmed that they bind in a manner that can be exploited to inhibit the target) and then developed into inhibitors that exhibit reasonable potency (typically low micromolar or better). Validation is usually accomplished by detecting activity in a biochemical assay and/or by obtaining information about the binding site (for example, from protein chemical shift perturbations, NMR competition experiments using known ligands, or observation of the bound ligand in an X-ray or NMR structure). The development of validated hits into leads can be accomplished using several strategies, previously described (Lepre, 2002; Lepre et al., 2002) as (1) combining, (2) elaborating, or (3) varying the molecular fragments. At one time, the various methods for detecting ligand binding were each mainly associated with a particular strategy (e.g. SAR by NMR (Shuker et al., 1996) was associated with fragment combination, and SHAPES (Fejzo et al., 1999) with fragment elaboration), but these distinctions have blurred as the field has evolved. We therefore find that classifying applications by the hit development strategy is more generally useful than classifying by detection method. This classification, however, is also imperfect since more than one strategy is often applied in a given application. This reflects the flexibility and ease with which these methods can be tailored to address the problem at hand. The following sections present examples of applications of each of the three strategies.

3.1 Applications of a combination strategy

The combination strategy consists of combining molecular fragments that bind to adjacent sites on the target. Combining multiple, weakly binding fragments into a more complex molecule can theoretically increase potency by orders of magnitude, since the binding energies are expected to be approximately additive (Shuker et al., 1996). In order to design

a molecule that combines the fragments, information about the relative orientations and distances between the bound fragments is very valuable, and is often obtained from NOE contacts (Fejzo et al., 1999; Li et al., 1999; Lugovskoy et al., 2002; Pellecchia et al., 2002a), chemical shift perturbation mapping experiments (Klaus and Senn, 2003; McCoy and Wyss, 2000; McCoy and Wyss, 2002; Medek et al., 2000; Wyss et al., 2002), or crystallography (Boehm et al., 2000; Carr and Jhoti, 2002; Fejzo, 2002; Hajduk et al., 2000b; Lepre et al., 2002; Lesuisse et al., 2002; Liepinsh and Otting, 1997; Nienaber et al., 2000). Structural information can also be inferred when fragments bind at overlapping sub-sites, so that the relative position of their functional groups is apparent from the molecular topology, a procedure called "fragment fusion" (Fejzo et al., 1999; Fejzo, 2002; Lepre et al., 2002).

The first reported applications of fragment linking employed the SAR by NMR method, using HSQC-based methods to detect binding and either X-ray crystallography or NOE-derived NMR structures to determine the bound fragment orientations. The early SAR by NMR examples were mostly proof-of-concept cases, and they have been extensively reviewed elsewhere (Lepre et al., 2004). Recent examples have been more pharmaceutically relevant, such as the design of Bcl-2 inhibitors (Oltersdorf et al., 2005).

A particularly extensive and detailed example of fragment linking appears in a series of papers (Liu et al., 2003a; Liu et al., 2003b; Liu et al., 2003c; Szczepankiewicz et al., 2003) describing the iterative construction of protein tyrosine phosphatase 1B (PTP1B) inhibitors. The goal was to find fragments with low charge that bind to the phosphotyrosine site while conferring selectivity over other phosphatases, particularly the close homolog T-cell PTPase (TCPTP). A diaryloxamic acid fragment was found to bind at the phosphotyrosine catalytic site with 93 μM affinity, and testing of analogs (using the variation strategy, *vide infra*) identified an active 1-aminonapththalene derivative (Figure 4-8). Crystallography revealed that the naphthyl bound in the pTyr pocket and was capable of being linked to a fragment bound in a second, less conserved pTyr binding site. Synthesis of a linked compound improved potency to 1 μM and increased pan-phosphatase selectivity. Improved ligands for the second pTyr site were found in a second NMR screen (Figure 4-9), leading to linked salicylate analogs with mid-nanomolar potencies and weak selectivity for PTP1B vs. TCPTP, but poor cell permeability. A third screen (Figure 4-10) was carried out using only monocarboxylic or non-carboxylate containing fragments in an attempt to improve cell permeability, leading to salicylate-linked isoxazole carboxylates with low micromolar potencies, moderate cell potency, and 30-fold or better selectivity for PTP1B vs. TCPTP and other phosphatases.

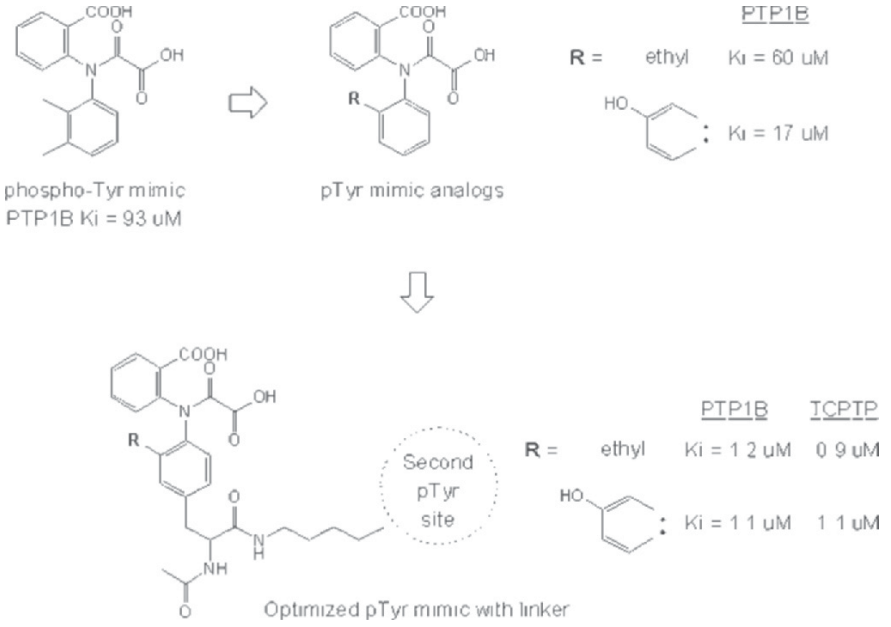


Figure 4-8. The phospho-tyrosine mimic identified in the primary NMR screen (upper left) was optimized by testing various analogs (upper right). A polymethylene linker was then attached (bottom) to access the second phospho-tyrosine binding site (Liu et al., 2003a). Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

Although high-resolution X-ray or NMR structures are the preferred means to obtain structural information about bound fragments, they are often difficult to obtain. Several alternative, lower-resolution methods have been developed to generate information to guide fragment linking.

Intermolecular transferred NOEs have long been used to identify close contacts between bound ligands and determine relative orientation (Fejzo et al., 1999); a recent reprise of this method is a variant of SAR by NMR called SAR by ILOEs (Becattini et al., 2004).

Amide chemical shift perturbations have been used quite elegantly to model the binding sites of fragments (Wyss et al., 2004) for systems in which the chemical shift assignments and a three dimensional structure of the apo protein are available. The ligand-induced amide chemical shift changes of ^{15}N-labeled NS3-4A protease (from hepatitis C virus) were used first to identify the binding sub-sites in the substrate groove of NMR hits, then to model the bound orientations of the ligands by fitting them into

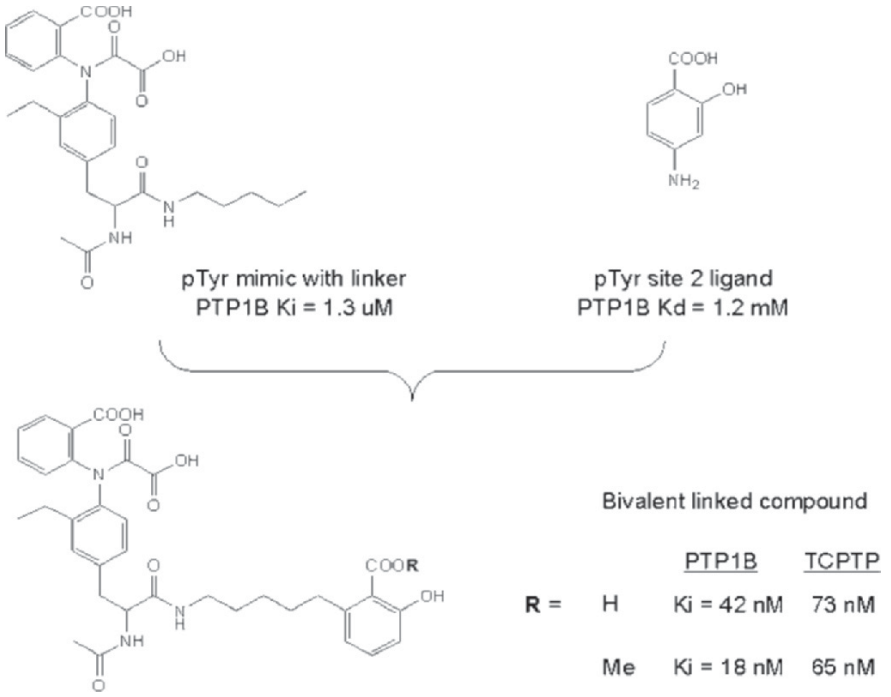


Figure 4-9. A second NMR screen identified fragments binding to the second pTyr site. The best of these fragments was attached to the pTyr mimic-linker compound from the first step (top row), yielding a bivalent linked compound (bottom) (Liu et al., 2003b). Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

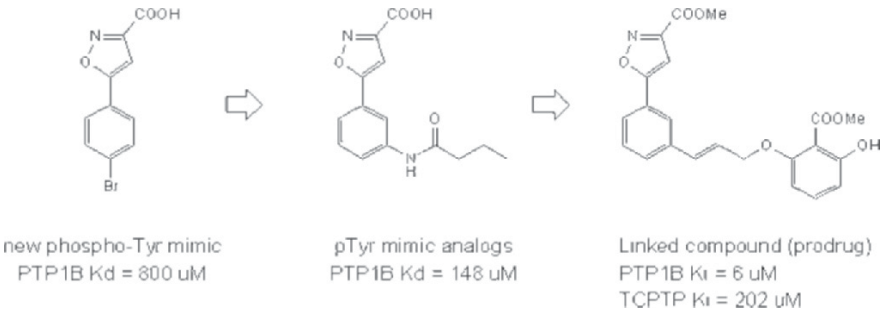


Figure 4-10. A third NMR screen identified a monoanionic pTyr mimic (left) that was optimized by testing analogs (center), and then connected to the previously identified salicylate fragment in the second pTyr site to make a bivalent linked compound (right) (Liu et al., 2003c). Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

j-surfaces derived from the perturbations (McCoy and Wyss, 2002), and these models were used to design a linked molecule with 100-fold greater potency.

Proximity information can also be derived from differential relaxation rate enhancements for fragments bound near spin labels or paramagnetic centers. While the SLAPSTIC method (Jahnke et al., 2001) uses spin-labelled protein to determine distances to bound ligands, labelled ligands are used more frequently than labelled proteins. Second-site screening using a spin-labelled first ligand has been applied to find inhibitors of tubulin (Jahnke et al., 2003) and Bcl-xL (Jahnke et al., 2003; Jahnke et al., 2000). More recently, paramagnetic Mn-ATP (McCoy et al., 2005) and TEMPO-adenine (Jahnke et al., 2005)] have been used to screen for allosteric inhibitors of kinases and estimate proximity to the ATP site using differential relaxation effects. A possible complicating factor with these labelling-based approaches is the unexpected binding of fragments to multiple sites on the protein, which will confound attempts to derive accurate distances.

The "NMR-DOC" (NMR Docking Of Compounds) method (Pellecchia et al., 2002a) uses amino-acid type specific isotopic labels to identify specific residues in an active site. Fragment binding to that site is detected using selective saturation transfer via the labelled residues. The conformations of bound ligands are then modeled using the crystal structure of the target and NOEs from the assigned residues to the ligands. For enzymes with proximal cofactor and substrate binding sites, the "NMR-SOLVE" (Structure Oriented Library Valency Engineering) method addresses the fragment linking problem by screening libraries of cofactors already attached to fragments that target the nearby substrate site.

3.2 Applications of an elaboration strategy

In the elaboration strategy, relatively simple primary hits are elaborated by adding chemical functionality, producing more complex molecules. The more elaborate molecules can make additional ligand:protein interactions, leading to higher potency. Structural information about the bound ligands is not necessary, but can be useful for deciding the type and location of functionality to be added.

One of the first reported methods using fragment elaboration was the SHAPES strategy, a flexible approach that follows ligand-detected NMR screening of drug-like molecules with successive rounds of activity assays and structure-based optimization (Fejzo et al., 1999; Fejzo, 2002; Lepre et al., 2002). A classical example of the elaboration strategy is the SHAPES screen of fatty acid binding protein (FABP-4) (Lepre et al., 2004; Lepre et al.,

2002). Based on the crystal structures of two primary screening hits with low micromolar potencies, a follow-up library of 134 elaborated analogs was screened calorimetrically, which in turn produced nine leads with low micromolar to nanomolar affinities. The crystal structures of five more bound ligands were subsequently solved, mapping out the essential ligand:protein interactions and defining the binding pharmacophore. A similar approach to finding FABP-4 inhibitors, also using an elaboration strategy guided by structural information, has been reported (van Dongen et al., 2002a). Other examples of fragment elaboration include the design of urokinase inhibitors (Hajduk et al., 2000b; Huth and Sun, 2002) and the discovery of new zinc-binding motifs for metalloprotease inhibitors (Klaus and Senn, 2003).

Virtual screening has proven to be a useful tool for selecting both primary screening compounds and elaborated analogs of hits. Even a relatively simple docking model can be quite effective for removing compounds that are sterically unsuited to bind a given target, improving the efficiency of the subsequent NMR screen. An example of virtual screening combined with an elaboration strategy is the discovery of inhibitors for the human hydroxysteroid dehydrogenase 3α-HSD (Jahnke et al., 2002), in which NMR screening found a small, weakly binding "reporter ligand" that was then iteratively elaborated through two rounds of compounds selected by virtual screening, producing leads with sub-micromolar potencies (Rudisser and Jahnke, 2002).

3.3 Applications of a variation strategy

In the variation strategy, selected portions of a primary screening hit are systematically modified. This can be accomplished in some cases by purchasing analogs (as in the discovery of isoxazole-based Jnk3 inhibitors starting from an imidazole hit (Fejzo, 2002; Lepre et al., 2002)), but more commonly requires synthesis of second generation compounds. These compounds might not be significantly more complex than the primary hits, but are intended to make more optimal interactions with the target. Structural information is not required but can be very useful for optimizing ligand:protein contacts.

Combinatorial chemistry has been effectively used to synthesize variants of NMR screening hits. The primary fragments should contain either a molecular core with multiple points of attachment for substituents (so they can be readily varied), or two cores connected by a combi-chem accessible linker, to permit easy replacement of the cores (such as the SHAPES Linking Library (Lepre, 2001[1]). An early example of the former is the development of ErmAM inhibitors using a triazine core (Hajduk et al., 1999a). The bare

triazine core (K_D = 1 mM) was found in the primary NMR screen and subsequently functionalized at two positions to yield a lead with 75 μM potency. Combinatorial exploration of 643 disubstitution variants produced several low micromolar inhibitors of ErmAM and ErmC.

Because of their modular topology, peptide-based inhibitors are particularly amenable to combinatorial synthesis. A fragment optimization approach can he used, in which one portion of the molecule is systematically replaced while the rest remains constant. An example of this is the optimization of X-linked IAP (XIAP) BIR3 domain inhibitors based on the SMAC tetrapeptide inhibitor AVPI, starting from the dipeptide fragment AV-NH_2 (K_D = 812 μM) (Park et al., 2005). Parallel synthesis was used to make analogs in which the Val residue was replaced with 5-membered heterocycles (such as thiazoles and imidazoles) capable of replicating the peptide:BIR3 hydrogen bonding interactions, and these analogs were then assayed for binding by NMR (Figure 4-11). Further optimization by adding substituents to occupy the Pro and Ile binding sites led to a non-peptidic inhibitor with submicromolar affinity.

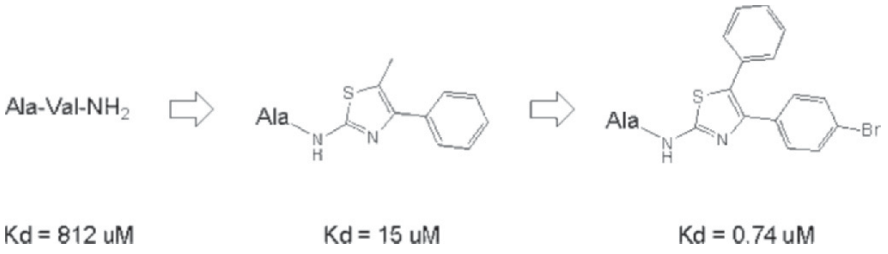


Figure 4-11. Design of non-peptidic XIAP inhibitors Starting from the peptide fragment Ala-Val-NH2 (left), heterocyclic replacements for the Val amide were tested, producing a thiazole core (center) that was subsequently optimized by adding a phenyl to fill the Pro-binding pocket and Br-phenyl to occupy the Ile-binding groove (right).

Another strategy for optimizing portions of a lead molecule is to perform a screen to find fragments that bind to the same protein sub-site as an undesired moiety. This is done by screening in the presence of saturating amounts of a molecule identical to the lead but lacking the portion to be optimized, leaving the sub-site open to bind new fragments. New fragments binding to that site are then synthetically linked to the original molecular core. An attractive feature of this approach is the ability to rapidly test potential reagents, allowing chemistry to focus upon only those fragments that bind to the site of interest.

An early demonstration of this concept was the optimization of adenosine kinase (AK) inhibitors (Hajduk et al., 2000c). A ^{15}N HSQC-based screen

with a truncated lead was carried out to replace an offending bromophenyl moiety with a functional group that would improve aqueous solubility and pharmacokinetic properties (Huth and Sun, 2002). Highly soluble but weak (K_D = 3 mM) indole and 2-phenylimidazole fragments were validated as bromophenyl pocket binders using chemical shift perturbation and NMR competition experiments. Linking of these fragments to the original scaffold yielded compounds with nanomolar *in vitro* potency.

Similarly, as shown in Figure 4-12, compounds blocking the association of LFA-1 (leukocyte function-associated antigen-1) with ICAM-1 (intracellular adhesion molecule-1) were optimized to replace the hydrophobic isopropylphenyl ring of a lead that was highly potent (IC_{50} = 44 nM) but had

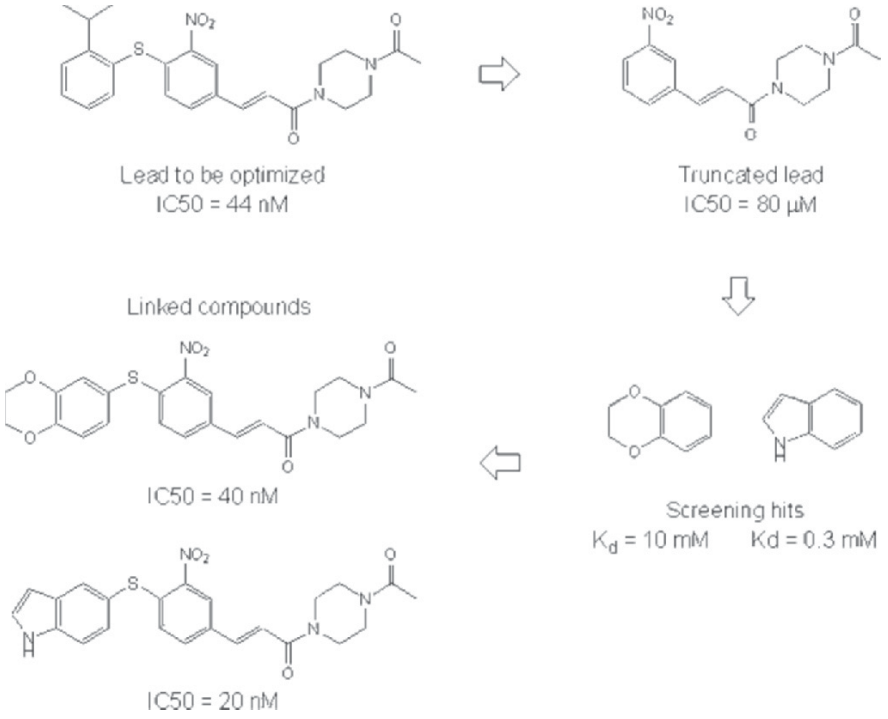


Figure 4-12. Fragment optimization of LFA-1/ICAM-1 inhibitors. The hydrophobic isopropylphenyl group was removed from the poorly soluble starting lead molecule (upper left) to make a truncated scaffold (upper right). Screening in the presence of the truncated scaffold identified two highly soluble candidates to replace the isopropylphenyl (lower right), and linked compounds with improved PK properties were synthesized (lower left). Reprinted with permission from Chemical Reviews (2004), 104 (8) 3641-3675. © 2004 American Chemical Society.

poor solubility and bioavailability (Huth and Sun, 2002). With the central hydrophobic site of the LFA-1 I domain allosteric site blocked by a truncated scaffold, a set of small, polar fragments was screened, producing a number of hits (e.g. indole, benzodioxane) with millimolar affinities. NOESY results were used to determine the orientations of the bound ligands and design linked compounds that were equipotent with the parent compound but had increased bioavailability.

4. CONCLUSIONS

In this review, we have made an effort to provide the most contemporary overview possible of NMR screening techniques and their application in pharmaceutical research. Since the inception of NMR screening as a distinct discipline only 9 years ago, it is difficult to describe any single approach as dated, however, some experimental techniques and follow-up strategies have been more widely adopted than others, and we have made an effort to provide a more rigorous description of these in both the experimental and applications section of this review. In addition to physical and technical descriptions of how these experiments may be carried out, we have provided, through recent examples, a rich context of applications in which these experiments have been applied, illustrating the potential for success when integrated into a strategy incorporating other highly enabling technologies such as virtual screening, enzymology and X-ray crystallography. As the field has evolved, and more examples have been presented of how these methods have been used in practice, it is clear that each target and drug design problem is unique. It is hoped that the descriptions provided and references to the literature therein will allow investigators to implement these methods in their own laboratories, as well as develop drug design strategies uniquely suited to their targets and goals.

REFERENCES

Becattini B., Sareth S., Zhai D., Crowell K. J., Leone M., Reed J. C. and Pellecchia M. 2004. Targeting apoptosis via chemical design: inhibition of bid-induced cell death by small organic molecules. Chem. Biol. **11**, 1107–1117.

Boehm H. J., Boehringer M., Bur D., Gmuender H., Huber W., Klaus W., Kostrewa D., Kuehne H., Luebbers T., Meunier-Keller N. et al. 2000. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. J. Med. Chem. **43**, 2664–2674.

Carr R. and Jhoti H. 2002. Structure-based screening of low-affinity compounds. Drug Discovery Today **7**, 522–527.

Claasen B., Axmann M., Meinecke R. and Meyer B. 2005. Direct observation of ligand binding to membrane proteins in living cells by a saturation transfer double difference (STDD) NMR spectroscopy method shows a significantly higher affinity of integrin alpha(IIb)beta3 in native platelets than in liposomes. J. Am. Chem. Soc. **127**, 916–919.

Dalvit C., Fogliatto G., Stewart A., Veronesi M. and Stockman B. 2001. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. J. Biomol. NMR **21**, 349–359.

Dalvit C., Pevarello P., Tato M., Veronesi M., Vulpetti A. and Sundstrom M. 2000. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. J. Biomol. NMR **18**, 65–68.

Fejzo J., Lepre C. A., Peng J. W., Bemis G. W., Ajay, Murcko M. A. and Moore J. M. 1999. The SHAPES strategy: An NMR-based approach for lead generation in drug discovery. Chem. Biol. **6**, 755–769.

Fejzo J., Lepre C. and Xie X. 2002. Applications of NMR screening in drug discovery. Curr. Top. Med. Chem. **2**, 1349–1364.

Hajduk P. J., Augeri D. J., Mack J., Mendoza R., Yang J., Betz S. F. and Fesik S. W. 2000a. NMR-based screening of proteins containing 13C-labeled methyl groups. J. Am. Chem. Soc. **122**, 7898–7904.

Hajduk P. J., Boyd S., Nettesheim D., Nienaber V., Severin J., Smith R., Davidson D., Rockway T. and Fesik S. W. 2000b. Identification of novel inhibitors of urokinase via NMR-based screening. J. Med. Chem. **43**, 3862–3866.

Hajduk P. J., Dinges J., Schkeryantz J. M., Janowick D., Kaminski M., Tufano M., Augeri D. J., Petros A., Nienaber V., Zhong P. et al. 1999a. Novel inhibitors of erm methyltransferases from NMR and parallel synthesis. J. Med. Chem. **42**, 3852–3859.

Hajduk P. J., Gerfin T., Boehlen J.-M., Haberli M., Marek D. and Fesik S. W. 1999b. High-throughput nuclear magnetic resonance-based screening. J. Med. Chem. **42**, 2525–2517.

Hajduk P. J., Gomtsyan A., Didomenico S., Cowart M., Bayburt E. K., Solomon L., Severin J., Smith R., Walter K., Holzman T. F. et al. 2000c. Design of adenosine kinase inhibitors from the NMR-based screening of fragments. J. Med. Chem. **43**, 4781–4786.

Hajduk P. J., Meadows R. P. and Fesik S. J. 1999c. NMR-based screening in drug discovery. Q. Rev. Biophys. **32**, 211–240.

Huth J. R. and Sun C. 2002. Utility of NMR in lead optimization: fragment-based approaches. Comb. Chem. High Throughput Screen. **5**, 631–643.

Hwang T.-L. and Shaka A. J. 1995. Water suppression that works. excitation sculpting using arbitrary waveforms and pulsed field gradients. J. Magn. Reson. Ser. A **112**, 275–279.

Jahnke W., Blommers M. J., Fernandez C., Zwingelstein C. and Amstutz R. 2005. Strategies for the NMR-Based identification and optimization of allosteric protein kinase inhibitors. Chembiochem. **6**, 1607–1610.

Jahnke W., Floersheim P., Ostermeier C., Zhang X., Hemmig R., Hurth K. and Uzunov D. P. 2002. NMR reporter screening for the detection of high-affinity ligands. Angew. Chem. Int. Ed. Engl. **41**, 3420–3423.

Jahnke W., Florsheimer A., Blommers M. J., Paris C. G., Heim J., Nalin C. M. and Perez L. B. 2003. Second-site NMR screening and linker design. Curr. Top. Med. Chem. **3**, 69–80.

Jahnke W., Perez L. B., Paris G., Strauss A., Fendrich G. and Nalin C. M. 2000. Second-Site NMR Screening with a Spin-Labeled First Ligand. J. Am. Chem. Soc. **122**, 7394–7395.

Jahnke W., Rudisser S. and Zurini M. 2001. Spin label enhanced NMR screening. J. Am. Chem. Soc. **123**, 3149–3150.

Johnson E. C., Feher V. A., Peng J. W., Moore J. M. and Williamson J. R. 2003. Application of NMR SHAPES screening to an RNA target. J. Am. Chem. Soc. **125**, 15724–15725.

Klaus W. and Senn H. 2003. Strategies for hit finding using NMR. In BioNMR in Drug Research, O. Zerbe, ed. (Weinheim, Wiley-VCH) pp. 417–437.

Klein J., Meinecke R., Mayer M. and Meyer B. 1999. Detecting binding affinity to immobilized receptor proteins in compound libraries by HR-MAS STD NMR. J. Am. Chem. Soc. **121**, 5336–5337.

Lepre C. A. 2001. Library design for NMR-based screening. Drug Discovery Today **6**, 133–140.

Lepre C. A. 2002. Strategies for NMR screening and library design. In BioNMR Techniques in Drug Research, O. Zerbe, ed. (Weinheim, Wiley-VCH) pp. 1349–1364.

Lepre C. A., Moore J. M. and Peng J. W. 2004. Theory and applications of NMR-based screening in pharmaceutical research. Chem. Rev. **104**, 3641–3676.

Lepre C. A., Peng J., Fejzo J., Abdul-Manan N., Pocas J., Jacobs M., Xie X. and Moore J. M. 2002. Applications of SHAPES screening in drug discovery. Comb. Chem. High Throughput Screen. **5,** 583–590.

Lesuisse D., Lange G., Deprez P., Benard D., Schoot B., Delettre G., Marquette J. P., Broto P., Jean-Baptiste V., Bichet P. et al. 2002. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: Application to the discovery of nanomolar inhibitors of Src SH2. J. Med. Chem. **45**, 2379–2387.

Li D., DeRose E. F. and London R. E. 1999. The inter-ligand overhauser effect: A powerful new NMR approach for mapping structural relationships of macromolecular ligands. J. Biomol. NMR **15**, 71–76.

Liepinsh E. and Otting G. 1997. Organic solvents identify specific ligand binding sites on protein surfaces. Nat. Biotechnol. **15**, 264–268.

Lipinski C. A., Lombardo F., Dominy B. W. and Feeny P. J. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **23**, 3–25.

Liu G., Szczepankiewicz B. G., Pei Z., Janowick D. A., Xin Z., Hajduk P. J., Abad-Zapatero C., Liang H., Hutchins C. W., Fesik S. W. et al. 2003a. Discovery and structure-activity relationship of oxalylarylaminobenzoic acids as inhibitors of protein tyrosine phosphatase 1B. J. Med. Chem. **46**, 2093–2103.

Liu G., Xin Z., Liang H., Abad-Zapatero C., Hajduk P. J., Janowick D. A., Szczepankiewicz B. G., Pei Z., Hutchins C. W., Ballaron S. J. et al. 2003b. Selective protein tyrosine phosphatase 1B inhibitors: Targeting the second phosphotyrosine binding site with non-carboxylic acid-containing ligands. J. Med. Chem. **46**, 3437–3440.

Liu G., Xin Z., Pei Z., Hajduk P. J., Abad-Zapatero C., Hutchins C. W., Zhao H., Lubben T.H., Ballaron S. J., Haasch D. L. et al. 2003c. Fragment screening and assembly: A highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor. J. Med. Chem. **46**, 4232–4235.

Lugovskoy A. A., Degterev A. I., Fahmy A. F., Zhou P., Gross J. D., Yuan J. and Wagner G. 2002. A novel approach for characterizing protein ligand complexes: Molecular basis for specificity of small-molecule Bcl-2 inhibitors. J. Am. Chem. Soc. **124**, 1234–1240.

Mayer M. and Meyer B. 1999. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew. Chem. Int. Ed. **38**, 1784–1788.

McCoy M. A., Senior M. M. and Wyss D. F. 2005. Screening of protein kinases by ATP-STD NMR spectroscopy. J. Am. Chem. Soc. **127**, 7978–7979.

McCoy M. A. and Wyss D. F. 2000. Alignment of weakly interacting molecules to protein surfaces using simulations of chemical shift perturbations. J. Biomol. NMR **18**, 189–198.

McCoy M. A. and Wyss D. F. 2002. Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. J. Am. Chem. Soc. **124**, 11758–11763.

Medek A., Hajduk P. J., Mack J. and Fesik S. W. 2000. The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands. J. Am. Chem. Soc. **122**, 1241–1242.

Meinecke R. and Meyer B. 2001. Determination of the binding specificity of an integral membrne protein by saturation transfer difference NMR: RGD peptide ligands binding to integrin aIIbb3. J. Am. Chem. Soc. **44**, 3059–3065.

Moore J. M. 1999. NMR screening in drug discovery. Curr. Opin. Biotechnol. **10**, 54–58.

Nienaber V. L., Richardson P. L., Klighofer V., Bouska J. J., Giranda V. L. and Greer J. 2000. Discovering novel ligands for macromolecules using X-ray crystallographic screening. Nat. Biotechnol. **18**, 1105–1108.

Oltersdorf T., Elmore S. W., Shoemaker A. R., Armstrong R. C., Augeri D. J., Belli B. A., Bruncko M., Deckwerth T. L., Dinges J., Hajduk P. J. et al. 2005. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature **435**, 677–681.

Park C. M., Sun C., Olejniczak E. T., Wilson A. E., Meadows R. P., Betz S. F., Elmore S. W. and Fesik S. W. 2005. Non-peptidic small molecule inhibitors of XIAP. Bioorg. Med. Chem. Lett. **15**, 771–775.

Pellecchia M., Meininger D., Dong Q., Chang E., Jack R. and Sem D. S. 2002a. NMR-based structural characterization of large protein-ligand interactions. J. Biomol. NMR **22**, 165–173.

Pellecchia M., Sem D. S. and Wuthrich K. 2002b. NMR in drug discovery. Nature Rev. Drug Discov. **1**, 211–219.

Peng J. W., Moore J. and Abdul-Manan N. 2004. NMR Experiments for lead generation in drug discovery. Prog. Nucl. Magn. Reson. Spectrosc. **44**, 225–256.

Peng J. W., Lepre C. A., Fejzo J., Abdul-Manan N. and Moore J. M. 2001. Nuclear magnetic resonance-based approaches for lead generation in drug discovery. Methods Enzymol. **338**, 202–230.

Roberts G. C. 2000. Applications of NMR in drug discovery. Drug Discovery Today **5**, 230–240.

Ross A., Schlotterbeck G., Klaus W. and Senn H. 2000. Automation of NMR measurements and data evaluation for systematically screening interactions of small molecules with target proteins. J. Biomol. NMR **16**, 139–146.

Ross A. and Senn H. 2001. Automation of measurements and data evaluation in biomolecular NMR screening. Drug Discovery Today **6**, 583–593.

Rudisser S. and Jahnke W. 2002. NMR and in silico screening. Comb. Chem. High Throughput Screen. **5**, 591–603.

Shuker S. B., Hajduk P. J., Meadows R. P. and Fesik S. W. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. Science **274**, 1531–1534.

Stockman B. and Dalvit C. 2002. NMR screening techniques in drug discovery and drug design. Prog. Nucl. Magn. Reson. Spectrosc. **41**, 187–231.

Szczepankiewicz B. G., Liu G., Hajduk P. J., Abad-Zapatero C., Pei Z., Xin Z., Lubben T. H., Trevillyan J. M., Stashko M. A., Ballaron S. J. et al. 2003. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. J. Am. Chem. Soc. **125**, 4087–4096.

Teague S. J., Davis A. M., Leeson P. D. and Oprea T. 1999. The design of leadlike combinatorial libraries. Angew. Chem. Int. Ed. **38**, 3743–3747.

van Dongen M., Uppenberg J., Svensson B., Lundbäck T., Åkerud T., Wikström M. and Schultz J. 2002a. Structure-based screening as applied to human FABP4: A highly efficient alternative to HTS for hit generation. J. Am. Chem. Soc. **124**, 11874–11880.

van Dongen M., Weigelt J., Uppenberg J., Schultz J. and Wikstrom M. 2002b. Structure-based screening and design in drug discovery. Drug Discovery Today **7**, 471–478.

Wyss D. F., Arasappan A., Senior M. M., Wang Y.-S., Beyer B. M., Njoroge F. G. and McCoy M. A. 2004. Non-peptidic small molecule inhibitors of the single-chain hepatitic C virus NS3 protease NS4A cofactor complex discovered by structure-based NMR screening. J. Med. Chem. **47**, 2486–2498.

Wyss D. F., McCoy M. A. and Senior M. M. 2002. NMR-based approaches for lead discovery. Curr. Opin. Drug Discov. Devel. **5**, 630–647.

Chapter 5

FRAGMENT-BASED SCREENING BY X-RAY CRYSTALLOGRAPHY

Valerio Berdini, Marc O'Reilly, Miles S. Congreve and Ian J. Tickle
Astex Therapeutics Ltd., 436 Science Park, Milton Road, Cambridge CB4 0QA, UK.

1. INTRODUCTION

Historically, the process of obtaining crystals of a target protein suitable for X-ray crystallographic structure determination has been the major obstacle to crystallographic screening. The subsequent generation of liganded protein structures has also often been a relatively low-throughput process, and thus only able to impact on the drug optimisation process to a limited extent. The crystal structures of progressively greater numbers of therapeutically validated targets are, however, now entering the public domain as recent advances in methodologies and automation have significantly accelerated the rate at which it is possible to progress 'from gene to structure'. Thus X-ray crystallography has evolved into a technology platform that is able to contribute to all phases of the drug discovery process in a rapid and timely fashion. This includes the primary screening phase of a project, where high-throughput crystallographic screening can be used to generate 'hits' to feed into the lead optimisation process. The fact that lead molecules can be visualised bound to their target protein at an early phase of the project, means that a directed, structure-based, approach can be taken to lead optimisation. Compared with more traditional lead optimisation processes, a structure-based approach reduces the volume of chemical space that must be explored and also dramatically aids in the interpretation of structure-activity relationships (SARs) within a class of ligands.

A novel approach to primary screening that has evolved in recent years is the use of fragment-based libraries. Typically fragments are small (100-250Da)

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 99–127.

organic molecules that exhibit low binding affinities (~100μM to 10mM) against target proteins, and, as such, would not usually be identified by more conventional screening techniques such as *in vitro* bio-assays. Screening for fragment binding may however be performed using protein X-ray crystallography, or by NMR (see preceding chapter). X-ray crystallography is able to elucidate the exact binding mode of fragments and thus provide medicinal chemistry design strategies to evolve a fragment into a potent, drug-sized selective lead compound. Despite the low affinity of initial fragment hits, useful fragments typically exhibit a high ligand efficiency, that is, a high value for the average free energy of binding per heavy atom (*i.e*. excluding hydrogens) (Hopkins *et al*., 2004). It is critical that during the development into the lead compound, this high ligand efficiency is maintained.

Some of the first experiments in which X-ray crystallography was used as a screening tool were reported by Verlinde *et al.* (1997) who exposed crystals of trypanosomal triosephosphate isomerase to cocktails (mixtures) of compounds in the search for inhibitors. More recently Nienaber *et al.* (2000) at Abbott Laboratories have described their *CrystaLEAD*™ process for X-ray-based screening of shape diverse fragment sets. The work of Blundell *et al.* (2002) with trypsin describes how a 353-fragment screen was performed and subsequently deconvoluted. The potential for identifying novel ligands which can then be evolved into more complex inhibitors using a fragment approach is exemplified by Lesuisse *et al.* (2002) in their studies with the Src SH2 domain, and by Hartshorn *et al.* (2005) and Gill *et al.* (2005) in their studies with p38α MAP kinase and other targets.

Scientists at Astex have integrated virtual screening, NMR and high-throughput X-ray crystallography into a fragment-based discovery platform – *Pyramid*™. Central to *Pyramid*™ is *AutoSolve*®, a web-based software application linked to an in-house developed LIMS system and *Oracle*® database that performs automated processing and analysis of all protein-ligand X-ray data.

Other companies such as Structural GenomiX (SGX) have also developed platforms for lead compound identification using high-throughput protein structure determination: SGX's *FAST*™ (Fragments of Active Structures Technology) enables the rapid identification of novel, potent, and selective small-molecule inhibitors of drug targets. Card *et al.* (2005) at Plexxikon have described their process *Scaffold-Based Drug Discovery*™ for the design of molecular scaffolds and ligands, based on X-ray analysis of co-crystals of protein and ligand: this is described fully in the following chapter.

As with many screening techniques, X-ray crystallographic fragment screening is most efficiently carried out in a cocktail format. A complete library, comprised of many hundreds of fragments, is partitioned into a set of

multi-fragment cocktails, each containing an equal number of components (typically 4 - 10). Individual crystals of the target protein are then soaked in these cocktails and X-ray diffraction data collected from the soaked crystals. The resulting electron density maps reveal whether one, or more, of the fragments have bound to sites on the protein. In the event of multiple fragments competing for the same site it may be necessary to deconvolute the cocktail by partitioning it into smaller cocktails and/or individual fragment soaks. Identification of the exact binding mode of a ligand can also be made easier by employing ligand libraries in which the ligands are decorated with electron-dense substituent groups, such as halogens (exemplified by SGX's *FAST*™).

Although fragment-based screening typically seeks to identify ligands bound to a predefined site on a protein (Figure 5-1*a*) it is often possible to identify alternative ligand binding sites and/or observe simultaneous fragment binding at spatially distinct sites on the protein (Figure 5-1*b*). Fragments that are observed to bind in relatively close proximity to one another provide data on the ways in which different functionalities can be

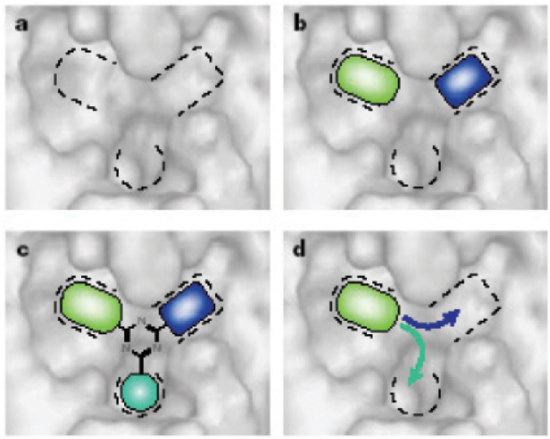


Figure 5-1a. A region of interest is defined on the protein surface; *b*: two fragments bound to distinct sites (pockets) within the region of interest; *c*: synthesis of a ligand that combines the fragments identified in *b*, but also accesses a previously unexplored pocket on the protein surface; *d*: a fragment (template molecule) bound to a single pocket on the protein suggests potential growth vectors.

combined (linked) within a single, ideally higher potency, ligand (Figure 5-1*c*). Analysis of the binding of individual fragments can help to define sensible growth vectors off a template molecule (Figure 5-1*d*). Such

analyses can be of great benefit if fragment binding induces conformational changes within a protein, revealing previously inaccessible, or structurally modified, binding pockets.

As alluded to previously, a judicious choice of fragments within a screening set can dramatically influence the performance (hit rate) of a screening set and also significantly ease the process of 'hit' identification and/or deconvolution.

2. FRAGMENT LIBRARIES

Of critical importance to the success or failure of a biophysical screen of fragments is the composition of the fragment libraries themselves. Two complementary approaches to library assembly can be envisaged. The first seeks to take advantage of the fact that fragments are small and can be chosen to probe the large number of potential interactions with the protein. Based on functional groups and scaffolds well represented in bioactive compounds, it should therefore be possible to represent 'drug-fragment space' with a relatively small number of compounds. The second approach is to assemble libraries targeted towards particular proteins or protein classes to maximise the chance of success for each individual target family. One can apply knowledge of known ligands and their key interactions with proteins to select fragments for acquisition or synthesis. Additionally, virtual screening can be employed to acquire compounds that dock well into the active site of the target or targets.

In general, fragments are selected to have molecular weights of 100-250Da and to be relatively simple with few functional groups, making them chemically suitable for rapid synthetic optimisation. Compounds smaller than 100Da might also be detectable in a crystallographic experiment, though with reduced confidence since these will generally be very weakly bound (Ki >> 1mM). Compounds larger than 250Da are less suitable, since the increased size and complexity is likely to reduce the likelihood of binding (Hann *et al*., 2001), and a larger compound that efficiently fits into an active site would be expected to have activity easily detectable in a biological assay, making a more standard high-throughput screening approach more suitable.

2.1 Physico-chemical properties of library members

It is important to consider calculated physicochemical properties when selecting fragments. There is a growing body of literature investigating the properties of small molecules that are required to make good lead compounds (Hann et al., 2001; Oprea, 2001). The now familiar Lipinski's 'Rule of 5'

(Lipinski et al., 2001) provides a useful framework for developing orally bio-available drug candidates. These rules have been further developed by others such as Veber et al. (2002) who showed that the number of rotatable bonds (NROT) is also an important parameter, with a maximum of 7 seeming optimal for oral bio-availability. Furthermore, there is literature indicating polar surface area (PSA) as another key property (Clark and Picket, 2000) - passively absorbed molecules with PSA > 110-140 Å are reported as likely to have low oral bio-availabilities. More recently, the term 'lead-like' was introduced for molecules identified from HTS campaigns that are suitable for further optimisation and that have properties somewhat 'scaled down' from Lipinski values (Oprea, 2001; Teague *et al*., 1999). All of these studies address the issues facing compounds discovered using conventional bioassay-based screening of drug-size compound libraries. However, when considering lower molecular weight fragments a different set of rules is likely to apply.

