Nicholas A. Meanwell   *Editor*

Tactics in Contemporary Drug Design

Springer

9
Topics in Medicinal Chemistry

Editorial Board:

Aims and Scope

Drug research requires interdisciplinary team-work at the interface between chemistry, biology and medicine. Therefore, the new topic-related series *Topics in Medicinal Chemistry* will cover all relevant aspects of drug research, e.g. pathobiochemistry of diseases, identification and validation of (emerging) drug targets, structural biology, drugability of targets, drug design approaches, chemogenomics, synthetic chemistry including combinatorial methods, bioorganic chemistry, natural compounds, high-throughput screening, pharmacological in vitro and in vivo investigations, drug-receptor interactions on the molecular level, structure-activity relationships, drug absorption, distribution, metabolism, elimination, toxicology and pharmacogenomics.

In general, special volumes are edited by well known guest editors.

In references *Topics in Medicinal Chemistry* is abbreviated *Top Med Chem* and is cited as a journal.

Nicholas A. Meanwell
Editor

Tactics in Contemporary Drug Design

With contributions by

U.M. Hanumegowda · J.G. Kenna · A.G. Leach ·
D.J. Leishman · N.A. Meanwell · T. Noeske · Z. Rankovic ·
A. Regueiro-Ren · M.W. Sinz · S.H. Stahl · M.A. Walker ·
R.J. Young

Springer

Editor
Nicholas A. Meanwell
Bristol-Myers Squibb
Wallingford
Connecticut
USA

Preface to the Series

Medicinal chemistry is both science and art. The science of medicinal chemistry offers mankind one of its best hopes for improving the quality of life. The art of medicinal chemistry continues to challenge its practitioners with the need for both intuition and experience to discover new drugs. Hence sharing the experience of drug discovery is uniquely beneficial to the field of medicinal chemistry.

The series Topics in Medicinal Chemistry is designed to help both novice and experienced medicinal chemists share insights from the drug discovery process. For the novice, the introductory chapter to each volume provides background and valuable perspective on a field of medicinal chemistry not available elsewhere. Succeeding chapters then provide examples of successful drug discovery efforts that describe the most up-to-date work from this field.

The editors have chosen topics from both important therapeutic areas and from work that advances the discipline of medicinal chemistry. For example, cancer, metabolic syndrome and Alzheimer's disease are fields in which academia and industry are heavily invested to discover new drugs because of their considerable unmet medical need. The editors have therefore prioritized covering new developments in medicinal chemistry in these fields. In addition, important advances in the discipline, such as fragment-based drug design and other aspects of new lead-seeking approaches, are also planned for early volumes in this series. Each volume thus offers a unique opportunity to capture the most up-to-date perspective in an area of medicinal chemistry.

<div align="right">

Dr. Peter R. Bernstein
Prof. Dr. Armin Buschauer
Prof. Dr. Gunda I. Georg
A. K. Saxena
Dr. Claudiu T. Supuran
Dr. John Lowe
Dr. Hans Ulrich Stilz

</div>

Preface

Contemporary drug discovery practices have evolved considerably over the last 30 years as our knowledge and understanding of the attributes and profiles of successful drugs have both deepened and become more sophisticated. These advances have contributed to the effort to enhance compound durability and improve overall success rates in an industry where drug candidate failure is still overwhelmingly the more common outcome of a development program. In 1990, the principal reasons for drug failure were distributed almost evenly between toxicity, lack of efficacy, and poor pharmacokinetic (PK) properties. Advances in preclinical profiling have enhanced the ability to predict human PK parameters to an extent that this has been reduced significantly as a source of drug failure. Despite this development, drug discovery output over the last 25 years has remained stubbornly low, with failure rates for small molecules not changing significantly as toxicity and lack of efficacy have emerged as the major sources of candidate demise. However, improved preclinical practices are more effectively selecting drug candidates suitable for testing a particular mechanistic hypothesis in Phase II clinical trials. Clinical success is dependent upon demonstrating exposure of a drug at the site of action, the engagement of the target and an expression of the pharmacological effect that is anticipated based on the underlying theory. A heightened awareness of drug physical properties, developability issues and side effects is reflected by the implementation of a broad range of in vitro and in vivo assays designed to remove candidate compounds with a higher potential to fail or to precipitate unacceptable effects in a clinical setting.

This volume of the *Topics in Medicinal Chemistry* series focuses on exploring tactical approaches to solving problems in drug design that affect candidate developability and quality and which are encountered frequently in drug development programs. In the chapter "Physical Properties in Drug Design" Robert J. Young summarizes the importance of controlling and modulating physical properties during the drug design and optimization phase. This is an area of fundamental importance that has been the subject of considerable scrutiny by the medicinal chemistry community over the last decade as it has focused an introspective lens on developing a deeper understanding of what factors drive compound durability.

Designing better molecules at the outset requires a fuller appreciation of the factors governing drug–target interactions and the several efficiency metrics that have been devised provide useful guideposts for the medicinal chemist. Ligand lipophilic efficiency (LLE or LipE) has emerged as one of the most prominent and effective means of assessing compound quality by providing a basic understanding of the contribution of the lipophilicity of a molecule to the affinity for its cognate target. By routinely applying this metric during the analysis of structure–activity relationships, the current trend in drug design of depending on flat, sp^2 center-rich, lipophilic molecules that derive much of their potency from non-specific entropic effects rather than more specific enthalpy-based drug–target interactions may be reversed.

Drug solubility is fundamentally dependent on the inherent physical properties of a molecule and contemporary practices are clearly producing poorly soluble compounds that require the application of sophisticated formulation strategies to facilitate drug delivery in vivo. Indeed, there is a burgeoning reliance on spray-dried dispersions to deliver drug candidates both pre-clinically and clinically, a reflection on the chemical space being explored in many discovery programs. This may be a function of the nature of a biological target or the design practices of the medicinal chemist or a combination of these phenomena. In the chapter "Improving Solubility via Structural Modification" Michael A. Walker reviews the importance of solubility in drug delivery and provides a synopsis of approaches that have been applied to the design of drug candidates with enhanced pharmaceutical properties.