Table 5-1 shows the average calculated physico-chemical properties of a set of 40 fragment hits identified by *Pyramid*™ against three different targets. Only diverse hits have been included in the analysis (in this context, diverse means the hits represent distinctly different opportunities for optimisation). The results of this limited study indicated that the fragments that were hits in our screening process against these targets obey, on average, a 'Rule of 3'™ in which molecular weight (MWT) is < 300, hydrogen bond

Table 5-1. Fragment screening hits - average calculated properties.

Target protein	**No. of Hits**	**'Rule of 3'™ properties**				**Other properties**	
		MWT	**HBA**	**HBD**	**CLogP**	**NROT**	**PSA**
Aspartyl proteinase	13	228	1.1	2.9	2.7	3.5	44
Serine proteinase	13	202	1.7	3.1	1.8	2.9	56
Kinase	14	204	2.5	2	1.6	1.7	61
'Rule of 3'™ guidelines		< 300	≤ 3	≤ 3	≤ 3		

donors (HBD) ≤ 3, hydrogen bond acceptors (HBA) ≤ 3 and ClogP ≤ 3 (Congreve *et al.*, 2003a). In addition, NROT ≤ 3 and PSA ≤ 60 might also be useful limits for fragment selection. This data implies that a 'Rule of 3'™ may be useful for constructing fragment libraries for efficient lead discovery.

A recent publication (Baurin *et al.*, 2004) supports the relevance of this 'Rule of 3'™ guideline. An NMR screening fragment collection of 1315 compounds was assembled using rigorous computational filters based on

calculated properties, functional groups and rings, followed by filtering by a medicinal chemistry team for 'drug-likeness'. This collection of 4 libraries, called SeeDs (Selection of Experimentally Exploitable Drug start points) adhered closely to the 'Rule of 3'™ and, in addition, the hits identified for two targets (CDK2 and HSP90) indicated a minimum level of size and complexity that might also be useful to consider. The findings support the conclusion that hit rates from a library of very simple fragments are, in general, greater than for larger libraries which contain more complex pharmacophores, as determined by 2-D 3-point pharmacophoric finger-printing.

As well as calculating properties for compound selection, a quality control process is needed to ensure that the final library members are useful for screening purposes. In particular, compounds that are shown experimentally to be insoluble under conditions similar to those to be used for crystallographic screening must be removed. It is also useful to study the behavior of the compounds to be included using NMR spectroscopy in a suitable buffer to establish that the compounds are pure, chemically stable over time, do not precipitate and do not aggregate and bind non-specifically to proteins as part of a quality control protocol. Since fragment libraries will tend to consist of only hundreds of compounds and will be used against many protein targets these detailed experiments are of importance to ensure that the fragment libraries are of the highest possible value and so that the results can be easily interpreted. Various approaches can be used to define a useful library of compounds for inclusion in fragment libraries. Some approaches that have been successfully employed by the authors are now outlined.

2.2 Drug fragment library

This library was designed to provide a diverse range of fragments that contain ring systems and side-chains often present in drug molecules. Initially, a small library of simple organic ring systems was identified that occur in drug molecules so that they would be likely to have reduced toxicity liabilities and are amenable to optimisation by medicinal chemistry practices. Previous analyses have demonstrated that only a relatively low number of organic ring systems (sometimes known as scaffolds or frameworks in other work) occur in the majority of drug molecules (Fejzo *et al*., 1999; Bemis and Murcko, 1996 and 1999). Simple ring systems were selected and are shown in Figure 5-2 In addition to those rings found in drugs, a set of simple carbocyclic and heterocyclic ring systems were chosen as additional fragments. The ring systems chosen are shown in Figure 5-3.

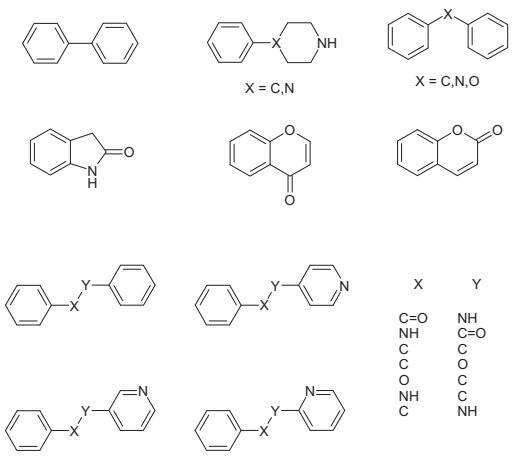


Figure 5-2. Drug ring systems.

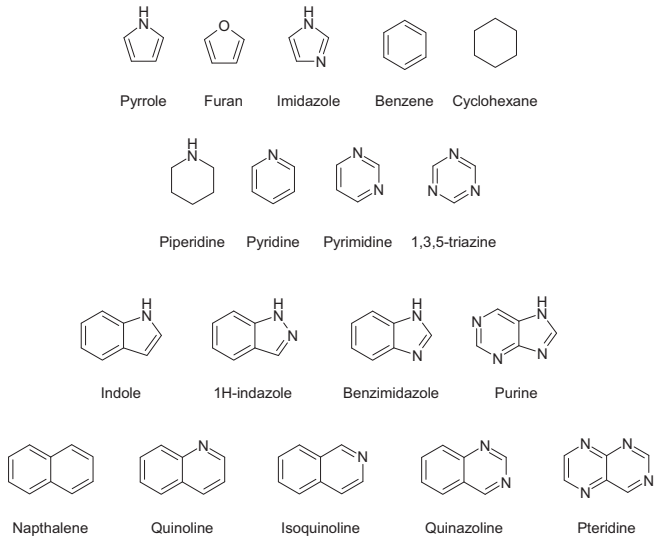


Figure 5-3. Simple carbocyclic and heterocyclic ring systems.

The next stage was to generate a virtual library by combining the ring systems shown above with a set of desirable side-chains, in order to generate the actual members chosen in the Drug Fragment library. The side-chains used in this process were divided into three categories:

1. Functional groups. This set of side-chains consists of those that are observed frequently in drug molecules. The side-chains are shown in Figure 5-4.
2. Lipophilic/secondary substituents. The properties of the fragments were further modified by introducing the substitution of side-chains from another set. Most of these side-chains are lipophilic and were intended to pick up hydrophobic interactions in a protein binding site. The secondary substituents for carbon atoms are shown in Figure 5-5.
3. Nitrogen-substituents. The final possibility for substituting side-chains onto a framework is a set of N-substituents. These are also shown in Figure 5-5.

The virtual library was then created by substituting each of the relevant side-chains onto each of the ring systems. Each carbon atom was substituted by the side-chains from the functional groups and by those from the secondary substituents. The nitrogen atoms were only substituted by the group of N-substituents. Each ring system was limited to one substituent at a

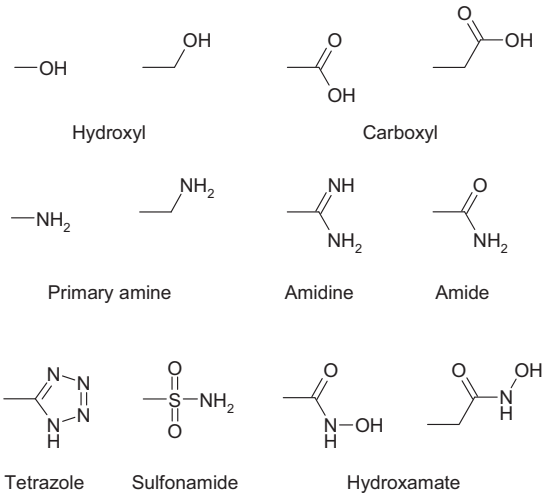


Figure 5-4. Functional group side-chains.

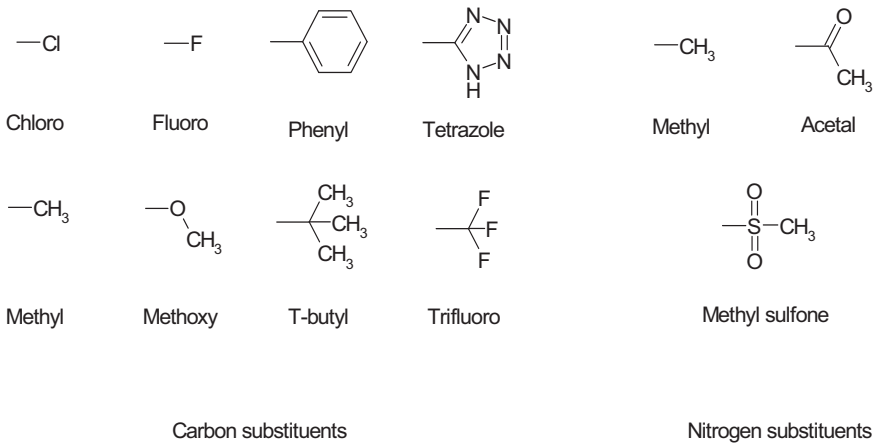


Figure 5-5. Additional substituents.

time, with the exception of benzene and imidazole. The latter two were substituted at all *ortho*, *meta* and *para* positions with all sets of functional and secondary side-chains.

The virtual library was generated as SMILES strings and the resulting compounds were searched for in a database of available compounds. A total of 4513 compounds were generated by the virtual enumeration stage of which 401 were available from commercial suppliers. Manual inspection then removed some compounds with extreme properties and this, coupled with the unavailability of some compounds, resulted in a final library of 327 compounds (the Drug Fragment library).

2.3 Privileged fragment library

In addition to known drug molecules, fragmentation of good quality lead molecules is also possible. We selected a broad range of drug targets (39 enzymes and 25 receptors) and a range of moieties that were considered to be 'privileged' from a medicinal chemistry perspective were then selected. Compounds that best mimicked these fragments were purchased or synthesised to assemble this 'Privileged Fragment library' or PFL. As an additional component of this library we added fragment hits from some of our *Pyramid* screens, because we considered that these molecules had shown some pedigree as valuable X-ray hits. Strict physicochemical property criteria were applied to ensure the average properties of the library were within the 'Rule of 3'™. Finally, compounds that did not fit our quality control limits (>90% pure by both LC/MS and ^{1}H NMR analysis), or were poorly soluble in DMSO at 2M and/or DMSO/aqueous buffer at 50mM were removed and replaced.

2.4 Targeted libraries

Virtual screening using GOLD (Jones et al., 1995 and 1997; Verdonk et al., 2003) was additionally used to construct a number of target-specific fragment libraries. The fragments were selected using a proprietary virtual screening platform which has been described elsewhere (Watson *et al*., 2003). In outline, a database called ATLAS (Astex Therapeutics Library of Available Substances) was constructed from chemical and library suppliers and stored as unique SMILES strings in an Oracle database. The database contains over 3.6 million discrete chemical entities that can be queried from a Windows environment using substructure filters and physical property filters (such as molecular weight, CLogP, polar surface area, 'Rule of 3'™ criteria, *etc*.). This querying will produce lists of available compounds that satisfy the requirements of the user, and can be docked against an active site for the protein target of interest. CORINA (Gasteiger *et al*., 2004) is the program used to convert the SMILES strings into 3D structures. A proprietary version of the docking program GOLD was used to perform virtual screening as described below. A variety of scoring functions were used, *i.e.* GoldScore (Verdonk *et al*., 2003) and ChemScore (Baxter *et al*., 1998) both with and without protein-based pharmacophores. The docking jobs were run on a Linux cluster. The scoring function that was used in the virtual screening was selected, case-by-case, based on docking results for a test set of ligands for which the binding mode was known.

Generally, multiple virtual screens were performed using different protein conformations in order to select the best compounds. The results from the virtual screening were also stored in Oracle and queried using a web-based interface. This interface allows the user to select fragments for visualisation using a number of filters that include: simple physical properties such as molecular weight; scoring functions such as ChemScore, GoldScore, DrugScore (Gohlke *et al*., 2000), predefined pharmacophores, *etc*; components of scoring functions; steric or electrostatic clashes; the formation of specific hydrogen bonds; and 2D substructure. Molecular visualisation is based on the Java-coded molecular visualisation program, AstexViewer™ (Hartshorn, 2002). Multiple ligands can be visualised interactively in the protein and/or the protein surface, and predicted binding modes can be compared with experimentally observed ones.

A Focused Kinase library was also constructed which contained motifs that were thought likely to bind to the ATP binding site of kinases. Analysis of the literature and patents identified a number of scaffolds that were often observed to bind in the ATP binding site of kinases and, in particular, formed hydrogen bonds with the backbone 'hinge' region (Gum *et al*., 1999; Dreyer *et al*., 2001). Enumeration of these scaffolds with drug-like side-chains was

used to construct virtual libraries of candidate fragments and available fragments were purchased from chemical suppliers. The library was additionally augmented with compounds from virtual screening of representative kinases. The version of the Focused Kinase library screened in this work contained 212 compounds.

3. COMPOUND SOAKING, DATA COLLECTION AND PROCESSING

The introduction of small fragments into a protein crystal, via the interstitial spaces and solvent channels, is often facilitated by the use of an organic solvent component (*e.g.* DMSO) in the cocktail soaking experiments. The exact choice of organic solvent is often highly dependent upon the physical characteristics and stability of the protein crystals, but given experimentation most protein crystals can normally be stabilised and soaked. The advent of cryo-protection during X-ray data collections means that complete, high quality, diffraction datasets can usually be collected from a single protein crystal. As such, only one or two protein crystals are typically required to obtain a cocktail soak dataset. Thus a relatively low number of crystals can be used to screen a large fragment library. The rate at which datasets can now be collected, at either in-house, or synchrotron, X-ray sources, has also greatly increased due to the introduction of charge-coupled device (CCD) X-ray detectors and robotic sample changers. Semi-automated data reduction and processing software have also reduced the need for manual intervention en-route to electron density map calculation and evaluation for protein-ligand complexes. The process of electron density map analysis and interpretation has been significantly accelerated by the implementation of software packages such as Astex's AutoSolve® (Blundell *et al.*, 2002). In the majority of cases, AutoSolve® is able to correctly identify the ligand in a cocktail, by automatic trial fitting of the possible ligand molecules to the difference electron density that is consistent with fragment binding. Although diffraction data better than 3Å resolution are sufficient for robust fragment identification, fitting, and refinement, data resolutions better than 2.5Å are preferable.

Fragment-based screening using PYRAMID™ was employed to identify 'hits' against the cell cycle-dependent kinase 2 (CDK2) and the MAP kinase p38α. CDK2 was screened against a set of cocktails constructed from a fragment library that was specifically targeted at the ATP binding cleft of kinases ('Focused Kinase Set', FKS: 212 fragments), as well as a small panel of virtual screening compounds, whilst p38α was screened against a diverse 'Drug Fragment Set' (DFS: 327 fragments). The method used to partition

the library into cocktails maximises the chemical diversity within each cocktail, in order to simplify later deconvolution (Hartshorn *et al.*, 2005). The data obtained for CDK2 and p38α are shown in Table 5-2. Examples of ligands identified from the FKS and DFS are shown in Figures 5-6*a-c*. A 'hit' from virtual screening is shown in Figure 5-6*d*. Details of the methods by which selected fragments from Figure 5-6 were evolved into potent inhibitors are described in a later section.

Table 5-2. Data for FKS and DFS screening against CDK2 and p38α.

Screening set	FKS	DFS
Target protein	CDK2	p38α
No. of fragments in set	212	327
No. of fragments per cocktail	4	4
No. of cocktails	54	79
No. of datasets collected	171	110
No. of single hit cocktails	20	3
No. of multiple hit cocktails	4	0
No. of unique fragment hits	29	3
% of unique fragment hits	13.7	1.0

For many of the cocktails, a specific ligand could clearly be identified in the difference density from its shape and non-bonded contacts with protein and solvent molecules, and it was only when this ligand was subsequently removed from the cocktail that secondary and/or tertiary binders could be identified. This effectively represents a binding competition experiment. In other instances, ambiguous ligand density was observed, suggesting simultaneous, partial or disordered binding of one or more ligands, and it was only on subsequent deconvolution into smaller cocktails and/or singletons that these multiple fragment hits could be unambiguously identified.

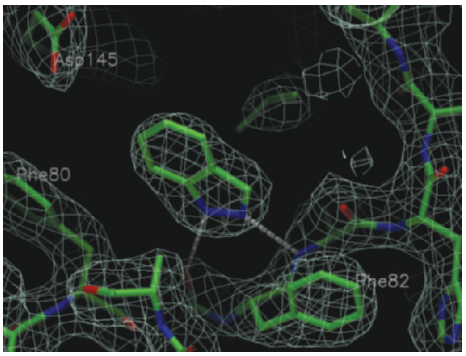


Figure 5-6a.

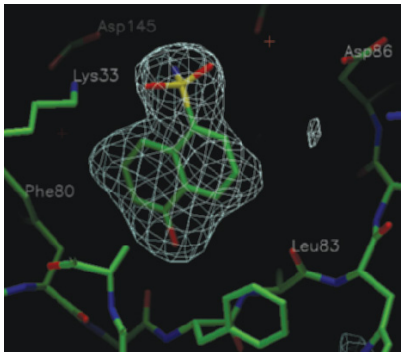
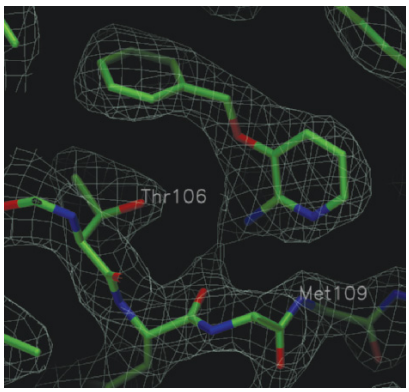


Figure 5-6b.



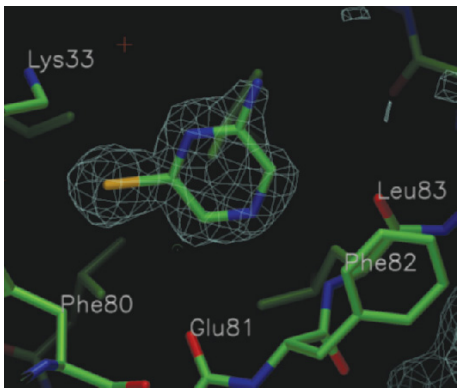


Figure 5-6c. *Figure 5-6d.*

Figure 5-6a: FKS screen 'hit' bound at the active site of CDK2 ($2F_o$-F_c map at 1σ); *b*: FKS screen 'hit' bound at the active site of CDK2 (F_o-F_c map at 3.5σ); *c*: DFS screen 'hit' in the active site of p38α ($2F_o$-F_c map at 1σ); *d*: CDK2 active site 'hit' identified by virtual screening (F_o-F_c map at 3σ).

Table 5-2 shows that four cocktails contained multiple hits. As diverse ligand species may be simultaneously introduced into a crystal there is the possibility of performing '*in situ*' chemistry within a crystal. In a study using CDK2, Congreve *et al.* (2003b) have described how a template fragment, that is cross-reactive with a degenerate library of potential substituent fragments, can be used to synthesise and screen a library of more complex ligands, *in situ*, within a protein crystal (see Figure 5-7), thus

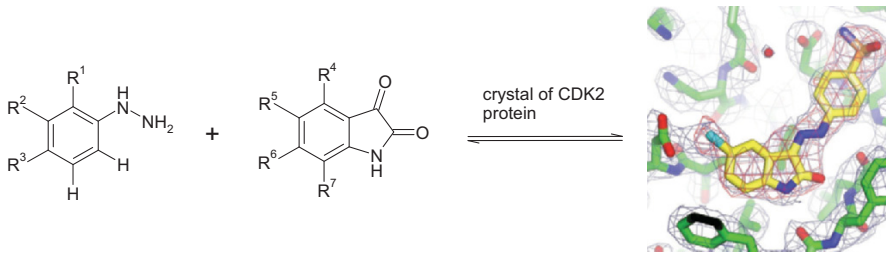


Figure 5-7. Example of Dynamic Combinatorial Crystallography (DCX™) showing the combination of two reactive sets of intermediates (in this case aryl hydrazines with variable groups R1-R3 and isatins with variable groups R4-R7) within a crystal of the kinase CDK2.

allowing the identification of the preferred product of the combinatorial library to be directly identified by protein-ligand crystallography.

4. PROTEIN-FRAGMENT INTERACTIONS IN CDK2

CDK2 is a representative of a series of Ser/Thr protein kinases involved in the cell cycle. CDK inhibitors are thought to have potential in anti-cancer therapy (Fischer *et al*., 2000; Knockaert *et al*., 2002; Sausville *et al*., 1999). The architecture of the adenosine 5'-triphosphate (ATP) binding site of protein kinases is structurally well understood and CDK2 behaves as a classic case study (Gill , 2004).

Crystals of CDK2 were soaked with our Focused Kinase Set and provided 29 hits out of 212 compounds (13.7% hit rate). A wide range of activities of the identified fragments has been measured with enzyme IC_{50}'s in the range from millimolar to low micromolar.

Despite the fact that the Focused Kinase Set was designed with compounds planned to bind the ATP binding site, in depth structural analysis of the hits showed a diverse set of interactions within the pocket. To facilitate the analysis of protein-fragment interactions, the CDK2 ATP binding pocket can be visualised as four main pharmacophoric regions (Figure 5-8).

The 'hinge' region (residues Glu81 to Leu83, delineated in yellow in Figure 5-8) lies between the two kinase lobes and plays the major role, with polar interactions occurring between the backbone of these two key amino-acids and the ligand. The central hydrogen bond between the NH of Leu83 and an acceptor of the ligand is the main interaction. The pocket, shaped into a narrow hydrophobic cavity by Leu134 and Ala31 side-chains, appears optimal for binding heteroaromatic rings (e.g. purines, pyrazines, indazoles), although a wide range of fragments possessing an acceptor have been observed in this pocket. From the screened set, 28 out of the 29 hits identified form this H-bond and the remaining hit uses a water to bridge itself with the backbone NH of Leu83.

The two carbonyls of Glu81 and Leu83 act as acceptors and they are potentially able to form two further H-bonds. A single fragment can pick up one of both the carbonyls simultaneously. In this second example, a strong network of three H-bonds become the core of the inhibitor activity.

It is worth noting that a 'weak' H-bond is often observed between an electron-deficient aromatic C-H group and one of the carbonyls (Pierce, 2002). This positive interaction can be seen in ATP itself. A substituted piperazine shows how two of these interactions are capable of driving the

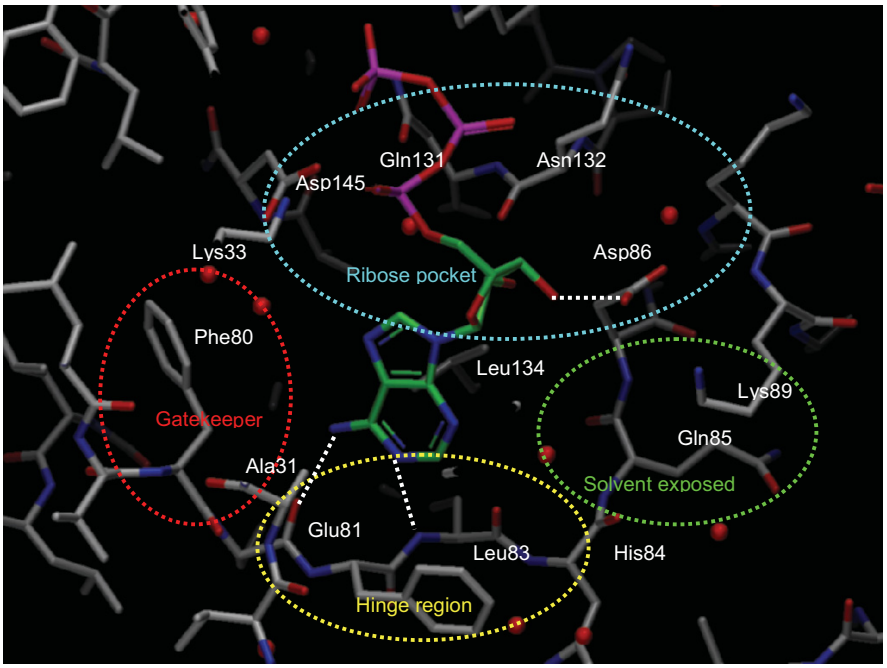


Figure 5-8. The CDK2 ATP binding site can be visualised as four separate regions.

overall binding mode; the interaction through the CH-N-CH motif is clearly shown in Figure 5-9.

Moving along the backbone to the 'hinge' region, a contiguous sequence of amino-acids (His84 to Lys89) forms a loop that exposes this part of the site to the outer surface of the kinase, towards the solvent (Figure 5-8, in green). This pocket is organised as a hydrophobic region defined by Ile10 side-chain and the backbone loop itself, together with several hydrophilic side-chains (Gln85, Asp86, Lys89) pointing towards the external surface. Hydrophobic groups with a polar head tend to show the highest affinity in this pocket. The part of the ligand occupying this region competes with water molecules to interact with the numerous donor and acceptors present, forming H-bonds. Polar and charged groups are favoured here (Anderson, 2003; Davis, 2001; Clare, 2001; Misra, 2004). The size and features of the fragment may induce the formation of a salt bridge between Asp86 and Lys89.

On the left hand side of the hinge there is the so-called 'gatekeeper' residue (Figure 5-8, in red), a key kinase amino-acid that modulates the contour of the ATP pocket, so advantage can be taken of this to design selective inhibitors. The gatekeeper in CDK2 is a rigid phenylalanine (Phe80) that together with the Val64 side-chain makes this a small

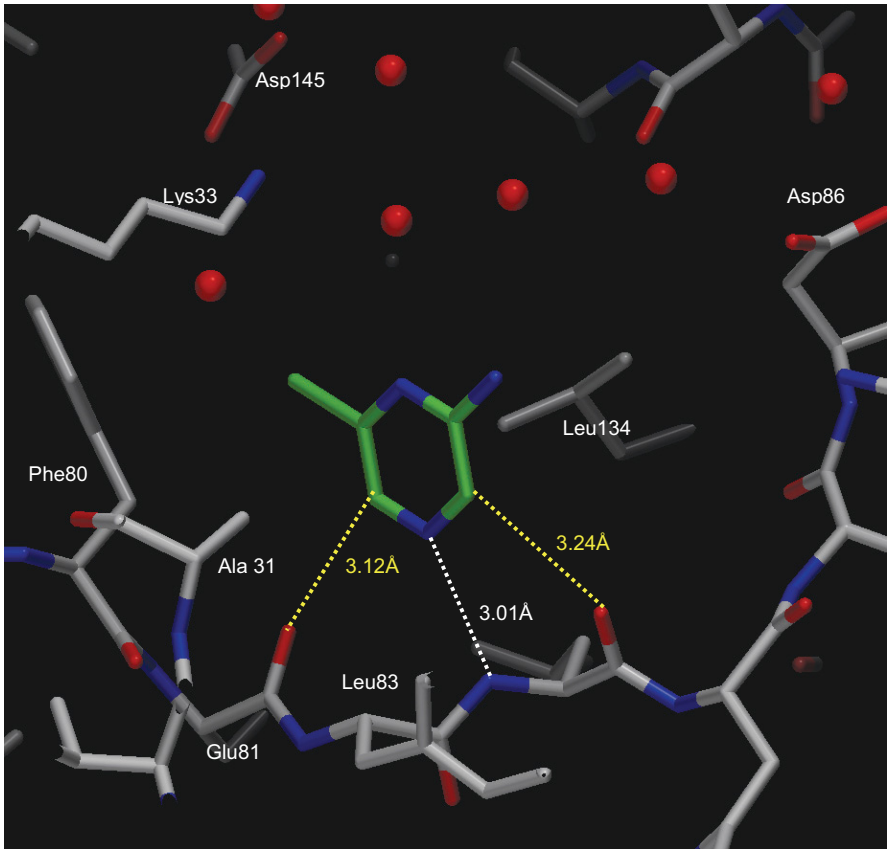


Figure 5-9. This substituted piperazine binds to the hinge residues forming a 'normal' H-bond (in white) with its N atom and the NH of Leu83. It also forms two 'weak' H-bonds using its electron deficient aromatic C-H bonds (in yellow) as donors that are accepted by the backbone O atoms of Glu81 and Leu83.

hydrophobic region. Lipophilic interactions have been observed in several protein-ligand complexes with aromatic and small alkyl groups and halides with the inhibitor stacking against the Phe80 aromatic ring (Dreyer, 2000; Furet, 2002; Pevarello, 2004; Tang, 2003). Indazole in the kinase set is one example of a compound stacking an aromatic ring against Phe80 (Figure 5-10).

The fourth region is the location where the ribose ring of adenosines sits together with phosphate and magnesium ions, if present (Figure 5-8, in cyan). It is a large hydrophilic pocket, dominated by four polar amino-acids (Asp145, Asp86, Gln131 and Asn132). In the absence of a ligand this is a highly solvated region. Positive charges and H-bond donor-containing

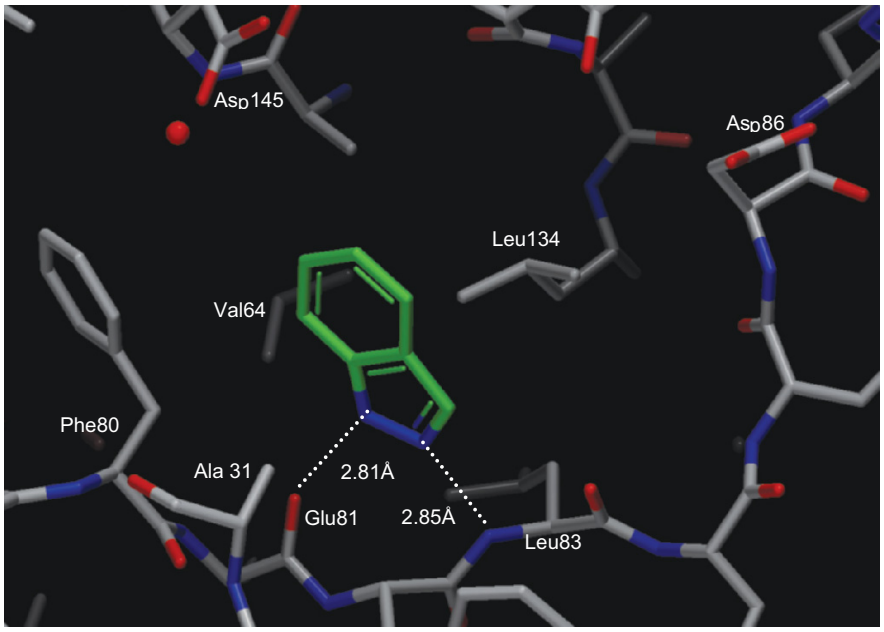


Figure 5-10. Indazole interacts with the hinge residues forming two H-bonds (in white) with its two N atoms. The aromatic ring occupies the hydrophobic region defined by Phe80 (gatekeeper) and Val64.

ligands can either form water bridges or displace the waters to form H-bond networks. For example, a naphthol sulphonamide uses a phenolic function to interact with the hinge, extending its polar sulphonamide group into this region and forming a water-mediated interaction with the side-chain of Asp86 (Figure 5-11).

It is worth noting that the Asp145 and Lys33 residues, which can either form a salt bridge or bind the ligand, are involved in an important rearrangement within the protein when cyclinA activates the kinase (Arris, 2000; Brown, 1999; Davies, 2002; Gibson, 2002). In the cyclin-bound complexes they are not available to form the same interactions with the ligand, so that Lys33 forms a salt bridge with the incoming Glu51 and Asp145 interacts with the NH of Gly147.

5. HITS-TO-LEADS OPTIMIZATION

P38α mitogen-activated protein (MAP) kinase is an intracellular serine/threonine (Ser/Thr) kinase that is activated by a range of environmental stimuli such as TNF-α, IL-1β and stress (Han *et al.*, 1994;

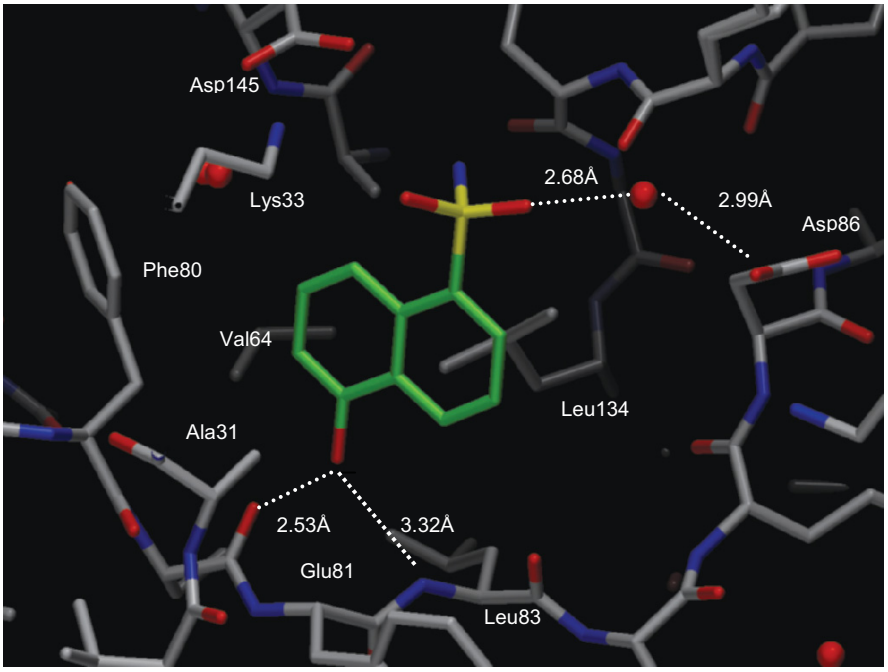


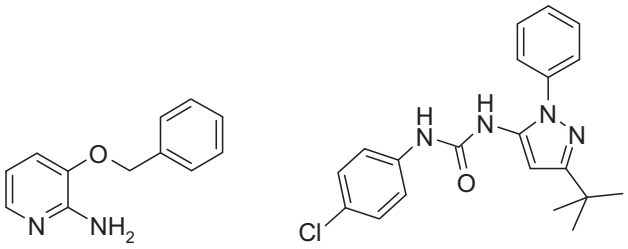
Figure 5-11. This naphthol sulphonamide uses its OH group to interact with the hinge forming two H-bonds and extends its polar sulphonamide group into the 'ribose pocket' to form a water-mediated interaction with the side-chain of Asp86.

Raingeaud *et al*., 1996a). Activation of p38α occurs through bis-phosphorylation by the dual-specificity Ser/Thr MAP kinases MKK3 and MKK6 on the Thr180-Gly181-Tyr182 motif located on the activation loop (Raingeaud *et al*., 1996a,b). In its activated state, p38α phosphorylates a range of intracellular protein substrates that post-transcriptionally regulate the biosynthesis of TNF-α and IL-1β. The pathophysiological consequence of excessive production of TNF-α and IL-1β is thought to be significant mediation of the progression of many inflammatory diseases such as inflammatory bowel disease (Badger *et al*., 1996), psoriasis (Feldmann *et al*., 1996), Crohn's disease (Rutgeerts *et al*., 1999), and rheumatoid arthritis (Foster *et al*., 2000).

In the preceding sections we have described how high-throughput X-ray crystallographic screening can be used to identify low-affinity fragment hits for a range of targets. Whilst these fragment hits have only low potency (μM-mM), they are deemed to exhibit 'high efficiency' binding given their low molecular weights (100-250Da) and limited levels of functionality. In this case study we describe the structure-guided chemistry strategy that was

employed to transform one fragment hit into potent compounds having 'lead-like' properties (Oprea, 2002). Our strategy relies on obtaining many protein-ligand co-crystal structures to guide iterative cycles of hit optimisation (Blundell *et al*., 2002b; Carr & Jhoti, 2002).

2-Amino-3-benzyloxypyridine 1 (Figure 5-12, IC_{50}=1.3mM) was identified from X-ray crystallographic fragment screening of non-phosphorylated p38α MAP kinase. Here we describe the key compounds synthesised, together with relevant details concerning the observed crystallographic binding modes of more advanced compounds, culminating in a novel and potent lead series.



1 **2**

Figure 5-12. Fragments hits 1 and 2.

Fragment 1 binds competitively to the previously described hinge region of the ATP binding site. As expected for such a low molecular weight compound, 1 demonstrated low *in vitro* potency against p38α (IC_{50}=1.3mM); however the synthetic tractability of this molecule made it a valuable starting point to initiate structure-guided fragment optimisation. The X-ray crystal structure of p38α MAP kinase complexed with fragment hit 1 was solved and demonstrates a clearly defined binding mode (Figure 5-13) where 1 interacts with the hinge region. The inhibitor makes H-bond interactions through the pyridyl N atom to the NH of Met109 and the 2-amino group to the His107 O atom. The benzyloxy group also makes a major interaction by filling the lipophilic specificity pocket.

Analysis of the X-ray structure of 1 suggested that our initial efforts to increase potency of this low-affinity fragment should focus on improving the hydrophobic interaction with the lipophilic specificity pocket. Appropriate space-filling of this pocket has been shown to confer large increases in affinity and selectivity (Chakravarty & Dugar, 2002; Cirillo *et al*., 2002; Jackson & Bullington, 2002; Lisnock *et al*., 1998). Substitution of the phenyl ring of 1 with a 2,6-dichloro substitution pattern proved optimal (4, IC_{50}=109μM), affording an increase in potency of over 10-fold (Table 5-3). A further improvement in affinity for p38α was gained through replacement

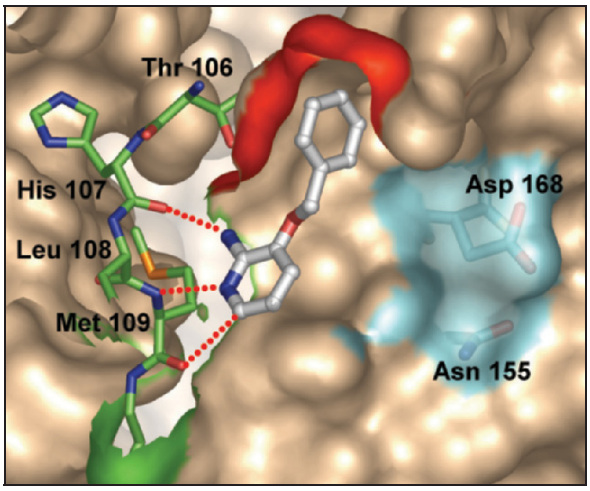
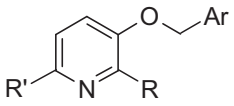


Figure 5-13. The p38α MAP kinase ATP binding site with Fragment 1 bound. The hydrophobic region 1 (red), which is not occupied by ATP, is a known key specificity pocket (Lisnock *et al*., 1998). Region 2 (green) is also hydrophobic and accommodates aliphatic moieties on inhibitors with terminal polar groups to modulate potency and physicochemical properties. The ribose-binding region (blue) contains the amino-acids that interact with the hydroxyl groups of the ribose of ATP. Fragment 1 is outlined in grey and shown making key H-bond contacts to hinge residues.

of the benzyl group of 1 with a 1-napthyl group (Figure 5-15). This substitution (5, IC_{50}=44μM) resulted in a 30-fold increase in potency over 1 by virtue of its more extensive hydrophobic interface in the specificity pocket. Despite the superior increase in activity conferred by the 1-naphthyl group, the 2,6-dichloro substituted compound 4 was advanced instead, since it offered the best balance of potency and physicochemical properties. Despite being involved by virtue of H-bonding to the O atom of His107, the 2-amino substituent of 1 was found to be unnecessary for intrinsic p38α activity as demonstrated by des-amino compounds such as 3 (IC_{50}=1mM), and was thus removed from subsequent analogues. The X-ray structure of 1 instead suggested that accessing the hydrophobic region 2 (Figure 5-13) with appropriate functionality might increase inhibitor activity. Indeed, substitution in the pyridine ring *para*- to the 2,6-dichlorobenzyloxy group was found to be tolerated and the *gem*-dimethylethanolamine derivative 7 gave an increase in activity compared with 4, albeit only four-fold.

To explore alternative routes for optimisation of fragment 1, we synthesised and solved the protein-ligand complexes of a number of key compounds described in the literature in order to understand more fully the

Table 5-3. Activities of 1 & 3-10 against p38α MAP kinase



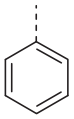
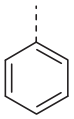
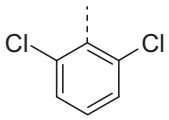
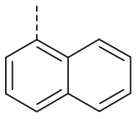
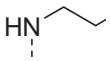
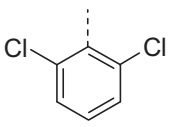
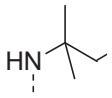
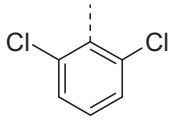
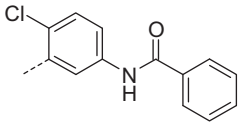
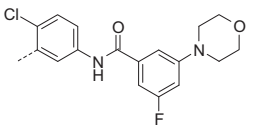
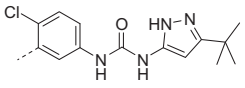
Compound	R	R'	Ar	IC_{50} (μM)†
1	NH_2	H		1300
3	H	H		1000
4	NH_2	H		109
5	NH_2	H		44
6	H			200

Table 5-3. (Continued)

7	H			24
8	H	H		30
9	H	H		0.065
10	H	H		0.35

†Average of 2 or more determinations.

interactions made by these potent inhibitors. Synthesis of the known urea 2 (Figure 5-12, IC_{50}=196nM) and subsequent soaking into p38α MAP kinase crystals revealed a unique binding mode for this compound (Dumas *et al*., 2002). A significant conformational rearrangement of the residues Asp168-Phe169-Gly170 (DFG motif) in the conserved activation loop of the kinase was induced, revealing a polar channel formed by Asp168 and Glu71 from the ATP binding site to a lipophilic pocket formed by an approximate 10Å movement of Phe169. Overlaying X-ray crystal structures of 2 with our inhibitors suggested a number of functional group substitution opportunities from the 2- and 5-positions of the phenyl ring of 1 to improve potency and selectivity (Figure 5-14).

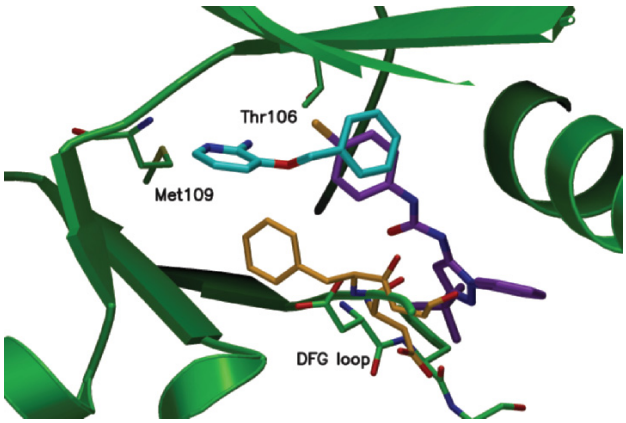


Figure 5-14. Compound 1 (cyan) superposed on compound 2 (purple). The significant shift in the DFG loop (green for 1, orange for 2) is highlighted. Unlike 1, 2 does not interact with the hinge region of the kinase.

Synthesis of a range of amide and urea analogues, including 8, 9 and 10 (Table 5-3) allowed us to exploit this activation loop movement. Benzamide 8 (IC_{50}=30μM) was tolerated, but a significant potency improvement was made with amide 9 (IC_{50}=65nM). The *tert*-butyl-substituted pyrazole urea 10 (IC_{50}=350nM) also served as a useful template for us to develop analogues with nanomolar potency against p38α MAP kinase. Large changes in the conformations of the conserved residues of the DFG motif are required for binding of these amide and urea derivatives (9 and 10, Figures 5-15*a*, *b*).

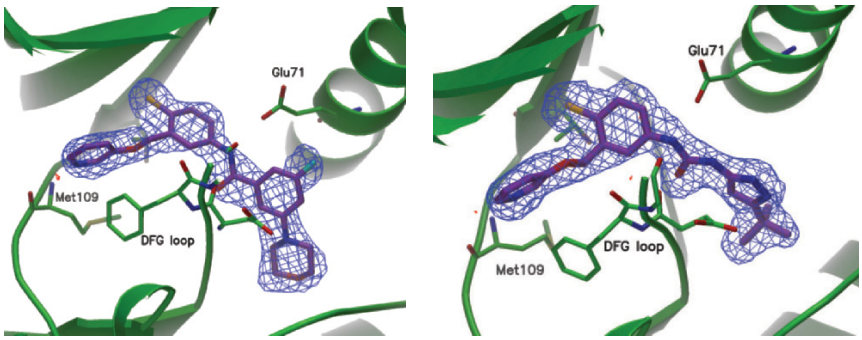


Figure 5-15a, b. Omit map with difference electron density (contoured at 3σ) for 9 and 10 bound to p38α.