Inhibitors and inducers of the cyctochrome P450 (CYP 450) enzymes, important components of drug metabolism, are perpetrators of drug–drug interactions. This is of particular importance in an era when patients take some drugs chronically and may be prescribed with additional drugs that can be victims in a drug–drug interaction. CYP 450 inhibition can be of sufficient severity that the exposure of a victim drug can increase to levels where toxicity is evident; alternatively, efficacy can be severely compromised when CYP 450 induction occurs. In the chapter "Tactics to Avoid Inhibition of Cytochrome P450s" Andrew G. Leach examines CYP 450 structure and function and summarizes some of the strategies and tactics that have been employed to avoid CYP 450 inhibition. The pregnane X and constitutive aryl hydrocarbon nuclear receptors are important mediators of CYP 450 induction that sense xenobiotics and activation of these receptors by drugs is also a potential source of drug–drug interactions. In the chapter "Avoiding PXR and CAR Activation and CYP3A4 Enzyme Induction" Michael W. Sinz examines examples of drug molecules that activate these receptors and the current understanding of structure–activity relationships. In this chapter, an emphasis is placed on a discussion of the changes in molecular structure of modulators that have been observed to impact affinity and function.

The chapters "Strategies for Minimisation of the Cholestatic Liver Injury Liability Posed by Drug-Induced Bile Salt Export Pump (BSEP) Inhibition," "Drug Discovery vs hERG," and "Drug-Induced Phospholipidosis: Prediction, Detection, and Mitigation Strategies" focus attention on three prominent off-target effects of drug candidates that can be a significant source of drug toxicity and lead to

termination of a drug development program. Much of the increase in our understanding of these problems has been based on clinical observations and an iterative cycle of feedback to drug discovery. This has led to the implementation of preclinical screens designed to identify problematic molecules and eliminate them from consideration early in the drug discovery process. Drug-mediated inhibition of the bile salt export pump (BSEP) can precipitate cholestatic liver injury and this is an area of emerging concern. In the chapter "Strategies for Minimisation of the Cholestatic Liver Injury Liability Posed by Drug-Induced Bile Salt Export Pump (BSEP) Inhibition" J. Gerry Kenna and his colleagues capture the current understanding of the biochemical pharmacology associated with this transporter and attempt to equate inhibitory potency with the potential for the observation of clinical problems. This is a particularly daunting task given the current incomplete understanding of inhibitors of this transporter. The implementation of screening for inhibitors of the human ether-à-go-go gene product, a rapidly activating potassium channel that mediates the repolarizing I_{Kr} current in the cardiac action potential and commonly known as hERG, was precipitated by clinical observations with the antihistamine terfenadine. Terfenadine is a prodrug of the carboxylic acid metabolite fexofenadine and coadminstration of this drug with the antifungal agent ketoconazole, a CYP P450 inhibitor, led to elevated levels of terfenadine which is a potent inhibitor of hERG channel function. Inhibition of the hERG channel leads to prolongation of the QT interval in the electrocardiogram which can take on a unique shape described as *torsades de pointes*. This can be manifested as a polymorphic ventricular tachycardia and lead to arrhythmias and, possibly, sudden death. In the chapter "Drug Discovery vs hERG" Zoran Rankovic and Derek J. Leishman discuss the structure of the hERG channel and describe the screening methodology that has become a mandatory part of the preclinical drug profiling landscape. They proceed to delineate the pharmacophore that has been developed for pore inhibitors and evaluate and summarize tactical approaches to structural modification of compounds that have been shown to reduce the propensity for hERG inhibition. Phospholipidosis is a disorder of lysosomal storage that is characterized by the accumulation of phospholipids in tissues. This phenomenon has been most prominently associated with lipophilic basic amines and many of structures that precipitate phospholipidosis overlap with the hERG inhibitor pharmacophore. In the chapter "Drug-Induced Phospholipidosis: Prediction, Detection, and Mitigation Strategies" Umesh M. Hanumegowda and Alicia Regueiro-Ren analyze this problem, summarize the physical chemical features associated with its occurrence and describe some of the strategies and tactics that have been successfully applied to mitigate the phenomenon in the design of drug candidates.

In the final chapter, "The Influence of Bioisosteres in Drug Design: Tactical Applications to Address Developability Problems," the application of effective structural mimics of a range of functional groups, commonly described as bioisosteres, to the process of drug design and optimization is discussed. The application of bioisosteres can be a useful approach to address many of the problems commonly encountered in contemporary drug discovery programs. In this chapter, these are annotated based on the nature of the problem encountered rather than by the more

traditional approach of grouping based on the nature of the functional group being emulated. However, the application of bioisosteres is very much contextual in nature with structural emulation highly dependent on the specific drug–target interactions under consideration.

I would very much like to express my gratitude to the authors for taking the time and care to compose and contribute their chapters to this enterprise. This volume will be successful if their thoughts and insights inspire ideas, concepts, approaches, and strategies that have utility in solving problems associated with drug optimization campaigns. Application of the principles that are described within these pages to solve specific problems is rarely completely repetitive in nature and, as most commonly practiced, requires some level of interpretation and improvisation to be successfully applied to a new context. This has been and will continue to be a natural part of the evolution of drug design practices and should be viewed as a source of intellectual stimulation rather than a limitation. Finally, I would like to thank Peter R. Bernstein for planting the seed for this book and his encouragement and support throughout the project.

Wallingford, CT, USA   Nicholas A. Meanwell
January 2014

Contents

Physical Properties in Drug Design

Robert J. Young

Abstract The physical properties of investigational molecules in drug discovery programmes have been the subjects of intense scrutiny, largely due to a propensity for the pursuit of examples where they are sub-optimal. This chapter introduces the noteworthy contributions that identified the shortcomings and then defines and discusses the key physical parameters (lipophilicity, pK_a and solubility) and contemporary developments in their measurement and use. These physical characteristics impact the passage of a drug molecule from the administered dose to the site of action, profoundly influencing its pharmacokinetics and pharmacology. In particular, lipophilicity has a major influence on various parameters used to assess the developability of experimental molecules; the additional impact of aromaticity or flatness in structures and differentiation between the roles intrinsic ($\log P$) and effective ($\log D$) are also illustrated. In conclusion, the combined influences of good properties in efficient molecules are presented as powerful indicators of quality.