The Phe169 side chain has again moved approximately 10Å to a new "DFG-out" conformation, whereby the Phe169 side chain now hinders access to the ATP binding site (Pargellis *et al*., 2002; Regan *et al*., 2002; Regan *et al*., 2003). The polar channel formed by Asp168 and Glu71 from the ATP binding site to this allosteric pocket is involved in a H-bond network to both the amide and urea functionalities. This movement of the Phe169 side chain exposes a largely lipophilic pocket into which the morpholine of amide 9 or the *tert*-butyl group of 10 may then insert. The morpholine and *tert*-butyl moieties respectively make substantial contacts with a number of neighbouring hydrophobic residues; both 9 and 10 still maintain contacts to the hinge and lipophilic specificity pockets.

The case study outlined exemplifies the potential of X-ray crystallographic screening and structure-guided fragment optimisation to generate useful lead compounds. We have identified a potent and selective series of p38α MAP kinase inhibitors starting from a mM-affinity fragment 1 by employing structure-guided chemistry. Initial attempts to improve potency were centred on the hydrophobic regions (Figure 5-13: red and green regions), gaining over a 50-fold improvement in kinase activity for 7 over 1. Attempts to improve the potency of the series into the nM range successfully exploited a binding pocket revealed by the large conformational movement of the conserved DFG binding loop. Compound 9 has drug-like physicochemical properties, demonstrates at least 100-fold selectivity over a panel of kinases, and was able to potently inhibit the production of TNF-α from LPS-stimulated THP-1 cells. The encouraging early *in vitro* data generated from this series warrants further investigation in a range of animal models of inflammatory disease.

6. SUMMARY

The application of the newly established fragment screening and hit optimisation technology, based on automated high-throughput X-ray crystallography, iterative structure-based design and synthetic medicinal chemistry, has enabled the discovery of novel potent inhibitors of kinases that have now reached the clinical trial stage of development in a remarkably short time-frame. Fragment-based screening is now clearly established as an important new tool for drug discovery.

ACKNOWLEDGEMENTS

We would like to thank Harren Jhoti for useful discussions during the preparation of this article. All figures were visualised and prepared using the AstexViewer™ program (Hartshorn, 2002).

Figures 5-1, 5-2 and 5-3 are reprinted from Drug Discovery Today Vol. 7(9): Carr, R. and Jhoti, H., Structure-based screening of Low-affinity compounds, pp. 522-527, Copyright(2002), with permission from Elsevier.

Figure 5-4 and 5-5 are reprinted with permission from J. Med. Chem. Vol. 48(2): Hartshorn, M.J., Murray, C.W., Cleasby, A., Frederickson, M., Tickle, I.J., and Jhoti, H., Fragment-based lead discovery using X-ray crystallography, pp. 403-413, Copyright(2005), American Chemical Society.

Figure 5-7 is reprinted with permission from Angewandte Chemie (International edition) Vol. 42(37): Congreve, M.S., Davis, D.J., Devine, L., Granata, C., O'Reilly, M., Wyatt, P., and Jhoti, H., Detection of ligands from a dynamic combinatorial library by X-ray Crystallography, pp. 4479-4482, Copyright(2003), Wiley-VCH.

Figures 5-13, 5-14 and 5-15 and Table 5-3 are reprinted with permission From J Med. Chem. Vol. 48(2): Gill, A.L., Frederickson, M., Cleasby, A., Woodhead, S.J., Carr, M.G., Woodhead, A.J., Walker, M.T., Congreve, M.S., Devine, L.A., Tisi, D., Magor, L.C.A., Davis, D.J., Curry, J., Anthony, R., Padova, A., Murray, C.W., Carr, R.A.E., and Jhoti, H., Identification of novel p38>Greek alpha< MAP kinase inhibitors using Fragment-based lead generation, pp. 414-426, Copyright(2005), American Chemical Society.

REFERENCES

Anderson M., Beattie J. F., Breault G. A., Breed J., Byth K. F., Culshaw J. D., Ellston R. P., Green S., Minshull C. A., Norman R. A. et al. 2003. Imidazo[1,2-A]pyridines: A potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hybridisation. Bioorg. Med. Chem. Lett. **13(18)**, 3021–3026.

Arris C. E., Boyle F. T., Calvert A. H., Curtin N. J., Endicott J. A., Garman E. F., Gibson A. E., Golding B. T., Grant S., Griffin R. J. et al. 2000. Identification of novel purine and pyrimidine cyclin-dependent kinase inhibitors with distinct molecular interactions and tumor cell growth inhibition profiles. J. Med. Chem. **43(15)**, 2797–2804.

Badger A. M., Bradbeer J. N., Votta B., Lee J. C., Adams J. L. and Griswold D. E. 1996. Pharmacological profile of SB 203580, a selective inhibitor of cytokine suppressive binding protein/p38 kinase, in animal models of arthritis, bone resorption, endotoxin shock and immune function. J. Pharmacol. Exp. Ther. **279**, 1453–1461.

Baurin N., Aboul-Ela F., Barril X., Davis B., Drysdale M., Dymock B., Finch H., Fromont C., Richardson C., Simmonite H. and Hubbard R. E. 2004. Design and characterization of libraries of molecular fragments for use in NMR screening against protein targets. J. Chem. Inf. Comput. Sci. **44(6)**, 2157–2166.

Baxter C. A., Murray C. W., Clark D. E., Westhead D. R. and Eldridge M. D. 1998. Flexible docking using Tabu search and an empirical estimate of binding affinity. Proteins **33**, 367–382.

Bemis G. W. and Murcko M. A. 1996. Properties of known drugs. 1. Molecular frameworks. J. Med. Chem. **39**, 2887–2893.

Bemis G. W. and Murcko M. A. 1999. Properties of known drugs. 2. Side chains. J. Med. Chem. **42**, 5095–5099.

Blundell T. L., Abell C., Cleasby A., Hartshorn M. J., Tickle I. J., Parasini E. and Jhoti H. 2002a. In Drug Design: Special Publication, ed. D. R. Flower, Royal Society of Chemistry, Cambridge, vol. 279 pp. 53–59.

Blundell T. L., Jhoti H. and Abell C. 2002b. High-throughput crystallography for lead discovery in drug design. Nat. Rev. Drug Discov. **1**, 45–54.

Brown N. R., Noble M. E., Lawrie A. M., Morris M. C., Tunnah P., Divita G., Johnson L. N. and Endicott J. A. 1999. Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity. J. Biol. Chem. **274(13)**, 8746–8756.

Card G. L., Blasdel L., England B. P., Zhang C., Suzuki Y., Gillette S., Fong D., Ibrahim P. N., Artis D. R., Bollag G. et al. 2005. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. Nat. Biotechnol. **23**, 201–207.

Carr R. and Jhoti H. 2002. Structure-based screening of low-affinity compounds. Drug Discovery Today **7**, 522–527.

Chakravarty S. and Dugar S. 2002. Inhibitors of p38 alpha MAP kinase. Annu. Rep. Med. Chem. **37**, 177-186.

Cirillo P. F., Pargellis C. and Regan J. 2002. The non-diaryl heterocycle classes of p38 MAP kinase inhibitors. Curr. Top. Med. Chem. **2**, 1021–1035.

Clare P. M., Poorman R. A., Kelley L. C., Watenpaugh K. D., Bannow C. A. and Leach K. L. 2001. The cyclin-dependent kinases cdk2 and cdk5 act by a random, anticooperative kinetic mechanism. J. Biol. Chem. **276(51)**, 48292–48299.

Clark D. E. and Picket S. D. 2000. Computational methods for the prediction of drug-likeness. Drug Discovery Today **5**, 49–58.

Congreve M., Carr R., Murray C. and Jhoti H. 2003a. A 'rule of three'™ for fragment-based lead discovery? Drug Discovery Today **8**, 876–877.

Congreve M. S., Davis D. J., Devine L., Granata C., O'Reilly M., Wyatt P. and Jhoti H. 2003b. Detection of ligands from a dynamic combinatorial library by X-ray crystallography. Angew. Chem. Int. Ed. **42**, 4479–4482.

Davies T. G., Bentley J., Arris C. E., Boyle F. T., Curtin N. J., Endicott J. A., Gibson A. E., Golding B. T., Griffin R. J., Hardcastle I. R. et al. 2002. Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. Nat. Struct. Biol. **9(10)**, 745–749.

Davis S. T., Benson B. G., Bramson H. N., Chapman D. E., Dickerson S. H., Dold K. M., Eberwein D. J., Edelstein M., Frye S. V., Gampe, R. T., Jr. et al. 2001. Prevention of chemotherapy-induced alopecia in rats by CDK inhibitors. Science **291**(5501), 134–137.

Dreyer M. K., Borcherding D. R., Dumont J. A., Peet N. P., Tsay J. T., Wright P. S., Bitonti A. J., Shen J. and Kim S. H. 2001. Crystal structure of human cyclin-dependent kinase 2 in complex with the adenine-derived inhibitor H717. J. Med. Chem. **44**, 524–530.

Dumas J., Hatoum-Mokdad H., Sibley R. N., Smith R. A., Scott W. J., Khire U., Lee W., Wood J., Wolanin D., Cooley J. et al. 2002. Synthesis and pharmacological characterization of a potent, orally active p38 kinase inhibitor. Bioorg. Med. Chem. Lett. **12**, 1559–1562.

Fejzo J., Lepre C. A., Peng J. W., Bemis G. W., Ajay, Murcko M. A. and Moore J. M. 1999. The SHAPES strategy: An NMR-based approach for lead generation in drug discovery. Chem. Biol. **6**, 755–769.

Feldmann M., Brennan F. M. and Maini R. N. 1996. Role of cytokines in rheumatoid arthritis. Annu. Rev. Immunol. **14**, 397–440.

Fischer P. M. and Lane D. P. 2000. Inhibitors of cyclin-dependent kinases as anti-cancer therapeutics. Curr. Med. Chem. **7**, 1213–1245.

Foster M. L., Halley F. and Souness J. E. 2000. Potential of p38 inhibitors in the treatment of rheumatoid arthritis. Drug News Perspect. **13**, 488–497.

Furet P., Meyer T., Strauss A., Raccuglia S. and Rondeau J. M. 2002. Structure-based design and protein X-ray analysis of a protein kinase inhibitor. Bioorg. Med. Chem. Lett. **12(2)**, 221–224.

Gasteiger J., Rudolph C. and Sadowski J. 2004. Automatic generation of 3D-atomic coordinates for organic molecules. Tetrah. Comput. Meth. **3**, 537–547.

Gibson A. E., Arris C. E., Bentley J., Boyle F. T., Curtin N. J., Davies T. G., Endicott J. A., Golding B. T., Grant S., Griffin R. J. et al. 2002. Probing the ATP ribose-binding domain of cyclin-dependent kinases 1 and 2 with O(6)-Substituted guanine derivatives. J. Med. Chem. **45(16)**, 3381–3393.

Gill A. L. 2004. New lead generation strategies for protein kinase inhibitors – fragment-based screening approaches. Mini. Rev. Med. Chem. **4(3)**, 301–311.

Gill A. L., Frederickson M., Cleasby A., Woodhead S. J., Carr M. G., Woodhead A. J., Walker M. T., Congreve M. S., Devine L. A., Tisi D. et al. 2005. Identification of novel p38? MAP kinase inhibitors using fragment-based lead generation. J. Med. Chem. **48(2)**, 414–426.

Gohlke H., Hendlich M. and Klebe G. 2000. Knowledge-based scoring function to predict protein-ligand interactions. J. Mol. Biol. **295**, 337–356.

Gum R. J., McLaughlin M. M., Kumar S., Wang Z., Bower M. J., Lee J. C., Adams J. L., Livi G. P., Goldsmith E. J. and Young P. R. 1998. Acquisition of sensitivity of stress-activated protein kinases to the p38 inhibitor, SB 203580, by alteration of one or more amino acids within the ATP binding pocket. J. Biol. Chem. **273**, 15605–15610.

Han J., Lee J. D., Bibbs L. and Ulevitch R. J. 1994. A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. Science **265**, 808–811.

Hann M., Leach A. R. and Harper G. 2001. Molecular complexity and its impact on the probability of finding leads for drug discovery. J. Chem. Inf. Comput. Sci. 41**(3)**, 856–864.

Hartshorn M. J. 2002. AstexViewer? A visualisation aid for structure-based drug design, J. Comp. Aid. Mol. Des. **16**, 871–881.

Hartshorn M. J., Murray C. W., Cleasby A., Frederickson M., Tickle I. J. and Jhoti H. 2005. Fragment-based lead discovery using X-ray crystallography. J. Med. Chem. **48(2)**, 403–413.

Hopkins A. L., Groom C. R. and Alex A. 2004. Ligand efficiency: A useful metric for lead selection. Drug Discovery Today **9**, 430–431.

Jackson P. F. and Bullington J. L. 2002. Pyridinylimidazole-based p38 MAP kinase inhibitors. Curr. Top. Med. Chem. **2**, 1011–1020.

Jones G., Willett P. and Glen R. C. 1995. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol. **245**, 43–53.

Jones G., Willett P., Glen R. C., Leach A. R. and Taylor R. 1997. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. **267**, 727–748.

Knockaert M., Greengard P. and Meijer L. 2002. Pharmacological inhibitors of cyclin-dependent kinases. Trends Pharmacol. Sci. **23**, 417–425.

Lesuisse D., Lange G., Deprez P., Bénard D., Schoot B., Delettre G., Marquette J.-P., Broto P., Véronique J.-B., Bichet P. et al. 2002. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: Application to the discovery of nanomolar inhibitors of Src SH2. J. Med. Chem. **45**, 2379–2387.

Lipinski C. A., Lombardo F., Dominy B. W. and Feeney P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **46**(**1–3**), 3–26.

Lisnock J., Tebben A., Frantz B., O'Neill E. A., Croft G., O'Keefe S. J., Li B., Hacker C., de Laszlo S., Smith A. et al. 1998. Molecular basis for p38 protein kinase inhibitor specificity. Biochemistry **37**, 16573–16581.

Misra R. N., Xiao H. Y., Kim K. S., Lu S., Han W. C., Barbosa S. A., Hunt J. T., Rawlins D. B., Shan W., Ahmed S. Z. et al. 2004. N-(Cycloalkylamino)acyl-2-aminothiazole inhibitors of cyclin-dependent kinase 2. N-[5-[[[5-(1,1-dimethylethyl)-2-oxazolyl]methyl]thio]-2-thiazolyl]-4-piperidine carboxamide (BMS-387032), a highly efficacious and selective antitumor agent. J. Med. Chem. **47(7)**, 1719–1728.

Nienaber V. L., Richardson P. L., Klighofer V., Bouska J. J., Giranda V. L. and Greer J. 2000. Nat. Biotechnol. **18**, 1105–1108.

Oprea T. I. 2001. Is there a difference between leads and drugs? A historical perspective. J. Chem. Inf. Comput. Sci. **41**, 1308–1315.

Oprea T. I. 2002. Current trends in lead discovery: Are we looking for the appropriate properties? J. Comput. Aided Mol. Des. **6**, 325–334.

Pargellis C., Tong L., Churchill L., Cirillo P. F., Gilmore T., Graham A. G., Grob P. M., Hickey E. R., Moss N., Pav S. and Regan J. 2002. Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. Nat. Struct. Biol. **9**, 268–272.

Pevarello P., Brasca M. G., Amici R., Orsini P., Traquandi G., Corti L., Piutti C., Sansonna P., Villa M., Pierce B. S. et al. 2004. 3-Aminopyrazole inhibitors of CDK2/cyclin A as antitumor agents. Part 1. Lead Finding. J. Med. Chem. **47(13)**, 3367–3380.

Pierce A. C., Sandretto K. L. and Bemis G. W. 2002. Kinase inhibitors and the case for CH…O hydrogen bonds in protein-ligand binding. Proteins **49(4)**, 567–576.

Raingeaud J., Gupta S., Rogers J. S., Dickens M., Han J., Ulevitch R. J. and David R. J. 1996a. Characterization of the structure and function of a novel MAP kinase kinase (MKK6). J. Biol. Chem. **271**, 2886–2891.

Raingeaud J., Whitmarsh A. J., Barrett T., Derijard B. and Davis R. J. 1247. MKK3- and MKK6-regulated gene expression is mediated by the p38 mitogen-activated protein kinase signal transduction pathway. Mol. Cell Biol. **16**, 1247.

Regan J., Breitfelder S., Cirillo P., Gilmore T., Graham A. G., Hickey E., Klaus B., Madwed J., Moriak M., Moss N. et al. 2002. Pyrazole urea-based inhibitors of p38 MAP kinase: From lead compound to clinical candidate. J. Med. Chem. **45**, 2994–3008.

Regan J., Capolino A., Cirillo P. F., Gilmore T., Graham A. G., Hickey E., Kroe R. R., Madwed J., Moriak M., Nelson R. et al. 2003. Structure-activity relationships of the p38? MAP kinase inhibitor 1-(5-tert-Butyl-2-p-tolyl-2H-pyrazol-3-yl)-3-[4-(2-morpholin-4-yl-ethoxy)naphthalen-1-yl]urea (BIRB 796). J. Med. Chem. **46**, 4676–4686.

Rutgeerts P., D'Haens G., Targan S., Vasiliauskas E., Hanauer S. B., Present D. H., Mayer L., van Hogezand R. A., Braakman T., DeWoody K. L. et al. 1999. Efficacy and safety of retreatment with anti-tumour necrosis factor antibody (Infliximab) to maintain remission in Crohn's disease. Gastroenterology **117**, 761–769.

Sausville E. A., Zaharevitz D., Gussio R., Meijer L., Louarn-Leost M., Kunick C., Schultz R., Lahusen T., Headlee D., Stinson S. et al. 1999. Cyclin-dependent kinases: initial approaches to exploit a novel therapeutic target. Pharmacol. Ther. **82**, 285–292.

Tang J., Shewchuk L. M., Sato H., Hasegawa M., Washio Y. and Nishigaki N. 2003. Anilinopyrazoles as selective CDK2 inhibitors: Design, synthesis, biological evaluation, and X-ray crystallographic analysis. Bioorg. Med. Chem. Lett. **13(18)**, 2985–2988.

Teague S. J., Davis A. M., Leeson P. D. and Oprea T. 1999. The design of lead-like combinatorial libraries. Angew. Chem. Int. Ed. **38(24)**, 3743–3748.

Veber D. F., Johnson S. R., Cheng H. Y., Smith B. R., Ward K. W. and Kopple K. D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. J. Med. Chem. **45(12)**, 2615–2623.

Verdonk M. L., Cole J. C., Hartshorn M. J., Murray C. W. and Taylor R. D. 2003. Improved protein-ligand docking using GOLD. Proteins **52**, 609–623.

Verlinde C. L. M. J., Kim H., Bernstein B. E., Mande S. C. and Hol W. G. J. 1997. In Structure-based Drug Design, ed. P. Veerapandian, Marcel Dekker, New York pp. 365–394.

Watson P., Verdonk M. L. and Hartshorn M. J. 2003. A web-based platform for virtual screening. J. Mol. Graph. Model. **22**, 71–82.

Chapter 6

SCAFFOLD-BASED DRUG DISCOVERY

Kam Y. J. Zhang, Michael V. Milburn and Dean R. Artis
Plexxikon Inc., 91 Bolivar Dr., Berkeley, California 94710, USA

1. INTRODUCTION

The high attrition rate in modern drug discovery places a premium on the identification of high quality lead compounds (Milne, 2001). These compounds should not only possess sufficient potency and selectivity within the target profile, but importantly good pharmacological properties as well. The advent of genomic sciences has exacerbated this problem by revealing many potential drug targets and often close relationships between those within a protein family (Drews, 2000). The resulting 'target-rich, lead-poor' pipeline in drug discovery has stimulated many novel methods of lead identification as alternatives to the traditional drug discovery approach (Rees et al., 2004).

Traditional drug discovery often starts with the identification of potent lead compounds for a pre-selected protein target from a large library of drug-like chemicals using high-throughput screening (HTS). Typically, the most potent compounds are selected for chemical optimization toward the specific molecular target. Recent advances in combinatorial chemistry have greatly expanded the size of compound libraries to millions of compounds. Robotics technology and process automation have made the high-throughput screening of these large compound libraries more tractable than ever before. However, the dramatic increase in the amount of raw screening data has not resulted in a commensurate increase in the pace of new drug discovery (Drews, 2000). There are many reasons for this limited success. First, many combinatorial libraries, although containing a large number of compounds, can lack fundamental chemical diversity since they were derived from a limited, related set of starting materials and built around a common core.

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 129–153.

Secondly, the most potent compounds identified during initial screening may not necessarily possess suitable drug-like properties (Lipinski et al., 1997) and their chemical optimization may not be productive.

Fragment-based approaches have been proposed to enable the discovery of compounds from unexploited chemical space (Erlanson et al., 2004a; Rees et al., 2004). In general, these methods have utilized "fragments" with molecular weight (MW) lower than 150 Dalton (D), which tend to bind target proteins very weakly (in the mM range) and thus can not be detected by biochemical assay-based HTS methods. Instead, biophysical methods, such as NMR or X-ray, are used to screen a small collection of basic chemical building blocks and subsequently expand, merge, or link them to increase potency and selectivity (Erlanson et al., 2004b; Nienaber et al., 2000; Shuker et al., 1996). The "SAR by NMR" method introduced by Fesik and co-workers has ushered in a new era in fragment-based drug discovery (Shuker et al., 1996). This method has been applied to the discovery of lead compounds for many drug targets, such as FKBP, adenosine kinase, stromelysin, human papillomavirus E2 DNA binding protein and protein tyrosine phosphatase 1B. A similar "SAR by X-ray" approach that screens a fragment library by soaking a cocktail of 8-10 fragments into a preformed protein crystal was introduced (Nienaber et al., 2000). A new class of orally bioavailable urokinase inhibitors was discovered using this method. A dynamic combinatorial X-ray crystallography (DCX) screening method and a fragment discovery approach called Pyramid, has been proposed by Jhoti and colleagues (Carr and Jhoti, 2002; Hartshorn et al., 2005; Rees et al., 2004). The Pyramid method screens a mixture of small fragments by soaking them into a preformed target protein crystal and using automated software to analyze the binding of the fragments in the co-crystal structure. Pyramid has been successfully used for the identification of inhibitors for p38, CDK2, thrombin, ribonuclease A and PTP1B (Gill et al., 2005; Hartshorn et al., 2005). Taking advantage of the existence of free cysteines in the active site of a protein target, a site directed ligand discovery (tethering) method has been proposed that screens a small library of disulfide-containing fragments against a protein target with either a native or engineered cysteine in or near the active site (Erlanson et al., 2000). This technique was subsequently improved by an in situ assembly (extended tethering) method whereby a cysteine residue in or near the active site is covalently linked via disulfide bond to a low affinity small molecule and then this complex is used to screen against a library of disulfide-containing fragments (Erlanson et al., 2003). This tethering technique has been applied in the discovery of inhibitors for thymidylate synthase and caspase-3. In spite of the tremendous successes from the fragment-based approaches, the difficulty of identifying weakly-binding fragments and elaborating or linking them into high-affinity compounds remains a formidable challenge (Erlanson et al., 2004a).

Instead of using basic chemical building blocks, i.e., fragments, as starting points for drug discovery, we have sought to use chemical scaffolds which are significantly larger than "traditional fragments" and are much richer in functional groups that could form key interactions with the target protein - thus providing a more robust anchoring point for subsequent chemical optimization through substitution. Chemical scaffolds have been the bounty of drug discovery, with numerous examples throughout the rich history of medicinal chemistry. For example benzodiazepine derived compounds are central nervous system (CNS) depressants used to relieve anxiety or treat insomnia and the β-lactam family of compounds are widely used antibiotics (Gordon et al., 1994); the ubiquitous group of first generation aryl acetic acid non-steroidal anti-inflammatory drugs (Roberts and Morrow, 2001); quinazolines (Shewchuk et al., 2000) and oxindoles (Sun et al., 1998) are scaffolds for kinase inhibitors; benzyl thiazolidine diones as insulin sensitizers (Sternbach, 2003). Many PDE4 inhibitors, such as cilomilast and roflumilast, for the treatment of asthma and chronic obstructive pulmonary disease (COPD) are derived from the dialkoxyphenyl (catechol) scaffold (Card et al., 2004). The advantage of using scaffolds as starting points for lead optimization, particularly in approaching targets in the same protein family, is the greater efficiency with which potency can be achieved through the synthesis of fewer compounds. However, within any given family of targets, the collection of pharmacologically proven scaffolds that can be used for drug discovery is very limited.

In order to overcome the limited repertoire of new compounds that can serve as scaffolds for drug discovery, we have developed a robust and efficient strategy for the identification of weakly-active, intermediate molecular weight chemical scaffolds that can be further developed into drug candidates that specifically bind to and either inhibit or regulate the activity of given target proteins. This approach enables us to discover novel classes of bioactive compounds that would have been missed by traditional high throughput screening methods because of their low affinity. Furthermore, this approach allows the optimization of compounds by adding the least amount of molecular weight to reach a desired potency, thus improving the chances of having better drug like properties.

2. THE SCAFFOLD-BASED DRUG DISCOVERY PARADIGM

The paradigm for scaffold-based drug design uses a starting library of about 20,000 compounds with a molecular weight in the range of 125 to 350 Daltons and a combination of biochemical and high-throughput

co-crystallography as the primary screening method. The central difference between the scaffold-based approach and traditional high throughput screening methods is that the typical compounds identified in the initial screen are only weakly active at high concentration, and this activity is only used initially as a threshold to focus the compound selection for co-crystallography. However, this also represents a major distinction with respect to modern fragment-based screening approaches, where fragment bioactivity is usually un-measurable and thus peripheral to the initial selection process. In the Scaffold-based approach, only bioactive compounds yielding co-crystal structures with a target are pursued. The initial design and synthetic efforts are guided by both the bioactivity and the co-crystal structure of the low affinity compound in complex with the target protein (Table 6-1). In practice, the scaffold based drug discovery approach can be divided into three steps: scaffold identification, scaffold validation and chemical optimization (Figure 6-1).

Table 6-1. Comparison of fragment-based, scaffold-based and high-throughput screening-based approaches.

Technology Applied:	Fragment-Based	Scaffold-Based	HTS-Based
# of Compounds	100-2000	20,000-40,000	100,000-1,000,000+
MW (Dalton)	< 150	125-350	250-600
Biochemical Screening	No	Yes	Yes
Automated Screening	No	Yes	Yes
Co-crystallization	Yes/No (as mixtures, soaked or tethered)	Yes	No
Automated cocrystallization	Yes/No (as above)	Yes	No
Automated Structure Analysis	Yes	Yes	No
Hit Identification	Structure-Based	Bioactivity and Structure-Based (e.g. < 200 μM)	Bioactivity, Potency-Based (e.g. <1 μM)

The first step is to identify scaffold candidates using a low-affinity screening of an intermediate molecular weight compound library and using high throughput co-crystallography to reveal the molecular basis underlying

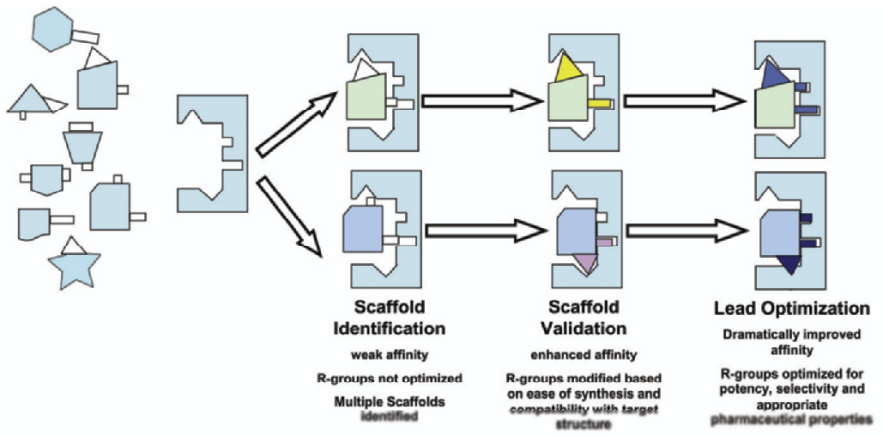


Figure 6-1. Scaffold-based drug discovery in a schematic illustration. First in scaffold identification, a library of scaffold-like compounds are screened by a biochemical assay against a family of closely related proteins and only those that hit multiple proteins at low-affinity are selected for further screening by co-crystallography. Second in scaffold validation, compounds derived from scaffold candidates with small substitutions are further screened by co-crystallography and only those that bind to the target protein with a conserved binding mode upon chemical modification are considered validated scaffold. Third in chemical optimization, the validated scaffold is turned into potent and selective inhibitors against multiple members of the protein family with predictable SAR and fewer compounds being synthesized.

the activity of the newly identified compounds. For this purpose, a core compound library has been designed to maximize the chance of discovering novel scaffolds by selecting a non-redundant and diverse set of intermediate molecular weight compounds from commercially available chemicals. The resulting co-crystal structures are analyzed and the putative scaffolds are examined for the number and types of interactions made with the target and the ease and likely productivity of substitutions to access different binding pockets within the target site.

The second step is to validate whether the newly identified compound can serve as a scaffold for further chemical optimization. For this purpose, a small number of derivatives of the scaffold candidate are synthesized, activities are measured and the structural basis of the activity and nature of the interactions with the target protein are analyzed by co-crystallography. A scaffold is considered validated when it forms stable interactions with the protein target, its binding mode is tolerant to small substitutions, and the SAR of the initial analogs are consistent with the crystallographically determined binding mode.

The third step is to optimize the validated scaffold into potent and selective inhibitors. Derivative compounds based on the new scaffold are

designed using the co-crystal structure of the scaffold as a guide and subsequently synthesized. Virtual libraries can be computationally enumerated using available reagents according to the synthetic schema. The binding poses of the compounds are predicted using scaffold-anchored docking based on the crystal structures, and can readily be further refined by molecular dynamics simulations. The compounds are then scored, and ranked according to their predicted ability to attain potency by forming favorable interactions with active site residues of the target protein.

The scaffold-based drug discovery paradigm produces multiple scaffolds for the lead optimization for a given protein target. This reduces the risk in lead generation and increases the probability of success in the preclinical and clinical development since multiple lead candidates derived from different chemical classes (scaffolds) can be used. This approach shortens the discovery time and requires the synthesis of fewer compounds to achieve a desired potency and selectivity profile. The same validated scaffold for a protein family can be optimized into selective inhibitors for a given target protein within the same family or developed into a pan-inhibitor against a subset of targets within the family. Several key components in the scaffold-based drug discovery process and their unique features will be described in more detail below.

2.1 Scaffold library construction and profiling

To facilitate the discovery of novel scaffolds and enable the use of biochemical screening tools, we have constructed a scaffold-library using the available compounds from 17 different chemical vendors. At the time the library was built, there were a total of almost two million compounds from the selected vendors. These compounds were analyzed to remove duplicates and to filter out these with reactive groups, then filtered to the molecular weight range of 120D-350D. This resulted in a collection of 275,555 compounds in the desired molecular weight range. These compounds were broken down into potential scaffold components – that is those smaller substructures within the molecule that might represent a new key binding motif for some prospective target. This was accomplished by fragmenting at rotatable bonds and thus dividing each compound into smaller substructures. The fragmentation yielded a total of 1,277,373 individual substructures, many of which were very similar to one another. These scaffold components were then clustered according to their chemical similarity into groups to diminish the effects of this redundancy. Compounds were then selected from across the component clusters to create a diverse set of putative scaffolds usable in a target independent fashion. A pruning

process was used to remove any compounds within a high similarity threshold (Tanimoto fingerprint similarity index >0.85) of one another within each cluster, since in theory one molecule could have been derived from a multiple scaffold component clusters. This novel approach has resulted in a core scaffold library consisting of 20,360 compounds, covering roughly 80% of the clustered scaffold component space, and composing the main library used for screening.

The resulting scaffold library contains compounds that possess a significant amount of "headroom" with respect to the drug-like properties based on the observations collectively referred to as Lipinski's rules (Lipinski et al., 1997) (Figure 6-2) and also fulfils the "Rule of Three" for fragment libraries (Congreve et al., 2003). Consequently, these compounds represent ideal starting points for an optimization program that will generally require some modest increases in molecular weight in order to achieve the potency and selectivity of the target profile.

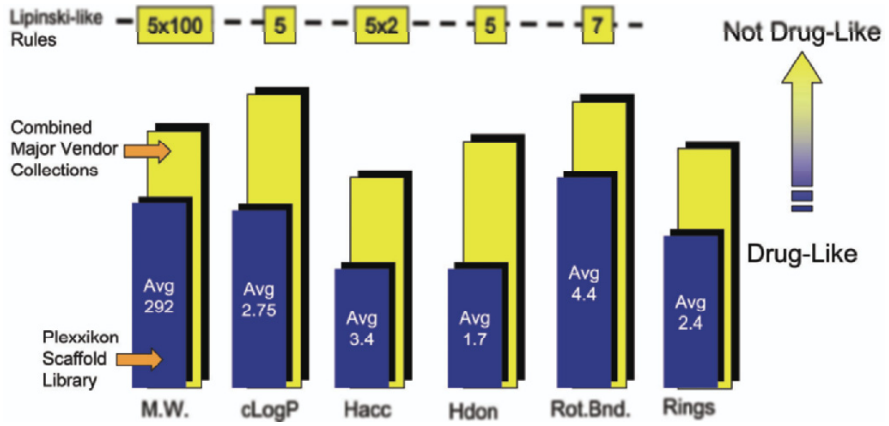


Figure 6-2. Physical properties of scaffold library. Several physical properties, molecular weight (M.W.), hydrophobicity (cLogP), hydrogen bond acceptors (Hacc), hydrogen bond donors (Hdon), number of rotatable bonds (Rot.Bnd.) and number of rings (Rings), are computed for our scaffold library and compared with that of the combined major vendor collections. The compounds in the scaffold library possess more favorable drug-like properties.

The average size of these scaffold-like compounds in the scaffold library is between 120-350 D, which is larger than the fragment-like basic chemical building blocks and smaller than the drug-like compounds in the traditional

HTS library (Figure 6-3). The relatively larger size of the compounds in the scaffold-library than the fragment library enabled the use of biochemical assays since these compounds tends to have functional groups. However, since they are significantly smaller than the drug-like molecules, their binding-affinity to the target protein tends to be weak and consequently higher concentrations of compounds have to be used in the assay.

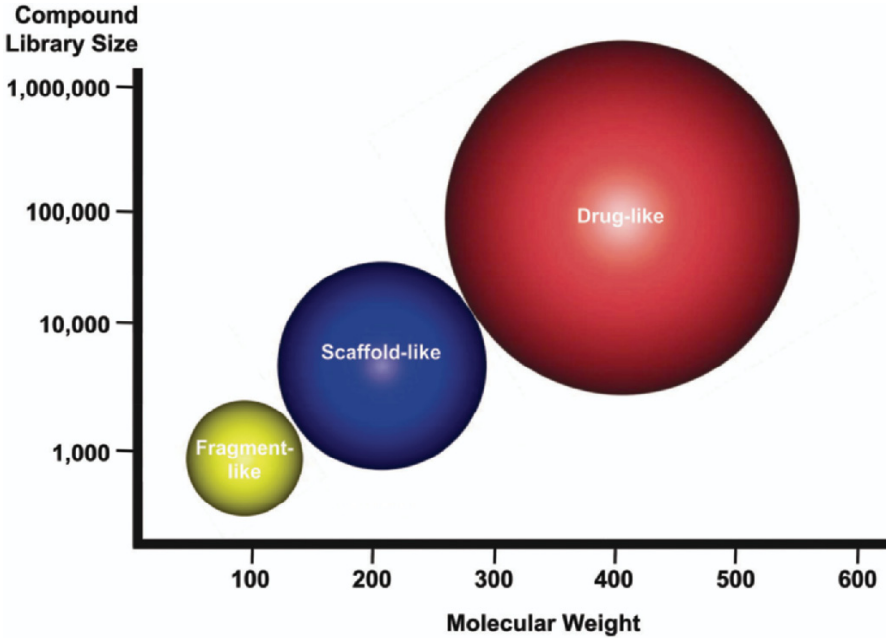


Figure 6-3. Comparing the size of scaffold library with fragment library and traditional compound library. Compounds in the scaffold library are on average larger than fragments and smaller than drug-like molecules. Consequently, the number of compounds in the scaffold library that are needed to represent a diversified chemical space is about one order of magnitude larger than the size of a fragment library and about one order of magnitude smaller than the size of a drug-like compound library.

2.2 Low-affinity biochemical screening as an initial filter

The compounds in the scaffold library can be detected by biochemical assays developed for traditional HTS since these compounds are larger than

basic chemical building blocks and possess functional groups to interact with protein targets. However, these compounds will generally not have all functional groups in ideal locations for any given binding site; some of the interactions will likely represent negative binding determinants. Consequently, high compound concentration (generally 100-200 μM) has to be used due to the relatively low binding affinity of these compounds. One benefit of the core scaffold library as described here has been the relatively high solubility of the compounds, due to the balance in the array of hydrophobic and polar functional groups in these compounds. Nevertheless, high compound concentration necessitates assays tolerant of modest DMSO concentration, and enhances the risk of false signal. One strategy that has proven to be of some assistance in dealing with the high false-positive rate associated with screening compounds at high concentration, has been to use several proteins within the target protein family as part of the scaffold screening process. Compounds that show low-affinity to a distribution of the selected family members are chosen for further screening by X-ray crystallography (Figure 6-4). The use of multiple members of the target family to select compounds has two advantages. The first is an increase in signal to noise ratio and a decrease in false positive rate. The second is that the selected compounds, if validated as a scaffold, can be used to generate lead compounds for multiple members of the target family and therefore increase the efficiency of overall lead generation.

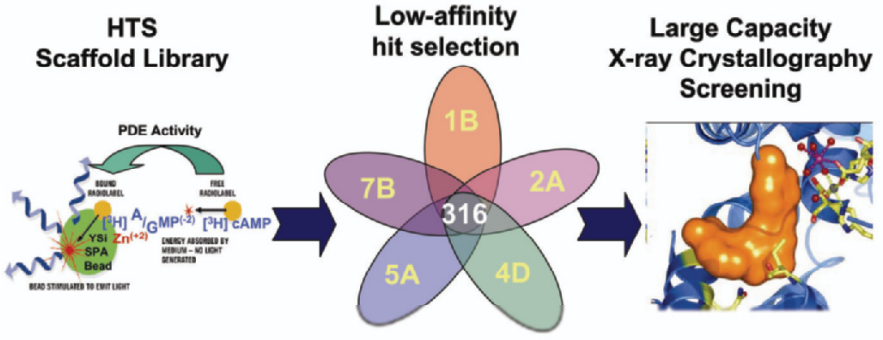


Figure 6-4. Low-affinity screening against multiple members of a protein family. Compounds in the scaffold library are screened against multiple members of a protein family using a biochemical assay in HTS format. For example in PDEs, a scintillation proximity assay is used to screen against PDE1B, PDE2A, PDE4D, PDE5A and PDE7A with compounds at 200 μM concentration. Only those 316 compounds that showed above 30% inhibition against at least three out of the five PDEs in the screening panel are selected for further screening by co-crystallography.

2.3 Automated high throughput co-crystallography as the second filter

The challenges to the efficiency of hit-to-lead optimization programs are significant, even when the focus is initially centered on relatively potent compounds (Gillepie and Goodnow, 2004). HTS approaches focused on low-affinity compounds must deal with a potential extreme prevalence of false positives. However, utilizing the great strides in robotics made in recent years, the true positive hits from the low-affinity biochemical screening can be very efficiently identified by high-throughput X-ray co-crystallography. Furthermore, because the scaffold candidate is validated through the determination of its detailed co-crystal structure with the target protein, the nature of the key interactions within the complex are immediately apparent, before any synthetic efforts have been committed. X-ray crystallography can provide a picture at atomic-level resolution of the interactions between small molecules and their protein targets. It has been traditionally used only in the lead optimization stage of structure-based drug discovery once potent lead compounds have been identified since it is a relatively time-consuming process (Verlinde and Hol, 1994). Technological advances in parallel cloning and expression optimization (Lesley, 2001), crystallization robotics (Stevens, 2000), automated data collection (Muchmore et al., 2000), automated phasing and model building (Adams and Grosse-Kunstleve, 2000; Holton and Alber, 2004) have recently enabled X-ray crystallography to be used as a primary screening technique in drug discovery (Blundell et al., 2002; Carr and Jhoti, 2002).

At the screening stage, co-crystallography allows the following key questions to be asked: Where does a compound bind to the protein? Does the compound make specific interactions with the target? Are appropriate R-group substitutions available? Does the compound bind as such that chemistry can be productive? Are multiple sites for substitution available? Would the chemistry lead to novel compounds? Based on the answers to these questions, the starting points that represent the highest probability of success can be selected from a large and diverse group.

Our high-throughput co-crystallography platform is illustrated in Figure 6-5. Within the past three years, more than 6019 crystals have been grown and mounted; 1445 X-ray diffraction datasets have been collected; 1326 co-crystal structures have been determined with an overall hit-rate of 38%. Our current run-rate is about 450 co-crystal structures annually, sufficient to drive the current scaffold discovery and lead generation efforts at Plexxikon. This vast array of co-crystal structure data has been analyzed computationally to guide our medicinal chemistry.

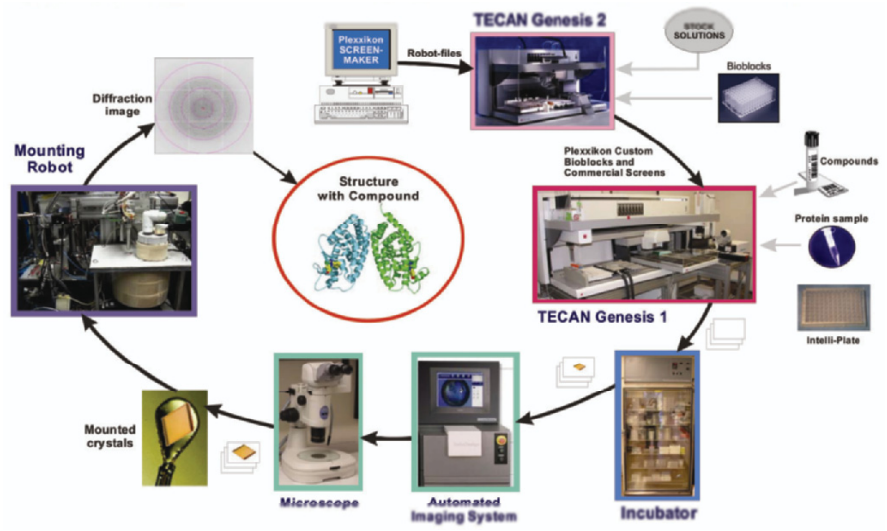


Figure 6-5. High throughput crystallography platform. Crystallization buffers (either random or grid screens) are made by a robot starting from stock solutions driven by a web-based software. Crystallization plates are setup automatically by a robot. Plates are examined by an automated imaging system. Diffraction data are collected at synchrotron facilities with mounting robots and automation control software. The first electron density for the bound small molecule is generated without human intervention from the diffraction dataset using automation scripts. The inhibitor compound is built into the electron density manually with some help from automated ligand fitting software. The structure refinement of the inhibitor-protein complex is completely automated.

2.4 Computational approaches for scaffold validation and lead optimization

In general, the keys to the success in scaffold-based drug discovery lie in the diversity of the data focused on a given target. The richness of the biological and structural data generated by the platform described above, in combination, enable an "information-rich" approach to the process of scaffold validation and subsequent lead optimization. With the availability of multiple co-crystal structures of diverse compounds in a given active site, many of the key interactions available within the site are visible in different contexts. The process of pursuing any individual scaffold then becomes one of optimizing an already fairly robust signal, rather than one of faint signal detection. Consequently, an efficient informatics approach should leverage as much as possible from the experimental data into the computational efforts, as opposed to more speculative techniques such as naïve docking or *de novo* design. In this context, informatics can help provide a framework in

which biological, structural and chemical ideas can be integrated and turned to highlighting the best opportunities for lead generation.

One key element of our informatics approach lies in the analysis of related binding sites for areas of potential localized binding energy. Over many years, a large number of different techniques have been used for this purpose, including such diverse approaches as GRID (Goodford, 1985) and SiteID (Tripos Associates, St. Louis, MO). Based on the crystal structure of the target protein with and without scaffold candidates (Figure 6-6) different types of properties, energetic or topological, can be calculated and mapped to a reference grid that can allow for the comparison of many different proteins. For instance, an interaction energy map can be generated with GRID by running a chemical probe around the protein surface in the active site. Similar interaction maps for some homologous proteins are also generated if their co-crystal structures with scaffold candidates are available. A consensus interaction map is created based on these individual interaction maps and thus revealing the common space for scaffold binding. Clustering

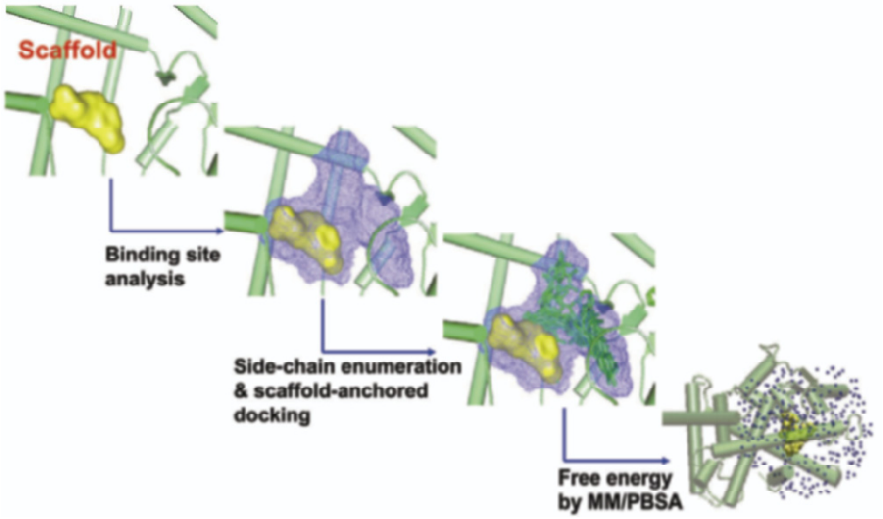


Figure 6-6. Scaffold-guided computational design. Starting from a co-crystal structure of the target (or surrogate) protein with a scaffold candidate, the inhibitor binding site on the target protein (usually the active site) is analyzed by GRID to generate an interaction energy surface. Then into this cavity, derivative compounds with enumerated substitutions on the scaffold candidate generated according to synthetic schema are docked into the binding site with the scaffold piece anchored. For a more accurate evaluation of the predicted derivative compound binding mode, a free energy calculation is performed using MM/PBSA method with explicit water molecules modeled.

of the multiplicity of grids for different chemical probes across members of a protein family can highlight interesting similarities and differences, which in turn represent tremendous potential opportunities for enhancing potency and selectivity during the various stages of scaffold-based discovery.

A small compound library can subsequently be enumerated with substitutions in the appropriate sites on the scaffold candidate and the resulting compound evaluated in the context of the various consensus interaction maps. For the purpose of this evaluation, the structures can be docked into the active site. In these cases, a proprietary fitting/minimization algorithm relying on the crystal structure of the scaffold as the anchor is generally used. The use of scaffold as an anchor in docking has greatly reduced the degrees of freedom and increased efficiency and accuracy of docking. As a crude visual measure of compound fitness, this type of approach is often sufficient.

For more detailed analysis, a measure of binding energy can be calculated for a selected subset of compounds. For these kinds of calculations we have relied on molecular dynamics simulations using AMBER7 (Case et al., 2002) and PARM94 force field (Cornell et al., 1995), with atomic partial charges of compounds generated semi-empirically using MOPAC (Stewart, 1985) through the Antechamber program. The complex binding free energy (ΔG) has been estimated using the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) methodology (Srinivasan et al., 1998). Generally, the structures from scaffold-guided docking can be used as the starting points for molecular dynamics simulations with some minor modifications, which allows for a significant amount of automation of the setup process for each simulation. The simulations are run in the presence of explicit water molecules, and an algorithm-directed selection from these waters is utilized as part of the MM/PBSA calculation. Use of this methodology has resulted in reasonable correlations between experimental and computed measures of binding, and these calculations have been used prospectively to evaluate synthetic candidates prior to synthesis. Such practical and very encouraging results have been obtained for diverse groups of compounds interacting with proteins in several different target families. In a little over half of the cases examined to date in-house, it has been possible to obtain a model which has proved useful for evaluating prospective analogs quantitatively (generally, with $r2 \geq 0.6$). In the remaining cases, the results from these simulations have ranged from qualitatively useful to substantively in error. Consequently, an important part of the initial target evaluation and analysis consists of benchmarking the system using this methodology to determine whether quantitative activity prediction can be achieved.

2.5 Synthetic chemistry based on validated scaffold increases efficiency

Medicinal chemistry is the most resource intensive step in the entire "hit-to-lead" process in the early stage of drug discovery. The traditional approach to lead optimization is driven by the structure-activity relationship (SAR) of potent lead candidates identified from HTS. Sometimes, the SAR can be confusing or even misleading for a variety of reasons. As a result, chemistry resources, inevitably in short supply, can be used to very unproductive ends. The SAR based optimization assumes the core piece in the lead candidate has a stable binding mode to the target protein and various substitutions off the same site on the core piece are probing the same location in the target protein. Only when this assumption holds true, can the SAR be used to guide the optimization process. When this assumption is invalid - for instance, when the binding mode of a lead series changes significantly with a particular substitution - the efficiency of the optimization process can be dramatically impaired.

The introduction of scaffold validation step into the "hit-to-lead" process is aimed at addressing the above problem and increasing the efficiency in lead optimization. The validated scaffold forms key interactions with the target protein and its binding mode is tolerant to substitutions. Therefore, lead optimization starting from a validated scaffold tends to yield a more predictable SAR and to take a shorter time to achieve desired potency improvement through the synthesis of fewer compounds. Various substitutions at different sites of the scaffold can be combined to achieve an additive SAR effect thereby greatly increasing the efficiency of optimization (Figure 6-7).

However, the most dramatic impact on the direction and focus of medicinal chemistry is really derived from the information-rich context provided by the rest of the platform. The "chemical intuition" in lead optimization can function at a different level in this environment. New synthetic avenues can be considered in the light of ease of chemistry to achieve a specific structural hypothesis, rather than as an end unto synthesis alone. This tends to focus progress in chemistry on efforts that are more likely to be both tractable and successful. The information content of each compound is magnified under these circumstances, as is the importance of each result.

This observation has been connected with some interesting efficiencies of process that have emerged. As projects progress, the need for crystallography in the lead series tends to dramatically ramp down, as more and more about the target is understood. This frees up crystallization

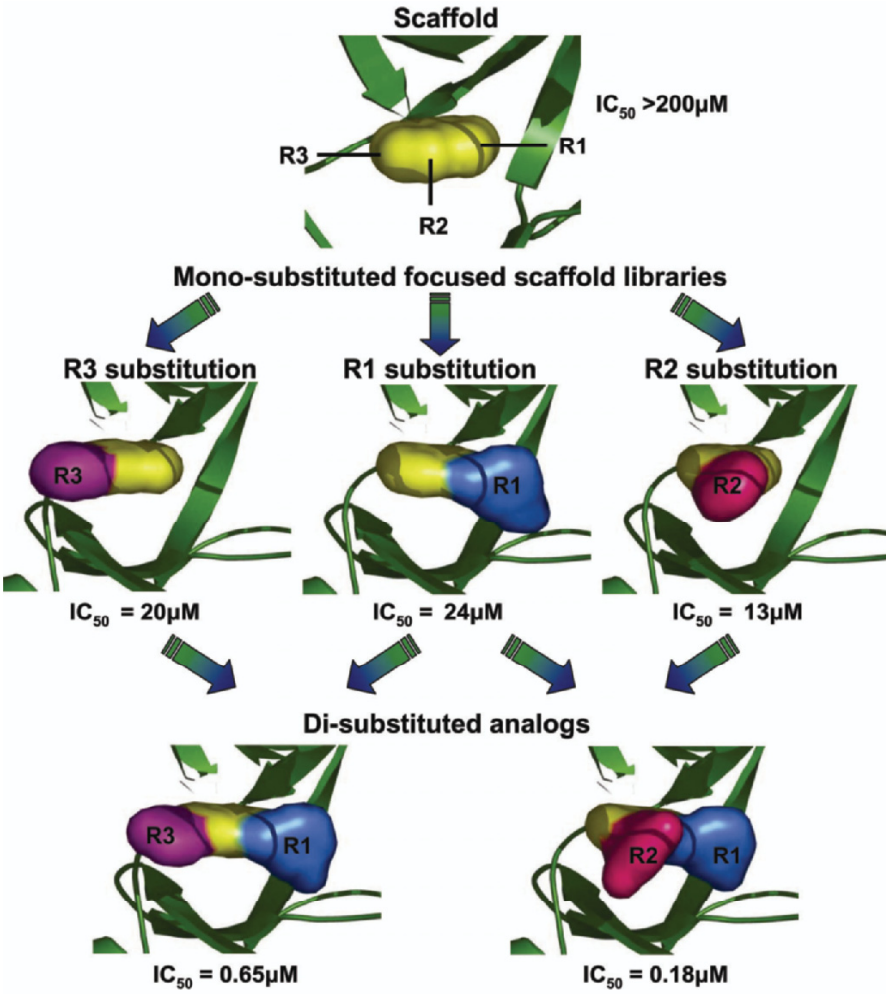


Figure 6-7. Scaffold-based chemical optimization is highly efficient. The scaffold validation process first identifies feasible substitution sites on the compound based on its co-crystal structure with the protein target (such as R1, R2 and R3 in the illustration). The co-crystal structures of these mono-substituted compounds and their IC50s validates the scaffold since its binding mode is tolerant to these small substitutions. Consequently, these mono-substituted compounds become great starting points for further chemical optimization of the scaffold into more potent inhibitors. A simple combination of the mono-substituted compounds (at R1, R2 or R3 positions) produces di-substituted compounds with predictable increases in potency. This is made possible by the validated scaffold that anchors the mono- and di-substituted compounds in the same binding orientation in the target protein binding site.

resources to impact projects through the identification of more novel opportunities. When quantitative activity predictions are part of an optimization process, it is often the outlier compounds (whether for better or worse) that offer some of the most interesting opportunities to advance a given project. This process helps focus crystallography on the compounds most likely to provide new insights.

One additional unexpected benefit of scaffold-based drug discovery has become apparent through evaluation of the pharmaceutical properties of several different series of compounds *in vivo*. In the typical process, the architecture of a new lead series is developed relatively early in the optimization process, when potency is still at a relatively modest level (Figure 6-8). At this point, the major changes to the framework of the lead series are in place, and the compounds are not very far from their final molecular weight. The subsequent steps of tuning potency and selectivity are then more akin to engineering, rather than architecture. As a result, the pharmaceutical potential of the lead series as a whole becomes evident earlier in the optimization process than might normally be expected. This has allowed for an earlier entry into evaluation for pharmacokinetic characteristics, and a more ready incorporation of these factors into the lead generation process.

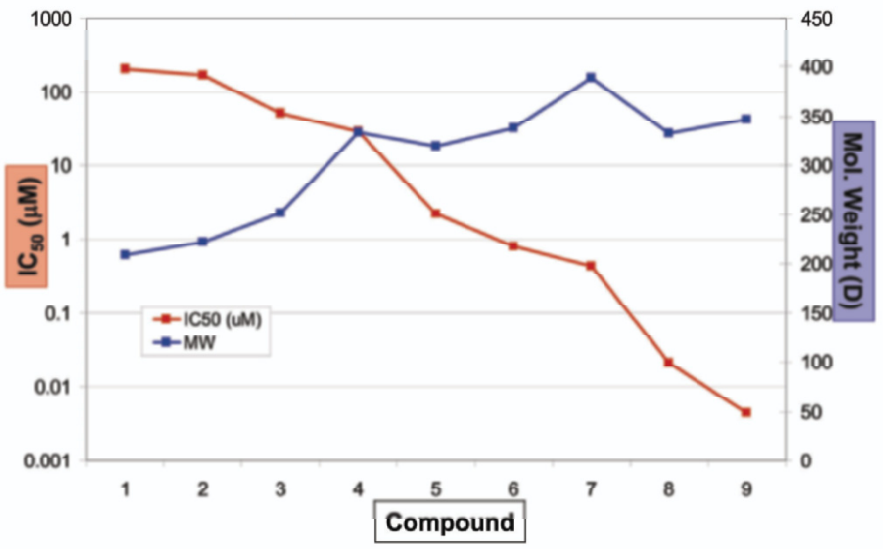


Figure 6-8. Potency improvement versus molecular weight gain. In a typical scaffold-based drug discovery process, a substantial molecular weight gain is incurred to reach modest potency increase at the scaffold validation stage (compounds #1-#4). Whereas a substantial potency improvement has resulted from slight molecular weight increase in the chemical optimization stage (compounds #5-#9).

3. THE DISCOVERY OF POTENT PDE4 INHIBITORS USING SCAFFOLD-BASED DRUG DISCOVERY

Cyclic nucleotide phosphodiesterases (PDEs) are enzymes that play a critical role in maintaining the cellular level of cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) (Beavo, 1995; Conti and Jin, 1999; Houslay, 1998), and regulating a myriad of physiological processes (Francis et al., 2001). There are 11 subfamilies of human PDEs, and the cAMP-selective PDE4 subfamily is comprised of 4 members (Conti et al., 2003; Houslay and Adams, 2003). PDE4B is of particular importance in the inflammatory responses of lymphocytes. Therefore, PDE4 is a good target for developing drugs treating asthma and chronic pulmonary obstructive disease (COPD). One of the shortcomings of the PDE4 inhibitors currently in clinical development is a narrow therapeutic window between efficacy and side effects (e.g. emesis). The side effects may be associated with the specific chemotypes and therefore, there is an urgent need to identify new classes of PDE4 inhibitors that exhibit less emetic effect at the effective therapeutic dose.

We have applied the scaffold-based drug discovery approach to identify novel scaffolds that can be further optimized into potent PDE4 inhibitors (Card et al., 2005). We first screened the core library of about 20,000 compounds against a representative subset of PDE family members, PDE1B, PDE2A, PDE4D, PDE5A and PDE7B, using a high throughput scintillation proximity assay (SPA) (Bardelle et al., 1999). A total of 316 compounds showed greater than 30% inhibition at 200 μM for three or more PDEs in the screening panel. These 316 compounds were all set up for co-crystallization with PDE4D and PDE4B; of those, 269 compounds were co-crystallized and 107 co-crystal structures were solved. A low affinity 3,5-dimethyl-1 H-pyrazole-4-carboxylic acid ethyl ester (PCEE, MW=168D, IC50 = 82 μM for PDE4D) revealed the characteristic features of a potential scaffold binding to PDE4D (Card et al., 2004; Zhang et al., 2004): the pyrazole ring is sandwiched in between the hydrophobic clamp formed by residues $F372^{4D}$ and $I336^{4D}$ (where the superscript identifies the protein to which the residue number refers to); the carboxylate of the pyrazole is hydrogen bonded to the invariant purine-selective $Q369^{4D}$ (Figure 6-9a).

To determine whether PCEE could serve as a scaffold for PDE4, the co-crystal structure of this scaffold candidate with PDE4D was analyzed and three potential sites of substitution were identified based on the ability of making favorable chemical interactions in the available space at the active site. Consequently, a small set of compounds with substitutions at the 1-, 3- or 5-positions of the pyrazole were synthesized and tested in the PDE assay.

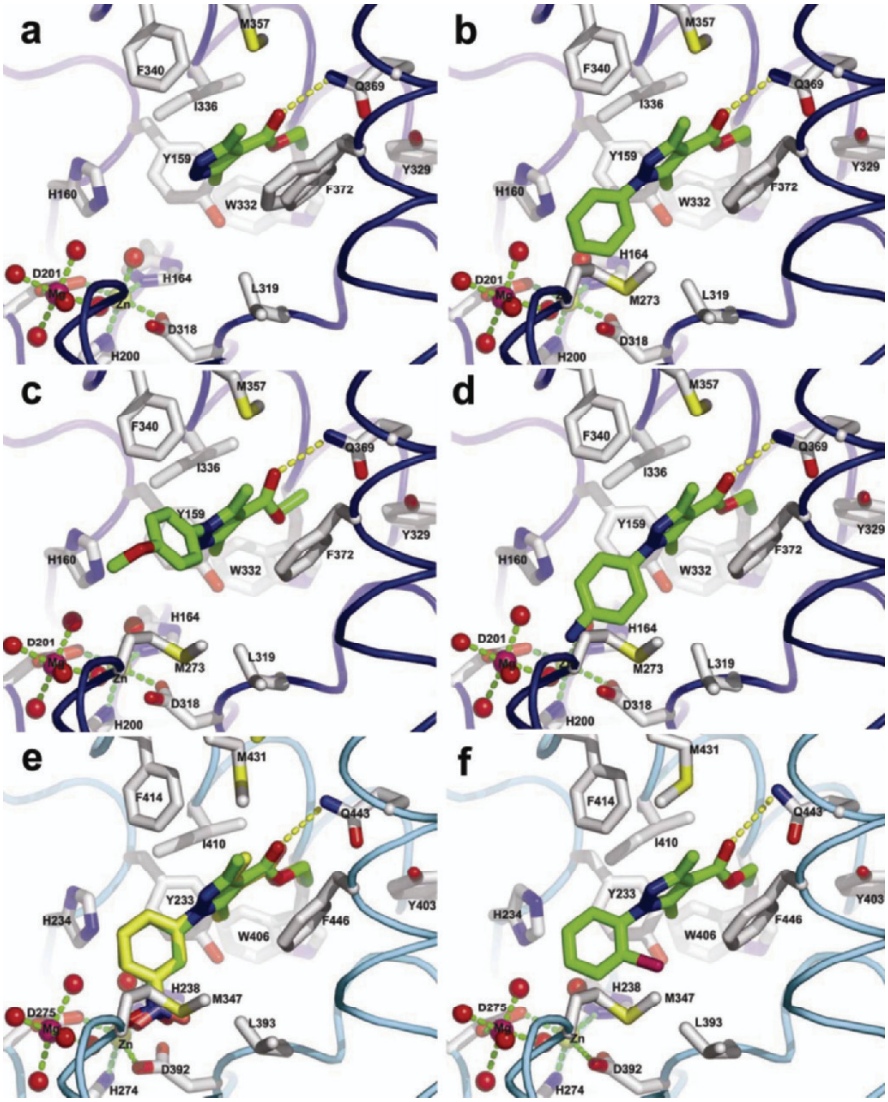


Figure 6-9. Crystal structures of the pyrazole scaffold and its derivatives in complex with PDE4B or PDE4D.

(a) Crystal structure of 3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4D, showing the pyrazole ring sandwiched in the hydrophobic clamp formed by F372 and I336. The conserved H-bond, seen in all pyrazole derivative co-crystal structures, between the NE2 atom of the invariant glutamine and the carboxylate group, is shown.

(b) The crystal structure of 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4D, showing the same interactions as its parent compound, and thus validating

the dimethyl pyrazole as a scaffold. The dimethyl pyrazole is sandwiched by F372 and I336 and the carbonyl oxygen forms an H-bond with Q369. The ethoxy group is tucked into the Q1 pocket.

(c) Crystal structure of 1-(4-methoxy-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4D. The methoxy-phenyl group rotated 180° degree to point away from the di-metal ions to avoid the repulsive interactions between the methoxy group and the di-metal ions.

(d) Crystal structure of 1-(4-amino-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4D. The amine group forms three H-bonds with three water molecules, two of which are coordinated to Mg^{2+}. However, this amine nitrogen is also in close proximity to the carbon atom in M273 which results in unfavourable interactions.

(e) Crystal structure of 3,5-dimethyl-1-(3-nitro-phenyl)-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4B and PDE4D. Only the backbone tube and the side chains from PDE4B are shown for clarity. The carbon atoms of pyrazole bound to PDE4B and PDE4D are shown in green and yellow respectively.

(**f**) Crystal structure of 1-(2-chloro-phenyl)-3,5-dimethyl-1H-pyrazole-4-carboxylic acid ethyl ester bound to PDE4B. The Cl-substitution at the ortho-position of the phenyl ring makes several hydrophobic contacts with residues M347, L393 and F446.

Comparison of various aryl substitutions at R1, R2 and R3 positions indicated that the phenyl substitution at R1 position is most potent. Therefore, we selected 3,5-dimethyl-1-phenyl-1H-pyrazole-4-carboxylic acid ethyl ester (PhPCEE, MW=244D) as a derivative scaffold for further optimization of potency.

We determined the co-crystal structure of PhPCEE in complex with PDE4D. This co-crystal structure has revealed that the phenyl substitution at the 1-position of the pyrazole ring does not change the binding mode of the pyrazole in PDE4, thus validating the PCEE moiety as a scaffold for PDE4 (Figure 6-9b). Furthermore, the 1-phenyl substitution gains several additional hydrophobic interactions with the conserved residues $H160^{4D}$, $M273^{4D}$ and $L319^{4D}$ in the active site pocket. These interactions resulted in the significantly increased potency of this compound in PDE4B (IC50 = 270 nM) from that of its parent compound (IC50 = 60 μM).

Because of the improved potency and the conserved binding mode of the PCEE scaffold in PhPCEE (Figures 6-9a and 6-9b), we chose PhPCEE as the new starting point for lead optimization. Based on the available reagent pool, more than 100 compounds were designed *in silico*. These compounds were docked into the PDE4B active site pocket with the scaffold part anchored at the observed binding pose. The binding energy was also estimated using the MM-PBSA method (Case et al., 2002; Srinivasan et al., 1998). After eliminating a large number of substitutions that were predicted to cause

undesirable interactions with residues in the active site, a total of 10 compounds were synthesized with a predicted increase in binding affinity due to the formation of favorable interactions.

We determined the co-crystal structures of four of the 10 synthesized molecules (Figures 6-9c-f): one with decreased potency, one with similar potency, and two with increased potency, when compared to the parent compound PhPCEE. The substitution of an OCH_3 group (MW=274D) has caused the methoxy phenyl group to flip in order to avoid close contacts with the di-metal ions (Figure 6-9c). Consequently, this compound has lost binding affinity by about 10 fold compared to its parent compound with an IC50 of 5.9 μM. A substitution at the para-position by a NH_2 group (MW=259D) is close to and pointing at the di-metal ions (Figure 6-9d) has hardly improved the binding affinity. In contrast, the pyrazole with a NO_2-substitution at the meta-position of the phenyl ring (MW=289D) formed H-bonds with $T345^{4B}$, $D392^{4B}$ and the two water molecules coordinating the Zn^{2+} (Figure 6-9e). The IC50 of this compound has decreased to 33 nM from that of 310 nM for its parent compound. Similarly, the Cl-substitution (MW=278D) at the ortho-position makes several hydrophobic contacts with residues $M347^{4B}$, $L393^{4B}$ and $F446^{4B}$ (Figure 6-9f). These extra hydrophobic interactions have increased the binding affinity about 10 fold from its parent compound to an IC50 of 56 nM.

These co-crystal structures of PhPCEE derivatives have proven that the computational design of compounds and the prediction of SAR to be accurate. The high accuracy of these computational predictions is mainly due to the use of the validated scaffold as a starting point and the availability of its co-crystal structure with the target protein.

4. SUMMARY

The power of the scaffold based lead discovery platform has been demonstrated through the identification and optimization of the pyrazoles as PDE inhibitors. The binding mode of the pyrazole carboxylic ester scaffold does not change when substituents are added. This predictable binding mode anchored by the low-affinity pyrazole carboxylic ester scaffold has greatly increased the efficiency of chemical optimization. All-in-all, a 4000-fold increase in potency was achieved starting from the initial pyrazole carboxylic ester scaffold (IC50 82 μM in PDE4D) to that of the 2-chlorophenyl and 3-nitrophenyl pyrazoles (IC50 20 nM in PDE4D) in two rounds of chemical synthesis and only a total of 21 compounds had to be synthesized. The total molecular weight added is only 121 Daltons, which

has resulted in a 4000-fold potency increase. Due to the relatively small size of the compound, additional substitutions can be made to further exploit many available areas of the active site to achieve higher potency and selectivity.

There are a number of unique features that differentiate our scaffold-based lead discovery approach from the fragment-based lead discovery approaches (Erlanson et al., 2004a; Erlanson et al., 2004b; Hartshorn et al., 2005; Nienaber et al., 2000; Rees et al., 2004; Shuker et al., 1996). First, the scaffold library contains compounds that are significantly larger than the basic building blocks with an average molecular weight of about 250D. This has increased the size of the scaffold library to about 20,000 compounds. Furthermore, these scaffold-like compounds contain functional groups that enable them to bind to the protein target at an affinity that could be detected by HTS method. To reduce the high false positive rate associated with low-affinity biochemical screening, we include other proteins closely related to the target protein (i.e. proteins from the same family) in the screening and only select compounds that show activity against multiple members from the screening panel. Finally, the scaffold validation step ensures that the binding mode is indifferent to substitution. This leads to more predictable SAR and more efficient chemistry in lead optimization.

In conclusion, we have developed a strategy that enables rapid and efficient design of novel leads for a given target protein. We have applied our scaffold-based lead discovery platform to rapidly discover the pyrazole carboxylic ester scaffold and to optimize it into potent PDE4 inhibitors. The robustness and efficiency of the scaffold based lead discovery method should make it widely applicable to expedite the lead discovery effort for many other targets for which known small molecule modulators are limited.

5. ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Peter Hirth and Prof. Joseph Schlessinger and Prof. Sung-Hou Kim for the formative vision that led to the scaffold-based drug discovery paradigm. We would also like to thank our many colleagues at Plexxikon for their fundamental contributions in building and implementing this paradigm. In particular, Gideon Bollag, Prabha Ibrahim, Chao Zhang, Sam Gillette, Ryan Bremer, Yong-Liang Zhu, Fernando Martin, Jim Arnold, Gaston Habets, Bruce England, Heike Krupka, Graeme Card, Weiru Wang and Abhinav Kumar were instrumental in aspects of the work presented here.

REFERENCES

Adams P. D. and Grosse-Kunstleve R. W. 2000. Recent developments in software for the automation of crystallographic macromolecular structure determination. Curr. Opin. Struct. Biol. **10**, 564–568.

Bardelle C., Smales C., Ito M., Nomoto K., Wong E. Y., Kato H., Saeki T. and Staddon J. M. 1999. Phosphodiesterase 4 conformers: Preparation of recombinant enzymes and assay for inhibitors. Anal. Biochem. **275**, 148–155.

Beavo J. A. 1995. Cyclic nucleotide phosphodiesterases: Functional implications of multiple isoforms. Physiol. Rev. **75**, 725–748.

Blundell T. L., Jhoti H. and Abell C. 2002. High-throughput crystallography for lead discovery in drug design. Nat. Rev. Drug Discov. **1**, 45–54.

Card G. L., Blasdel L., England B. P., Zhang C., Suzuki Y., Gillette S., Fong D., Ibrahim P. N., Artis D. R., Bollag G. et al. 2005. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. Nat. Biotechnol. **23**, 201–207.

Card G. L., England B. P., Suzuki Y., Fong D., Powell B., Lee B., Luu C., Tabrizizad M., Gillette S., Ibrahim P. N. et al. 2004. Structural Basis for the Activity of Drugs that Inhibit Phosphodiesterases. Structure (Camb) **12**, 2233–2247.

Carr R. and Jhoti H. 2002. Structure-based screening of low-affinity compounds. Drug Discovery Today **7**, 522–527.

Case D. A., Pearlman D. A., Caldwell J. W., Cheatham T. E., Wang J., Ross W. S., Simmerling C., Darden T., Merz K. M., Stanton R. V. et al. 2002. AMBER 7. San Francisco, University of California.

Congreve M., Carr R., Murray C. and Jhoti H. 2003. A 'rule of three' for fragment-based lead discovery? Drug Discovery Today **8**, 876–877.

Conti M. and Jin S. L. 1999. The molecular biology of cyclic nucleotide phosphodiesterases. Prog. Nucleic Acid Res. Mol. Biol. **63**, 1–38.

Conti M., Richter W., Mehats C., Livera G., Park J. Y. and Jin C. 2003. Cyclic AMP-specific PDE4 phosphodiesterases as critical components of cyclic AMP signaling. J. Biol. Chem. **278**, 5493–5496.

Cornell W. D., Cieplak P., Bayly C. I., Gould I. R., Merz J. K. M., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. W. and Kollman P. A. 1995. A second generation force field for the simulation of proteins, nucleic acid, and organic molecules. J. Am. Chem. Soc. **117**, 5179–5197.

Drews J. 2000. Drug discovery: A historical perspective. Science **287**, 1960–1964.

Erlanson D. A., Braisted A. C., Raphael D. R., Randal M., Stroud R. M., Gordon E. M. and Wells J. A. 2000. Site-directed ligand discovery. Proc. Natl. Acad. Sci. U. S. A. **97**, 9367–9372.

Erlanson D. A., Lam J. W., Wiesmann C., Luong T. N., Simmons R. L., DeLano W. L., Choong I. C., Burdett M. T., Flanagan W. M., Lee D. et al. 2003. In situ assembly of enzyme inhibitors using extended tethering. Nat. Biotechnol. **21**, 308–314.

Erlanson D. A., McDowell R. S. and O'Brien T. 2004a. Fragment-based drug discovery. J. Med. Chem. **47**, 3463–3482.

Erlanson D. A., Wells J. A. and Braisted A. C. 2004b. Tethering: Fragment-based drug discovery. Annu. Rev. Biophys. Biomol. Struct. **33**, 199–223.

Francis S. H., Turko I. V. and Corbin J. D. 2001. Cyclic nucleotide phosphodiesterases: Relating structure and function. Prog. Nucleic Acid Res. Mol. Biol. **65**, 1–52.

Gill A. L., Frederickson M., Cleasby A., Woodhead S. J., Carr M. G., Woodhead A. J., Walker M. T., Congreve M. S., Devine L. A., Tisi D. et al. 2005. Identification of Novel p38alpha MAP Kinase Inhibitors Using Fragment-Based Lead Generation. J. Med. Chem. **48**, 414–426.

Gillepie P. and Goodnow R. A., Jr. 2004. The Hit-to-Lead Process in Drug Discovery. Annual Reports in Medicinal Chemistry **39**, 293–302.

Goodford P. J. 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J. Med. Chem. **28**, 849–857.

Gordon E. M., Barrett R. W., Dower W. J., Fodor S. P. and Gallop M. A. 1994. Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. J. Med. Chem. **37**, 1385–1401.

Hajduk P. J., Boyd S., Nettesheim D., Nienaber V., Severin J., Smith R., Davidson D., Rockway T. and Fesik S. W. 2000. Identification of novel inhibitors of urokinase via NMR-based screening. J. Med. Chem. **43**, 3862–3866.

Hajduk P. J., Dinges J., Miknis G. F., Merlock M., Middleton T., Kempf D. J., Egan D. A., Walter K. A., Robins T. S., Shuker S. B. et al. 1997. NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. J. Med. Chem. **40**, 3144–3150.

Hartshorn M. J., Murray C. W., Cleasby A., Frederickson M., Tickle I. J. and Jhoti H. 2005. Fragment-Based Lead Discovery Using X-ray Crystallography. J. Med. Chem. **48**, 403–413.

Holton J. and Alber T. 2004. From the Cover: Automated protein crystal structure determination using ELVES. Proc. Natl. Acad. Sci. U. S. A. **101**, 1537–1542.

Houslay M. D. 1998. Adaptation in cyclic AMP signalling processes: A central role for cyclic AMP phosphodiesterases. Semin. Cell Dev. Biol. **9**, 161–167.

Houslay M. D. and Adams D. R. 2003. PDE4 cAMP phosphodiesterases: Modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization. Biochemistry J **370**, 1–18.

Lesley S. A. 2001. High-throughput proteomics: Protein expression and purification in the postgenomic world. Protein Expr. Purif. **22**, 159–164.

Lipinski C. A., Lombardo F., Dominy B. W. and Feeney P. J. 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **23**, 3–25.

Liu G., Xin Z., Pei Z., Hajduk P. J., Abad-Zapatero C., Hutchins C. W., Zhao H., Lubben T. H., Ballaron S. J., Haasch D. L. et al. 2003. Fragment screening and assembly: A highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor. J. Med. Chem. **46**, 4232–4235.

Milne G. M. 2001. Accelerating R&D productivity. Paper presented at: IBC's Drug Discovery Technology World Congress. Boston, MA, USA.

Muchmore S. W., Olson J., Jones R., Pan J., Blum M., Greer J., Merrick S. M., Magdalinos P. and Nienaber V. L. 2000. Automated crystal mounting and data collection for protein crystallography. Structure Fold. Des. **8**, R243–246.

Nienaber V. L., Richardson P. L., Klighofer V., Bouska J. J., Giranda V. L. and Greer J. 2000. Discovering novel ligands for macromolecules using X-ray crystallographic screening. Nat. Biotechnol. **18**, 1105–1108.

Rees D. C., Congreve M., Murray C. W. and Carr R. 2004. Fragment-based lead discovery. Nat. Rev. Drug Discov. **3**, 660–672.

Roberts L. J. and Morrow J. D. 2001. Analgesic-Antipyretic and Anti-inflammatory Agents. In Goodman and Gilman's 'The Pharmacological Basis of Therapeutics'. Hardman J. G., Limbird L. E. and Gilman A. G. (eds) (New York, McGraw-Hill) pp. 710–713.

Shewchuk L., Hassell A., Wisely B., Rocque W., Holmes W., Veal J. and Kuyper L. F. 2000. Binding mode of the 4-anilinoquinazoline class of protein kinase inhibitor: X-ray crystallographic studies of 4-anilinoquinazolines bound to cyclin-dependent kinase 2 and p38 kinase. J. Med. Chem. **43**, 133–138.

Shuker S. B., Hajduk P. J., Meadows R. P. and Fesik S. W. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. Science **274**, 1531–1534.

Srinivasan J., Cheatham T. E., Cieplak P., Kollman P. A. and Case D. A. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J. Am. Chem. Soc. **120**, 9401–9409.

Sternbach D. D. 2003. Modulators of peroxisome proliferator-activated receptors (PPARs). Annual Reports in Medicinal Chemistry **38**, 71–80.

Stevens R. C. 2000. High-throughput protein crystallization. Curr. Opin. Struct. Biol. **10**, 558–563.

Stewart J. J. P. 1985. MOPAC Manual. Colorado Springs, CO, United States Air Force Academy.

Sun L., Tran N., Tang F., App H., Hirth P., McMahon G. and Tang C. 1998. Synthesis and biological evaluations of 3-substituted indolin-2-ones: A novel class of tyrosine kinase inhibitors that exhibit selectivity toward particular receptor tyrosine kinases. J. Med. Chem. **41**, 2588–2603.

Szczepankiewicz B. G., Liu G., Hajduk P. J., Abad-Zapatero C., Pei Z., Xin Z., Lubben T. H., Trevillyan J. M., Stashko M. A., Ballaron S. J. et al. 2003. Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. J. Am. Chem. Soc. **125**, 4087–4096.

Verlinde C. L. and Hol W. G. 1994. Structure-based drug design: Progress, results and challenges. Structure **2**, 577–587.

Zhang K. Y. J., Card G. L., Suzuki Y., Artis D. R., Fong D., Gillette S., Hsieh D., Neiman J., West B. L., Zhang C. et al. 2004. A Glutamine Switch Mechanism for Nucleotide Selectivity by Phosphodiesterases. Mol. Cell **15**, 279–286.

Chapter 7

BIOPHYSICAL METHODS

Mechanism of Action studies

Chun-Wa Chung and Peter N. Lowe
GlaxoSmithKline Research and Development, Gunnels Wood Road, Stevenage, Hertfordshire SG1 2NY, United Kingdom

1. INTRODUCTION

The purpose of this chapter is to outline the processes involved in designing an effective and efficient mechanism of action (MOA) strategy, and to introduce the principles behind some of the techniques commonly employed for this purpose and the considerations in their usage.

The term 'mechanism of action' has many meanings but here we define it as the essential information about the target-ligand interaction that permits a compound to be moved with confidence from one stage of the drug discovery process to the next. Consequently, the information required from MOA investigations largely depends upon the stage of drug discovery pipeline at which these studies are initiated and the degree of characterisation the compound has already undergone. When a compound is first identified as a putative modulator of a target, determining its MOA may merely refer to confirming that it interacts directly with the target, rather than being an 'assay artifact'. As a compound is progressed, more detailed questions about the precise nature and kinetics of the interaction may be posed, and MOA studies are focused on properties that assess a compound's drug-like potential.

Therefore the first consideration when designing an efficient MOA process is to determine the level of information required. Often the following questions may be asked about a compound's MOA:

- Does it bind the target directly? With what affinity?
- What is the site of interaction? Is there direct competition with a substrate or cofactor, or is there an allosteric mechanism?

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 155–199.

- What is the stoichiometry of the interaction?
- What are the atomic details of the interaction? Is the binding covalent, what residues are involved and are there opportunities to modify the compound to gain potency and specificity?
- Are only equilibrium parameters required or will kinetic measurements be useful in determining therapeutic benefit?

The information desired and practical limitations (e.g. protein availability, throughput, assay sensitivity and the availability of tool compounds) clearly govern the techniques that may be employed. Commonly, several methods, of different principle, will be available, and choices should be actively made to select the technique, or combination of techniques, which are most appropriate. These choices are system dependent and require a clear understanding of the system and techniques themselves. Fortunately, common fundamental principles dictate how the parameters of interest can be determined; these principles are discussed in section 2. The choice between techniques is then determined by the theoretical and practical limitations of each method as outlined in section 3. The key to a successful MOA approach is to understand the drawbacks of each technique and to try and compensate for these using careful experimental design and, particularly with more complex systems, several complementary methods.

Many of these considerations are true for any MOA approach and are not unique to a fragment based method. However, the low molecular weight (MW) and low affinities of leads typically associated with this method may challenge some techniques more than others; where appropriate these challenges will be highlighted.

2. PRINCIPLES

Protein-ligand interactions are generally monitored either under equilibrium (or maybe steady-state) conditions, in which the concentrations of specific species are measured in a time-independent manner, or under pre-equilibrium conditions, in which the formation of specific species are measured as a function of time. Under equilibrium conditions, the concentrations of the equilibrated species are defined by one or more equilibrium dissociation constants (K_d). As each K_d is composed of a combination of kinetic rate constants, it is also possible to derive K_d values from the appropriate kinetic parameters. In this sense, a full set of kinetic parameters gives more information than the equilibrium parameters.

To use direct measurements to obtain both equilibrium & kinetic parameters requires fitting the experimental data to a model which is

assumed to represent the physical processes taking place. Whilst a good fit between experiment and model data cannot prove a model is correct, merely showing that it is consistent with the data, the lack of a fit can be used as effective evidence that the ligand:macromolecule interaction does not occur in the assumed manner. This information might be of particular use for tailoring experiments in a complex system, to decide for example the best approach to structure determination with multicomponent complexes, where the presence of co-factors or substrate may be necessary for ligand binding.

2.1 Considerations for equilibrium measurements

To measure equilibrium concentrations of reactants to derive K_d requires:

- Sufficient time to reach equilibrium, at which time they are independent of the order in which reagents are added.
- A suitable model, including the stoichiometry of the interaction (see below).
- A method to measure the absolute concentration of free or bound species, or the ratio of bound to free concentrations, without perturbing the equilibrium. The latter is achieved by:
  - having separately measurable signals from each species,
  - deconvoluting the contributions to the total signal from each species, or
  - rapidly separating bound from free species such that each can be independently measured. (This is not a suitable approach for weak or rapidly dissociating ligands.)
- In general, where there is a possibility that not all of a component is fully active, the fraction that is active is required to give true K_d values.

The signal that is measured can be an intrinsic property of the ligand or target, *e.g.* nuclear spin (NMR), mass (MS, mass spectrometry), refractive index (SPR) or intrinsic fluorescence, or one can modify the target or ligand to give a probe or reporter, *e.g.* a fluorescent or spin label or radioactive tag.

Some techniques such as ES-MS measure the ratio of bound to free species, whereas others, such as SPA, measure only bound signals. Another subset of methods measure the sum of signals from both bound and free which are deconvoluted only by obtaining a full saturation curve.

2.2 Stoichiometry

The stoichiometry of interaction is a useful measurement for understanding mechanism. In most simple systems, a 1:1 stoichiometry is

expected, and hence any other stoichiometry suggests a different or more complex mode of interaction. For more complex systems, *e.g.* a ligand binding to an oligomeric protein, or a protein with multiple sites, other stoichiometries are mechanistically quite acceptable. The stoichiometry, however, is one of the more challenging measurements to make as it generally requires accurate knowledge of both the total ligand and macromolecule concentrations. The absolute errors need to be less than 20% for each to be able to unambiguously distinguish a 1:1 from a 2:1 complex. Furthermore, the relevant concentrations required are generally only of those species capable of forming the interaction, *i.e.* the functionally "active" species. For small molecules, the problems are related to the presence of impurities and potentially insolubility of the compounds and methods such as NMR can provide the necessary resolution and quantitation required. For macromolecules, extinction coefficients, based on experimental averages, prove more accurate estimates of concentration than relative values based on dye binding. For proteins, amino acid analysis remains the gold standard. It should be noted that the mass based detection methods, either direct via mass spectroscopy, or indirect via refractive index and SPR, do not require the concentrations of the ligand and macromolecule to be known for stoichiometries to be determined. However, SPR does rely on assumptions relating refractive index to mass for the determination. These have been shown to be a good approximation for proteins, but less so for small compounds, especially those containing a number of heteroatoms, such as chlorine or bromine.

2.3 Kinetic measurements

The measurement of kinetics is often performed by recording the same types of signal as those for equilibrium measurements, as a function of time after the reaction has been initiated. Special technology, such as stop-flow and quench-flow, are required to measure reaction rates with half-lives of less than a few seconds.

It is difficult to generalize exactly how to obtain kinetic information as this depends upon several factors relating to the specific system. These are principally: the affinity range, the absolute rates, the molecular nature of the process under investigation and the availability of appropriate reporters and competing ligands.

The interpretation of kinetic data can often be more complex than equilibrium data and there is a much wider choice of potential models to consider. For a 1:1 binding process, the simplest kinetic model (Scheme 1 in Figure 7-1) includes an association rate constant (k_{ass}) and a dissociation rate

Scheme 1: Simple Biomolecular Association

$$A + B \underset{k_{diss}}{\overset{k_{ass}}{\rightleftharpoons}} A.B$$

$$K_d = \frac{[A]\,[B]}{[A.B]} = \frac{k_{diss}}{k_{ass}}$$

$$k_{obs} = k_{ass}[A] + k_{diss}$$

Scheme 2: Biomolecular Association followed by rate-limiting "isomerization"

$$A + B \underset{k_{diss}}{\overset{k_{ass}}{\rightleftharpoons}} A.B. \xrightarrow{k_{isom}} A.B^{*}$$

$$k_{obs} = \frac{k_{isom}[A]}{(K_d + [A])}$$

Where k_{diss} = dissociation constant of the bimolecular complex (A.B), k_{ass} is second order association rate constant for formation of A.B and k_{isom} is the unimolecular rate constant for a change of A.B to a stabler complex, $A.B^{*}$. k_{obs} = measured rate constant under pseudo-first order conditions

Figure 7-1. Kinetics of bimolecular association

constant (k_{diss}) the ratio of which gives K_d. k_{ass} is usually obtained by measuring the effect of concentration on the observed rate constant (k_{obs}) for complex formation under pseudo-first order conditions. Under these conditions, one interacting partner (A) is in large molar excess over the other (B) so that its free concentration stays large and near constant. A plot of the concentration of A versus k_{obs} would then give a straight line with slope equal to k_{ass} and the Y-intercept equal to k_{diss} . The error in the intercept can be very large and k_{diss} can also, and often has to, be determined independently by preparing a complex and then inducing dissociation under conditions in which rebinding is prevented. This is either through extensive rapid dilution, or by adding a competing ligand which does not affect the signal being monitored.