Keywords Physical properties, Lipophilicity, Hydrophobicity, pKa, Solubility, log P, log D, Chromatographic hydrophocbity measurements, Aromaticity, Property forecast index, Drug efficiency, Ligand efficiency, Ligand lipophilicity Efficiency, Developability classification system, Thermodynamics, Structure property relationships, Developability, Attrition, Permeation, Cytochrome P450, hERG, Promiscuity, Plasma protein binding

Contents

R.J. Young (✉)
MDR Medicinal Chemistry, GlaxoSmithKline R&D, Stevenage, Hertfordshire, UK SG1 2NY
e-mail: Rob.J.Young@gsk.com

Abbreviations

ADME (or ADMET)	Absorption, distribution, metabolism and elimination
BBB	Blood–brain barrier
BEI	Binding efficiency index
CHI	Chromatographic hydrophobicity index
CLND	Chemiluminescent nitrogen detection
cmr	Calculated molar refraction
DCS	Developability classification system
DMPK	Drug metabolism and pharmacokinetics
FaSSIF	Fasted state simulated intestinal fluids
FeSSIF	Fed state simulated intestinal fluids
GSE	General solubility equation
GSK	GlaxoSmithKline
hERG	Human ether-a-go-go-related gene
HSA	Human serum albumin
IAM	Immobilised artificial membrane
ITC	Isothermal titration calorimetry
LE	Ligand efficiency
LLE	Ligand lipophilicity efficiency
MPbAP	Melting point based absorption potential
OW	Octanol/water
PAMPA	Parallel artificial membrane permeation assays
PFI	Property forecast index

QED	Quantitative estimate of drug-likeness
QSAR	Quantitative structure activity relationships
QSPR	Quantitative structure property relationships
SGF	Simulated gastric fluid
SILE	Size-independent ligand efficiency

1 Introduction: The Realisation of Physical Property Issues

Physical properties have returned to the forefront of contemporary thinking and practice in medicinal chemistry [1]; it is logical to think that any drug molecule must possess some level of water solubility to enable systemic exposure via aqueous media to reach its target and that its physical make-up must have complementary features to engage with that target. The aqueous environment of the digestive tract, blood and intracellular environments is opposite to the more lipophilic environments of the sites of drug action, where recognition will require a particular combination of polar and hydrophobic interactions. The appreciation of such concepts is not new; Hansch wrote, in 1987, "without convincing evidence to the contrary, drugs should be made as hydrophilic as possible without loss of efficacy" [2]. Of many contributions to the field, Hansch was a pioneer in the application of physical aspects of medicinal chemistry in drug discovery, particularly aspects of physical organic chemistry applicable to the generation of quantitative structure activity relationships (QSAR), which were particularly in vogue from the 1970s onwards (Fig. 1) [3]. During these years, drug discovery was largely driven by mimicking of the ligands or substrates of the pharmacological target; this was often complemented by the use of natural products as the starting points. The early 1990s saw the advent of protein crystallography and allied computational methods, which provided strong structural bases for designing effective molecules with optimal binding characteristics. However, what followed arguably contributed to a perfect storm, caused by changes in practices, cultures and inflation of properties. Increasingly, the first exposure of a test compound to an intact tissue or whole animal test, for many years the bedrocks of pharmacological evaluation, was much further down a screening cascade – primary screening was often performed on an isolated recombinant protein, whereby the focus shifted to a binding or inhibition constant rather than a functional efficacy. Such assays enabled higher throughput screening and large numbers of compounds generated by combinatorial and high-throughput chemistry methods. This led to the generation of productivity metrics, wherein quantity and potency were believed to drive success, distracting attention away from the physical quality of synthetic molecules. The quality of leads surfacing from high-throughput screening campaigns exacerbated this trend, whereby the most potent molecules were pursued regardless of properties or ligand efficiency. It was not until towards the end of the 1990s that Lipinski's seminal publication, proposing the Rule of 5 [4], made the first attempt to alter the trend, by suggesting restraint in the size, lipophilicity and hydrogen-bonding
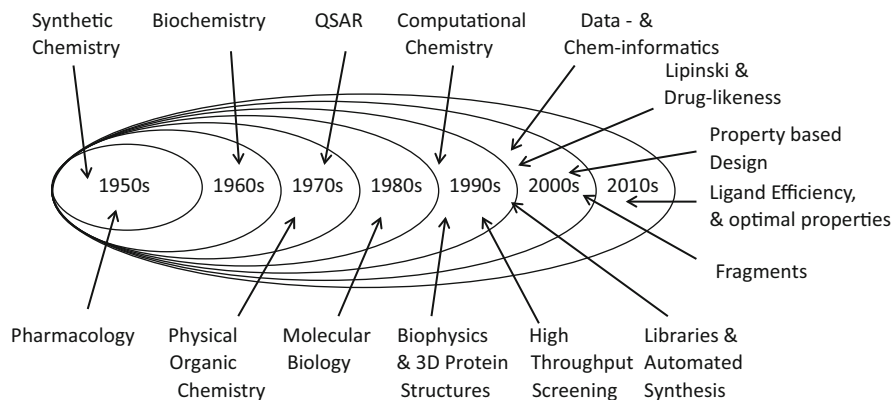
Fig. 1 The evolution of technologies in medicinal chemistry over the past six decades, updated from Hann's original version [3] and gratefully reproduced with his permission

character of experimental molecules. The combination of both structure- and property-based drug design began to take hold in the early 2000s, but many candidate molecules were relatively large and lipophilic.

1.1 Molecular Obesity and Obsession with Potency

The watershed publication of Leeson and Springthorpe [5] provided a particular focus on the impact of excessive lipophilicity on attrition (Fig. 2); subsequent studies and analyses cited in the following sections have corroborated these findings and/or have provided insight on particular consequences of sub-optimal physical properties. This lipophilic inflation was succinctly described by the term "molecular obesity" by Hann [6] in a review exploring the origins of excessive lipophilicity, which was diagnosed as an obsession with potency that is commonly and easily achieved by resorting to more lipophilic compounds, leading to a premature demise in development. As this chapter develops, the impact of this trend on various developability parameters becomes clear; lipophilicity and solubility (a highly desirable characteristic for a drug molecule to have) are intrinsically intertwined but their impacts have much wider connotations. This renewed focus on physical properties and the advent of fragment-based drug design in recent years [7–9] appear to suggest the era of ligand efficiency [10] combined with optimal physical properties will define the drug discovery processes of the second decade of the twenty-first century [11]. Candidate quality resulting from lead optimisation stages in the drug discovery process has also come under increased scrutiny; the importance of physical quality [1, 12], structure–property relationships [13] and ligand lipophilicity efficiency [14] are important concepts that are developed in this
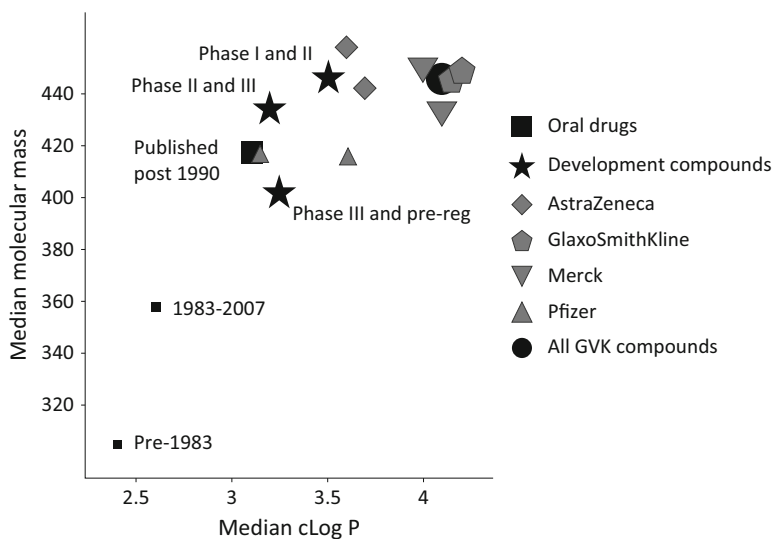
Fig. 2 The increase in molecular weight and clog *P* in contemporary drug discovery adapted from Leeson and Springthorpe's analysis [5] of the overall trends in median clog *P* and molecular mass in compounds from four pharmaceutical companies from Prous/GVK databases (hence two spots for each). The figure compares the data for historical oral drugs and compounds in development to those in current patents from the four companies indicated; the size of markers indicates the average number of Lipinski violations [4]

chapter. In all stages of drug discovery, it is becoming apparent that ligand efficiency and good properties are just as important as potency [15].

1.2 Med Chem Inflates Properties: The Body Doesn't Change

Given that the average lipophilicity of investigational drug molecules has increased in recent years [5], the average aqueous solubility will, inexorably, have declined [16 and Sect. 2.3]. However, something that has not changed is that some dissolution is necessary irrespective of the route of administration, most of all if oral dosing is required. The human body has evolved over millions of years to deal with xenobiotics and, whilst there can be particular variations between individuals – e.g. cytochrome P450 polymorphisms and genetic absence of particular metabolising enzymes [17] – human physiology has not overtly changed to an extent that would be accommodating to more lipophilic molecules [18]. In particular, the natural clearance mechanisms of metabolism and/or conjugation have a primary role to solubilise fatty or aromatic regions of molecules to facilitate more rapid exclusion, be this as a physiological role or in the removal of a xenobiotic. Whilst there are drugs on the market that occupy the extremes of the normal

Fig. 3 Distribution of activity values for 261 oral drug molecules adapted from Gleeson et al. [15], showing the surprisingly low average values (mean value of 7.3 and a median of 7.7)

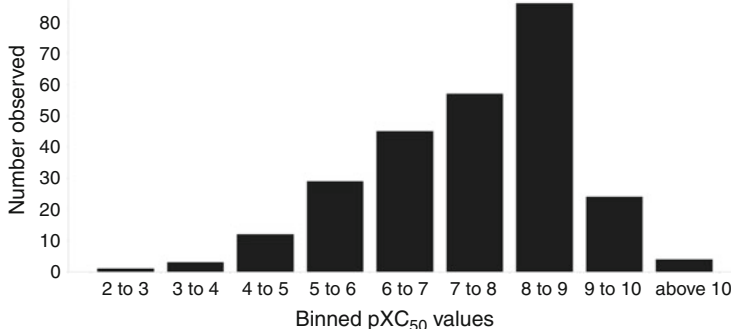distribution curve of physical properties [19], the vast majority lie within a fairly well-defined region where they are, to paraphrase Hansch, "hydrophilic enough, whilst possessing sufficient activity for a pharmacological effect at their target". The affinities of licensed drug molecules are surprisingly low on average (Fig. 3), indicative of how balancing each of potency, pharmacokinetics and pharmaco-dynamic efficacy is the key to successful medicinal chemistry. The environment in which these pharmacokinetic and pharmacological properties are manifested, the human body, is not overtly changing in its fundamental biochemistry. These biochemical mechanisms are the culmination of the evolutionary process, which has honed an exquisite system to deal with xenobiotics and regulate endogenous active molecules. It could be argued that successful drugs are those honed to subjugate particular aspects of the systems, having an effect without overtly affecting the general operations. The physical characteristics of molecules are the most useful metrics available to describe "drug-like space", in terms of parameters such as size, weight, charge and lipophilicity [20] that can be used to quantify the behaviour and interactions of the molecule [1]. A culmination of the many descrip-tors used to parameterise drug likeness has been the development of a quantitative estimate of drug-likeness (QED) [21].

2   Physical Properties: What Are They and Which Are Important?

This subsection gives details on lipophilicity/hydrophobicity, pK_a and solubility, the most important and relevant physical parameters in drug discovery, in some detail. These parameters are interdependent; how such relationships are manifested is described with illustrative examples. The measurement of physical properties [22] is a rapidly evolving arena in terms of techniques, technology and automation; these has been well reviewed [23, 24], so the focus of the following sections is more on particular aspects pertinent to drug design.

2.1 Lipophilicity/Hydrophobicity

In a human body, the physical characteristics of compartments encountered by drug molecules are defined within the extremes of aqueous and hydrocarbon-rich (lipid) environments. Indeed, pharmacology and pharmacokinetics are driven by the particular behaviour of molecules in such surroundings and their distribution between them. Chemists are familiar with the partitioning of molecules between water and immiscible organic solvents – the basis for quantifying this behaviour in drug discovery [25].