A 1:1 binding process, however, can be a lot more complex than this one step interaction described above and may involve other equilibria or rate-limiting steps. A common example would be a rapid initial binding equilibrium followed by a slower "isomerization" of the ligand: macromolecule complex (Scheme 2). Here, a plot of concentration of A versus k_{obs} would give a hyperbolic curve with a maximum value of k_{obs} at saturating concentrations of A. That saturating rate is the first order rate constant (unimolecular) for the "isomerization" step (Fersht, 1999).

If it is not possible to vary the concentration of A over a sufficient range it is impossible to distinguish these two types of mechanism, and hence the kinetic information is open to misinterpretation. Obtaining the precise and correct molecular interpretation of kinetic data can be a considerable challenge, but the key kinetic parameters, k_{ass} and k_{diss} can give clear and unique clues to mechanism.

For a typical simple small molecule ligand binding, association rate constant (k_{ass}) values between ~10^7-10^9 $M^{-1}s^{-1}$ can be considered as being essentially diffusion controlled. If the value of k_{ass} is significantly lower than this, it may indicate that there is a rate limiting isomerization. This may occur before ligand-binding or only once the complex is formed. There may be a rate-limiting structural change in the ligand and/or in the protein. An unexpectedly slow k_{ass} might also indicate a modification of the protein structure, *e.g.* covalent reaction or denaturation. Note that a low k_{ass} will also be observed if the concentration of ligand is lower than expected, due to incorrect knowledge of its concentration, impurities, slow isomerization, or poor solubility.

Variations in K_d are most commonly determined by changes in the dissociation rate constant, k_{diss}. A disconnect between the two is likely to indicate a more complex mechanism than anticipated.

A common error is to confuse rates and rate constants. Thus a very high affinity ligand might take hours to bind yet be a "fast" binder. This occurs because measurements are often made at concentrations around the K_d for the interaction. If that K_d is well below nM then even with the fastest diffusion-controlled binding the measured rate will be "slow".

2.4 Specific issues associated with low MW and low affinity ligands

When screening at concentrations above approximately 50μM, it is common for compounds that are either inactive or have already saturated a binding site to show additional effects in many types of assay. These can be caused *inter alia* by interference, non-specific binding, effects due to insolubility or impurities *etc*. Low affinity binding is generally only

measured at high ligand concentrations, so elucidating the MOA of a low affinity fragment can be problematic as the relative significance of impurities, non-specific binding and insolubility can be greater for these fragments than for high affinity ligands.

For high affinity ligands, one common criterion for "specific" binding is to show saturable binding at 1:1 stoichiometry. For the reasons stated above, this can be much more difficult to demonstrate with low affinity ligands. Multiple binding may also be tolerated in the initial stages of a fragment based approach if there is access to crystallographic data, for example, so that there is confidence that specificity and potency may be readily achieveable with iterative chemistry. The key is really evidence of a progressible MOA, the criteria of which are dependent upon the system under evaluation. There are some MOAs, however, that are unlikely to be acceptable. Besides the more obvious one of assay interference, or interaction with substrates/ligands rather than the macromolecule itself, various types of protein structural modification may also be unacceptable. These include covalent reaction and protein denaturation/unfolding. A related example of the latter, which has recently been identified, involves the interaction of macromolecules with the surface of aggregates of small molecules (McGovern et al 2002, McGovern et al 2003, Ryan et al 2003). Several criteria were used to identify such undesirable interactions including the reduction in effect on elevating [macromolecule], inclusion of detergent or BSA. This example highlights the difficulties in designing foolproof exclusion strategies when there are many possible undesirable MOAs; consequently it often proves simpler to confirm compounds of interest have a simple desirable mechanism, *e.g.* time-independent, single site, substrate competitive. However, it should be borne in mind that many effective drugs have complex mechanisms (Swinney 2004).

3. METHODS

3.1 Fluorescence

Fluorescence is a process in which electrons within a molecule are excited by light at one wavelength, and then relax, emitting light at a longer wavelength. The difference in wavelengths is termed the Stokes Shift. The measured fluorescence intensity is dependent upon the intensity of the exciting light, the extinction coefficient, the concentration of the fluorophore, the fluorescence lifetime and the quantum efficiency of the process.

The fluorophores can be intrinsic or extrinsic. For the macromolecule, the intrinsic fluorescence of tryptophan residues is most commonly monitored. However, the small molecule ligand might also possess intrinsic fluorescence. Alternatively, the ligand can be tagged with an extrinsic fluorophore. The latter often allows the selection of tags with more favourable fluorescence properties (wavelengths and intensity) but has the disadvantage of potentially altering the properties of the parent ligand, such as solubility, affinity and specificity.

The change in several types of fluorescence property can be used to monitor the interaction of macromolecules with ligands (Pope et al 1999). The most commonly used variants are:

1. *Fluorescence intensity* (sometimes abbreviated to FLINT)

The fluorescence intensity can increase or decrease upon interaction, as a consequence of changes in quantum efficiency and/or life time. These can be induced by differences in environment, *e.g.* hydrophobicity, or changes in structure, *e.g.* changes in the extent of conjugation. The change can also result from differences in the equilibrium proportion of isomers or conformers between bound and free species.

2. *Fluorescence anisotropy and polarization* (FP)

When a fluorophore is excited with plane-polarized light, the degree to which the polarization is retained is dependent upon the fluorescence lifetime and the local and global mobility of the fluorophore. When a low molecular weight fluorescent ligand binds to a protein, there is likely to be a decrease in global mobility, and if the molecule is flexible, a decrease in local mobility. These combine to give an increase in the retention of polarization upon fluorophore binding to protein. Additionally, there might be a change in fluorescence lifetime and intensity which may contribute to the polarisation that is retained. Fluorescence anisotropy and fluorescence polarization are simply two different ways of converting the raw data into a number which reflects the degree to which polarization is retained. The former has important advantages in mechanistic work (see below).

3. *Fluorescence resonance energy transfer* (FRET) and *homogeneous time-resolved FRET* (HTRF)

FRET is the radiationless transfer of energy through space between a donor and acceptor fluorophore. Significant FRET only occurs over short distances, as the magnitude of FRET is proportional to $1/\text{distance}^6$. The

magnitude also depends upon factors that are intrinsic to the donor and acceptor fluorophores: quantum yield of donor, extinction coefficient of acceptor and spectral overlap of emission spectrum of donor and absorbance spectrum of acceptor. The donor is a fluorophore, but the acceptor can either be a "quencher" which just absorbs radiation or a higher wavelength fluorophore. In either case, the result is a reduction in fluorescence intensity at the donor emission wavelength, but in the latter case there will be re-emission at the acceptor emission wavelength. In HTRF, the same process occurs except that advantage is taken of choosing the donor to be a long lived fluorophore, typically a lanthanide. In this instance, the acceptor emission fluorescence can be measured after a time delay, thereby removing many artefactual effects of short-lived fluorophores in the ligands being evaluated.

4. *Fluorescence lifetime*

The fluorescence lifetime is the length of time that a fluorophore remains in the excited state and gives information on the processes that lead to relaxation of the excited electron. These are dependent upon mobility, environment and structural changes. In general, intensity and polarization measurements determine a single parameter, which represents a weighted average of all the species present. In contrast, lifetime measurements can distinguish several fluorophores within a mixture if they have significantly different lifetimes. For example, when monitoring intrinsic protein fluorescence, the effect of a ligand on individual tryptophan residues may be resolved. Confocal fluorescence techniques (see below) can give highly precise lifetime measurements.

5. *Fluorescence Correlation Spectroscopy (FCS)*

FCS is a powerful technique that allows the fluorescence properties of individual molecules to be monitored rather than a single ensemble average. Confocal fluorescence microscopy is combined with sophisticated correlation-based analysis to look at low concentrations of fluorescent species in very small volumes, so the properties of individual molecules, such as brightness, polarization and lifetime can be simultaneously measured. This allows a detailed and precise picture of the distribution of these properties across the molecules in the sample to be obtained.

3.1.1 Summary of the MOA Information that can be obtained from fluorescence measurements

- Proof that a ligand binds to a macromolecule and the affinity of the interaction.
- Under appropriate conditions, the stoichiometry of the interaction can be measured.
- Through competition experiments, the site of binding can be identified.
- The kinetic parameters defining binding and dissociation can be measured.

3.1.2 Considerations for usage

Fluorescence would normally be measured under conditions of low absorbance to avoid inner filter effects (see below), and hence is always performed at relatively low concentrations of fluorophore. The different types of fluorophore are considered in more detail here.

1. Intrinsic protein fluorescence

Intrinsic protein fluorescence would normally be performed in the range of 10 to 5000nM protein, dependent upon the number of tryptophans present and their individual fluorescence intensity. The intensity and wavelength of tryptophan emission is highly sensitive to the environment and can increase or decrease upon ligand interaction. Fluorescence is normally excited at around 295nm which is above the peak absorbance wavelength. This reduces issues with inner filter effects from the protein or ligand, and reduces the contribution from tyrosine fluorescence. The peak emission intensity is usually between 330 and 350nm.

Most commonly changes in fluorescence intensity are measured. However, very useful information can also be obtained from fluorescence lifetime measurements, especially if the contributions of individual tryptophans can be resolved, or by FRET using donation to a bound ligand with significant absorbance around 330-350nm.

To demonstrate and quantify binding, the ligand is titrated into a fixed concentration of protein until no further change in fluorescence occurs. This data is then fitted to a model defining binding. An important control experiment, which can identify several types of artifact, is to titrate the ligand into a solution of L-tryptophan at the same concentration as the total tryptophan content within the protein. Clearly, it is essential that the ligand neither has significant absorbance at 295nm, or fluorescence at about 340nm.

2. Intrinsic ligand fluorescence

The range of intrinsic fluorescence brightness is large, but in general concentrations around 1000nM might be chosen. Fluorescence intensity changes associated with macromolecule binding are the easiest to measure. In the absence of a large intensity change, fluorescence polarization and lifetime can however give very precise measurements of the proportion of complex formed. Data obtained from a titration of macromolecule (protein or nucleic acid) into a fixed concentration of ligand is fitted to an appropriate isotherm. Information on whether other molecules share the same binding site can then be obtained by titrating these into a mixture of macromolecule and the fluorescent ligand and monitoring reduced fluorescence due to competitive displacement.

3. Tagged ligand fluorescence

For ligands tagged with bright fluorophores (*e.g.* fluorescein, tetramethylrhodamine, Cy5b, Cy3b, Alexa dyes *etc*) it is common to use concentrations between 0.1 and 5nM, though higher concentrations can be advantageous to reduce interference from other molecules being analyzed. For those with low intensity changes upon binding, a feature of many of the brightest commercial fluorophores, fluorescence polarization is most appropriate. For screening unknown ligands, most commonly they would be titrated into a mixture of macromolecule and tagged ligand with the macromolecule concentration chosen to be at around its K_d value. This will allow detection and quantitation of the affinity of those ligands that bind at the same site (competitive) with the tagged ligand. However, binding to other sites may not be detected. The low concentrations of tagged ligand that can be used, when combined with a high affinity, mean that low protein concentrations can be employed. This makes the assay very sensitive to competitive inhibitors irrespective of affinity. A drawback is that there can be issues associated with enhanced interaction of low concentrations of components with surfaces.

3.1.3 Problems and how to overcome them

1. Impurities

Contamination with fluorescent substances is a common problem when dealing with intrinsic fluorescence. Water and buffers should not be left in contact with plastic. Ligands are often associated with fluorescent contaminants. To ensure that the fluorescence ascribed to being intrinsic to

a ligand is in fact not due to a contaminant, a check should be made to show consistency between the absorbance and fluorescence excitation spectra.

2. Intensity change in a fluorescence polarization experiment

Unlike fluorescence intensity, which is an extensive property dependent upon the concentration, the fluorescence anisotropy of a single species is an intensive property, independent of concentration. However, the ability to measure anisotropy is dependent upon measuring a fluorescence intensity signal. The measured fluorescence anisotropy of a solution is the sum of the fluorescence anisotropies of each individual component weighted according to the relative fluorescence intensities of each component. The advantages of using anisotropy over polarization is that only anisotropy is additive in this direct way, and thus is more appropriate to measurements using curve fitting, particular when there is an accompanying intensity change. Unfortunately many current instruments only output polarization values (usually quoted in mP "units"). However, the two are related by:

Anisotropy = 2P/(3-P), where P is polarization expressed in (mP/1000)

In any fluorescence polarization experiment it is essential that simultaneous measurements of the Total Intensity and Anisotropy are made, where

$$\text{Total Intensity} = I_{\text{parallel}} + (2 \text{ x gfactor x } I_{\text{perpendicular}})$$

$$\text{Anisotropy} = \frac{I_{\text{parallel}} - (\text{ gfactor x } I_{\text{perpendicular}})}{I_{\text{parallel}} + (2 \text{ x gfactor x } I_{\text{perpendicular}})}$$

I_{parallel} and $I_{\text{perpendicular}}$ are the measured intensities parallel and perpendicular to the excitation beam and gfactor is an instrument-dependent constant.

For binding reactions accompanied by an intensity change, an appropriate equation that takes these changes into account must be used to calculate Kd from the changes of fluorescence anisotropy. It should be noted that if the change in intensity is excessive, a binding isotherm cannot be reliably obtained from the anisotropy data and it may be advisable to use the intensity changes themselves to calculate binding parameters. One should also always be aware of intensity changes unrelated to the binding event

being monitored by anisotropy *e.g.* due to addition of a fluorescent ligand or through fluorescent impurities.

3. Inner filter, quenching and other effects

If the system has too high an absorbance, light will not reach the region of the sample in which the fluorescence is being measured and hence the fluorescence intensity is lower than anticipated. This is a greater problem in a cuvette than in most plate systems, as the latter have short path lengths. When assessing the intrinsic fluorescence of ligands, particular care should be taken to ensure that the concentration of compound is chosen to be sufficiently low to avoid inner filter effects (< 0.1A at every wavelength is a good guide), yet to be sufficiently high to be greater than background fluorescence, Raleigh scattering and Raman bands.

Alterations of fluorescence intensity, and in some cases anisotropy, can occur through electronic and physical interaction between the fluorophore and assay components other than the protein interest. It is therefore important to run appropriate controls and blanks to ensure that one is not mislead.

With intrinsic and tagged-ligand fluorescence it is often useful to obtain fluorescence excitation and emission spectra both in buffer (preferably at more than one pH) and in an organic solvent (ethanol or chloroform). This gives early information as to whether the compounds fluorescence is highly environment sensitive and the level of change that may be expected from compound binding.

4. Issues with low MW or low affinity ligands

When using intrinsic fluorescence of low affinity ligands, there is a high protein requirement to define the shape of binding curves. In addition, high concentrations of low affinity ligands are likely to cause interference unless the ligand has no absorbance above ~290nm. Fluorophore-tagged ligands are well suited to testing the ability of low affinity compounds to compete. In order to make the assay system robust to high concentration screening, ideally, a relatively weak tagged-ligand is chosen, such that relatively high concentrations of fluorophore and macromolecule can be used. Then under conditions such that fluorophore concentration gives the highest intensity without any inner filter effects and [fluorophore] ≈ [macromolecule] ≈ K_d, the system will retain sensitivity to competitors yet be resistant to interference.

3.1.4 Pros and Cons of fluorescence

Pros

- High sensitivity and low reagent requirements.
- Relatively quick and easy to perform experiments.
- Precise determination of K_d, and in some cases stoichiometry.
- In many cases automatable, and transferable to high density plate formats.
- Information on the site of binding, but usually only when competitive with a known ligand.
- The measurement is in solution.
- When using intrinsic fluorescence there is no requirement for labeling, and hence the MOA of the compound of interest, rather than an analogue, can be examined.
- Fluorescence anisotropy is a relatively simple parameter which can relate directly to the MW change upon formation of a ligand:macromolecule complex.

Cons

- Tagged ligand fluorescence requires chemical synthesis and can alter the properties of the ligand or the ligand:macromolecule interaction.
- Intrinsic protein fluorescence is not possible with all ligands, as some may interfere through overlapping absorbance or fluorescence.
- Intrinsic compound fluorescence often uses excitation and emission wavelengths that are susceptible to interference by competitive ligands making competition experiments difficult.
- It is unusual to get information on site of binding for ligands that do not compete with a known ligand.
- It is impossible to predict whether a ligand, protein or tagged-ligand has suitable solution and fluorescence properties without experimentation.
- The magnitude of a fluorescence intensity change associated with binding has no relationship with affinity, or the MW change.
- Fluorescence measurements have the potential to be affected by many types of artifact, which may fool the unwary! The most common issues relate to inner filter effects, insolubility, impurities, instability, surface effects (*e.g.* binding to plastic plates), and intensity changes during a polarization experiment.

3.2 Nuclear magnetic resonance (NMR)

NMR spectroscopy is a technique that allows the characteristics of magnetically active nuclei with nuclear spin > ½ to be probed. Typically,

properties such as chemical shift, relaxation times, scalar and dipolar couplings, *etc.* on nuclei such as ^{1}H, ^{19}F, ^{15}N and ^{13}C are investigated. Many of these parameters are sensitive to changes associated with ligand binding and so provide a rich source of information on the interaction. In addition, one of the key advantages of NMR is that it is possible to design an array of experiments that allow the properties of interest to be studied in a readily accessible manner, by tailoring the sequence of radio-frequency pulses used to excite the acquisition spectrum. The versatility of this technique means that there is a huge array of NMR experiments that could be used to monitor binding interactions and many of these are covered in greater detail in Chapter 4 of this book. For comparative purposes, however, it is instructive to discuss some generic aspects of this technique in the context of other biophysical methods for MOA analysis.

It is useful to divide NMR methods into two distinct categories, ligand based methods and protein based methods (Guenther et al 2004, Peng et al 2004).

3.2.1 Ligand based methods

Ligand based methods only monitor the resonance signals of the ligand and the effect of the target protein on the ligand resonances. When a ligand interacts with a protein this interaction may be detected using a wide variety of experiments. All of these rely on the fact that either the chemical shift and/or the relaxation parameters of the ligand resonances are perturbed on binding. Ligand based experiments are therefore difference experiments where the parameters of interest are monitored in the presence and absence of the target protein.

A disadvantage of ligand-based experiments is that generally no information can be inferred about the site of interaction, so specific and non-specific interactions cannot be differentiated. To ensure that the ligand binds at the desired site a third experiment must be performed, where the desired site is made unavailable for binding. This may be achieved by using a competitor for the active site or by using a mutant protein no longer competent to bind ligand at this site. It is not uncommon for the ligand perturbations to be only partially reversed during the competition experiment. This may be interpreted as the presence of more than one site of binding, or only partial competition via an allosteric mechanism; in either case there is ambiguity as to the precise site of interaction. Another potential problem with ligand based experiments is that many rely on the rapid exchange between bound and free ligand and the sensitivity of having a large excess of free ligand. For this subset of experiments, such as trNOE, STD, Waterlogsy, compounds that bind tightly will not be detected as the

exchange between the bound and free states is too slow to transfer the required information to the only species monitored; tight binders may therefore appear as false negatives.

3.2.2 Protein based methods

Protein based methods rely on changes in the properties of the protein resonances on ligand binding. If the protein resonances can be resolved and assigned to specific atoms in the structure then the site of binding can be inferred by the localisation of these changes. By far the most popular manifestation of this method is the use of ^{15}N labelled protein and the acquisition of ^{1}H-^{15}N HSQC spectra. In this two-dimensional spectrum a single peak is observed for each directly bonded proton-nitrogen atom pair. The spectrum is therefore readily interpreted as a residue-resolved map that probes the environment of the entire protein backbone (except for proline). In favourable cases, specific ligand binding causes only a distinct subset of resonances to be perturbed and allows not only the binding event to be monitored, but the site and affinity of the interaction to be measured as well. Binding with affinities generally between pM and mM can be detected, though it will not be possible to determine the Kd of highly potent ligands due to low sensitivity of NMR and the tight binding limit. The high information content, ease of interpretation, speed of data acquisition and robust nature of this experiment makes ^{1}H-^{15}N chemical shift mapping a powerful MOA tool in suitable systems.

Unfortunately, ligand binding can produce extensive chemical shift changes, especially when protein movements occur, making it difficult to pinpoint the precise region of binding. Often the magnitude of a scaled chemical shift is used for analysis, but as the size of the change is not directly proportional to the distance from the binding site, this does not overcome the issues with the diffuse changes that may occur. The generation of ^{15}N-labelled protein is also not possible for all systems. For large proteins the ability to resolve and assign individual resonances can also become problematic, so some of the advantages of this method are lost.

3.2.3 Summary of the MOA information that can be obtained from NMR

- Confirmation of interaction between protein and ligand.
  In favourable cases:
- The equilibrium dissociation constant (K_d) for the binding reaction.
- The site of binding. A significant advantage of protein-based methods is that it is possible to identify the site of ligand binding without any

assumptions about the putative site, *i.e.* without the need for a competitive ligand.
- Qualitative kinetic information.

3.2.4 Considerations for usage

For ligand based methods it is essential to ensure that specific binding is monitored. This normally requires the ability to demonstrate the abolition of binding in the presence of a competitive ligand.

For protein-based experiments, typically labelled protein must be used.

It is possible to gauge the affinity of the binding from both ligand- and protein-based methods, but it is not the easiest way for measuring K_d as a time and labour intensive titration is often necessary. In some instances, NMR may give some qualitatively indication of on and off rates, however it is not an appropriate method to measure kinetic parameters unless these are unusually slow.

Typical sample volumes are in the ~300-500µl range with ligand and protein concentrations <100µM now being tractable. However, the exact requirements will be system- and experiment-dependent.

In order to increase the throughput of NMR methods, it is possible to test pools of compounds. A number of factors should be considered when designing efficient pools for ligand- and protein-based experiments. First, the tolerance for organic load (both compound and solvent); typically, pools of ~10 compounds at ~1mM each can be accommodated. Second, the likelihood of finding a binding component in each pool should not be too high. For protein-based experiments each component in an active pool must be individually tested for deconvolution. If the hit rate is high, pooling does not represent an efficient approach. For ligand-based experiments the considerations are slightly different as the binding components are directly detected within the experiment, so no subsequent deconvolution experiments are required. However, as the concentration of ligand is typically far in excess of that of protein in these experiments there is a danger of saturation by the most potent ligands, so low affinity binding components in the presence of tighter binders will be false negatives. Finally, the structural diversity of the compound pool is a factor. For protein-based methods pools of *minimum* diversity should be considered to minimise the number of pools that need to be deconvoluted. By contrast, for ligand-based experiments,

pools of *maximum* chemical diversity are desirable to allow sufficient signals from each ligand to be monitored.

3.2.5 Problems and how to overcome them

1. Use of spin labels to introduce site specific information into ligand based approaches

A generic problem with many of the ligand based methods is that they give no specificity or binding site information. This can be particularly problematic when adopting a fragment-based approach when high concentrations of compounds are used and non-specific binding is likely. One solution is to combine the use of spin-labels with ligand based detection methods. In the simplest manifestation, a specific spin-labelled residue is introduced into the protein proximal to the site of interest. Ligands binding close to the labelled residue will experience unusually large relaxation effects due to the presence of the paramagnetic centre, which will be reversed when the spin label is reduced (*e.g.* using ascorbic acid). Although this is essentially still a difference method, the quality of the information is normally very good with little room for ambiguity, even for very weak binding ligands where low populations are bound; populations of 1% can be easily visualised. The speed, low sample consumption and clarity of this method make it very attractive for screening low affinity interactions. However, it is important to bear in mind that the proximity that is detected is non-directional and may be due to low affinity non-specific interactions on the surface of the protein away from the site of interest, so active site blocking experiments are strongly advised. A second manifestation of this approach is the use of a spin-labelled first ligand which has already been characterised to bind to a subsite of interest. Compounds can be screened to find fragments that bind simultaneously and close to this labelled first ligand. In favourable situations even the relative orientation of the two co-binders may be determined. This is especially useful in a fragments based approach, especially, in the absence of crystallographic data, as guidance in the linkage of fragments that can bind simultaneously is desirable (Jahnke 2002).

2. Issues with low MW or low affinity ligands

For low affinity compounds, solubility at high concentrations is required. NMR is an excellent tool for low MW and low affinity ligands as it is often possible to measure a number of experiments on the same sample in order to

gain confidence in the presence of a binding interaction and eliminate potential artifacts.

3.2.6 Pros and Cons of NMR

Pros

- Versatile and information-rich. Experiments can be uniquely tailored to address specific questions.
- Solution-based homogenous system which can give information on a large variety of interactions over long time periods, under a wide range of conditions, *e.g.* temperature, solvent, buffer, concentrations.
- Non-destructive technique where samples can be recovered.
- Protein based methods can give direct access to the site of binding, however ligand-based experiments (except those with spin-labels) require competition experiments.
- The versatility of NMR also allows a wide range of ligand binding to be detected from very weak interactions (e.g high mM) to those of high affinity.
- The two major issues with progressing fragment based leads, namely, confirming that they interact in a specific manner and knowing their site of interaction in order to develop them are well served by NMR methods.

Cons

- Low sensitivity and high requirements for ligand and macromolecules compared to catalytic or fluorescent methods.
- The magnitude of the changes in the measured signal is often not proportional to affinity and can be ligand/macromolecule dependent.
- Throughput is modest as it cannot be converted to a parallel or plate based format, though pools of compounds can be used to increase throughput.
- Versatility leads to many possibilities in experimental design.

3.3 Surface plasmon resonance (SPR)

In SPR one binding partner is immobilised onto the surface of a solid support. Putative binders are introduced to the immobilised partner and the interaction is monitored by changes in refractive index at the surface due to the co-localisation of the ligand with the partner (Lofas 2003, Rich and Myszka 2004, van der Merwe 2001) (Figure 7-2).

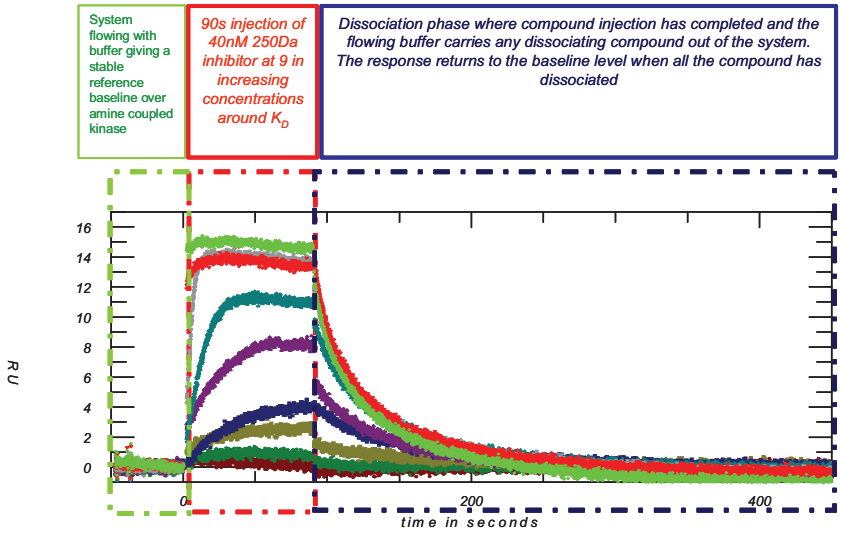


Figure 7-2. Sensorgram of BIAcore experiment where nine different concentrations of a 40nM 250Da kinase inhibitor were sequentially injected over an immobilized surface of an amine coupled kinase

There are several optical phenomena that allow this localisation to be monitored, the most established of which is surface plasmon resonance as used by manufacturers such as BIAcore. The BIAcore apparatus is a flow system, in which putative binders (analytes) are injected over an immobilised ligand on a chip and a real time binding response, or sensorgram, is recorded. As a binder is introduced to an immobilised ligand the signal increases at a rate characterised by the association process, reaching a plateau at the equilibrium level determined by K_d and the concentrations of binder and immobilised ligand. As the injection ends, the complex on the surface dissociates and any dissociating analyte is washed away by the continuous flowing buffer. The sensorgram therefore falls to its original level at a rate characterised by the dissociation process. SPR therefore allows both equilibrium and kinetic parameters to be obtained from a single injection. The equilibrium constant can be calculated in two distinct ways. The first approach is to combine the kinetic rate constants for the forward and backward reactions. Second, and more directly, it is obtained by measuring the effect of changing concentration of analyte on the response at steady-state and fitting to a binding isotherm.

Two experimental configurations are commonly used to monitor ligand:macromolecule interactions by SPR. The first of these is direct binding, in which the target protein is immobilised in a functional form onto the surface, by direct covalent coupling, or by capture using some other previously immobilized species, such as an antibody or streptavidin. When

a ligand is injected over the immobilised protein its binding can be monitored in real time as an increase in the refractive index at the surface. This configuration gives the most information as the kinetics and equilibrium characteristics of binding can be determined. The second configuration is a competition experiment in which a known ligand, often a peptide or oligonucleotide, is immobilised onto the surface. Injection of the target protein at an appropriate concentration (~K_d) results in a binding response that is easily measured on modern BIAcore instruments. Putative inhibitors are assayed by co-injection with a constant concentration of target protein. If they are able to interfere with ligand binding this will give a reduced rate of binding of the analyte and a lower binding response. Titrations of the inhibitor give a dose response curve that can be fitted using an appropriate competition equation to determine the affinity of the inhibitor.

3.3.1 Summary of the MOA information that can be obtained from SPR

- For the direct binding experiment, kinetic and equilibrium parameters are accessible, *e.g.* k_{ass}, $k_{diss,,}K_d$. By assuming that the response at saturating concentration of analyte is directly proportional to the mass at the surface, it is also possible to calculate a stoichiometry of the interaction.
- Information about the site of binding may be obtained using the appropriate competition experiments or if an active site blocked or mutant protein is used.
- For competition experiments only the equilibrium parameter, K_d, for the antagonist can be obtained.

3.3.2 Considerations for usage

Immobilization of one component is a requirement for SPR and can be problematic. The immobilised protein must be functional, stable (over hours) and at sufficiently high levels to permit the detection of the binding signals. For direct binding experiments the active protein should be present at >1000 Response Units, and preferably significantly greater than this. The immobilisation process is often protein specific, but it has been found that inclusion of a saturating concentration of ligand can enhance the stability of proteins and increase the chances of success. A rapidly dissociating ligand is chosen such that it is simply washed away after the immobilisation. Exploring the nature of the running buffer can also improve the quality of the data and help maintain the integrity of the immobilised protein. Degassing and filtering the running buffer and all injectants is desirable to ensure the smooth running of the instrument. It is essential that analytes are

soluble and stable in the running buffer prior to being injected over the immobilised surface in order to extract meaningful data.

Whilst it is possible to extract rates and thus also K_d from a single injection at one concentration, it is wise to perform replicates at a range of concentrations (ideally chosen to cover the range between <5 K_d and >5 K_d). It is common for single sensorgrams to be influenced by transient bubbles and these erratic traces can easily be identified as spurious and rejected if a full characterisation is completed. It is also advisable to check that the K_d value derived from fitting the on and off rates from the sensorgrams are consistent from those derived from fitting the equilibrium levels of the sensorgrams; an inconsistency may suggest an inappropriate binding model or the need for more careful experimental design.

Non-specific binding can be a problem, especially at high compound concentrations. Subtraction of the response from a suitable reference surface may be able to compensate for this, but it is difficult to produce an ideal reference surface and high non-specific binding can easily mask any true binding event.

Although, it is possible to obtain the kinetics of compound binding from direct binding experiments, this is practically quite challenging. The relatively high immobilisation levels necessary to detect small molecule binding makes it likely that mass transport issues limit the kinetic parameters measured to those associated with the diffusion of the molecules to the surface layer rather than give the true association rates of the interaction. Typically association rate constants $>10^7$ M^{-1} s^{-1} cannot be distinguished.

Whilst it is true that very small amounts of material are immobilised onto each chip and a chip can be used many times with suitable regeneration, there is often the need to develop and optimise an ideal assay protocol and this should be factored into considerations of protein consumption. It should also be noted that despite the microfluidic nature of the BIAcore, relatively large volumes are necessary for each sample injections, *e.g.* 50-150µl of each analyte concentration assuming an injection time of 30-90s at 100µl/min. This means that the analyte consumption, especially for monitoring low affinity interactions and when replicate injections are performed is comparable to those of 'low' sensitivity techniques such as NMR.

3.3.3 Problems and how to overcome them

1. Matching the refractive index of the injectant with the running buffer

In order to attribute the response changes in the sensorgram to binding events it is necessary to eliminate all other sources of refractive index

change upon compound injection. It is therefore necessary to match the refractive index of the running buffer to that of the injectant solution as closely as possible; ideally the components to be injected should be directly dissolved in the running buffer to be used. For small molecules this is often not possible as many compounds are solubilised in DMSO prior to their dilution into aqueous media. The hygroscopic nature of DMSO adds an additional variability to solvent matching, so even when the running buffer is nominally matched to the compound solutions the mismatch can be greater than the expected small molecule binding signal. A solution is to use a series of solutions that span the anticipated errors in solvent matching to generate a solvent correction curve, from which mismatches can be corrected (Karp et al., 2005).

2. Measuring 'True' kinetics

In order to maximise the chance of being able to derive the true interaction kinetics rather than mass transport limited rates, the level of ligand immobilised must be as low as is practically feasible and the flow rate as high as possible (Schuck and Minton, 1996). These considerations also reduce the error due to rebinding on dissociation rates. For macromolecule:macromolecule interactions it is possible to introduce small molecule inhibitors into the running buffer during the dissociation phase to block any vacant sites and prevent re-binding.

3. Issues with low MW or low affinity ligands

As competition experiments do not directly probe the size of the inhibitors there are no issues with the use of this experiment for low MW or affinity ligands, beyond the normal solubility and specificity considerations associated with all techniques.

For direct binding experiments, the magnitude of the signal is directly proportional to molecular weight, so the sensitivity of the SPR method can be compromised by a fragment based approach. Although compounds as low as ~150Da have been successful characterised, the combination of low MW and low affinities do present significant challenges to SPR when attempting to characterise binding directly. Problems such as non-specific binding are directly visualised and it may be difficult to deconvolute this from the specific binding that may also be occurring. This is particular the case when the solubility threshold of compounds is reached and only partial dose response curves can be obtained; A full dose-response curve (binding isotherm) that can be well-fitted to the desired mathematical model increases confidence in the reliability of the information obtained.

Ideally samples should be completely soluble, however saturated solutions can be filtered and the limiting concentration of the ligand determined using methods such as NMR. This is essential to obtain quantitative data, to avoid misleading results arising from non-specific binding of aggregates, and to prevent the physical blockage of the microfluidics by insoluble particulates.

3.3.4 Pros and Cons of the technique

Pros

- As changes in mass, which is an intrinsic property of interactions with an immobilized component, is measured, all binders could be monitored.
- For direct binding experiments, it is possible to access kinetic and equilibrium data on compound binding without the need for labelling or the use of a competitive tool compound.
- There is no intrinsic lower or upper limit to the affinities that can be characterised.
- The magnitude of the response signal is directly related to the mass of the binding ligand. It is possible to obtain stoichiometries from the saturating response levels.

Cons

- One component of the interaction must be immobilised so the measurements do not take place in solution.
- It is necessary to immobilise in a functional manner and find conditions under which protein integrity can be maintain over long periods. This is system dependent and can be very time consuming.
- It is necessary to regenerate the surface prior to reinjection of another sample of analyte. In practice, however, for rapidly dissociating ligands this can be trivial.
- In a typical assay, the same immobilised protein is repetitively challenged. This is a linear process and any failure of the surface, due to bubbles, protein denaturation, compound precipitation, *etc.* during a run will result in all subsequent compound injections being invalid. Furthermore, the cost of replacing the chip surface is high.
- Sensitivity is dependent on mass of the binding component.
- In practice, the range of affinities for which both kinetic & steady-state paramaters are measurable is limited. For example with very high affinity interactions, low concentrations of analyte are required to give a dose-response curve for the steady-state binding signal, and hence even with high association rate constants the measured rate of association may be too low to give analyzable data.

3.4 Isothermal titration calorimetry (ITC)

Calorimetry is the measurement of the heat produced (exothermic) or heat taken in (endothermic) during a reaction, such as the binding of a ligand to a macromolecule (O'Brien et al 2001). The magnitude of the associated heat (enthalpy) change is directly proportional to the amount of reaction product.

To measure ligand:macromolecule binding in an isothermal titration calorimeter, typically ligand solution is loaded into an automatic syringe which repetitively delivers small aliquots (*e.g.* 5µl) into a stirred cell (volume about 1.4ml), containing a solution of macromolecule. During the first injection the amount of ligand:macromolecule complex formed is highest, and the heat change is therefore greatest. With subsequent injections there are fewer unoccupied sites on the macromolecule and hence the extra amount of ligand:macromolecule complex formed is lower and the magnitude of the heat change decreases. Ultimately all sites are occupied and in an ideal situation no further heat change occurs. The total heat change that occurs during each injection is measured, where necessary appropriate blanks are subtracted and the data fitted to a relevant model describing the reaction taking place.

3.4.1 Summary of the MOA information that can be obtained from ITC

For a ligand binding to a macromolecule the curve fitting can directly produce very precise estimates of:

- The equilibrium association constant (K_a) for the binding reaction.
- The measured stoichiometry (n), *i.e.* the number of binding sites on the macromolecule for the ligand.
- The enthalpy change (ΔH) for the binding reaction.
- Through the use of the following relationships, the total free energy change (ΔG), the total entropy change (ΔS) and the equilibrium dissociation constant (K_d) for the binding reaction at temperature, T can be obtained:

$$\Delta G = -RT\ln K_a$$
$$\Delta G = \Delta H - T\Delta S$$
$$K_a = 1/K_d$$

ITC is effectively the only technique in which ΔG, ΔH, ΔS, n and K_d can be simultaneously obtained from a single experiment. The thermodynamic parameters give detailed information on the affinity and the energetics that determine that affinity. The apparent stoichiometry information can be used

either to establish the number of binding sites per molecule, when the concentrations are precisely known, or to establish the concentration of binding sites, if the absolute stoichiometry and concentration of the ligand is known.

3.4.2 Considerations for usage

1. Protein consumption

To accurately, and simultaneously, determine ΔG, ΔH, ΔS, n and K_d , the initial molar concentration to be used in the cell (typically macromolecule) should be determined by $10<c<100$, where c is the product of K_a and the concentration. Thus, for a reaction with $K_a=10^5 M$, (K_d = 10μM) and using a Microcal ITC instrument which requires a minimum of 1.8 ml of macromolecule solution to fill the cell, the cell concentration should ideally lie between 100 and 1000 μM. This represents a requirement of between 180 and 1800nmol per complete titration, *i.e.* for a protein with MW 50kDa between 9 and 90mg per experiment! For a weak affinity interactor *e.g.* Kd of 500μM, the concentration of macromolecule should lie between 5000 and 50000μM and at least 450mg is ideally required.

The protein requirement is dramatically reduced if a) higher affinities are being measured, b) only ΔH or c) less precision is required. Thus, for a Kd of 1μM less than 0.5mg of protein may suffice.

With very high affinities, the minimal concentration required becomes dependent upon the magnitude of ΔH as it is essential to obtain a heat change significantly above any backgrounds.

In all cases the molar concentration in the syringe is required to be at least 5-fold greater, and in some cases much higher, than that in the cell in order to achieve a full titration curve of species.

The titrations can also be performed in reverse, *i.e.* macromolecule in syringe titrated into a solution of the small molecule ligand in the cell. This has little impact on the total amount of reagents required but may be advantageous if the ligand is of limited solubility.

3.4.3 Problems and how to overcome them

1. Entropically driven reactions

The measured heat change is not only dependent upon the molar enthalpy change and the concentrations of reactants but also on the temperature and heat capacity of the system. Thus, at any selected temperature it is possible that the heat change is zero. This could be confused with lack of binding.

Under such circumstances, the temperature can be either increased or decreased, or a competition experiment can be performed. In the latter, two ITC titrations are performed. First a "gold standard" compound is titrated into macromolecule to obtain its affinity and enthalpy change. Second, in a separate titration the "gold standard" compound is titrated into the macromolecule, pre-equilibrated with the test ligand (at a concentration chosen to give between 50% and 99% saturation with that ligand). Provided that the test ligand and "gold standard" are mutually exclusive (*e.g.* bind at the same site) the site of action and affinity and enthalpy change for the test ligand can be determined.

2. Heat changes other than the binding reaction

ITC measures all heat changes taking place within the cell, and not only those associated with the desired interaction. These unintended heat changes may not always be easily predicted. They include chemical reactions (a common example being the slow, yet highly exothermic oxidation of dithiothreitol) and physical processes such as heats of dilution of any assay component or changes in aggregation. It is advisable to run two control titrations in which either the ligand is omitted from the syringe or the macromolecule is omitted from the cell. If the heat changes in either control are high compared to the heat changes for the ligand/macromolecule interaction, then the reliability of the binding measurement may be in doubt. A decision is required as to whether to subtract one or other, or both blanks from the test run. The decision is dependent upon one's assumptions as to the cause of the various heat changes.

3. Interpretation of thermodynamic parameters

In situations where there is a change in ionization upon binding, the measured enthalpy change is dependent upon the nature of the buffer used, and specifically upon its heat of ionization. This can be used to advantage to "amplify" a measured heat change, but also needs to be taken into account when determining thermodynamic parameters.

4. Common issues associated with use of small molecule binders

A common problem is limited solubility of the ligand. There can be issues associated with changes in aggregation status, slow rates of binding, or inadequate representations of the concentration. Further, it is common practice to dissolve compounds in DMSO. As the heat of dilution of DMSO is very high, it is essential that there is no DMSO mismatch between the

solution injected and the solution in the cell. It is common practice to exhaustively dialyse the protein against a buffer of choice and then to dissolve the ligand in the dialysate. Then, if required, the protein solution is supplemented with DMSO so as to avoid any solvent or buffer mismatch.

5. Issues with low MW or low affinity ligands

There are no issues associated with the low molecular weight of a fragment-based approach as the signal is independent of the molecular weight of the ligand.

For direct binding, measurements of low affinity interactions require very high protein and ligand concentrations to define the shape of the binding isotherms. Thus solubility at high concentrations, especially in the injection syringe is essential. In addition there are increased problems due to the heat of dilution of DMSO when high concentrations of ligand and hence DMSO, are used.

3.4.4 Pros and Cons of the technique

Pros

- This method provides one of the most accurate and precise determinations of n and K_d for unlabelled ligands in solution.
- ITC is effectively the only technique in which ΔG, ΔH, ΔS, n and K_d can be simultaneously obtained from a single experiment.
- The signal is independent of MW.
- The measurement is in solution with no requirement for labelling, or purity.

Cons

- High requirements for both macromolecule and ligand.
- The magnitude of the measured signal is directly dependent upon the ΔH and conversely has no direct dependence upon the affinity.
- Interference by unanticipated reactions, such as chemical (*e.g.* oxidation) or physical (*e.g.* aggregation) processes.

3.5 Protein mass spectrometry

Desolvated gas-phase ions derived from proteins and their ligand complexes are detected and quantified by mass spectrometry (Benesch and Robinson 2006, Hill et al 2001). Most commonly electrospray ionization (ESI) is

used. The ionization process can be subtly adjusted so as to vary the amount of internal energy imparted and hence vary the stability of the ionized complex. In one extreme only strong covalent bonds are retained, whereas under mild conditions the protein is trapped with bound ligands, tightly associated water molecules and counterions, where intramolecular (secondary structure) and intermolecular (quaternary structure) bonds are preserved. The extent to which the ionized protein complex resembles that in aqueous solution has been debated, but under the right circumstances it is clear that the complex formation monitored by mild ESI-MS parallels that in solution.