Hydrophilic (from the Greek hydro = water and philos = loving) is the term used to describe compounds with a propensity for aqueous solubility. The etymologically opposite term is hydrophobic (Greek phobos = fear) – and hydrophobicity is a commonly used term in this sense, although the IUPAC definition of hydrophobicity is "the association of non-polar groups or molecules in an aqueous environment which arises from the tendency of water to exclude non-polar molecules". The physiological opposite of the aqueous environment is represented by the hydrocarbon-rich chains of lipids (fats); the term lipophilic (Greek lipos = fat, grease) is used to describe the compounds preferring to reside in such an environment, often interchangeably with hydrophobic. IUPAC defines lipophilicity as "the affinity of a molecule or a moiety for a lipophilic environment". The missing term in this etymological quartet, lipophobic, is rarely encountered.

The partitioning of molecules between aqueous buffers and an immiscible solvent has been known to have relevance to drug action for many years and many experimental combinations have been reported [26]. It was the Hansch group in 1964 [27] that proposed *n*-octan-1-ol and aqueous buffer as the standard system with most relevance to drug discovery [28].

The *partition coefficient*, *P* or usually expressed as log *P* ($\log_{10} P_{OW}$), is measured from the equilibrium concentrations of a sample dispersed between the layers of *n*-octan-1-ol and aqueous buffer, after the mixture has been well shaken and then allowed to reach equilibrium.

Log *P*, the ***intrinsic lipophilicity/hydrophobicity*** **of a compound, is a constant**.

If a molecule has no ionisable centres, partitioning is straightforward to comprehend as the concentration ratio between the phases does not vary with the pH of the aqueous solution. However, most drug molecules do have one or more ionisable centres, and pH profoundly influences the distribution of neutral and charged forms between the phases. Yet all compounds have a log *P* value, representing the partitioning of the unionised form of the molecule. This is also a constant and can be measured with a buffer pH significantly below (acids) or above (bases) the pK_a of the ionisable centre in the particular molecule. A second asymptote observed under the opposite conditions (pH $\gg pK_a$ for acids and pH $\ll pK_a$ for amines) equates to the log *P* of the fully ionised form. Of course, the distribution behaviour becomes more complex for two or more ionisable centres, as log *P* may be hypothetical as a neutral species might not exist at any pH; however the value can be estimated by extrapolation.

Partition coefficient, P = concentration in octanol/concentration in water

$$P = \frac{[\text{solute}]_{\text{octanol}}}{[\text{solute}]_{\text{water}}}$$

Taking ionisable centres into account leads to the *distribution coefficient*, D, which represents the distribution of all species of a molecule between the aqueous phase and immiscible solvent at a given pH.

Log D, **the *effective lipophilicity/hydrophobicity*, is a variable depending on pH** – and that pH should always be quoted. It is a common misconception in that log D is merely "log P measured at pH 7.4" – the following section defines the parameter and the physical basis of its variation.

Charged species inherently have more affinity with the aqueous phase, rather than the neutral fatty hydrocarbon environment of the immiscible solvent (octan-1-ol being the standard model). Indeed, a single charged centre imparts a preference for the aqueous phase of typically 500- to 10,000-fold (depending on the nature of the ionophore). This behaviour is familiar to synthetic chemists in the workups of reactions where an acid or amine is the product. In the case of an organic amine, such a molecule can readily be dissolved in aqueous solutions as its ionised (protonated) form at low pH and neutral organic impurities or reaction by-products can be extracted with an immiscible lipophilic solvent. By raising the pH of the aqueous solution to alkaline values above the pK_a, the form of the amine is the free base; this now has greater solubility in the immiscible lipophilic solvent, which can be used to extract the molecule in this unionised form. The opposite case holds for acids: at low pH (below the pK_a) the protonated (unionised) form of the acid is more soluble in the immiscible lipophilic solvent; at a pH above the pK_a the ionised form predominates and this has greater aqueous solubility.

This is illustrated in the following graphs plotting distribution, defined as D, for a base and an acid between octanol and water with varying pH: the examples used are sumatriptan (a 5-hydroxy tryptamine analogue) and acetyl salicylic acid (aspirin). The difference in partitioning of the unionised and ionised forms of molecules between the octanol and aqueous phase such that, to a good approximation, only the unionised form substantially distributes into the octanol layer and the ionised form substantially into the water. Table 1 and Fig. 4 show that distribution closely tracks pH and that once the pH is more than about 2 log units away from the pK_a, then distribution reflects the partitioning (vide supra) between one phase and the other.

Distribution, D, is defined as the distribution of all relevant species between the two phases, described as follows for an acid (HA) and a base (B).

Table 1 The distribution (ratio of concentrations in octanol: aqueous) versus ionisation of sumatriptan and acetyl salicylic acid (aspirin) across the pH range from 1 to 11

Sumatriptan: base $pK_a = 9.6$, log $P = 0.8$				Acetyl salicylic acid: $pK_a = 3.5$, log $P = 1.2$			
pH	Distribution	Log D	% Ionised	pH	Distribution	Log D	% Ionised
11	6.0680	0.78	3.829	11	0.00159	−2.800	100
10	4.5131	0.65	28.475	10	0.00159	−2.799	100
9.5	2.7935	0.45	55.7312	9	0.00164	−2.786	99.9997
9	1.2672	0.10	79.924	8	0.00209	−2.681	99.9968
8	0.1552	−0.81	97.550	7	0.00659	−2.181	99.9684
7	0.0164	−1.78	99.749	6	0.052	−1.288	99.6848
6	0.0022	−2.65	99.975	5	0.487	−0.312	96.9347
5	0.00079	−3.10	99.9975	4	3.81	0.581	75.9747
4	0.00065	−3.19	99.9997	3.5	7.93	0.899	50.0
3	0.00063	−3.20	100	3	12.04	1.081	24.0253
2	0.00063	−3.20	100	2	15.36	1.186	3.0653
1	0.00063	−3.20	100	1	15.80	1.199	0.3152
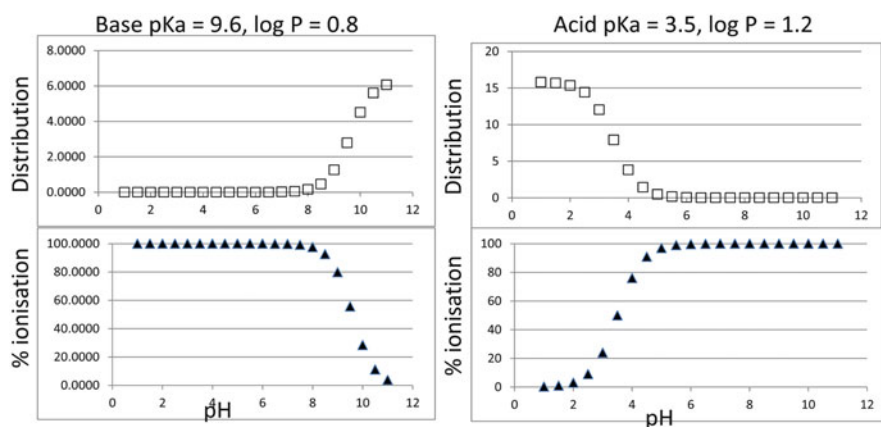
Fig. 4 Distribution versus ionisation for a base with $pK_a = 9.6$ and log $P = 0.8$ and an acid with $pK_a = 3.5$ and log $P = 1.2$

Thus,

$$D = \frac{[HA] + [A^-]_{octanol}}{[HA] + [A^-]_{water}} \qquad D = \frac{[B] + [BH^+]_{octanol}}{[B] + [BH^+]_{water}}$$

where the equilibrium can be summarised as follows (for an acid, AH), showing the impact of pK_a:
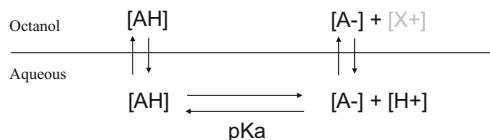
And for a base:

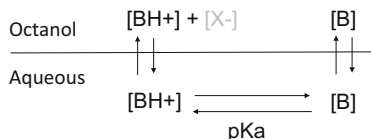Log D is predictable at a given pH and can be calculated using the equation:

$$\log D_{pH} = \log\left\{10^{(\log P)} + 10^{[\log P - \Delta P + ch*(pK_a - pH)]}\right\} - \log\left\{10\left(1 + 10^{[ch*(pK_a - pH)]}\right)\right\}$$

where ch (charge) $= 1$ for bases -1 for acids, $\Delta P =$ difference in partition of ionised and neutral species.

ΔP is variable and represents a 500- to 10,000-fold contribution of the charge. ΔP values for ionophores are as follows: acids, 4.2; aliphatic amines, 3.1; aromatic amines (e.g. aniline), 2.5; and aromatic embedded nitrogen (e.g. pyridine), 4.0.

The extractions of acids or bases into aqueous solutions are generally achieved above pH 12 or below pH 1, respectively; these are extremes beyond the range that would typically be encountered by a drug molecule (see Sect. 3). Indeed, at the site of action, a drug is likely to be in an environment where the pH is not too far removed from pH 7.4 (often referred to as physiological pH, the pH of blood plasma and other biological aqueous environments), which is close to neutrality and in a region where many acidic or basic motifs typically incorporated into drugs would be partially ionised. Thus there can be big changes in the distribution (log D) of such molecules with only minor changes in pH, for example, the pH of the ileum (where drug is absorbed) is about 6.5 going into the bloodstream at pH 7.4. In other sections, log D_{pH}, the *effective hydrophobicity* of a molecule at a given pH, is shown to have a significant impact on many aspects of drug action, particularly those where a kinetic distribution process is important (i.e. how much of the molecule is present in the ionised or unionised form). Thus, the impact of log D is most important on interactions important in ADME, i.e. drug absorption, distribution, metabolism and elimination (Sect. 4). In contrast, log P appears to be more important in molecular recognition and promiscuity (Sect. 4.2.4) [29].

The most relevant and important log D value for most scenarios in medicinal chemistry is log $D_{pH7.4}$ (or log $D_{7.4}$), although distribution at other pH values is sometimes pertinent. For example, log $D_{6.5}$ (pH of the duodenum) is used in models of intestinal absorption and log $D_{2.0}$ (pH of the stomach) can influence oral drug dissolution in the gastric fluids (Sect. 2.3.2). A log D value can be measured at any
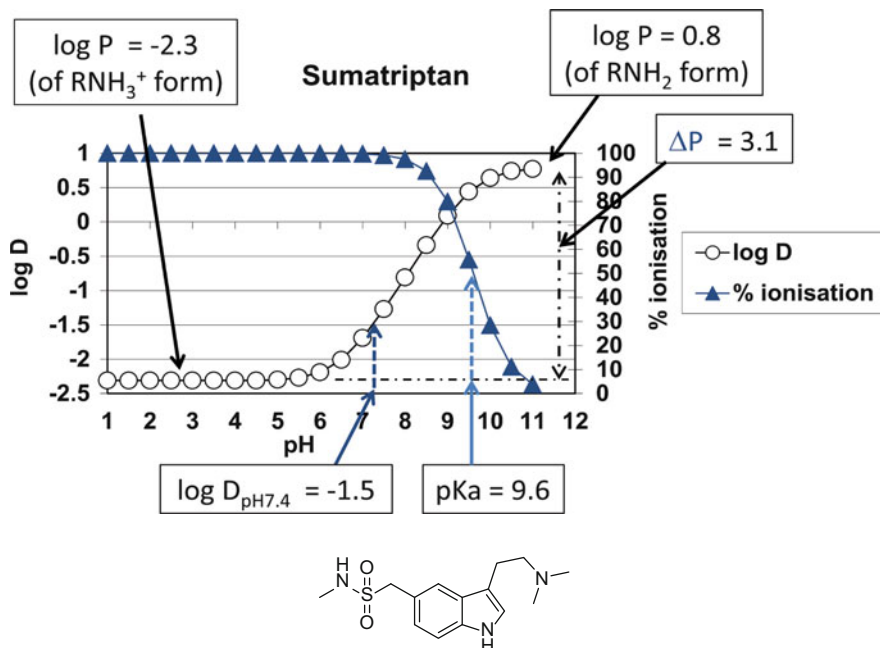
Fig. 5 log D and ionisation versus pH profiles of sumatriptan, illustrating the concepts of log D_{pH}, log P, ΔP and pK_a for this basic compound, wherein: pK_a = 9.6 (50% ionisation). log P of unionised form (the upper asymptote) = 0.8, log P of ionised form = −2.3, log $D_{pH7.4}$ (physiological pH) = −1.5

pH, however at the extremes, away from the pK_a, the value measured becomes an asymptote, which is the log P for either the ionised or unionised form (Figs. 5 and 6). Understanding a molecule's behaviour across the pH range is important, as highly basic (pK_a > 10) compounds are rare and few organic acids have a pK_a < 2; this will mean that at some point local pH will have a profound effect on the ionisation state and thus distribution of the molecule.