3.5.1 Summary of the MOA information that can be obtained from mass spectrometry

- Confirmation that a ligand binds to a macromolecule and the stoichiometry of interaction, including direct observation of the number of macromolecule subunits in a quaternary structure.
- Within a series, some broad estimates of the affinity of the interaction may be made.
- Competition experiments can identify whether two ligands bind competititvely at the same site or bind simultaneously to two sites and if the former occurs with a known site ligand to identify the site of binding.
- The existence of two ligands that bind simultaneously can be observed.
- The existence of covalent modification or chemical structure changes can be monitored.
- The mass identification of the bound species from a mixture can be made.
- In very favourable circumstances, the presence of conformational changes may be inferred, especially if combined with deuterium exchange methods.

3.5.2 Considerations for usage

A primary consideration is macromolecule homogeneity, both in terms of primary structure and tightly bound cofactors and ions. Homogeneity becomes more important with higher MW macromolecules as it becomes increasingly difficult to resolve apo and complex masses.

1. Denaturing mass spectrometry

If only covalent interactions are to be monitored the protein can be run under denaturing conditions, *e.g.* acidified aqueous organic solvent. The material can be then desalted by gel filtration or by hydrophobic

chromatography enabling the best quality spectra to be obtained. For many proteins this will allow precise determination of the molecular weight of any covalent bound ligand.

2. Native (non-covalent) mass spectrometry

To study non-covalent protein-ligand interactions, it is essential that the protein is fully buffer exchanged and that the system is run in a solvent that is both MS compatible and protein friendly. Often this is satisfied by 50-250mM ammonium acetate at the desired pH. For proteins that require high ionic strengths to maintain integrity and solubility it may be difficult to find conditions under which MS spectra of sufficient quality can be obtained.

3.5.3 Problems and how to overcome them

Significant work is required to ensure that the ionization conditions are optimized and to discriminate between genuine specific site binders and non-specific binders. Without "gold standards" this can be difficult, or indeed impossible, to achieve rigorously.

In order to obtain "clean" spectra, sufficient energy is imparted to strip the protein from loosely bound buffer ions *etc.* as well as loosely associated ligand. If the affinity of the ligand of interest is too low, even specifically bound ligand will be stripped. Conversely at too high a concentration of ligand multiple binding might be seen, even though one site is specific whereas another is not. This non-specific site may be a different site on each complex, emphasising that even for the spectrum of a 1:1 complex that the same mode of binding or indeed binding site may not be occupied by each of the molecules of that mass.

The lack of correlation between measured signal and affinity is a problem, particularly as the affinity of different types of interactions is differentially affected by the ESI process. Going from aqueous solution to the gas phase may increase the strength of electrostatic bonds but may reduce hydrophobic interactions. A weak electrostatic ligand may thus be more readily detected than a strong hydrophobic one. This problem is exacerbated by the possibility that individual complexes may have different ionisation potentials that differentially modulate the magnitude of the spectrum of the complex ion.

It can be difficult to identify ideal conditions to detect the native ion for each compound complex and to avoid false negative it may be necessary to adjust the ionisation parameters for each compound tested.

The protein should be of a sufficiently high homogeneity if accurate stoichiometry of complexes is required. This is even more critical for

identification of the bound ligand from the increase of mass of the protein upon complex formation.

3.5.4 Pros and Cons of the technique

Pros

- Directly monitors the existence of complexes.
- Directly observes the stoichiometry of interactions.
- Direct demonstration of covalent interaction and mass characteristics of modification.
- Identifies the mass of the species from a mixture that is competent to bind.
- Demonstration of whether pairs of ligand are mutually exclusive or can simultaneously bind.
- As a universal property is measured no labeling is required.
- The exact mass of bound species can be determined.

Cons

- Relatively high protein requirements are typically required, despite the inherent sensitivity of MS.
- The relationship between solution affinity and binding as observed in the gas phase is ill-understood, and likely to depend upon the precise nature of the bonds and forces maintaining binding, *e.g.* electrostatic interactions may be enhanced by the absence of solvation.

3.6 Affinity Chromatography

In affinity chromatography a ligand or macromolecule is immobilized on a stationary phase, typically in an hplc column, and the binding partner introduced in the mobile phase. Interaction results in retardation. The retention time is dependent upon the concentration of the immobilized species and the affinity. The shape of the peak is also determined by the on and off-rate constants. With an immobilized macromolecule, ligands can be detected by uv-vis absorbance, fluorescence or mass spectrometry.

3.6.1 Information that can be obtained

- Proof that a ligand binds to a macromolecule, relative affinities, and information on kinetics.
- Evidence for competition can be seen under appropriate conditions.

3.6.2 Considerations for usage

The immobilization process must not unduly perturb the binding interaction. The concentration of immobilized species must be of the same order of magnitude as the Kd for the interaction to get significant retardation. If the affinity is too high, the retention time becomes long and the peak shape deteriorates such that measurements cannot be easily made.

3.6.3 Problems and how to overcome them

1. Immobilisation

Loss of activity of the macromolecule might be prevented by immobilization in the presence of a reversible ligand.

2. Protein stability and long retention time/affinity

The protein must have sufficient stability to allow continuous flow for many hours. If retention times are too long because of high affinity, either the immobilized concentation must be reduced or a modifier, *e.g.* organic solvent, salt, *etc*. can be introduced into the mobile phase.

3. Demonstration of specific binding

It can be difficult to set up an appropriate control to show that the ligand interacts specifically with the target rather than with the matrix.

3.6.4 Pros and Cons of the technique

Pros

- Directly monitors the interaction between the ligand and isolated macromolecule.
- Excellent for weak affinity (provided that a high macromolecule concentration is immobilized.)
- When combined with MS analysis this provides positive identification of the active species of ligand, *e.g.* from an impure sample.

Cons

- Immobilization is difficult to control. Only a limited range of affinities is accessible under isocratic conditions.

- Relatively high protein requirements.
- A lot of work is often necessary to set up a system, though once set up and if the protein is stable, the throughput can be high.

3.7 Enzyme catalytic assay and functional assays

Here, the measurement of ligand:macromolecule binding is inferred through monitoring a functional property of a macromolecule *e.g.* for an enzyme its catalytic activity, or through interaction of the macromolecule with another species.

3.7.1 Information that can be obtained

- Affinity Kd, stoichiometry, kinetic parameters, are all inferred using competition with known ligands (*e.g.* substrate or macromolecule).

3.7.2 Considerations for usage

Steady–state enzyme kinetic experiments are commonly performed under conditions of excess substrate over enzyme, under which Michaelis-Menten kinetics are usually observed. Alternatively, excess enzyme over substrate can be used to create conditions under which only a single turnover occurs. In that situation, the product production can usually be represented by a single exponential which can give extra precision and can benefit subsequent fitting to more complex models. It has an advantage in giving very short assay times which might suit unstable systems, but has the disadvantage that stop-flow technology might be required to capture the data. Enzyme catalysis can be used to follow the kinetics of the approach to the steady state which gives detailed kinetic information regarding the rate-limiting intermediate steps, which is lost when performing steady-state analysis.

3.7.3 Pros and Cons of the technique

Pros

- These techniques are often quick, simple and cheap to run.
- They are readily miniaturisable and highly automatable and thus low in ligand and protein consumption.
- They involve direct monitoring of the modulation of functional activity which is normally the desired outcome of compound binding.

Cons

- The measurement only gives indirect evidence for interactions and the MOA determined is heavily reliant on the use of the appropriate model.

3.8 Other methods

There are many other methods to detect and quantify ligand binding that may be appropriate for a system of interest. As it is not practical to discuss these in detail, some of these are highlighted briefly below to further reinforce the often common considerations and also their uniqueness to MOA studies.

1. Gel filtration chromatography or ultrafiltration

This technique requires the separation of bound and free species either by gel-filtration or by ultrafiltration, followed by quantitation. To reduce the perturbation of the equilibrium by the separation step, this must be more rapid than the dissociation rate. Rapid gel-filtration is commonly performed using mini spin columns. The major problems are ligand or protein binding to the gel filtration medium or membrane, and ligand insolubility. These may be helped by the changing the matrix or modifying the elution conditions, *e.g.* addition of detergents. However, these difficulties might prove difficult to overcome and may be ligand dependent. It is wise to ensure that the level of recovery from the separation step is high so these issues are quickly identified and do not lead to misleading conclusions. For low affinity ligands, high protein concentrations are required to provide sufficient complex formation to be detected. The high off-rates typically associated with low affinity complexes means that separation techniques are not generally useful for characterising weak interactions.

2. Equilibrium dialysis

A given concentration of ligand is allowed to equilibrate across a semi-permeable membrane on one side of which the target macromolecule is placed. After equilibrium is attained, the concentrations of the ligand on both sides of the membrane are measured (*e.g.* by LCMS). The concentration on the "no protein" side represents the "free" concentration, the concentration on the other side is the sum of bound and free. The protein concentrations used must be of the same order as K_d and the ligands must be freely permeable and not stick to the container or membrane. The method directly monitors the existence of complexes without the use of labels, though gives little indication as to whether the interaction is specific. As before high

protein concentrations are required for low affinities which may make it less applicable to a fragment based approach.

3. Radioactive ligand binding

In this technique, either a radiolabelled ligand binds to a macromolecule, or radiolabelled macromolecule binds to a ligand and then the bound and/or the unbound radiolabel concentration is measured. Distinction of bound from unbound can be made by gel-filtration or ultrafiltration of an equilibrated mixture. To obtain measurements at equilibrium, scintillation proximity assay (SPA) is preferred. The unlabelled partner must be immobilized onto a SPA bead. A scintillation signal is obtained from bound, but not from free radiolabel. Hence the bound concentration can be measured in presence of free without separation. As a competition assay, it is very suitable for determining the K_d of a ligand and establishing the location of the binding site. Radiolabelling of macromolecule can be performed by reaction of lysine amino groups with N-succinimydylpropionate-2,3-^{3}H. In terms of radiolabel, ^{3}H is better than ^{33}P and ^{125}I and the former has a reduced pathlength. Non-specific binding can be an issue so using high radiolabel purity is best. Colour quenching can be corrected with a tartrazine quench curve provided that the colour is evenly distributed within the well, *i.e.* for soluble coloured compounds. Alternatively, or additionally, a control can be performed in which the unknown ligand, suspected of quenching, is added to assays in which ^{3}H is bound to the beads in an irreversible manner.

4. Analytical ultracentrifugation (AUC), size analysis and light scattering

Ligand binding is detected by change in the molecular size and shape determined either by analytical ultracentrifugation (AUC), size exclusion chromatography (SEC) or static or dynamic light scattering. From these changes, mechanistic information can be derived as to whether a ligand changes the oligomeric state, and whether ligand complexes are binary or of a higher order. Changes in aggregation might indicate an undesirable MOA, or might give information useful for crystal structure determination. Using AUC and SEC, with specific detection of ligand, can give confirmation of interaction with target and to what oligomeric state. However, these studies can be time-consuming, use large amounts of reagents and for light scattering be prone to dust and insolubility issues.

5. X-ray crystallography

Although not commonly considered a technique for characterizing the MOA of compounds the advances in X-ray crystallography now mean it is feasible to rapidly determine complexes in an amenable crystallographic system in a matter of days if not faster. The 3D (3-dimentional) structure of a compound in the active site of a protein automatically addresses a number of MOA questions: it confirms the direct interaction with the target, the site of interaction, and the details of the interaction and explains possible SAR (Structure-Activity-Relationship). Whilst a positive result from X-ray can be immensely informative, the failure to obtain a complex structure can be multi-factorial and offers little information. The major disadvantage of the technique is therefore it is only possible to gather information on the subset of compounds that form crystalline complexes. It should also be noted that even for this subset of compound it offers no quantitative information about solution affinities.

6. Label free biosensors

The Biacore implementation of SPR described above is now a well-established use of a biosensor to measure molecular interactions. Recently, a number of other biosensors, relying on different physical principles have, or are about to, become available commercially. A recent review has discussed their potential in the small molecule discovery area (Comley 2005). The technologies developed by Akubio (Resonant acoustic profiling), Corning (Resonant waveguide grating sensor) and SRU Biosystems (Guided mode resonance) (Cunningham et al. 2004) have the potential to overcome a drawback of the Biacore instrumentation as parallel screening on a large scale becomes possible. All give signals which correlate with increases in mass associated with the binding of small molecules to immobilized protein targets.

7. Enhanced protein stability

The ability of ligand binding to enhance protein stability is a well recognised phenomenon. The degree of stabilization can be systematically probed by observing the increased resistance of a protein to chemically or physically denaturing conditions, such as urea or temperature (Matulis et al 2005). Many methods are able to monitor the degree of denaturation, *e.g.* enzyme activity, NMR, CD and most recently, extrinsic fluorescence using a probe that binds selectively to unfolded protein. However, whilst these are useful tools there are caveats. Non-specific binding can also result in stabilization and specific binding can result in unexpected destabilization

under denaturating conditions (Horn and Shoichet 2004) These caveats may be especially pertinent for a low affinity approach, where relatively large concentrations of ligands are used and these may produced a significant non-specific stabilising effect, as seen with additives such as arginine, glycerol, *etc.*

4. EXAMPLES OF MOA STRATEGIES

4.1 Phosphotyrosine mimetics for SH2 Domains

4.1.1 Background

SH2 domains are common domains in intracellular proteins designed to sense the phosphorylation of specific tyrosine residues of protein partners. They often lie at control points in intracellular pathways and the ability to differentiate between tyrosine and phosphotyrosine (pY) is key to their role in regulation.

Finding a drug-like phosphotyrosine mimetic for incorporation into small molecule SH2 domain inhibitors has been a critical and difficult goal in targeting SH2 domains. One way of tackling this problem has been to directly screen small acid mimetics for their ability to bind at the pY pocket of SH2 domains and then grow from these fragments to give compounds with the required potency and selectivity (Xu et al 1995, Lesuisse et al 2002).

4.1.2 Defining the issues

Although phosphorylation of the tyrosine is essential for recognition and makes an extensive network of hydrogen bonds in the binding pocket, phosphotyrosine itself does not have high affinity for the protein. For example the phosphotyrosine affinity for Src SH2 is only ~1mM, despite its extensive H-bonding interactions with the protein. Any MOA must therefore be robust enough to cope with ligand concentrations in the mM range and the compounds themselves must be soluble in this concentration range. Additionally as the interest is only in fragments that bind in the phosphotyrosine (pY) pocket and not in any other proximal or allosteric pockets a MOA approach should ideally pinpoint the site of binding to this pocket.

4.1.3 Successful approaches

A multitude of biophysical methods have been successfully used to screen pY mimetics to SH2 domains. These include the following:

1. Fluorescent ligand competition experiments - monitoring the displacement of fluorescently labelled pY recognition peptide at mM compound concentrations (Cousins-Wasti et al 1996).
2. SPA – monitoring the decrease in signal as chemically tritated protein is prevented from binding to scintillate beads when mM concentrations of compound are present (Mandine et al 2001).
3. SPR/Biacore – where a pY recognition peptide is immobilised to the surface and mM compounds are co-injected with the SH2 domain.
4. Non-covalent ES-MS – where the direct binding of compounds are detected by the presence of the mass of compound+protein (Bligh et al 2003).

All these techniques have been able to detect the binding of low affinity compounds to SH2 domains. However, they do not allow the site of action to be precisely localised to the pY binding site (pY itself cannot be used as a probe due to its low intrinsic affinity). The only two methods that have been successfully used to determine the specific interaction of fragments to the pY pocket are protein based NMR experiments (^{1}H-^{15}N HSQC) and X-ray methods (Figures 7-3 and 7-4). These two methods have the additional advantage of being able to determine the binding site without the use of a competition assay and therefore there is no *a priori* assumption of the location of binding.

This example highlights several aspects to consider when MOA studies are being initiated. First, many of these techniques rely on the ability to carry out a competition assay to detect binding, and therefore require a probe for the desired site. In most cases it was necessary to use a probe that spanned more than the desired pY pocket in order to have the right properties, *e.g.* affinity. For some methods, the probe ligand had to be further modified by the appropriate attachment of a fluorescent, or radioactive label without compromising affinity and specificity. Second, although many techniques can be used to find pY mimetics, each has its own practical and theoretical detection limitations and it is therefore important to consider how these limitations effect the overall strategy. For example, methods where a signal related to formation of the complex is detected and the magnitude of this varies from one ligand to the next, *e.g.* non-covalent MS, NMR, X-ray, *etc*. have no absolute detection limits. This means that for some ligands, the

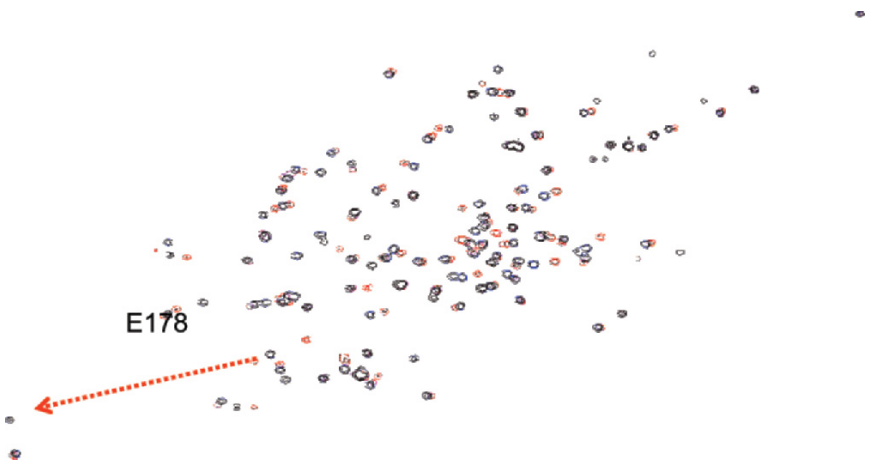


Figure 7-3. NMR spectrum of SRC SH2 domain with shift of E181 on binding highlighted

presence of 1% complex may be sufficient for detection, whereas for others even 10-20% complex formation may not be sufficient, as considerations such as the ionisability of the complex, the chemical shift change on complexation, or the ability of the compound to soak into the crystallographic system are dominant scaling factors. Both the screening and the compound selection strategy should reflect these limitations. Third, where a very large compound collection can be tested it may be adequate to use just one method and accept that some molecules of comparable affinity will be missed. However, where only a smaller screening pool can be assayed it may be desirable and necessary to combine a number of methods to maximise the number of hits detected. Multiple assays may also build confidence in validating hits at an early stage, especially for low affinity compounds that are on the borderline of detectability. Another way of building confidence is the inclusion of several representatives of the same chemotypes in the screening collection, or having these available in following up hits.

4.2 Nucleic-acid utilizing enzymes

4.2.1 Background

Nucleic acid utilizing enzymes such as helicases, topoisomerases, integrases and in particular polymerases have proven to be fruitful targets for drugs, particularly in the anti-microbial, anti-viral and oncology areas.

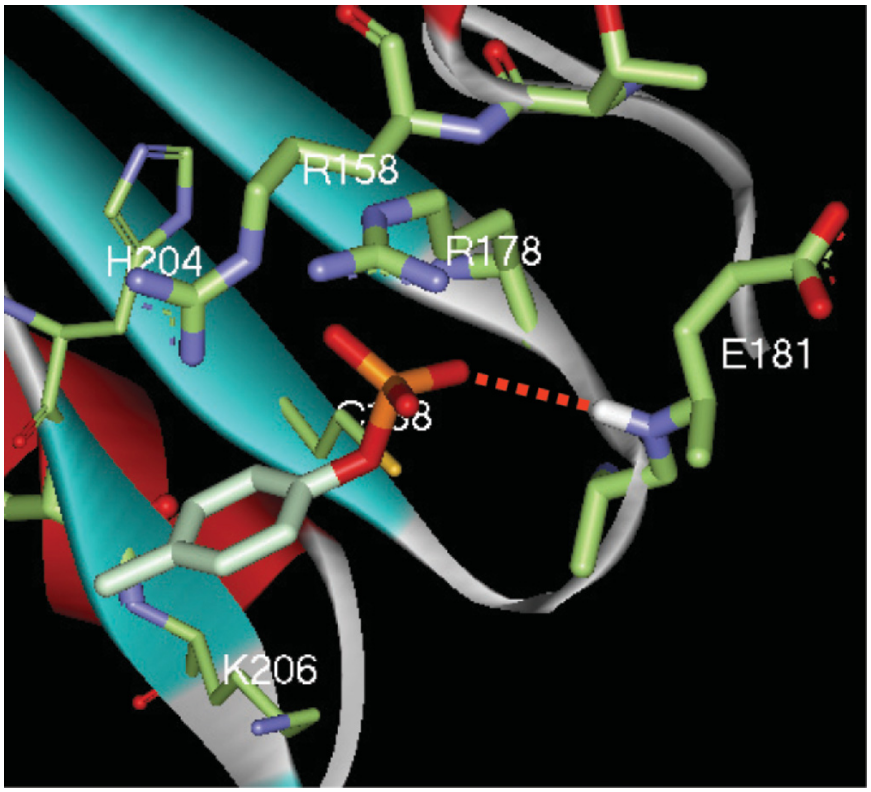


Figure 7-4. The phosphotyrosine binding site in the SRC SH2 domain

4.2.2 Defining the Issues

With SH2 domains, it is anticipated that all useful molecules will bind in the phosphotyrosine pocket, and hence ligands will all be of a "competitive" mechanism. Furthermore, the protein is relatively inflexible and allosteric effects are unlikely. Helicases and polymerases typify targets that are mechanistically more complex. They bind multiple natural ligands *e.g.* nucleic acids of variable sequence, nucleoside triphosphates and metal ions. The reaction mechanism involves several partial reactions, translocation of substrates and significant conformational mobility. Additionally, these types of targets typically bind drug-like ligands not only at sites where natural ligands bind but also at other "allosteric" sites. Thus, novel ligands will frequently involve an interaction that is not competitive. A further complexity is that functional assays generally have a complex nature, making traditional enzyme kinetics difficult to interpret. Moreover, the

presence of nucleic acid, with its amphipathic structure and base stacking presents a species which has a propensity to interact with many classes of small molecules. There is thus a requirement to demonstrate that small molecule ligands are affecting function via interaction with enzyme itself, rather than via the nucleic acid. Some of these issues ere exemplified by an elegant mechanistic analysis of hits from HTS against hepatitis C helicase, using a combination of biophysical methods (Sarver et al 2002).

4.2.3 Successful approaches

The complexity of such enzymes often requires a combination of biophysical and biochemical procedures to answer the key mechanistic questions. These are outlined below with solutions that have proved useful

1. *Does the compound interact with enzyme or nucleic acid ?*

- SPR: Biotinylated nucleic acid is immobilized on the chip, and compounds are tested for their ability to bind to it (Figure 7-5).

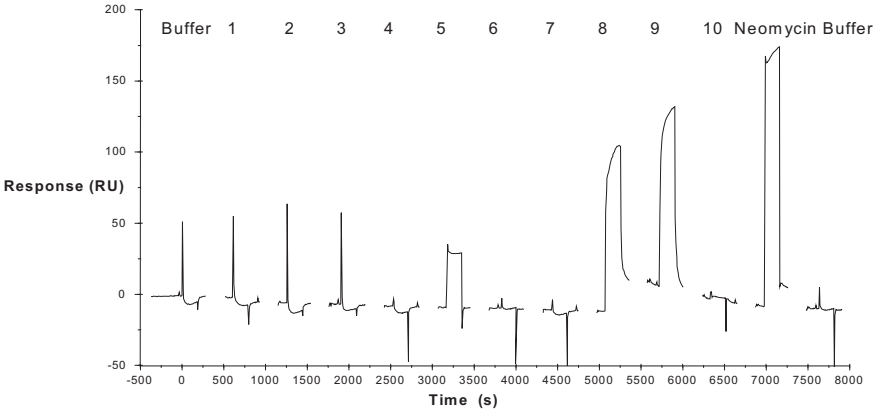


Figure 7-5. Example of testing a number of compounds over immobilised nucleic acid using SPR, Neomycin is a standard DNA binder. Compounds 5, 8 and 9 bind to nucleic acid, whereas the other compound do not.

- Ligand NMR: The effect of nucleic acid and/or enzyme on ligand signals can be used to establish to which component the compound binds. Relaxation-edited NMR of a ligand in the presence of helicase was used by Sarver and colleagues to demonstrate interaction with the enzyme, though a control experiment with an unrelated protein showed similar binding. As the system was in a slow exchange regime, quantitation of

affinity by NMR was insufficiently precise to establish that the observed binding was consistent with the K_d anticipated from the effect of the ligand on function.

- Fluorescence: Picogreen (Molecular Probes) shows enhanced fluorescence upon interaction with nucleic acid. Compounds that affect this enhancement are likely to be nucleic acid binders. This method is sensitive, high throughput and appears sensitive to both interchelators and electrostatic interactors.
- Stabilization of target by ligand-binding: Sarver et al. (2002) used changes in the kinetics of isothermal denaturation of helicase to attempt to demonstrate binding of hits to the helicase. Interestingly, some compounds caused an increase in the rate of denaturation, whereas specific compound binding is anticipated to give rise to an overall stabilization of the enzyme.
- ITC: Positive proof that the compound binds to enzyme, or conversely binds to nucleic acid can be obtained. Starket et al, (2002) shows an experiment which demonstrates high affinity binding of a small molecule ligand to helicase DNA substrate.

2. Does the compound interact with enzyme in a mechanistically acceptable manner?

- Fluorescence: Compound binding to enzyme might be monitored using changes consequent upon interaction of either intrinsic tryptophan fluorescence of the enzyme or intrinsic compound fluorescence. In an ideal situation the K_d and stoichiometry can be determined to assess whether these parameters appear acceptable.
- ITC: Demonstration that compound binds to enzyme with the appropriate affinity and a mechanistically acceptable stoichiometry.
- X-Ray: Confirmation of interaction with the target protein in a specific binding pocket.

3. What is the site of interaction ?

- X-ray: This can be particularly informative as such enzymes can have multiple ligand sites and identification of the site of interaction can facilitate understanding of SAR as well as provide clues to the mechanism of action. Sarver et al. (2002) tried to use crystallography to identify the site of binding of the compound that showed an interaction with helicase by NMR. The soaked crystal took up the compound as witnessed by concentration of the colour, but although the structure could be solved, no specific binding site was found. This was consistent

with "multiple, nonspecific sites of interaction", supported by the observed binding to an unrelated protein.

- Fluorescence: Competition experiments between ligands can indicate whether a compound binds at the same site as another ligand. Nucleic acids tagged with fluorophores, and small molecule ligands with intrinsic fluorescence can make useful probes.
- Radioactivity: If a "gold standard" of sufficient affinity can be made in radiolabelled form, SPAs may be used to ascertain whether ligands bind at the same site, or have some allosteric effect on that site.

5. CONCLUSIONS

Mechanism of action can be studied using a range of biochemical and biophysical methods. For simple systems (e.g. those in which there is a single binding site and novel ligands are competitive ligands), a single method might be sufficient and the choice will depend upon pragmatic considerations. For more complex systems, it is likely that no single method will give complete confidence, and hence it is highly desirable to create a strategy involving a combination of techniques and approaches which have been selected to be complementary, such that one technique provide insights in areas where another technique may be weaker. Even inherently simple systems, such as SH2 domains can have issues when low affinity ligands are being considered and again the use of a combination of methods of different principle can most rapidly give confidence on the mechanistic acceptability of a ligand, and hence the selection of the most appropriate compounds for progression.

REFERENCES

Benesch J. L. P. and Robinson C. V. 2006. Mass spectrometry of macromolecular assemblies: Preservation and dissociation. Curr. Opin. Struct. Biol. **16**, 245–251.

Bligh S. W. A., Haley T. and Lowe P. N. 2003. Measurement of dissociation constants of inhibitors binding to Src SH2 domain protein by non-covalent electrospray ionization mass spectrometry. J. Mol. Recognit. **16**, 139–147.

Bolbach G. 2005. Matrix-assisted laser desorption/ionization analysis of non-covalent complexes: Fundamentals and applications. Curr. Pharm. Des. **11**, 2535–2557.

Cladera J. and O'Shea P. 2001. Generic techniques for fluorescence measurements of protein-ligand interactions; real time kinetics and spatial imaging. In Protein-Ligand Interactions: Structure and spectroscopy. Harding S. E. and Chowdry B. Z. (eds) pp. 169–200.

Comley J. 2005. Label-free detection: New biosensors facilitate broader range of drug discovery applications. Drug Discovery World **6**, 63–74.

Cousins-Wasti R. C., Ingraham R. H., Morelock M. M. and Grygon C. A. 1996. Determination of Affinities for lck SH2 Binding Peptides using a Sensitive Fluorescence Assay: Comparison between the pYEEIP and pYQPQP Consensus Sequences Reveals Context-Dependent Binding Specificity. Biochemistry **35**, 16746–16752.

Cunningham B. T., Li P., Schulz S., Lin B., Vaird C., Gerstenmaier J., Genick C., Wang F., Fine E. and Laing L. 2004. Label-free assays on the BIND system. J. Biomol. Screen. **9**, 481–490.

Eccleston J. F., Hutchinson J. P. and White H. D. 2001. Stopped-flow techniques. In Protein-Ligand Interactions: Structure and spectroscopy. Harding S. E. and Chowdry B. Z. (eds) pp. 201–237.

Fersht A. 1999. Structure and mechanism in protein science. A guide to Enzyme Catalysis and Protein Folding. W. H. Freeman, New York.

Guenther U. L., Fischer C. and Rueterjans H. 2004. Nuclear magnetic resonance-based drug discovery. Methods and Principles in Medicinal Chemistry **21**(Molecular Biology in Medicinal Chemistry), 242–268.

Hill T. J., Lafitte D. and Derrick P. J. 2001. Electrospray ionization mass spectrometry. In Protein-Ligand Interactions: Hydrodynamics and calorimetry. Harding S. E. and Chowdry B. Z. (eds) pp. 311–346.

Horn J. R. and Shoichet B. K. 2004. Allosteric inhibition through core disruption. J. Mol. Biol. **336**, 1283–1291.

Jahnke W. 2002. Spin labels as a tool to identify and characterize protein - ligand interactions by NMR spectroscopy. ChemBioChem. B 167–173.

Karp N. A., Edwards P. R. and Leatherbarrow R. J. 2005. Biosens. Bioelectronics. **21**, 128–134.

Lesuisse D., Lange G., Deprez P., Benard D., Schoot B., Delettre G., Marquette J. P., Broto P., Jean-Baptiste V., Bichet P. et al. 2002. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: Application to the discovery of nanomolar inhibitors of Src SH2. J. Med. Chem. **45**, 2379–2387.

Lofas S. 2003. Modern drug discovery **6**(5), 47–49. SPR screening. Surface plasmon resonance is increasingly useful in the study of biomolecular associations.

Mandine E., Gofflo D., Jean-Baptiste V., Sarubbi E., Touyer G., Deprez P. and Lesuisse D. 2001. Src homology-2 domain binding assays by scintillation proximity and surface plasmon resonance. J. Mol. Recognit. **14**, 254–260.

Matulis D., Kranz J. K., Salemme F. R. and Todd M. J. 2005. Thermodynamic stability of carbonic anhydrase: Measurements of binding affinity and stoichiometry using thermoFluor. Biochemistry **44**, 5258–5266.

McGovern S. L., Caselli E., Grigorieff N. and Shoichet B. K. 2002. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. J. Med. Chem. **45**, 1712–1722.

McGovern S. L., Helfand B. T., Feng B. and Shoichet B. K. 2003. A specific mechanism of nonspecific inhibition. J. Med. Chem. **46**, 4265–4272.

O'Brien R., Chowdhry B. Z. and Ladbury J. E. 2001. Isothermal titration calorimetry of biomolecules. In Protein-Ligand Interactions: Hydrodynamics and calorimetry. Harding S. E. and Chowdry B. Z. (eds) pp. 263–286.

Peng J. W., Moore J. and Abdul-Manan N. 2004. NMR experiments for lead generation in drug discovery. Prog. Nucl. Magn. Reson. Spectrosc. **44**, 225–256.

Pope A. J., Haupts U. M. and Moore K. J. 1999. Homogeneous fluorescence readouts for miniaturized high-throughput screening: theory and practice. Drug Discovery Today **4**, 350–362.

Rich R. L. and Myszka D. G. 2004. Why you should be using more SPR biosensor technology. Drug Discovery Today: Technologies **1**, 301–308.

Ryan A. J., Gray N. M., Lowe P. N. and Chung C. 2003. Effect of detergent on "Promiscuous" inhibitors. J. Med. Chem. **46**, 3448–3451.

Sarver R. W., Rogers J. M., Stockman B. J., Epps D. E., DeZwaan J., Harris M. S. and Baldwin E. T. 2002. Physical methods to determine the binding modes of putative ligands for hepatitis C virus NS3 helicase. Anal. Biochem. **309**, 186–195.

Schuck, P. and Minton, A. P. 1996. Analysis of Mass Transport – Limited Binding Kinetics in Evanescent Wave Biosensors. Anal. Biochem. **240**, 262-272.

Swinney D. C. 2004. Opinion: Biochemical mechanisms of drug action: What does it take for success? Nature Rev. Drug Discov. **3**, 801–808.

Van Der Merwe P. A. 2001. Surface plasmon resonance. In Protein-Ligand Interactions: Hydrodynamics and calorimetry. Harding S. E. and Chowdhry B. Z. (eds) pp. 137–170.

Xu R. X., Word J. M., Davis D. G., Rink M. J., Willard D. H., Jr. and Gampe R. T., Jr. 1995. Solution structure of the human pp60c-src SH2 Domain Complexed with a Phosphorylated Tyrosine Pentapeptide. Biochemistry **34**, 2107–2121.

Chapter 8

ILLUSTRATION OF CURRENT CHALLENGES IN MOLECULAR DOCKING

An Application of Docking and Virtual Screening to Thymidine Kinase

Marcel L. Verdonk, Richard D. Taylor, Gianni Chessari and Christopher W. Murray

Astex Technology Ltd, 436 Cambridge Science Park, Milton Road, Cambridge CB4 0QA, United Kingdom

1. INTRODUCTION

Structure based drug design is now an established approach in drug discovery. Computational methodologies are used to facilitate structure based drug design at various stages of the process. One of the most important and routinely adopted methods is molecular docking. Molecular docking refers to the prediction of the binding mode of a specified compound within the active site of the protein target of interest (Taylor et al., 2002). A related application is virtual screening (VS) where a database of compounds is docked against a protein active site and a scoring function is used to rank the compounds in the database (Kitchen et al., 2004). This ranking is used in conjunction with other properties to identify compounds for synthesis or acquisition as part of an ongoing drug discovery program.

This chapter will focus on the application of molecular docking and VS to structure based drug design and will focus on the current challenges that limit the accuracy and utility of these methods. The key challenges will be outlined in the next section and the remainder of the chapter will be concerned with a practical example of docking against thymidine kinase (TK), using the docking program GOLD (Jones et al., 1995; Jones et al., 1997). This example has been chosen because it illustrates many of the difficulties associated with accurate molecular docking and VS. We describe

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 201–221.

a number of theoretical and pragmatic strategies that we have found useful in addressing the challenges associated with docking against TK.

2. CURRENT CHALLENGES FACING MOLECULAR DOCKING

2.1 Representation of the ligand

Before docking it is necessary to provide an appropriate 2-dimensional (2D) and 3-dimensional (3D) representation of the ligand. The key issues in the 2D representation of the ligand are its charge and tautomeric state (C/T state). This problem is made even more difficult because a ligand can change its C/T state in response to the environment of the protein. In individual dockings, a skilled modeler would consider all relevant states and dock each state against the protein, but in VS applications this approach is not practical. In this work a rule-based approach is used to enumerate plausible alternative states of the ligand and these states are docked individually. The best scoring state is then the one chosen for further analysis.

Ligands often change conformation when binding to proteins. Many programs take into account the torsional flexibility of the ligand by allowing specific bonds to be rotatable (or alternatively perform rigid docking of a pre-determined set of ligand conformations). The docking program we use here, GOLD (Jones et al., 1995; Jones et al., 1997), also allows exploration of different ring conformations through ring flipping. Other aspects of the valence geometry (e.g. bond angles) are generally fixed during docking primarily because of the need to reduce the number of degrees of freedom in the searching algorithm. This can lead to problems for moderately flexible parts of the fixed geometry such as the C-C=O bond angle between an aromatic ring and an exocylic carbonyl group. An initial geometry is required for the ligand and this must be constructed quickly and robustly for VS applications. Deficiencies in the generality or accuracy of the initial starting geometry will adversely affect molecular docking.

2.2 Representation of the protein

The problem of C/T states of the protein must also be considered. In particular, alternative protonation states are possible for histidine, aspartic acid, glutamic acid and the main chain N-terminal amine, depending on their environment. A partial solution is to assign protonation states manually based on examination of the active site and any known ligand binding modes

(as in the TK example below), although this approach will not work where the protonation state changes with different ligands. For an uncharged histidine, there is also the possibility of different tautomers, and for asparagine, glutamine and histidine the flipped sidechain alternatives are a possibility since it is usually not possible for crystallographers to assign these unambiguously. Here again the partial solution is to assign the "correct" states manually based on examination of contacts in the active site.

One of the key challenges facing docking today is the problem of protein flexibility, because most docking programs assume a rigid receptor. However, proteins are known to change conformation in response to ligand binding (Davis et al., 2003). Large-scale movements are particularly difficult to predict, except in situations where the key conformation states are associated with the mechanism of the protein and are known to occur in related proteins (e.g., flap closure in aspartic proteases (Bursavich et al., 2002) or the movement of the activation loop in kinases (Nagar et al., 2002)). Smaller movements of sidechains are easier to predict *de novo* and some docking methods attempt to do this automatically. An even more limited consideration of protein flexibility is implemented in GOLD and allows the rotation of terminal bonds involving hydrogen bond donating groups (e.g., the C-OH bond in serine residues).

A pragmatic solution to protein flexibility in docking applications is to dock against distinct, experimentally determined protein conformations and choose the best ligand solution obtained against this ensemble. More automated strategies based on this idea have been developed (Claussen et al., 2001; Knegtel et al., 1997). The draw back to these approaches is that they will not help if the ligand induces new protein conformations that are significantly different from ones previously observed. For TK, we have adopted the pragmatic multiple conformation approach to account for the effects of protein flexibility together with GOLD's default treatment of terminal hydrogen bond donating groups.

2.3 Water mediation

Water molecules often mediate favourable interactions between the protein and the ligand. In this case, for the purpose of molecular docking, the water can often be thought of as part of the protein. However such water molecules can also be displaced by ligands and the energetics of displacement will be depend, at least simplistically, on whether the loss of water-protein interactions is sufficiently compensated by the ligand-protein interactions plus the gain in entropy of the displaced water. The challenge for molecular docking is to determine for a particular ligand whether

potential mediating waters should be included in the docking or whether they will be displaced.

A pragmatic solution would be to dock against forms of the protein in which the relevant waters were (i) all omitted, (ii) some omitted, or (iii) all included in the binding site model. The simplest approach would then be to take the lowest energy ligand solution against this ensemble of protein-water states. Here, in the TK example, we have adopted a more rigorous and automated approach to the treatment of water mediation and the method is described in detail below.

2.4 Scoring

Docking functions minimize a scoring function and the ideal scoring function would have a deep, broad global minima at, or close to the experimentally observed structure for all classes of ligands and their associated protein targets. Unfortunately current scoring functions used in molecular docking are inadequate. In large scale tests of docking performance, state of the art docking programs can reproduce the experimental binding modes of a diverse test set of complexes in about 70-80% of cases (Friesner et al., 2004; Nissink et al., 2002). However, docking tests of this kind are biased since the challenges of receptor flexibility, C/T states of the ligand/protein and the valence geometry of the ligand are usually removed from the docking problem because of the way the test sets are constructed and used. The real performance of scoring functions in docking codes therefore falls far short of the ideal.

A related issue is the application of scoring methods to rank results from VS. Here the ideal scoring function would take the binding mode produced from molecular docking and give an accurate estimate of the free energy of binding. In fact this is an impossible task because this free energy must include a consideration of all states, not just the lowest energy one. A subsequent chapter in this book discusses ways to estimate binding free energies using simulation methods (see Chapter 9). In practical applications of VS, a crude scoring function is used to score the single binding mode produced by molecular docking. Popular choices are molecular mechanics functions, empirical scoring functions and knowledge-based functions (Ajay et al., 1995). Notwithstanding the wide choice of scoring methods available, the development of improved scoring functions probably remains the single most important challenge in molecular docking and VS. Methods that have been suggested in the literature to improve the ranking of compounds include consensus scoring (Charifson et al., 1999), re-scoring dockings produced with one scoring function using a second scoring function

(Hoffmann et al., 1999), the use of pharmacophores (Fradera et al., 2000; Hindle et al., 2002) and H-bonding constraints (Cole, 2002).

In the application to TK, the program GOLD is used to perform the docking/VS. Trial docking runs have been performed to determine the best available scoring function for this application, and as a result the Chemscore function was used to drive the dockings and the Goldscore function was used to rank compounds.

2.5 Searching and speed

Many different searching strategies have been used for molecular docking and it is beyond the scope of this work to review them (Taylor et al., 2002). Instead the balance between speed and accuracy is considered here. For single docking applications, the speed of docking is not usually an issue and accuracy should be the main criterion in deciding on the docking protocol. In VS applications, however, speed can be an issue although this depends heavily on the size of the VS being performed and the computing resources available. For any particular application, we would recommend that researchers use the most accurate protocol that can be supported with the available computational resource, and use computationally cheaper methods to reduce the database size.

Here, we used the 'Default 3' GOLD Genetic Algorithm settings (see below). Using these settings, for our TK test set, the docking of each C/T state of a compound takes approximately 1 minute on an Intel Xeon 2.4GHz PC running Linux. This means that, on our 110-node PC cluster, we could dock over 50,000 compounds per day (assuming 3 C/T states per compound), which we believe is sufficient for a typical drug discovery application.

2.6 Validation

In the previous subsections the challenges facing molecular docking and VS have been outlined and where appropriate, pragmatic approaches to addressing those challenges have been discussed. Table 8-1 summarizes the challenges and indicates our strategy to address those challenges in the following application to TK. Finally, we would stress the importance of validating a docking protocol on the protein of interest wherever possible. Such validation can drive the refinement of the docking protocol to give superior results and allows the molecular modeler to make predictions with confidence.

Table 8-1. The challenges faced in molecular docking and VS applications together with the approach adopted in this article for docking against TK.

Docking challenges	*Approach adopted for TK*
2D ligand representation	
Tautomers	Rule-based enumeration of possible states
Protonation	Rule-based enumeration of possible states
3D ligand representation	
Torsional flexibility	Allowed for "rotatable" bonds
Ring flexibility	Ring corner flipping allowed
Bond distances/angles	Fixed in Corina geometry for docking
2D protein representation	
Tautomers	Manually assigned; unchanged in docking
Protonation states	Manually assigned; unchanged in docking
Sidechain flipping	Manually assigned; unchanged in docking
3D protein representation	
Main chain flexibility	Not treated but "probably" not relevant
Side chain flexibility	Two known states of protein considered
Rotation of terminal groups	Terminal OH and NH_3 groups are rotated
Water mediation	Relevant waters identified and switched on and off as variables during docking
Scoring	
Docking function	Chemscore; Hydrogen-bond constraints
Ranking function in VS	Goldscore
Searching/speed	GOLD genetic algorithm with default 3 settings; Hydrogen-bond constraints.
Validation	Adopted docking protocol is based on performance in trial dockings

2.7 Thymidine kinase example

TK is a crucial enzyme in the salvage pathway of thymidine-5'-triphosphate (TTP), which is the precursor of the thymidine incorporated in DNA. A well-known class of drugs that act on TK are the antiviral compounds that target the herpes simplex virus type-1 (HSV-1). These compounds bind selectively to viral TK, where they get phosphorylated. This mono-phosphorylated species is then further phosphorylated by several other kinases to its triphosphate, which can either act as a competitive inhibitor to TTP of viral DNA polymerase, or can become incorporated into the viral DNA.