By way of example the full log D and ionisation versus pH profiles of sumatriptan and acetyl salicylic acid (aspirin) are shown in Figs. 5 and 6. In comparison to the distribution graphs in Fig. 4 note how the logged distribution function now goes below the zero before reaching an asymptote at a lower pH (base) or higher pH (acid).

In practice, many drug compounds do have a second ionisable centre – perhaps they may even be zwitterionic at physiological pH. In this case, the log P value is a theoretical concept, estimated by extrapolation – with maxima and minima in the log D–pH variation possible – depending on relative pK_a values and the nature of the charge. Most oral drugs will have weakly basic and/or weakly acidic centres and the behaviour of the neutral or fully charged form may be the key in a given scenario; for example, charged species do not tend to have good cell penetration. Intravenous administration or prodrugs are amongst common delivery options for highly charged or zwitterionic molecules.
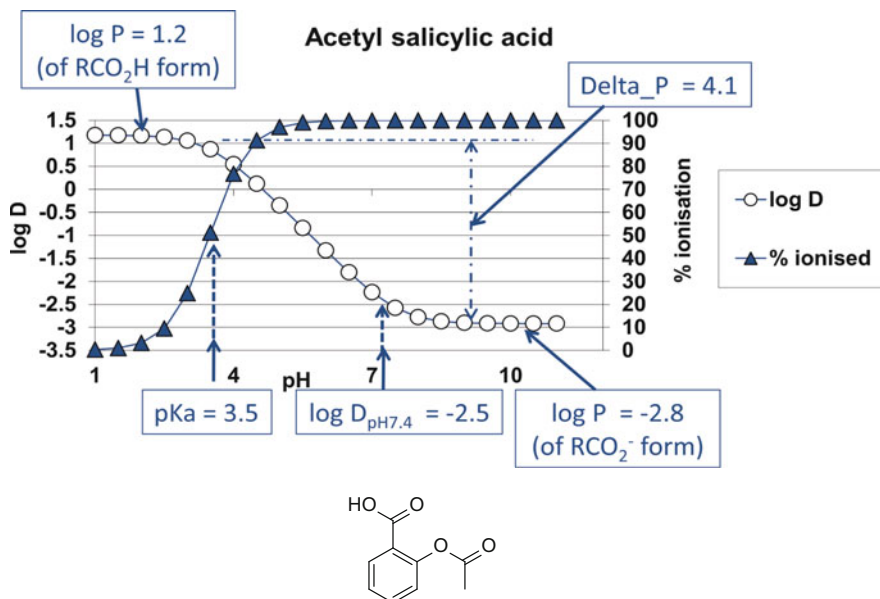
Fig. 6 log D and ionisation versus pH profiles of acetyl salicylic acid, illustrating the concepts of log D_{pH}, log P, ΔP and pK_a for this basic compound, wherein: pK_a = 3.5 (50% ionisation), log P of unionised form (the upper asymptote) = 1.2, log P of ionised form = −2.8, log $D_{pH7.4}$ (physiological pH) = −2.5

2.1.1 Small Hydrophilic Molecules: A Forgotten Strategy for Oral Medicines?

One approach exploiting a peculiar physical property niche, which appears to be little utilised in modern drug discovery, is the use of "small hydrophilic molecules". Historically, this approach enabled the discovery of many effective oral agents, whereby paracellular intestinal absorption is facilitated by permeation through the tight aqueous spaces between the cells lining the small intestine [30]. The molecules will necessarily be small (MW < 250 or calculated molar refraction less than 10) and tend to be basic or zwitterionic, with a distribution coefficient less than one (i.e. negative log D values, thus being more soluble in water – see Sect. 2.1). The β1-antagonist atenolol [31] is a commonly quoted example of molecule exploiting this manner of administration, whilst physical properties in this space were deliberately targeted in the discoveries of the thrombin activatable fibrinolysis inhibitor UK-396,082 [32] and the anti-migraine serotonin 1B/1D receptor agonist zolmitriptan [30] (Fig. 7).

2.1.2 Methods for the Measurement and Prediction of Lipophilicity

The measurement of lipophilicity using *n*-octan-1-ol and aqueous buffer has been the accepted standard in drug discovery and related disciplines for many years;
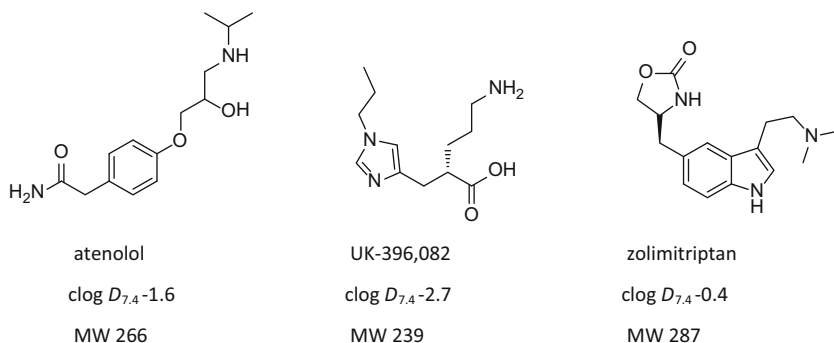
Fig. 7 Small, hydrophilic, oral drugs absorbed by the paracellular route