TK is a good example to illustrate most challenges commonly faced in docking and VS. Firstly, there is a degree of protein flexibility in TK, the key movement being a side chain flip in Gln125 (see Figure 8-1). In TK complexes with purine-type binders, the NH_2 group of Gln125 forms a (rather long) hydrogen bond with the backbone carbonyl of Met121; We will refer to this protein conformation as 'conformer A'. In complexes with pyrimidine-type binders, the Gln125 side chain is flipped around, and its

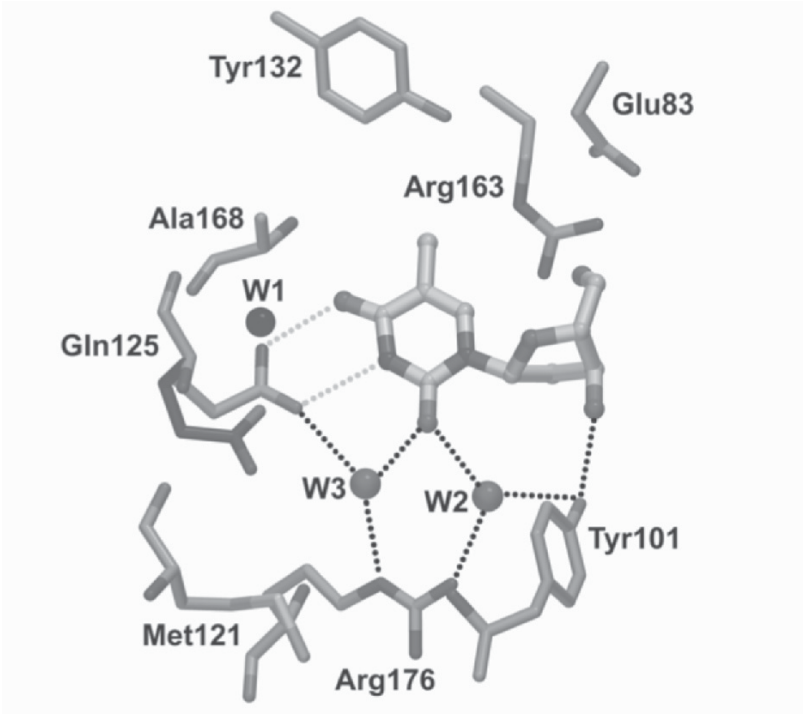


Figure 8-1. TK complex with thymidine (PDB entry 1kim, where the protein is in conformer B). The carbon atoms in the Gln125 side chain for protein conformer A and water molecule W1 are shown in black. The hydrogen-bond pattern is shown for the 1kim complex. The two hydrogen bonds used to define the hydrogen-bond constraint are shown in light grey.

NH_2 group now forms a (rather long) hydrogen bond with the backbone carbonyl of Ala168; We will refer to this protein conformation as 'conformer B'.

There are also three water molecules in the binding site that can either be displaced by the ligand or Gln125, or that can mediate through indirect hydrogen bonds between protein and ligand (see Figure 8-1). Water molecule W1 is only present in conformer A complexes and is displaced by the Gln125 NH_2 group in the conformer B complexes. Water molecules W2 and W3 only occur in conformer B complexes and are displaced by the ligand in conformer A complexes.

For the TK actives, it is relatively straightforward to derive their C/T state from the crystal structures, but other C/T states for these compounds are known. For example, the guanine-like compounds can exist as the lactam form or the lactim form, and in an unbiased docking/VS experiment both states need to be considered.

3. METHODOLOGY

3.1 Test set preparation

To establish the ability of GOLD to reproduce the binding modes of compounds binding to TK, we used a test set of 15 compounds for which the binding mode is known from X-ray crystallography (see Table 8-2). The structures were superimposed, and separate files were kept for protein and ligand in this new frame of reference.

SMILES strings were generated for all 15 'actives'. In addition, a library of 'inactives' was created by selecting compounds from our in-house database of commercially available compounds, ATLAS (Watson et al., 2003). In order to assess the ability of a scoring function that scores 3D

Table 8-2. Experimentally observed states (Obs) and success rates[a)] (Corr) for the predictions for the 15 TK actives of the protein conformer, water occupancies and compound C/T state. Ligands are named according to the PDB entry of the TK complex.

	conformer		*water W1*		*water W2*		*water W3*		*C/T*[b)]
Ligand	Obs	Corr	Obs	Corr	Obs	Corr	Obs	Corr	Corr
1e2k	B	100	off	100	on	100	on	100	100
1e2m	B	100	off	100	on	100	on	100	100
1e2n	B	100	off	100	on	100	on	100	100
1e2p	B	100	off	100	on	100	on	100	80
1ki2	A	96	off	4	off	100	off	100	100
1ki3	A	52	off	48	off	100	off	96	100
1ki4	B	100	off	100	on	100	on	100	100
1ki6	B	100	off	100	on	100	on	80	100
1ki7	B	88	off	88	on	100	on	100	92
1ki8	B	96	off	96	on	100	on	100	100
1kim	B	100	off	100	on	100	on	100	100
1qhi	A	84	on	84	off	52	off	84	84
1vtk	B	100	off	100	on	100	on	100	100
2ki5	A	64	on	64	off	100	off	64	76
3vtk[c)]	B	96	off	100	off	0	on	100	96
Overall		92		86		90		95	95

a) Percentage of the 25 possible combinations of docking runs AA1-AA5 and docking runs AB1-AB5 for which we correctly predict the protein conformer, water occupancy or C/T state.

b) The success rate at predicting the C/T state of the ligand. The 'correct' C/T state was assigned via a visual inspection of the complex.

c) The 3vtk complex is the only example in our test set where the compound binds to protein conformer B and one of the water molecules W2 and W3 is missing. The resolution of this crystal structure is 3.0Å, which means it will have been very difficult to locate water molecules in the complex.

protein-ligand interactions, we believe it is important to eliminate, as much as possible, differences between actives and inactives in lower dimensions. For example, larger compounds typically give higher scores than smaller compounds, because they have a larger surface area to interact with the protein. Hence, a large number of unspecific interactions formed by a larger compound, can easily outweigh a small number of highly specific interactions formed by a smaller compound. To prevent such artificial enrichments, we created a focused compound library containing 1500 compounds with similar 1D properties to the TK actives, following the methodology we described elsewhere (Verdonk et al., 2004).

3.2 Target preparation

We prepared two copies of the TK binding site, differing only in their conformation of the Gln125 side chains, and representing conformer A and conformer B (see Figure 8-1). Both conformers were built from Protein Data Bank (PDB, (Berman et al., 2000)) entry 1e2n. Three of the TK actives (PDB entries 1ki4, 1ki8 and 1qhi) extend into the cavity near Tyr132, and cause this side chain to twist and further open up the pocket. To accommodate these compounds, we replaced the co-ordinates (after superposition) of the Tyr132 side chain in the 1e2n complex, by those from the 1qhi complex. We used this modified 1e2n structure to represent conformer B; For conformer A, the co-ordinates of the Gln125 side chain were also replaced (after superposition) by those from PDB entry 1qhi (a representative of protein conformer A).

Three water molecules were added to both representations of the binding site. Water molecule W1 was taken from PDB entry 1qhi, water molecules W2 and W3 were taken from entry 1e2n.

Hydrogen atoms were added to the structures, ensuring that protonation and tautomeric states were correct. All atoms within 6Å of non-hydrogen atoms in any of the superimposed ligands were included in the definition of the binding site, excluding those atoms that form part of the adjacent ATP binding site.

3.3 Docking

Using a set of predefined rules, appropriate C/T states were generated for all compounds, and a SMILES string was created for each of these states. For the TK actives, this gave up to six C/T states per compound. 3D models were built automatically for each C/T state, using Corina (Gasteiger et al., 1990). The individual C/T states for each compound were then docked using GOLD against conformer A and conformer B of the protein, allowing the

three water molecules to rotate and toggle on and off; the Chemscore function was used to drive the dockings.

Allowing two conformers of the protein, three water molecules to spin around and toggle on/off, and multiple C/T states for each compound increases the search problem significantly, and therefore warrants longer search times. However, for the validation exercise to be useful, we need to use search settings that are realistic for VS. Hence, we used the 'Default 3' search settings (Verdonk et al., 2003), which are slightly slower than our standard VS settings, but still provide sufficient throughput. This search setting performs up to 10 dockings per compound C/T state, applying 30,000 Genetic Algorithm (GA) moves per docking. A GOLD run is terminated early if the top three dockings are within 1.5Å of each other. More details on the search settings used are available on request.

3.4 Scoring functions

The Chemscore function performs well at producing good quality binding modes for TK compounds, and was used to generate all the dockings presented in this chapter. The Goldscore function proved more effective at ranking compounds, and was therefore used to re-score the binding modes predicted by the Chemscore function.

The overall functional form of the Chemscore function we used in this work to drive the dockings is:

$$\begin{aligned} CS\ Fitness = \ & \Delta G_o + \Delta G_{hbond} S_{hbond} + \\ & \Delta G_{metal} S_{metal} + \Delta G_{lipo} S_{lipo} + \\ & \Delta G_{rot} H_{rot} + E_{clash} + E_{int} \end{aligned} \qquad 8.1$$

where S_{hbond}, S_{metal} and S_{lipo} are scores for hydrogen-bonding, acceptor-metal and lipophilic interactions, respectively; H_{rot} is a score representing the loss of conformational entropy of the ligand upon binding to the protein; E_{clash} is the protein-ligand clash energy, and E_{int} the ligand internal energy; the ΔG terms are coefficients derived from a multiple linear regression analysis on a training set of 82 protein-ligand complexes from the PDB (Eldridge et al., 1997). The functional form of all these terms has been described elsewhere (Baxter et al., 2000; Eldridge et al., 1997).

The Goldscore function is a molecular-mechanics-like function with four terms:

$$GS\ Fitness = S_{hb\_ext} + S_{vdw\_ext} + S_{hb\_int} + S_{vdw\_int} \qquad 8.2$$

where S_{hb_ext} is the protein-ligand hydrogen-bond score and S_{vdw_ext} is the protein-ligand Van der Waals score. S_{hb_int} is the contribution to the *Fitness* due to intramolecular hydrogen bonds in the ligand; this term is switched off in all calculations presented in this work (this is the GOLD default, and generally gives the best results); S_{vdw_int} is the contribution due to intramolecular strain in the ligand. A detailed description of the functional form of the Goldscore function is given elsewhere (Jones et al., 1995; Jones et al., 1997).

Recently, we incorporated into GOLD the ability to model interstitial water molecules (Verdonk et al., 2005). To do this, terms were added to both the Goldscore and the Chemscore functions. Hence, the score can be written as:

$$Fitness = \sigma_{P-L} + \sum_{w} o(w) \cdot \left(\sigma_{p} + \sigma_{PL-W}(w)\right) \qquad 8.3$$

where σ_{P-L} is the original score (i.e. Goldscore or Chemscore) for a given binding mode of the ligand; $o(w)$ is the occupancy of water molecule w and is either equal to 1 if the water is switched on, or zero if it is switched off; σ_{p} represents the free energy penalty associated with the loss of rigid-body entropy; $\sigma_{PL-W}(w)$ is the summation over the interactions formed by water molecule w with all ligand atoms, protein atoms, and other water molecules w' for which $o(w') = 1$; the summation is over all water molecules. The performance of this function and the justification for the terms will be described elsewhere.

As our rule-based enumeration protocol generates multiple C/T states for each compound, ideally an estimate of their relative stabilities is required. As such properties are very computationally expensive to calculate, and, in our view, not always reliable, we have assumed that the stabilities of all modeled C/T states of a compound are the same. For similar reasons, we also assumed that the relative energies of the two protein conformations (A and B) are the same.

For the scoring and ranking of compounds, we used the following modified versions of Goldscore and Chemscore functions given in equations 8.1 and 8.2, respectively.

$$Chemscore = CS\ Fitness - E_{int} \qquad 8.4$$

$$Goldscore = GS\ Fitness - S_{hb\_int} - S_{vdw\_int} \qquad 8.5$$

We have subtracted the intramolecular terms from the two scoring functions because, unless they are also calculated for the ligand in solution (Verdonk et al., 2004), these terms have arbitrary reference states.

3.5 Hydrogen-bond constraints

All the TK actives form two hydrogen bonds with the Gln125 side chain. Using such information as a constraint during docking can increase the VS hit rate by penalizing compounds that cannot form the hydrogen-bond motif and by forcing compounds that can form the hydrogen bonds to do so, even at the cost of higher clash energies.

Here, we tested the hydrogen-bond constraints as implemented in GOLD on our TK example. GOLD uses fitting points to place the ligand in the binding site (Jones et al., 1995; Jones et al., 1997; Jones et al., 1999). When hydrogen-bond constraints are set in GOLD, its search algorithm ensures that the hydrogen-bond fitting points of the selected donor(s) and/or acceptor(s) are always included in the list of protein fitting points. The selected fitting points are also given increased weights in the ligand-fitting procedure. The implementation of hydrogen-bond constraints in GOLD has been described in detail elsewhere (Cole, 2002).

4. RESULTS

All C/T states of the TK actives and focused library compounds were docked against the two states of the protein, using the Chemscore function. For each compound, the binding mode for the best-scoring protein-conformer-C/T-state combination was selected and re-scored (after a local SIMPLEX optimization) with the Goldscore function. To eliminate the variation in the success rates due to the stochastic nature of the search algorithm, we docked both the set of actives and the focused library five times against each of the two protein conformers. In the remainder of this chapter, we will refer to these five docking runs of the **A**ctives against protein conformer **A** as runs AA1 to AA5, those of the **F**ocused library compounds (i.e. the non-actives) against protein conformer **B** as runs FB1 to FB5, etc. All results presented are without the use of hydrogen-bond constraints, unless specifically stated.

4.1 Docking

Our docking protocol described above not only predicts the binding mode of each compound, but also its preferred C/T state, its optimal protein

conformer (A or B) and the presence or absence of the three water molecules. Table 8-2 gives a summary of the success rates for correctly predicting the protein conformers, compound C/T states, and water occupancies.

Our approach is able to select the correct protein conformer (92% correct), the correct C/T state (95% correct) and the correct water mediation pattern (water occupancies are correct in 90% of the cases) for most compounds. In 75% of the cases, the protein conformer, ligand C/T state, and all water occupancies are predicted correctly simultaneously.

Table 8-3 shows the performance of our docking protocol at reproducing the X-ray binding modes of the TK actives. For 13 out of the 15 compounds,

Table 8-3. Success rates[a)] for the predictions of the binding modes of the TK actives.

	RMSD (Å)			
Entry	<2.5	<2.0	<1.5	<1.0
1e2k	100	100	100	40
1e2m	100	100	100	40
1e2n	100	100	100	60
1e2p	80	80	40	0
1ki2	100	36	0	0
1ki3	100	96	24	0
1ki4	100	100	100	100
1ki6	100	100	80	80
1ki7	96	96	96	96
1ki8	100	100	100	100
1kim	100	100	100	60
1qhi	100	100	84	48
1vtk	100	100	100	60
2ki5	64	28	0	0
3vtk	100	100	100	44
Overall	96	89	75	49

a) Percentage of the 25 possible combinations of docking runs AA1-AA5 and docking runs AB1-AB5 for which the GOLD solution of the best-scoring C/T-state-protein-conformer combination is within the given RMSD cut-off. Success rates below 50% are shown on a dark grey background, those between 50% and 75% on a light grey background, and those better than 75% on a white background.

we reliably reproduce the X-ray binding mode within 2.0Å. This is a good result compared to the expected 70-80% success rate we obtain for a large test set (Nissink et al., 2002; Verdonk et al., 2003), especially since we have increased the number of degrees of freedom significantly. An example of a correctly predicted binding mode is shown in Figure 8-2. It is clear that for this compound the X-ray binding mode is reproduced quite precisely. The correct protein conformer (B) is selected, the correct water molecules are

switched on, and all hydrogen bonds observed in the X-ray structure are reproduced.

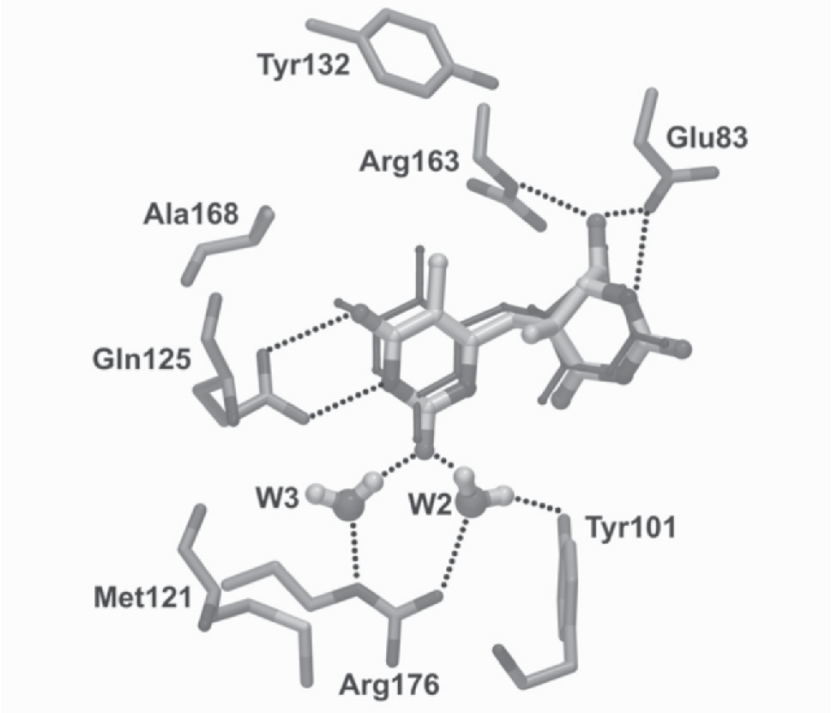


Figure 8-2. Example of a TK compound, (R,R)-6–(6-hydroxymethyl-5-methyl-2,4-dioxo-hexahydro-pyrimidin-5-ylmethyl)-5-methyl-1H-pyrimidin-2,4-dione, for which our docking protocol accurately reproduces the X-ray binding mode. For the predicted binding mode of the compound, the carbon atoms are shown in light grey. The ligand as observed in its X-ray structure (PDB entry 1e2n) are shown in dark grey.

Our protocol struggles to dock ganciclovir (1ki2) and acyclovir (2ki5) correctly. The overall binding modes produced for these compounds are roughly in line with their X-ray structures, but the key interaction motif is not reproduced. It is not entirely clear why this is the case, but ligand-induced fit does not appear to be the key problem here, as the structures are quite similar and docking against the native structures of these compounds still produces incorrect dockings. The most likely explanation for these failures appears to be a shortcoming in the scoring function. On visual inspection, the predicted binding mode is very acceptable, and it is understandable that the scoring function struggles to discriminate between

this solution and the experimental binding mode; the difference in scores is less than 1.0 kJ/mol.

Figure 8-3 compares the quality of the predicted binding modes obtained using two conformations of the protein simultaneously to those obtained against a single conformation. It is clear that in terms of water occupancies, compound C/T state and binding mode quality, the best performance is obtained when we use the combined results of dockings against conformer A and conformer B. Docking against conformer A alone gives dramatically poorer results. The drop-off in success rates is significantly less severe for docking against conformer B alone. This is largely because 11 out of the 15 TK actives bind to protein conformer B, and only 4 to protein conformer A.

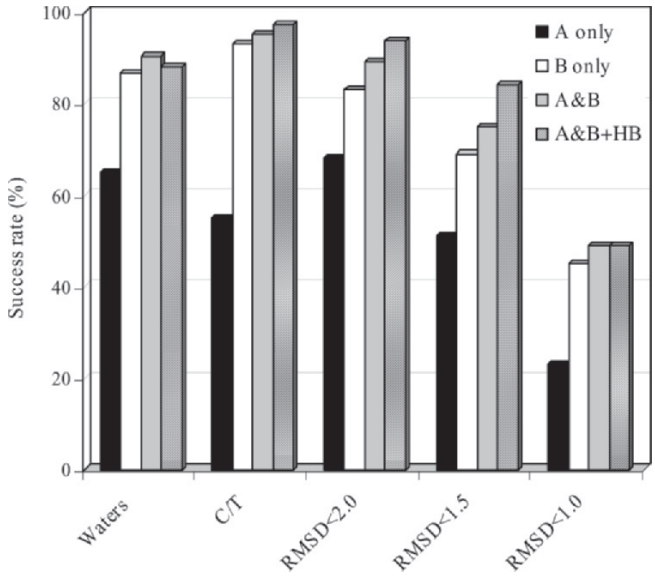


Figure 8-3. Overview of performance in terms of water occupancies (Waters), compound C/T state, and binding mode prediction at three RMSD cut-offs. Results are shown for docking against protein conformer A (black bars), against protein conformer B (white bars) and when results of docking against conformer A and conformer B are combined (grey bars). Results are also shown when two hydrogen-bond constraints were switched on to reflect the hydrogen bonds formed to Gln125 by all TK actives (dashed grey bars).

Various literature studies have reported docking success rates against TK, usually reporting on the same test set of 10 TK complexes (Kellenberger et al., 2004; Li et al., 2004). In these studies, the 1kim PDB entry (conformer B) is almost invariably used to dock against, and all water molecules are removed

prior to docking. It is therefore not surprising that most authors report that their protocol has problems docking purine-like compounds, as these bind to the alternative conformer (conformer A) of the protein. If we compare the performance on the conformer B complexes, our approach performs as well as any other protocol described in the literature (on such small test sets, it is difficult to quantify relative docking performance).

4.2 Virtual screening

Although the Chemscore function produced very good results when used to drive the dockings and to select the preferred protein conformer and ligand C/T state, it performs very poorly when it is used to rank compounds in a VS application. This is clear from the enrichment curve shown in Figure 8-4: There is virtually no enrichment of actives in the top part of the database, obtained with the Chemscore function. On the other hand, using the Goldscore function to rank the compounds gives good enrichments. We believe that, for TK, the Chemscore function may over-reward non-specific hydrophobic interactions, whereas the key recognition motif occurs via hydrogen bonding. In the remainder of this chapter, all enrichment results will be based on Goldscore.

The enrichments obtained when we use dockings against conformer A and against conformer B simultaneously are reasonable, particularly if we consider that we are comparing the TK actives to a set of compounds with similar 1D properties. The protocol we used to generate the focused library has selected 100 compounds focused around each of the 15 TK actives. As a result the focused library contains a number of thymidine-like and purine-like compounds, which clearly makes this test set quite challenging. Figure 8-5 gives an example of high-ranking thymidine-like focused library compounds that were selected around thymidine. It is clear that these compounds are similar to thymidine and we cannot exclude the possibility that they could have TK activity themselves.

A number of literature studies have reported enrichment plots for TK actives (Halgren et al., 2004; Kellenberger et al., 2004). The highest enrichments described in the literature were obtained using GLIDE, on a test set of 10 TK actives, where 20% of the actives (i.e. two actives) were retrieved in the top 1% of the database (Halgren et al., 2004). This enrichment is comparable to our results (we retrieve 17% of the actives,

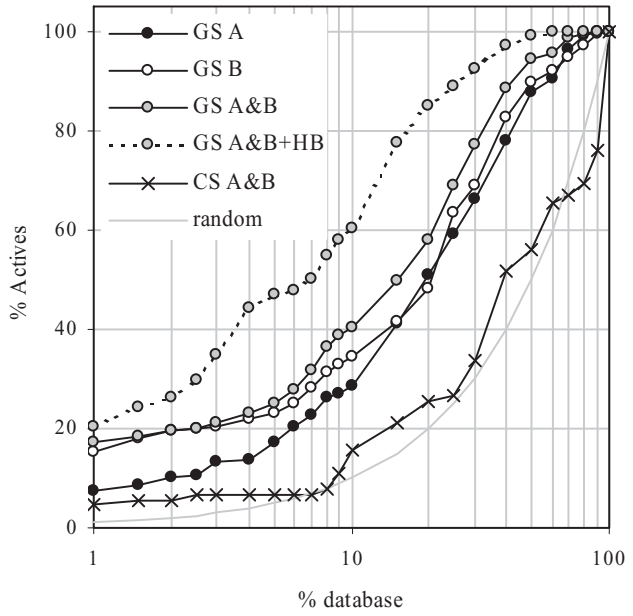


Figure 8-4. Enrichment curves for the various docking and scoring protocols described in this chapter. Chemscore (CS) results are shown for docking against both protein conformers. Goldscore (GS) results are shown for docking against protein conformer A alone, conformer B alone and for docking against both conformers. Results are also shown when hydrogen-bonding constraints are used (docking against both protein conformers). Results shown are averages over all combinations of the AA1-AA5, AB1-AB5, FA1-FA5 and FB1-FB5 runs. The grey line represents the fraction of the actives expected at random.

i.e. 2.5 actives out of 15, in the top 1%), but the enrichment curve reported for GLIDE is somewhat steeper. However, it is extremely difficult to compare our results to those reported in the literature because (i) the numbers of actives are too small to make a statistically meaningful comparison; (ii) our set of inactive compounds has similar 1D properties to the TK actives, whereas other studies have used random compound libraries or libraries with molecular weights roughly in line with the TK actives; (iii) we have included toggled waters, different C/T states for both the active and inactive ligands and side chain sampling, which increases the complexity for both docking and ranking.

Figure 8-4 also shows the enrichment curves when the compounds are docked against single conformers of the protein. Docking against conformer A alone gives significantly lower enrichments than our combined docking

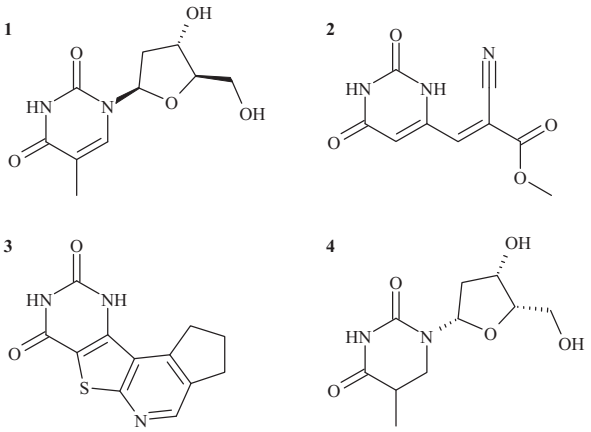


Figure 8-5. 2D diagrams of thymidine (**1**) and of three examples of high-ranking pyrimidine-like compounds in the focused library (**2**-**4**) that were selected around thymidine.

protocol. However, docking against conformer B alone gives enrichments close to those obtained when we use dockings against conformers A and B simultaneously. Again, this is largely caused by the fact that the majority of actives (11 out of 15) bind to the B conformer of TK.

Using the two protein conformers does have a favourable effect on the ranks of the TK actives. For example, 5-bromovinyldeoxyuridine (PDB entry 1ki8) ranks at 10% of the database when it is docked against its 'native' protein conformer (B). When docked against conformer A, however, its rank drops to 28%. Using dockings against conformers A and B simultaneously ranks the compound at 10% of the database again. Similarly, penciclovir (PDB entry 1ki3), ranks at 16% when docked against its 'native' protein conformer (A), at 33% against conformer B, and at 20% when dockings against both conformers are used. With a few exceptions (where ranks against conformer A are similar to those against conformer B), this is consistently the case, indicating that the docking protocol is able to identify actives that bind to protein conformer A as well as actives that bind to conformer B.

Finally, in a separate run, we used two hydrogen-bond constraints during the docking, forcing compounds to attempt to form the two hydrogen bonds that all TK actives form with Gln125 (see Figure 8-1). This has a favourable effect on the quality of the binding modes produced for the actives (see Figure 8-3). The hydrogen-bond constraints also improve the enrichments obtained in the VS experiment (Figure 8-4, dotted blue line), and we now retrieve nearly half of the TK actives in the top 5% of the database.

5. CONCLUSIONS

We have discussed the current challenges facing protein-ligand docking and VS and we have illustrated how many of these challenges can be addressed in an application to TK. On this test system, for most of the 15 complexes in our test set, we were able to predict ligand C/T state, protein conformer, water mediation, and ligand binding mode. In a VS experiment, our protocol retrieved 17% of the actives in the top 1% of a database containing compounds with 1D properties, similar to the actives.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Jason Cole from the Cambridge Crystallographic Data Centre (CCDC) for helping us set up the docking runs using hydrogen-bond constraints, and Drs Willem Nissink and Robin Taylor, also from CCDC, for their help with the validation of the water code in GOLD.

REFERENCES

Ajay and Murcko M. A. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. J. Med. Chem. **38**, 4953–4967.

Baxter C. A., Murray C. W., Waszkowycz B., Li J., Sykes R. A., Bone R. G. A., Perkins T. D. J. and Wylie W. 2000. New approach to molecular docking and its application to virtual screening of chemical databases. J. Chem. Inf. Comput. Sci. **40**, 254–262.

Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N. and Bourne P. E. 2000. The protein data bank. Nucleic Acids Res. **28,** 235–242.

Bursavich M. G. and Rich D. H. 2002. Designing non-peptide peptidomimetics in the 21st century: Inhibitors targeting conformational ensembles. J. Med. Chem. 45:541–558.

Charifson P. S., Corkery J. J., Murcko M. A. and Walters W. P. 1999. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J. Med. Chem. **42,** 5100–5109.

Claussen H., Buning C., Rarey M. and Lengauer T. 2001. FlexE: Efficient molecular docking considering protein structure variations. J. Mol. Biol. **308**, 377–395.

Cole J. C. 2002. Onward development of GOLD - selective dockings using known protein properties. In Designing drugs and crop protectants: Processes, problems and solutions, Ford, M. G. Livingston, D. Dearden, J. and Van de Waterbeemd, H. ed. Blackwell Publishing, Oxford, UK pp. 140–143.

Davis A. M., Teague S. J. and Kleywegt G. J. 2003. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design, Angew. Chem. Int. Ed. Engl. **42**, 2718–2736.

Eldridge M. D., Murray C. W., Auton T. R., Paolini G. V. and Mee R. P. 1997. Empirical scoring functions.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Comput. Aided Mol. Des. **11,** 425–445.

Fradera X., Knegtel R. M. A. and Mestres J. 2000. Similarity-driven flexible ligand docking Proteins, **40**, 623–636.

Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., Shelley M., Perry J. K. et al. 2004. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. **47**, 1739–1749.

Gasteiger J., Rudolph C. and Sadowski J. 1990. Automatic generation of 3D-atomic coordinates for organic molecules. Tetrah. Comput. Meth. **3,** 537–547.

Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T. and Banks J. L. 2004. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J. Med. Chem. **47**, 1750–1759.

Hindle S. A., Rarey M., Buning C. and Lengaue T. 2002. Flexible docking under pharmacophore type constraints. J. Comput. Aided Mol. Des. **16**, 129–149.

Hoffmann D., Kramer B., Washio T., Steinmetzer T., Rarey M. and Lengauer T. 1999. Two-stage method for protein-ligand docking. J. Med. Chem. **42**, 4422–4433.

Jones G., Willett P. and Glen R. C. 1995. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol. **245**, 43–53.

Jones G., Willett P., Glen R. C., Leach A. R. and Taylor R. 1997. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. **267**, 727–748.

Jones G., Willett P., Glen R. C., Leach A. R. and Taylor R. 1999. Further development of a genetic algorithm for ligand docking and its application to screening combinatorial libraries. In Rational drug design, novel methodology and practical applications, Parrill A. L. and Reddy M. R. ed. American Chemical Society, Las Vegas, USA pp. 271–291.

Kellenberger E., Rodrigo J., Muller P. and Rognan D. 2004. Comparative evaluation of eight docking tools for docking and virtual screening accuracy Proteins, 57**,** 225–242.

Kitchen D. B., Decornez H., Furr J. R. and Bajorath J. 2004. Docking and scoring in virtual screening for drug discovery: Methods and applications. Nature Rev. Drug Discov. **3,** 935–949.

Knegtel R. M. A., Kuntz I. D. and Oshiro C. M. 1997. Molecular docking to ensembles of protein structures. J. Mol. Biol. **266**, 424–440.

Li H., Li C., Gui C., Luo X., Chen K., Shen J., Wang X. and Jiang H. 2004. GAsDock: A new approach for rapid flexible docking based on an improved multi-population genetic algorithm. Bioorg. Med. Chem. Lett. **14**, 4671–4676.

Nagar B., Bornmann W. G., Pellicena P., Schindler T., Veach D. R., Miller W. T., Clarkson B. and Kuriyan J. 2002. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). Cancer Res. **62**, 4236–4243.

Nissink J. W., Murray C., Hartshorn M., Verdonk M. L., Cole J. C. and Taylor R. 2002. A new test set for validating predictions of protein-ligand interaction Proteins, **49**, 457–471.

Taylor R. D., Jewsbury P. J. and Essex J. W. 2002. A review of protein-small molecule docking methods. J. Comput. Aided Mol. Des. **16**, 151–166.

Verdonk M. L., Berdini V., Hartshorn M. J., Mooij W. T., Murray C. W., Taylor R. D. and Watson P. 2004. Virtual screening using protein-ligand docking: Avoiding artificial enrichment. J. Chem. Inf. Comput. Sci. **44**, 793–806.

Verdonk M. L., Cole J. C., Hartshorn M., Murray C. W. and Taylor R. D. 2003. Improved protein-ligand docking using GOLD Proteins. **52**, 609–623.

Verdonk M. L., Chessari G., Cole J. C., Hartshorn M. J., Murray C. J., Nissink J. W. M., Taylor R. D., and Taylor R. 2005. Modeling Water Molecules in protein-ligand docking using GOLD. J. Med Chem. **48**, 6504–6515.

Watson P., Verdonk M. L. and Hartshorn M. J. 2003. A web-based platform for virtual screening. J. Mol. Graph. Model. **22**, 71–82.

Chapter 9

SCORING FUNCTIONS

From Free-energies of Binding to Enrichment in Virtual screening

Luca A. Fenu[1], Richard A. Lewis[2], Andrew C. Good[3], Michael Bodkin[4] and Jonathan W. Essex[1]

1 School of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ. UK 2 Computer-Aided Drug Discovery WKL-136.3.94, Novartis Pharma AG CH-4002 Basel. Switzerland 3 Bristol-Myers Squibb, Research Parkway, Wallingford CT 06492. USA 4 Eli Lilly Research Centre, Erl Wood Manor, Windlesham, Surrey GU20 6PH, UK

1. INTRODUCTION

In Medicinal Chemistry, the potency of a drug is often characterised by its association constant with the protein target, which is in turn related to the free energy of binding. There are currently a number of ways to obtain estimates of this value, of which the most physically realistic are techniques involving methods such as molecular dynamics (MD) or Monte Carlo (MC) simulations, and statistical mechanics tools such as free energy perturbation (FEP) (Kollman 1993) to extract the relevant information from the simulation trajectories. However, these methods are expensive and, even with the modern availability of cheap supercomputing power, require running times in the order of days per each ligand.

Docking is the main instrument in structure-based virtual screening (SVS), where datasets of hundreds of thousands or even millions of lead-like molecules are 'screened' against a protein target to allow a subset to be identified for future synthesis and/or testing. It is worth mentioning that not all virtual screening is structure-based, as the parallel field of ligand-based virtual screening (LVS) is also well developed. In the case where there are known ligands for a protein, but no information is available on the three-dimensional structure for the protein, LVS is the obvious choice. LVS makes use of various descriptors, which can be one-dimensional, for example molecular weight, two-dimensional, such as a molecule's substructure, or

H. Jhoti and A. Leach (eds.), Structure-based Drug Discovery, 223–245.

three-dimensional, such as electronic or shape features. By means of these descriptors, LVS is able to select molecules which have chemical features similar to known ligands, and are therefore likely to bind to the same protein. The interested reader can find more information can find more information in the reviews by Kubinyi (Kubinyi 1997a; Kubinyi 1997b). In pharmaceutical research, virtual screening experiments have a runtime of between a few days and a few weeks. Even with a dedicated Linux cluster or other high-performance computation facilities, the time available to consider each ligand is limited to a few minutes, at the most. Docking fits nicely in this timeframe, although most algorithms can give more accurate results at the expense of longer computation time, essentially spent in more exhaustive sampling. Even if, in principle, calculating accurate binding affinities for the docked structures, using for example the simulation methods described above, is desirable, our primary interest is in obtaining a ranking of the binding affinities, allowing us to prioritise testing and synthesis of the most promising molecules.

Docking is customarily divided into two parts. During the first, we sample the conformational space of a ligand/protein complex, trying to identify the 'true' mode of binding for the ligand. To achieve this, a great number of conformations are generated for each molecule under consideration. These are then 'posed' inside the receptor pocket in various orientations. A function able to return an energy value for each three-dimensional structure is used to estimate the interaction between ligand and receptor, and is called a *scoring function*. Orientations which score well are kept, whereas the remainder are dropped. In the second phase, the different poses, belonging to the same or different ligands, are ordered according to their computed score. The scoring function used here may be more elaborate than that used in the first phase. In general, we expect the resulting scores to be correlated with binding free energies. Further information regarding small molecule docking algorithms and programs may be found in this book, or in other reviews (Halperin et al 2002; Sotriffer et al 2004; Taylor et al 2002).

Scoring functions obviously play an important role in both phases of docking. In the first, the scoring function drives the conformational and orientational sampling toward the minimum of the underlying energy surface. Because of this, we would like to have a function able to select the right binding modes, and allow us to explore efficiently the search space. Ideally, in the second phase, to obtain a correct ranking, we would like the scores returned to correlate well with experimental free energies of binding. Unfortunately, although the sampling and selection of poses is generally handled with a remarkable degree of accuracy, the final ranking of the ligands is more difficult. This is partly due to the fact that whereas the insertion of a ligand can be effectively modelled with simple functions that

take into account shape complementarity and some directional interactions, such as hydrogen bonds, free energies of binding depend strongly on competition with the solvent. Some consideration of the entropy changes on ligand binding is also needed. Because of this, the use of two different scoring functions during the two phases is often advocated.

The components required for a scoring function can be developed by considering the fundamental physics of intermolecular interactions (Atkins 1998; Leach 2001; Ajay and Murcko 1995). The important electrostatic interactions between the charges on the protein and ligand (Sotriffer et al 2004) may be modelled using a simple Coulomb expression, moderated by dielectric screening arising from any intervening molecules. For more accurate work, a full multipole expansion of the electrostatics may be included. Repulsion/dispersion interactions may be treated using either a simple Lennard-Jones 12-6 potential, with the option of the more accurate Buckingham potential that incorporates an exponential repulsion term. Hydrogen bonding is the result of the interaction between an electronegative atom (acceptor) and a hydrogen atom covalently bound to another electronegative atom (donor). The hydrogen atom possesses a considerable positive partial charge, favouring electrostatic interactions with the acceptor. There may, however, be an additional covalent contribution to the hydrogen bond. Hydrogen bonds are often attributed as giving specificity to the binding process, since in general, all hydrogen bonding sites in a protein-ligand complex should be satisfied for optimum binding to be observed (Bohm and Klebe 1996). Interactions involving π electron systems, including π-stacking and those involving cations or hydrogen bond donors, may be modelled using electrostatic and repulsion-dispersion approximations, although charge transfer effects may also need to be incorporated. The hydrophobic effect is commonly invoked to explain the preferential association between non-polar molecules, or areas of molecules, to minimize water contact. It reflects a complicated interplay of enthalpic and entropic effects involving the molecules and the aqueous solvent. This list of intermolecular forces is not, of course, exhaustive, as many other forces can play a role in binding, such as the internal energy of the ligand, which in order to bind, may have to assume a different conformational from the one it holds in solution. There are examples of studies that identify the most stable conformation of free and bound species to evaluate the free energy of binding, via a statistically-mechanically correct weighting of the conformational ensemble (Mardis et al 2001; Luo and Gilson 2000). This approach is too expensive for SVS, so scoring functions usually approximate this contribution with rule-based dihedral counts or internal van der Waals-like potentials (Giodanetto et al 2004; Jones et al 1997).

Different approaches to scoring have led to different kinds of scoring functions, which are usually classified in a tree-like structure, with three

principal branches, plus one more 'hybrid' type. The first three branches are those of force-field, empirical and knowledge-based scoring functions. For each of these, we will give a brief explanation of their rationale, problems encountered up until now, and the most recent attempts at solving these problems. For detailed information on the actual implementation of these approaches, the reader is referred to the original papers or a number of excellent reviews (Halperin et al 2002; Sotriffer et al 2004) or comparisons (Wang et al 2004; Wang et al 2003).

2. FORCE FIELD SCORING FUNCTIONS

This class of scoring function exploits currently existing molecular mechanics force fields to estimate the enthalpy of binding. Components of the AMBER (Cornell et al 1995), CHARMM (Brooks et al 1983) and other force fields are routinely used as scoring functions in various docking programs (Meng et al 1992; Kuntz et al 1982). As the calculation of non-bonded terms is expensive when applied to ligand/protein docked poses, it is often the case that values on a grid are pre-computed and then interpolated (both linearly or with higher degree polynomial expressions) to obtain the correct value at the site where the guest atom is located (Meng et al 1992). The van der Waals interactions are a commonly used component (Jones et al 1997; Kuntz et al 1982; Ewing et al 2001), as well as methods to take into account the desolvation energy and the long-range shielding effect of water (Shoichet et al 1999), either using a distance dependent dielectric constant (Vieth et al 1998) or a complete Poisson-Boltzmann treatment (Jeffrey and Saenger 1991). This last term is very expensive to compute, and has been substituted by faster approaches (Majeux et al 1999), such as GBSA (Zou et al 1999), recently recast in a pairwise approximation by Liu and coworkers (Liu et al 2004), which brings the time needed for a ligand GBSA evaluation within the time frame for virtual screening (0.5 s/pose, although this is not fast enough to be used as a docking function).

Another important term often included is one that takes into account the internal conformational energy of the ligand. Although experiments on ligand strain energy have been rather rare and sparse until now (Greer et al 1994), it is commonly accepted that when a ligand binds inside a binding site it will do so in a low energy conformation (Bostrom et al 1998). The implementation of this term is usually based on look-up tables for dihedral angles compiled from small molecule datasets.

A fresh approach to force field scoring has come in recent year from Pearlman and Charifson (Pearlman 1999; Pearlman and Charifson 2001), who developed and tested **OWFEG** (**O**ne-**W**indow **F**ree **E**nergy **G**rid),

a grid approximation of the free-energy perturbation method. In **OWFEG**, a short MD simulation is used to build a grid of free energy values relative to some probes (neutral, charged, methyl) around the active site. The score is then obtained by linear interpolation of these values for the docked pose. Both K_i prediction (Pearlman 1999) and SVS experiments ("Pearlman and Charison" 2001) have been performed using this method.

2.1 Problems and Solutions

The main problem with force-field scoring functions is that their computation is quite time consuming. Also, parameters that are initially derived for other simulation methods, such as MD or MC calculations, may not be suited for single-point energy estimates; correlations with experimental data may be poor, especially when the compounds ranked do not belong to the same class. Also, poor charge modelling seems to have a detrimental effect. The obvious solution to these problems seems to be the development of *ad hoc* potentials and charge models for specific use in scoring functions. Another palliative could be the introduction of new non-bonded terms, to estimate interactions until now neglected or 'merged' with others. As an example, we cite some recent papers (Raha et al 2005; Raha and Merz 2005; Raha and Merz 2004a; Raha and Merz 2004b) where a quantum-mechanically-based semi-empirical hamiltonian approach to scoring functions is used to estimate the metal-ligand interaction contribution to free energy of binding. Although currently too slow for virtual screening, there is the clear opportunity for such methods to be more widely used, if they provide information that simpler scoring functions cannot.

3. EMPIRICAL SCORING FUNCTIONS

Another possible approach to scoring is exemplified by the class of so-called empirical scoring functions. Their rationale is that the free energy of binding can be decomposed into a series of contributions, each of which has a clear and intuitive chemical explanation. This assumption is, strictly speaking, incorrect (Mark and van Gunsteren 1994), since whereas the free energy of binding is a function of state, the energy components into which it may be divided are not. However, the approximation is useful to a certain extent. By using terms with clear and intuitive explanations, we may be able to model specific interactions, such as π-stacking, that are difficult to capture using alternative approaches. Probably because of their lower computational requirements, the very many terms that can be used, and the different ways

in which they can be implemented, empirical scoring functions are present in many SVS methods. First attempts at empirical scoring involved the building of a knowledge base from the analysis of group-to-group interaction in 200 complexes of known experimental binding affinity (Andrews et al 1984). These first methods did not require any knowledge of the three-dimensional structure of the complex, as they simply were counts over interactions that were presumed to exist in the complexes used as reference.

Later, atom-to-atom approaches were used in the same fashion (Kuntz et al 1999). Adding knowledge of the three-dimensional structures of the training set leads to a greater level of detail in the analysis of interactions between host and ligand (Bohm 1994; Head et al 1996). A partitioning scheme can be developed in which the free energy of binding is given by a sum of interaction terms, each multiplied by its own coefficient. In mathematical terms, we can write:

$$\Delta G=\Sigma_i f_i \Delta G_i \qquad (1)$$

Here, the f_i are the coefficients, and the ΔG_i are the free energies associated with each term, each reflecting an interaction thought to be important for binding. Other, non-linear, functional forms have been investigated (Giordanetto et al 2004), but results seem to show that non-linear terms do not sufficiently improve predictive power to justify the computational overhead and loss of physical meaning on moving away from familiar terms. The coefficients are generally obtained through optimisation/regression techniques in which the scores given by the equation are fitted to known experimental binding affinities. A notable exception to this procedure comes from the work of Smith et al. (Smith et al 2003), in which the scoring function is instead parameterised using enrichment of true binders interspersed among a number of decoy molecules.

Common terms in empirical scoring functions include those for hydrogen bonds, ionic interactions, hydrophobic interactions, and the internal energy of the ligand. This makes modern empirical scoring functions similar to a non-bonded force field potential, with parameters provided by fitting to known experimental binding affinities. As these functions are additive by design, larger ligands tend to score better than small ones. In reality, however, larger ligands will suffer from the fact that, to bind, more degrees of freedom are frozen by confinement in the receptor pocket. For this reason, the entropy change usually disfavours binding. Terms that try to estimate this entropy change, predominantly through rotatable bonds counts or similar quick methods, are therefore added. Another approach to account for ligand size is to scale the scores obtained. As an example, Pan et al. suggested multiplying the score by the square root of the number of heavy atoms, or the ligand's molecular weight (Pan et al 2003).

3.1 Problems and Solutions

Despite their successes in modelling such complex interactions with relatively simple equations, empirical scoring functions have their limitations. A major drawback of any regression-based scoring function is the dependence on the size, composition and generality of the training set used to derive the weights. This is commonly recognized by workers in the field, and to counter this, attempts are being made to create common training and validation sets, containing diverse proteins and ligands. A section of this chapter is devoted to the description of these efforts. It must also be noted that creating scoring functions 'tailored' for a certain class of protein or ligand is becoming more common (Bohm and Stahl 1999). An example of this is Laederach's work with carbohydrates (Laederach and Reilly 2003) focused on affinity prediction. The other problem with empirical scoring functions is that they, in common with force field scoring functions, often lack terms to account for some interactions. For example, metal-ion interactions are often neglected or approximated by simple coulombic terms, without any reference to charge transfer and coordination. Unfortunately, to gather sufficient information to parameterise such term can be difficult, so their contribution to binding is merged into other terms. For example, if metal ion interactions are neglected, and we do not have proteins that use metals to bind the ligand in our training set, then our function will obviously fail when applied to such a protein. Alternatively, if we have a few cases in our training set, but no term to account for them, the fitting algorithm may try to compensate for this lack by increasing the coulombic contribution. This will lead to an overall overestimate of coulombic interaction. Even in the case of metal-proteins, this function would not behave correctly, as this modest overestimate is not sufficient to address the missing metal ion term. The answer, in this case, is to develop and add new terms, and once again to select accurately the training set so to reflect the application field. Another problem inherent in the empirical scoring function approach is that the training sets used to derive such functions do not contain “negative” data (i.e. unfavourable inter- or intra- molecular interactions, as these are very rarely observed in X-ray structures). Finally, it is of the utmost importance to remember that our models are just that, and as such limited in scope and applicability.

3.2 Recent Advances

A number of novel empirical scoring function have appeared recently: some are simple re-incarnations of earlier functions, for example, Verdonk et al. (Verdonk et al 2003) re-implemented ChemScore (Eldridge et al 1997)

within GOLD (Jones et al 1997). At the same time they conducted studies on the influence of using different scoring functions for the sampling and ranking parts of the docking process. The CCDC/Astex docking test set was divided into fragment-like and drug-like ligands. Overall, they confirmed that using GoldScore for both sampling and ranking gives the more accurate results, in contrast to the previous assumption that using two different scoring functions for the two phases could improve the quality of the poses. However, the optimum compromise between speed and accuracy was indeed to use ChemScore for sampling, and re-rank the outcomes with GoldScore. In this way, sampling was up to three times faster, and the results were nonetheless purged of false positives thanks to the better selectivity of the GoldScore function. Verdonk et al. (Verdonk et al 2005), in a more recent publication, allowed for water molecules to be switched on and off at their experimentally determined position during docking. GOLD's GoldScore and ChemScore functions were modified to consider the loss of the water's rigid-body entropy upon binding, with a constant penalty term σ^p, obtained from training against 58 complexes, and tested against 225. The method managed to predict correctly water displacement/mediation for over 90% of cases in both training and test sets.

Mancera et al. (Todorov et al 2003; Mancera et al 2004), in their program EasyDock (Todorov et al 2002), used PLP (Gehlhaar et al 1995) to drive sampling, and ScreenScore (PLP+FlexX's own score) (Stahl and Rarey 2001) to re-rank the outcomes. The new program GEMDOCK (Yang and Chen 2004), uses a genetic algorithm GA for docking, and a novel empirical scoring function. Composed of three terms, it contains coulombic interactions with simplified charges, PLP-like hydrogen bond terms and internal energy. This final term is in fact a constant penalty to restrict poses to a pre-defined docking box. The program has been tested for binding-mode recognition against 100 protein/ligand complexes selected from the PDB.

A new version of GLIDE has also been presented (Friesner et al 2004; Halgren et al 2004). This program uses a series of filters, of increasing complexity, to select good binding modes. It includes a scoring function derived from ChemScore (Eldridge et al 1997), extended with a few more terms. Both re-docking of known inhibitors (Friesner et al 2004) and SVS (Halgren et al 2004) experiments are presented. The authors claim that, overall, GLIDE's performances are equal or superior to those of GOLD and FlexX.

An opportunity to improve the performance of scoring functions lies in the method by which a function is parameterised. Smith et al. (Smith et al 2003) have built a novel scoring function, selecting terms from Gschwend's TEC, a toolkit containing a variety of 2 and 3-D descriptors for protein-ligand interactions. The novelty of the method lies not in the terms selected,

namely hydrogen bonds, coulombic interactions, a contact/clash count, and a buried hydrophobic surface term. Here, the parameterisation method attempts to mimic a SVS experiment, with a training set of twenty proteins and 1000 ligands selected from the WDI. The ligands were used as decoys to hide the known ligand, while a genetic algorithm adjusted the parameters to optimize the enrichment factors.

The functional form of an empirical scoring function can be changed as in, for example, Giordanetto et al. (Giordanetto et al 2004). They assembled and tested a series of scoring functions combining different terms obtained from QXP (McMartin and Bohacek 1997) descriptors, solvent accessible surface area and protein-side chain entropy. The different combinations were then trained against 100 PDB protein-ligand complexes (actually, their affinity/inhibition values) using a variety of methods (partial least squares, genetic algorithms, neural networks), and tested against 24 diverse complexes. Other forms of the scoring function, other than the simple linear combination of terms, were examined by the genetic algorithm, but with no significant improvement. Moreover, as in the case of neural networks, the non-linear terms do not always have a clear physical meaning.

4. KNOWLEDGE-BASED SCORING FUNCTIONS

The third family of scoring function has been developed thanks to the huge increase in high resolution structures (Sussman et al 1999) obtained via X-ray diffraction, solid state NMR, and a number of other related methods. These scoring functions are called knowledge-based, as their fundamental idea is to extract statistical information about the ligand/protein binding modes and to correlate these to the free energy of binding using statistical mechanics. Knowledge based pseudo-potentials are usually derived by accumulating radial distribution functions for selected pairs of protein-ligand atoms. The inverse formulation of the Boltzmann equation (see below), or a derivative thereof, is then used to extract the corresponding potentials of mean force from these distributions. To apply these potentials to scoring, the distances between the protein and ligand atoms in each pose is calculated, and the total score obtained by summing each of the energies derived from the potential curves using these distances.

$$E_{ij} = -kT \ln \rho_{ij}(r) - kT \ln Z \qquad (2)$$

Here, E_{ij} represents the interaction energy between atom *i* on the protein and atom *j* on the ligand; $\rho_{ij}(r)$ is the radial distribution function between the two atom-types at distance r, and Z refers to the partition function. A very

good and detailed explanation of the rationale behind this method can be found in the recent review by Sotriffer et al. (Sotriffer et al 2004).

By not imposing any particular functional form to the pseudo-potential curves, we are implicitly accounting for any possible kind of interaction that could occur, irrespective of whether it is enthalpic or entropic.

Besides the functional form used to derive the potentials from the populations, the differences between various incarnations of knowledge based function are in the amount of data used for parameterisation, and in the number of atom types. There is no doubt that the more protein-ligand complexes used in parameterisation, the better. With a larger knowledge-base, the statistical information extracted will be of better quality; one would expect, for example, smoother curves, and improved sampling for rare interactions. The same holds true for the diversity of the parameterisation set, which makes for a more robust scoring function. The number of atom-types, however, does not follow the same premise "the more, the better". Add too many, and the data is needlessly fragmented – there may not be sufficient information to build pseudo-potential curves between certain atom pairs, with the data therefore effectively being wasted. There is also the risk of implicitly over-fitting the function. Conversely, if there are insufficient atom-types, different interactions will be incorrectly grouped, giving potential curves that, although smoother, may have unphysical features: For example, a single pseudo potential curve with two minima could arise from the superposition of two single-minima curves. Alternatively, it may be signalling the presence of some angular-dependent phenomena, which radially-symmetric potentials cannot incorporate effectively. A balance must therefore be struck, and inevitably different researchers will make different decisions. Good practice is arguably to choose well-known and well-defined atom-types, as Gohlke and coworkers did with DrugScore (Gohlke et al 2000a; Gohlke et al 2000b), which uses 17 atom-types taken from the Sybyl mol2 file format. This is ideal as it reduces the difficulties of reliable ligand atom-typing. However, others (Muegge and Martin 1999; DeWitte and Shakhnovich 1996) used more than 17 atom types as their choice to represent the diversity of pairwise interactions in their parameterisation set, or adopted custom-defined atom-types: for example, Muegge et al., with their PMF scoring function (Muegge 2000), chose atom-types based on chemical features and not the atom's hybridisation state. All these scoring functions were parameterised with protein-ligand coordinate sets extracted from the PDB database (Sussman et al 1999), although DrugScore used Relibase (Hendlich et al 2003; Gunther et al 2003) to identify high-quality independently validated structures.

The first applications of knowledge-based scoring functions to drug design were strictly focused to HIV protease (Mizutani et al 1994;

Verkhivker et al 1995), the only proteins at the time for which there were sufficient structural data to develop a scoring function. Although promising, these early attempts did not return a generally applicable function. SMoG (DeWitte and Shakhnovich 1996; DeWitte et al 1997) was the first drug-design program to include a knowledge-based pseudo-potential. Later attempts introduced new terms to overcome some limitations of the methods due to insufficient sampling, or the neglect of solvent exposure. Mitchell et al. proposed BLEEP (Mitchell et al 1999a; Mitchell et al 1999b), where the functional form for the shell density is different, and the knowledge-based function is complemented by a van der Waals term to compensate for the low occurrence of short range interactions in crystals. The PMF scoring function (Muegge 2000; Muegge 2001; Muegge 1999; Muegge 2004) from Muegge, included a term to take into account the volume occupied by the ligand. This function was tested against the FKBP protein (Muegge 1999), and it was also modified to improve virtual screening performance, by substituting the knowledge-based curves with a van der Waals terms to account for short-range steric effects otherwise neglected. The PMF function contains a volume correction factor, to account for the volume around each ligand atom occupied by other ligand atoms. Its effect on the predictive power was also investigated (Muegge 2001) by the authors; apparently it only helps the function to fit to experimental affinities, but does not affect relative ranking of the ligands. In DrugScore (Gohlke et al 2000), from Gohlke and coworkers, a knowledge-based one-body potential scaled to the size of the solvent-accessible surface (SAS) of the protein and the ligand that becomes buried upon complex formation, was added to the PMF-like pairwise pseudo-potentials. Weighting coefficients scaled the two terms to optimise K_i reproduction. This results in DrugScore effectively being a mixed knowledge-based/empirical scoring function. Notably, the DrugScore potentials were 'anti-aliased' to avoid spurious spikes in the pseudo-potential curves due to misplacements in the crystal structures. The maximum radius considered was limited to 6 Å (12 in PMF), under the justification that only direct contacts, and not water-mediated ones, were important. Recently Ishchenko et al. released a new version (Ischenko and Shakhnovich 2002) of SMoG (SMoG2001), parameterised on 725 structures, and tested against 119, reporting performances similar to DrugScore and better than PMF and SCORE1 (LUDI (Bohm 1994)). They attribute improvements mainly to a better description of the reference state used to normalise the radial distribution function.

The use of the inverse Boltzmann distribution is not, however, the only way that information can be extracted from known structures. A very different approach is shown by Deng et al. (Deng 2004), who created the Structural Interaction Fingerprint method (SIFt). This technique, based on

the conversion of 3D structure into 1D binary fingerprints, not only makes for improvements in organisation and analysis of the structures, either experimental or generated by docking, but more importantly for our interest, it also returns superior enrichment performances when used to screen a database of ligands. The same authors in a follow-up paper (Chuaqui et al 2005) show how p-SIFt, a probabilistic interpretation of the SIFt method, can be used to generate target-specific knowledge-based filters for virtual screening.

4.1 Recent Advances

Recently, an additional scoring function was added to GOLD, when Verdonk et al. (Verdonk et al 2004) implemented a derivative of DrugScore. This implementation of the scoring function was tested for its ability to drive the docking algorithm, and found to be unsatisfactory compared to ChemScore and GoldScore. This is probably due to the lack of directional terms, which make it difficult for the scoring function to select the correct binding mode. DrugScore(CSD), a re-implementation of DrugScore using distance-dependent pair potentials obtained from small molecule crystal structures, was also unveiled (Velec et al 2005). The higher resolution of these structures compared to protein-ligand complexes led to substantial improvement in recognition of near-native binding mode being observed by the authors. Furthermore, the new scoring function shows a better correlation with binding affinities.

In addition to new implementations of existing scoring functions (Ischenko and Shakhnovich 2002), Feher et al. (Feher et al 2003) proposed a new knowledge based scoring function (BHB), based on descriptors for **B**uriedness, **H**ydrogen-bonding, and **B**inding energy. The approach is very different from traditional knowledge-based scoring functions, because the knowledge-base in this case is harvested not to build radial distribution function-like potentials, but is instead looking for characteristic binding modes and contacts. A list of 'good' interactions with the protein's residues must be compiled prior to running the virtual screening experiment, and the scoring function can also be tailored specifically for the target in a few hours, if sufficient structural data from known actives is available. This new BHB function claims a 12-fold enrichment improvement over GoldScore when re-scoring GOLD generated poses on a set of PDB protein-ligand complexes.

Muryshev et al. (Muryshev et al 2003) developed a three-stage, mixed knowledge-based/empirical parameterisation protocol: First, pairwise pseudo-potentials are extracted from a set of complexes using the inverse Boltzmann law; then, the potentials obtained are approximated by smooth

functions *via* multi-variate regression; finally, a re-calibration against known affinity data corrects some of the adjustable parameters, such as the energy and radius of hydrophobic interactions, introduced during the second stage. The function generated this way shows better performances than DOCK, GOLD, FlexX in predicting binding modes.

The strongest criticism of knowledge-based scoring functions, apart from doubts on their statistical-mechanical meaningfulness, is that not all interactions can be approximated efficiently by pairwise terms. Moreover, many interactions are highly directional, whereas radial distribution function-derived pseudo-potentials have spherical symmetry. Although a certain degree of directionality comes from the interplay of different pairwise interactions, the issue has not yet been fully addressed. However, there seems to be little concern in the literature, since these functions are cheap to evaluate, and already work reasonably well within the lead-identification phase of drug-design. With the constant increase in resolved structures, the reliability and general applicability of these functions is likely to improve, and the possibility (re-)opens for functions optimised on certain protein-ligand classes, or even for different alternative non-pairwise representations of the intermolecular interactions.

5. CONSENSUS SCORING

Since none of the scoring function classes presented above is clearly superior to the others, a more pragmatic approach to obtain good results is to combine their separate judgements. This approach, called 'consensus scoring', was pioneered by Charifson et al. (Charifson et al 1999); two or more scoring functions are applied to the same set of poses, and only structures that perform well in more than a pre-defined share of the said scoring functions are retained. Consensus scoring, analysed in detail by Wang and Wang (Wang and Wang 2001), is effective in reducing the number of false positives, which may be able to trick one function, but not all, into believing they are actives. Because of this, performance is particularly good when the functions included in the jury are different, in the sense that they describe different aspects of the binding. For example, a scoring function oriented towards the careful detection of hydrogen bonds will complement a more hydrophobic function, whereas the union with a more similar function may not be so productive. An example of this is given by Stahl and Rarey (Stahl and Rarey 2001) with ScreenScore, built from the consensus of PLP and FlexX. It is also generally found that for consensus scoring to be successful the individual scoring functions need to be of comparable quality, with similar standard errors. Moreover, as good as

consensus functions are at getting rid of false positives, they do nothing to enable us to recover false negatives which slip through most or all of our scoring nets and lie at the bottom of our ranked list. In these cases, the description of the protein-ligand interactions is most likely deficient, so there is still room for new terms to be added.

5.1 Recent Advances

The consensus scoring field has been very active recently, with a number of interesting papers exploring new solutions and extending and/or confirming other workers' ideas and suggestions. Recently (Verdonk et al 2004), Verdonk et al. tested a number of different consensus combinations and confirmed Wang's (Wang and Wang 2001) insight that rank-by-number outperforms rank-by-rank and rank-by-vote strategies in consensus scoring. Wang (Wang et al 2002) and coworkers again, using three scoring functions built upon their previous work (SCORE (Wang et al 1998)) created a new consensus scoring method (X-CSCORE), calibrated to reproduce binding affinities of 200 complexes, and tested against thirty more.

Guo et al. (Guo et al 2004) built a consensus method using five commercially available functions, previously cited and all available within Sybyl: ChemScore, G-score, F-score, PMF-score and DrugScore. The five were combined not as a simple jury, but using multi-linear regression. The training set was composed of 53 inhibitors of *Torpedo Californica* AChE from PDB entries with known affinity re-docked into a Human AChE. 16 (different) compounds were used as a test set.

Similarly, Jacobsson (Jacobsson et al 2003) applied different training methods to 'multidimensional' consensus scoring. Partial least squares, discriminant analysis, bayesian classification, and rule-based methods were applied, to improve discrimination between active and inactive compounds. Their analysis goes beyond the 'crude' measure of enrichment, by explicitly defining four more measures for assessing scoring function performance, in addition to the usual enrichment factor.

Naive bayesian machine-learning algorithms are used by Klon and coworkers (Klon et al 2004) to improve enrichment in high-throughput docking of databases, by selecting the important features from the top ranked structures with extended connectivity fingerprints, and correlating these to their high-scoring compounds. The method works by re-ranking the docking outcome and retrieving compound similar to those in the top of the list from the remainder. It is unclear (to us) though, if this would have a detrimental effect on identifying 'novel' leads, unlike any known inhibitor *i.e.* does this approach inhibit scaffold hopping?

6. COMPARISONS BETWEEN SCORING FUNCTIONS

Bissantz et al. (Bissantz et al 2000; Bissantz et al 2003) first had the idea of comparing different scoring functions on the same test set. From then on, the issue of comparison between different functions has gained importance. Recently a number of papers tackled the issue from different viewpoints. These are not just sterile comparisons head-to-head, but instead a useful method to identify common deficiencies in modern scoring functions. Wang et al. (Wang et al 2003) tested 11 freely and commercially available scoring functions by reproducing the affinity of 100 protein-ligand complexes. Recently, the same authors (Wang et al 2004) extended the test, by using 14 different scoring functions and binding affinities of 800 complexes from the PDBBind (Wang et al 2005) database. Simulated virtual screening experiments were also performed, on HIV-1 protease, carbonic anhydrase, and trypsin complexes. Ferrara et al. (Ferrara et al 2004) assessed nine scoring functions for their ability to recognise the correct ligand orientation in 189 complexes from the LigandPDB (Roche et al 2001). Krovat and coworkers (Krovat and Langer 2004) applied LigandFit to the test case of renin: 10 inhibitors were mixed with 990 drug-like compounds, docked and then analysed with seven scoring functions: LigScore1, LigScore2, PLP1, PLP2, JAIN, PMF, LUDI. The consensus combination PLP1+PLP2+PMF performed the best, presenting all known inhibitors within the top 8%.

Other papers, although not strictly focused on scoring functions, but on docking in general, still offer useful insights: Kontoyianni et al. (Knotoyianni et al 2004) investigated the sensitivity of five docking programs (FlexX, GOLD, DOCK, LigandFit, GLIDE) to the nature of the active site. 69 receptors (belonging to 14 families) were used as targets. Perola and coworkers (Perola et al 2004) compared three docking programs, GLIDE, GOLD and ICM. The test focused not on reproduction of binding affinities or ranking of known inhibitors, but instead on the recognition of the correct binding mode; a mixed test set, composed of PDB and Vertex *in-house* complexes (100 of each) was used. The influence of the nature of the active site, and the effect of energy minimization was assessed. A further test compared the three programs on HIV-1 protease, p38 MAP Kinase and IMPDH. GLIDE was deemed the top-performer. DOCK, FlexX, GLIDE, GOLD, Slide, Surflex, and QXP have been tested by Kellenberger et al. (Kellenberger et al 2004) in a 100 X-ray small-ligand structure reproduction experiment, and in a fictitious SVS experiment. The three programs (GOLD, GLIDE and Surflex) performed well at both tasks. The authors criticised docking performance on the basis of active site features, and identified common failures. In Vigers et al. (Vigers and Rizzi 2004), once again, it is

shown how FlexX and GOLD correctly select the right ligand orientation, but fail to consistently rank a series of ligands. A statistical multiple active site correction (MASC) is presented. This correction lowers scores for ligand which are found to score well not only with the target of interest, but another selected set of seven to nine diverse protein structures, on the assumption that these ligands are 'tricking' the scoring function into believing that they are 'universal binders'. The correction is shown to improve ranking, but unfortunately it is specific to the docking and scoring algorithm used, and so must be re-generated for any new combination.

Concerning structure-based virtual screening techniques, Verdonk et al. (Verdonk et al 2004) advocate the use of 'focused' libraries, instead of general 'drug-like' libraries, when setting up SVS test beds, to avoid 'artificial' enrichment due to diversity between the 'true' ligand and decoys.

7. NEW TEST AND TRAINING SETS

With an ever increasing number of scoring functions, and comparisons between them, the need has emerged for a common test set. The ideal sets should include many diverse proteins, be representative of the 'protein space', independent of the bias inherent in the current PDB. This bias is due to a combination of some proteins being more readily crystallized, while other proteins are simply more topical, for example HIV-1 protease. The ligands to which the proteins are complexed should also be diverse, and arguably all should be drug-like (Lipinski et al 2001; Lipinski et al 1997) molecules. This has been an issue until recently, as many scoring functions were parameterised against complexes containing cofactors, or peptidic ligand, or sugars. In the past, many scoring function authors simply published the selection procedure or just the PDB entries for the proteins used during the training or testing phases, leaving the reader to repeat the work in order to gather similar data. With the increasing size of the PDB, knowledge-based scoring functions have become possible, and much effort has been put into selecting PDB entries with suitable features for such training sets. The need for a standard 'set' against which to test new scoring functions is however clear. This need was fulfilled by CCDC and Astex Therapeutics who in 2002 published a set of 305 proteins (Nissink et al 2002), containing all the PDB entries from the GOLD (Jones et al 1997) and ChemScore (Eldridge et al 1997) test-sets, plus others. This was not just a re-presentation of data though, as the original structures had been completely revised starting from the experimental electron density data. Moreover, tautomeric states for ligands and proteins have been assessed. This set was welcomed by users and developers, and in recent publications (Verdonk et al

2003; Mancera et al 2004) it has been used in the way its authors hoped, namely as a test or training set (Bindewald and Skolnick, 2005). The second, more recent attempt to produce a standard dataset is that of PDBbind (http://www.pdbbind.org), by Wang et al. (Wang et al 2005; Wang et al 2004) This set consists of 800 protein-ligand complexes culled from the PDB. The complexes all share consistent features, such as high-resolution structure, drug-like ligands, and availability of experimental affinity data. Ligand and proteins have been separated and saved in different files, and all are searchable through a web interface, upon (free) subscription. Also, submission of new affinity data collected from literature is possible, so this set is likely to grow and become a digest of the PDB for drug-developers. The PDBbind set has been used by its own authors for the comparative evaluation of 14 scoring functions (Wang et al 2004).

8. CONCLUSIONS

In this chapter, a broad review of the scoring field in the context of docking small molecules to protein targets has been presented. We have tried to cover the basics of the field, while at the same time presenting recent advances and trends. The three main classes of force-field, empirical and knowledge-based scoring functions have been presented, accompanied by sections explicitly devoted to consensus scoring, comparison between scoring functions, and collection of structural data with the specific purpose of scoring function testing.

REFERENCES

Ajay and Murcko M. A. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. J. Med. Chem. **38(26)**, 4953–4967.

Andrews P. R., Craik D. J. and Martin J. L. 1984. Functional-group contributions to drug receptor interactions. J. Med. Chem. **27(12)**, 1648–1657.

Atkins P. W. 1998. Physical chemistry. 6th ed. Oxford University Press.

Bindewald E. and Skolnick J. 2005. A scoring function for docking ligands to low-resolution protein structures. J. Comput. Chem. **26(4)**, 374–383.

Bissantz C., Bernard P., Hibert M. and Rognan D. 2003. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? Proteins: Structure, Function, and Genetics **50(1)**, 5–25.

Bissantz C., Folkers G. and Rognan D. 2000. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J. Med. Chem. **43(25)**, 4759–4767.

Bohm H. J. and Klebe G. 1996. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs? Angew. Chem. Int. Ed. Engl. **35(22)**, 2589–2614.

Bohm H. J. and Stahl M. 1999. Rapid empirical scoring functions in virtual screening applications. Med. Chem. Res. **9(7-8)**, 445–462.

Bohm H. J. 1994. The development of a simple empirical scoring function to estimate the binding constant for a protein ligand complex of known 3-dimensional structure. J. Comput. Aided Mol. Des. **8(3)**, 243–256.

Bostrom J., Norrby P. O. and Liljefors T. 1998. Conformational energy penalties of protein-bound ligands. J. Comput. Aided Mol. Des. **12(4)**, 383–396.

Brooks B. R., Bruccoleri R. E., Olafson B. D., States D. J., Swaminathan S. and Karplus M. 1983. Charmm - a program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. **4(2)**, 187–217.

Charifson P. S., Corkery J. J., Murcko M. A. and Walters W. P. 1999. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J. Med. Chem. **42(25)**, 5100–5109.

Chuaqui C., Deng Z. and Singh J. 2005. Interaction profiles of protein kinase-inhibitor complexes and their application to virtual screening. J. Med. Chem. **48(1)**, 121–133.

Cornell W. D., Cieplak P., Bayly C. I., Gould I. R., Merz K. M., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. W. and Kollman P. A. 1995. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J. Am. Chem. Soc. **117(19)**, 5179–5197.

Deng Z., Chuaqui C. and Singh J. 2004 Structural interaction fingerprint (SIFt): A novel method for analyzing three-dimensional protein-ligand binding interactions. J. Med. Chem. **47(2)**, 337–344.

DeWitte R. S. and Shakhnovich E. I. 1996. SMoG: De Novo design method based on simple, fast, and accurate free energy estimates.1. Methodology and supporting evidence. J. Am. Chem. Soc. **118(47)**, 11733–11744.

DeWitte R. S., Ishchenko A. V. and Shakhnovich E. I. 1997. SMoG: De novo design method based on simple, fast, and accurate free energy estimates.2. Case studies in molecular design. J. Am. Chem. Soc. **119(20)**, 4608–4617.

Eldridge M. D., Murray C. W., Auton T. R., Paolini G. V. and Mee1 R. P. 1997. Empirical scoring functions.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Comput. Aided Mol. Des. **11(5)**, 425–445.

Ewing T. J. A., Makino S., Skillman A. G. and Kuntz I. D. 2001. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. J. Comput. Aided Mol. Des. **15(5)**, 411–428.

Feher M., Deretey E. and Roy S. 2003. BHB: A simple knowledge-based scoring function to improve the efficiency of database screening. J. Chem. Inf. Comput. Sci. **43(4)**, 1316–1327.

Ferrara P., Gohlke H., Price D., Klebe G. and Brooks C. L. 2004. Assessing scoring functions for protein-ligand interactions. J. Med. Chem. **47(12)**, 3032–3047.

Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., Shelley M., Perry J. K. et al. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. **47(7)**, 1739–1749.

Gehlhaar D. K., Verkhivkerl G. M., Rejtol P. A. and Freer S. T. 1995. Molecular recognition of the inhibitor Ag-1343 by Hiv-1 Protease - conformationally flexible docking by evolutionary programming. Chem. Biol. **2(5)**, 317–324.

Giordanetto F., Cotesta S., Catana C., Trosset J.-Y., Vulpetti A., Stouten P. F. W., and Kroemer R. T. 2004. Novel scoring functions comprising QXP, SASA, and protein side-chain entropy terms. J. Chem. Inf. Comput. Sci. **44(3)**, 882–893.

Gohlke H., Hendlich M. and Klebe G. 2000. Knowledge-based scoring function to predict protein-ligand interactions. J. Mol. Biol. **295(2**), 337–356.

Gohlke H., Hendlich M. and Klebe G. 2000. Predicting binding modes, binding affinities and 'hot spots' for protein-ligand complexes using a knowledge-based scoring function. Perspect. Drug Discovery Des. **20(1)**, 115–144.

Greer J., Erickson J. W., Baldwin J. J. and Varney M. D. 1994. Application of the 3-dimensional structures of protein target molecules in structure-based drug design. J. Med. Chem. **37(8)**, 1035–1054.

Gunther J., Bergner A., Hendlich M. and Klebe G. 2003. Utilising structural knowledge in drug design strategies: Applications using relibase. J. Mol. Biol. **326(2)**, 621–636.

Guo J. X., Hurley M. M., Wright J. B. and Lushington G. H. 2004. A docking score function for estimating ligand-protein interactions: Application to acetylcholinesterase inhibition. J. Med. Chem. **47(22)**, 5492–5500.

Halgren T. A., Murphy R. B., Friesner R. A., Beard H. S., Frye L. L., Pollard W. T. and Banks J. L. 2004. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J. Med. Chem. **47(7)**, 1750–1759.

Halperin I., Ma B., Wolfson H. and Nussinov R. 2002. Principles of docking: An overview of search algorithms and a guide to scoring functions. Proteins: Structure, Function, and Genetics **47(4)**, 409–443.

Head R. D., Smythe M. L., Oprea T. I., Waller C. L., Green S. M. and Marshall G. R. 1996. VALIDATE: A new method for the receptor-based prediction of binding affinities of novel ligands. J. Am. Chem. Soc. **118(16)**, 3959–3969.

Hendlich M., Bergner A., Gunther J. and Klebe G. 2003. Relibase: Design and development of a database for comprehensive analysis of protein-ligand interactions. J. Mol. Biol. **326(2)**, 607–620.

Ishchenko A. V. and Shakhnovich E. I. 2001. SMall molecule growth (SMoG2001): An improved knowledge-based scoring function for protein-ligand interactions. J. Med. Chem. **45(13)**, 2770–2780.

Jacobsson M., Lidén P., Stjernschantz E., Boström H. and Norinder U. 2003. Improving structure-based virtual screening by multivariate analysis of scoring data. J. Med. Chem. **46(26)**, 5781–5789.

Jeffrey G. A. and Saenger W. 1991. Hydrogen bonding in biological structures. Springer-Verlag.

Jones G., Willett P., Glen R. C., Leach A. R. and Taylor R. 1997. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. **267(3)**, 727–748.

Kellenberger E., Rodrigo J., Muller P. and Rognan D. 2004. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. Proteins: Structure, Function and Bioinformatics **57(2)**, 225–242.

Klon A. E., Glick M., Thoma M., Acklin P. and Davies J. W. 2004. Finding more needles in the haystack: A simple and efficient method for improving high-throughput docking results. J. Med. Chem. **47(11)**, 2743–2749.

Kollman P. 1993. Free-energy calculations - applications to chemical and biochemical phenomena. Chem. Rev. **93(7)**, 2395–2417.

Kontoyianni M., McClellan L. M. and Sokol G. S. 2004. Evaluation of docking performance: Comparative data on docking algorithms. J. Med. Chem. **47(3)**, 558–565.

Krovat E. M. and Langer T. 2004. Impact of scoring functions on enrichment in docking-based virtual screening: An application study on renin inhibitors. J. Chem. Inf. Comput. Sci. **44(3)**, 1123–1129.

Kubinyi H. 1997. QSAR and 3D QSAR in drug design.1. Methodology. Drug Discovery Today **2(11)**, 457–467.

Kubinyi H. 1997. QSAR and 3D QSAR in drug design.2. Applications and problems. Drug Discovery Today **2(12)**, 538–546.

Kuntz I. D., Blaney J. M., Oatley S. J., Langridge R. and Ferrin T. E. 1982. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. **161(2)**, 269–288.

Kuntz I. D., Chen K., Sharp K. A. and Kollman P. A. 1999. The maximal affinity of ligands. Proc. Natl. Acad. Sci. U. S. A. **96(18)**, 9997–10002.

Laederach A. and Reilly P. J. 2003. Specific empirical free energy function for automated docking of carbohydrates to proteins. J. Comput. Chem. **24(14)**, 1748–1757.

Leach A. R. 2001. Molecular modelling: principles and applications. 2nd ed. Prentice Hall.

Lipinski C. A., Lombardo F., Dominy B. W. and Feeney P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **46(1-3)**, 3–26.

Lipinski C. A., Lombardo B. W., Dominy P. J. and Feeney 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. **23(1-3)**, 3–25.

Liu H. Y., Kuntz I. D. and Zou X. Q. 2004. Pairwise GB/SA scoring function for structure-based drug design. J. Phys. Chem. B **108(17)**, 5453–5462.

Luo R. and Gilson M. K. 2000. Synthetic adenine receptors: Direct calculation of binding affinity and entropy. J. Am. Chem. Soc. **122(12)**, 2934–2937.

Majeux N., Scarsi M., Apostolakis J., Ehrhardt C. and Caflisch A. 1999. Exhaustive docking of molecular fragments with electrostatic solvation. Proteins: Structure Function and Genetics **37(1)**, 88–105.

Mancera R. L., Kallblad P. and Todorov N. P. 2004. Ligand-protein docking using a quantum stochastic tunneling optimization method. J. Comput. Chem. **25(6)**, 858–864.

Mardis K. L., Luo R. and Gilson M. K. 2001. Interpreting trends in the binding of cyclic ureas to HIV-1 protease. J. Mol. Biol. **309(2)**, 507–517.

Mark A. E. and Van Gunsteren W. F. 1994. Decomposition of the free-energy of a system in terms of specific interactions - implications for theoretical and experimental studies. J. Mol. Biol. **240(2)**, 167–176.

McMartin C. R. S. and Bohacek 1997. QXP: Powerful, rapid computer algorithms for structure-based drug design. J. Comput. Aided Mol. Des. **11(4)**, 333–344.

Meng E. C., Shoichet B. K. and Kuntz I. D. 1992. Automated docking with grid-based energy evaluation. J. Comput. Chem. **13(4)**, 505–524.

Mitchell J. B. O., Laskowski R. A., Alex A. and Thornton J. M. 1999. BLEEP - Potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data. J. Comput. Chem. **20(11)**, 1177–1185.

Mitchell J. B. O., Laskowski R. A., Alex A., Forster M. J. and Thornton J. M. 1999. BLEEP - Potential of mean force describing protein-ligand interactions: I. Generating potential. J. Comput. Chem. **20(11)**, 1165–1176.

Mizutani M. Y., Tomioka N. and Itai A. 1994. Rational automatic search method for stable docking models of protein and ligand. J. Mol. Biol. **243(2)**, 310–326.

Muegge I. Y. C. and Martin 1999. A general and fast scoring function for protein-ligand interactions: A simplified potential approach. J. Med. Chem. **42(5)**, 791–804.

Muegge I. 2000. A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. Perspect. Drug Discovery Des. **20(1)**, 99–114.

Muegge I. 2001. Effect of ligand volume correction on PMF scoring. J. Comput. Chem. **22(4)**, 418–425.

Muegge I., Martin Y. C., Hajduk P. J. and Fesik S. W. 1999. Evaluation of PMF scoring in docking weak ligands to the FK506 binding protein. J. Med. Chem. **42(14)**, 2498–2503.

Muegge I. 1999. Protein-ligand docking with a knowledge-based scoring function. Abstracts of Papers of the American Chemical Society **218**, U493-U493.

Muegge I. A. 2004. PMF scoring revisited. Abstracts of Papers of the American Chemical Society **228**, U515-U515.

Muryshev A. E., Tarasov D. N., Butygin A. V., Butygina O. Y., Aleksandrov A. B. and Nikitin S. M. 2003. A novel scoring function for molecular docking. J. Comput. Aided Mol. Des. **17**, 597–605.

Nissink J. W. M., Murray C., Hartshorn M., Verdonk M. L., Cole J. C. and Taylor R. 2002. A new test set for validating predictions of protein-ligand interaction. Proteins: Structure, Function, and Genetics **49(4)**, 457–471.

Pan Y. P., Huang N., Cho S., MacKerell A. D. and Jr. 2003. Consideration of molecular weight during compound selection in virtual target-based database screening. J. Chem. Inf. Comput. Sci. **43(1)**, 267–272.

Pearlman D. A. and Charifson P. S. 2001. Improved scoring of ligand-protein interactions using OWFEG free energy grids. J. Med. Chem. **44(4)**, 502–511.

Pearlman D. A. 1999. Free energy grids: a practical qualitative application of free energy perturbation to ligand design using the OWFEG method. J. Med. Chem. **42(21)**, 4313–4324.

Perola E., Walters W. P. and Charifson P. S. 2004. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. Proteins: Structure, Function, and Bioinformatics **56(2)**, 235–249.

Raha K. and Merz K. M. 2004. A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding. J. Am. Chem. Soc. **126(4)**, 1020–1021.

Raha K. and Merz K. M. 2005. Large-scale validation of a quantum mechanics based scoring function: predicting the binding affinity and the binding mode of a diverse set of protein-ligand complexes. J. Med. Chem. **48(14)**, 4558–4575.

Raha K. and Merz K. M. 2004. Zinc mediated ligand binding: a quantum mechanics based approach. Abstracts of Papers of the American Chemical Society **227**, U1016-U1017.

Raha K., van der Vaart A. J., Riley K. E., Peters M. B., Westerhoff L. M., Kim H., Merz K. M. and Jr. 2005. Pairwise decomposition of residue interaction energies using semi-empirical quantum mechanical methods in studies of protein-ligand interaction. J. Am. Chem. Soc. **127(18)**, 6583–6594.

Roche O., Kiyama R. and Brooks C. L. 2001. Ligand-protein database: Linking protein-ligand complex structures to binding data. J. Med. Chem. **44(22)**, 3592–3598.

Shoichet B. K., Leach A. R. and Kuntz I. D. 1999. Ligand solvation in molecular docking. Proteins: Structure, Function, and Genetics **34(1)**, 4–16.

Smith R., Hubbard R. E., Gschwend D. A., Leach A. R. and Good A. C. 2003. Analysis and optimization of structure-based virtual screening protocols (3). New methods and old problems in scoring function design. J. Mol. Graph. Model. **22(1)**, 41–53.

Sotriffer C., Klebe G., Stahl M. and Bohm H. J. 2004. Docking and Scoring Functions/Virtual Screening, in Burger's Medicinal Chemistry and Drug Discovery. pp. 281–333.

Stahl M. and Rarey M. 2001. Detailed analysis of scoring functions for virtual screening. J. Med. Chem. **44(7)**, 1035–1042.

Sussman J. L., Abola E. E., Lin D., Jiang J., Manning N. O. and Priluskyl J. 1999. The protein data bank - Bridging the gap between the sequence and 3D structure world. Genetica **106(1-2)**, 149–158.

Taylor R. D., Jewsbury P. J. and Essex J. W. 2002. A review of protein-small molecule docking methods. J. Comput. Aided Mol. Des. **16(3)**, 151–166.

Todorov N. P., Mancera R. L., Kallblad P. and Monthoux P. 2002. EasyDock: A new docking program for high-throughput screening and binding-mode search. Abstracts of Papers of the American Chemical Society **224**, U345-U345.

Todorov N. P., Mancera R. L. and Monthoux P. H. 2003. A new quantum stochastic tunnelling optimisation method for protein-ligand docking. Chem. Phys. Letters **369(3-4)**, 257–263.

Velec H. F. G., Gohlke H. and Klebe G. 2005. DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. J. Med. Chem. **48(20)**, 6296–6303.

Verdonk M. L., Cole J. C., Hartshorn M. J., Murray C. W. and Taylor R. D. 2003. Improved protein-ligand docking using GOLD. Proteins: Structure, Function, and Genetics **52(4)**, 609–623.

Verdonk M. L., Chessari G., Cole J. C., Hartshorn M. J., Murray C. W., Nissink J. W. M., Taylor R. D. and Taylor R. 2005. Modeling water molecules in protein-ligand docking using GOLD. J. Med. Chem. **48**, 6504–6515.

Verdonk M. L., Berdini V., Hartshorn M. J., Mooij W. T. M., Murray C. W., Taylor R. D. and Watson P. 2004. Virtual screening using protein-ligand docking: Avoiding artificial enrichment. J. Chem. Inf. Comput. Sci. **44(3)**, 793–806.

Verkhivker G., Appelt K., Freer S. T. and Villafranca J. E. 1995. Empirical free-energy calculations of ligand-protein crystallographic complexes.1. Knowledge-based ligand-protein interaction potentials applied to the prediction of human-immunodeficiency-virus-1 protease binding-affinity. Protein Eng. **8(7)**, 677–691.

Vieth M., Hirst J. D., Kolinski A. and Brooks C. L. 1998. Assessing energy functions for flexible docking. J. Comput. Chem. **19(14)**, 1612–1622.

Vigers G. P. A. and Rizzi J. P. 2004. Multiple active site corrections for docking and virtual screening. J. Med. Chem. **47(1)**, 80–89.

Wang R. X. and Wang S. M. 2001. How does consensus scoring work for virtual library screening? An idealized computer experiment. J. Chem. Inf. Comput. Sci. **41(5)**, 1422–1426.

Wang R. X., Lu Y., Fang X. and Wang S. 2004. An extensive test of 14 scoring functions using the PDBbind refined set of 800 protein-ligand complexes. J. Chem. Inf. Comput. Sci. **44(6)**, 2114–2125.

Wang R. X., Lu Y., Fang X. and Wang S. 1998. SCORE: A new empirical method for estimating the binding affinity of a protein-ligand complex. J. Mol. Model. **4(12)**, 379–394.

Wang R. X., Fang X., Lu Y. and Wang S. 2004. The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. J. Med. Chem. **47(12)**, 2977–2980.

Wang R. X., Fang X., Lu Y. and Wang S. 2005. The PDBbind database: Methodologies and updates. J. Med. Chem. **48(12)**, 4111–4119.

Wang R. X., Lai L. H. and Wang S. M. 2002. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. J. Comput. Aided Mol. Des. **16(1)**, 11–26.

Wang R. X., Lu Y. P. and Wang S. M. 2003. Comparative evaluation of 11 scoring functions for molecular docking. J. Med. Chem. **46(12)**, 2287–2303.

Yang J. M. and Chen C. C. 2004. GEMDOCK: A generic evolutionary method for molecular docking. Proteins: Structure, Function, and Bioinformatics **55(2)**, 288–304.

Zou X. Q., Sun Y. X. and Kuntz I. D. 1999. Inclusion of solvation in ligand binding free energy calculations using the generalized-born model. J. Am. Chem. Soc. **121(35)**, 8033–8043.

SUBJECT INDEX