the history of this is well documented [26, 28]. As the Medicinal Chemistry Project in Hansch's group at Pomona College developed, accurately measured lipophilicity values were added to the BioByte database, from which calculated log P values (clog P, BioByte or Daylight log P) were derived [33]. This assiduously curated database of lipophilicity values and the derived fragment contributions to them remains the basis for the gold standard in predicting hydrophobicity values [34]. These predictions are based on the premise that log P can be estimated from the summation of the contributions of its non-overlapping molecular fragments (small clusters of atoms with single or multiple covalent linkages). Such fragments and their log P contributions were determined by statistical analysis of the empirical data. There are two variations on this theme represented by the Hansch–Fujita π-constants (whereby for substituent $- X$, $\pi_{R-X} = \log P_{R-X} - \log P_{R-H}$) [27] or subsequent methods based on fragment contributions (sometimes written as f), for example, those due to Hansch–Leo [34] or Rekker–Mannhold [35]. Thus, in general use, the Hansch–Fujita constant is used to modify the parent log P value (by rearrangement of the equation to give $\log P_{R-X} = \pi_{R-X} + \log P_{R-H}$); it will depend on the site of any substituent – as illustrated in Table 2. The fragment contributions are summed for any given molecule, and thus $\log P = \Sigma a_i n_i$ where a represents the individual fragments and n the number of times they occur; the values are different between models as differing correction factors are used in the final calculation. These factors correct for any apparent electronic or steric effects within the log P calculation. Contemporary computational methods make this kind of analysis easier and increasingly accurate, providing particular structural intricacies are recognised with the databases. Common rules of thumb would include an increase of approximately 0.5 units for each methylene ($-CH_2^-$) linkage in a molecule or about 1.7 units for a phenyl ring; further examples are given in Table 3. These contributions are additive, but the impact of the environment onto which a substituent is attached can influence the outcome.

Table 2 Hansch–Fujita functional groups substituent π values, illustrating their contributions to the log P value of a given molecules; thus, for example, toluene, clog P is $0.56 + 1.76 = 2.32$ (i.e. $210\times$ more soluble in octanol than water) and benzamide, clog P is $1.76 - 1.49 = 0.5$ (i.e. ~$2\times$ more soluble in octanol than water)

Substituent	π Meta	π Para	π Benzene
–H	0.00	0.00	0.00
–CH_3	0.51	0.52	0.56
–Cl	0.76	0.70	0.71
–Br	0.94	1.02	0.86
–OH	−0.49	−0.61	−0.67
–OCH_3	0.12	−0.04	−0.02

Table 3 Illustrative examples of fragmental contributions f and $f\phi$ used in log P calculations

	Rekker constant f		Hansch–Leo constant f	
Fragment	Aliphatic f	Aromatic $f\phi$	Aliphatic f	Aromatic $f\phi$
–F	−0.462	0.399	−0.38	0.37
–Cl	0.061	0.922	0.06	0.94
–Br	0.27	1.131	0.2	1.09
–OH	−1.491	−0.343	−1.64	−0.4
–O–	−1.581	−0.433	−1.81	−0.57
–COOH	−0.954	−0.093	−1.09	−0.03
–NH_2	−1.428	−0.854	−1.54	−1
–NH–	−1.825	−0.964	−2.11	−1.03
–NO_2	−0.939	−0.078	−1.26	−0.02
–$CONH_2$	−1.97	−1.109	−2.18	−1.26
>C=O	−1.703	−0.842	−1.9	−0.32
–CF_3	0.757	1.331	0.88	–
–C≡N	−1.066	−0.205	−1.28	−0.34
–Ph	–	1.903	1.59	1.76

2.1.3 Octanol Water Methods: Issues with Overly Lipophilic Compounds

Hydrophobicity data generated in octanol–buffer partitioning/distribution experiments has come under scrutiny in recent years due to a perceived disconnect between measured and calculated values [36]. It was particularly noticeable that few measured values above four were recorded, in spite of calculations indicating much higher estimated values. The work of Hill [16] indicated that calculated log $D_{7.4}$ values correlated better with measurements for compounds with reasonable levels of measured solubility (>200 μM); conversely, for compounds in this set with poor solubility (<30 μM), very poor correlation was evident (Fig. 8). For compounds with solubility measured in the intermediate range between these levels, the observed trends were, as expected, between the two extremes.

There has been some discussion as to why the octanol water model breaks down with insoluble compounds; certainly octanol has some miscibility with water and,
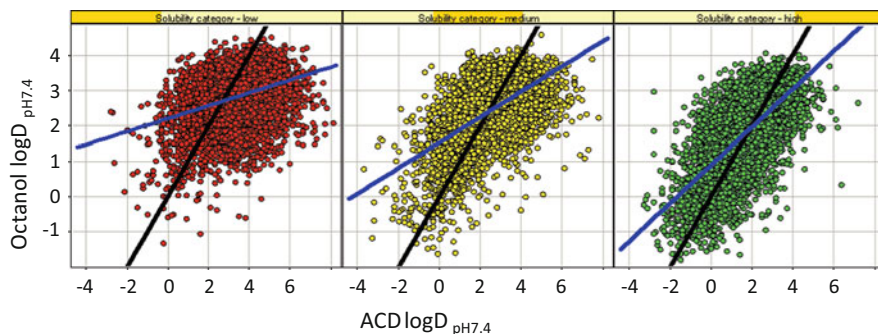
Fig. 8 Plot of measured versus calculated log $D_{pH7.4}$ values, reproduced with permission from [16], separated by solubility category. **a** (*red*): low-solubility compounds (<30 μM). **b** (*yellow*): intermediate solubility compounds (31–200 μM). **c** (*green*): high-solubility compounds (>200 μM)

as an alcohol, has some hydrogen-bonding capability, both issues that could compromise measurements. A study of the distribution behaviour of acids by Wenlock et al. [37] highlighted the issues of ion pairing in compromising the quality of log $D_{7.4}$ measurements. Additionally, it is highly likely that for highly crystalline or highly lipophilic molecules, their intrinsic solubility in both water and octanol will be very low and aggregation could be an issue by preventing proper equilibration between the disrupted phases. Some of these are perhaps moot points; quality drug-like molecules should be soluble and thus should behave better in the octanol/water system, as illustrated by the soluble examples in Fig. 8c. Modern drug discovery practices have thus necessitated a change in the way lipophilicity is estimated, due to the shortcomings of the traditional method with compounds in the era of molecular obesity.

2.1.4 Measurements Using Chromatographic Methods

The retention time of a compound on a C-18 column eluted with an acetonitrile/ aqueous buffer is fundamentally linked to its lipophilicity [38]; a log D value can be calculated from the so-called chromatographic hydrophobicity index (CHI), which equates to the percentage of acetonitrile corresponding to the retention time of the compound within the linear gradient typically employed [39]. The CHI value can be converted to a chromatographic log D value [29] using the simple relationship of Chrom log D_{pH} = CHI_{pH} × 0.0857 − 2. Accordingly, Chrom log D has a theoretical range from −2 (0% MeCN) to 6.57 (100%) – although, in practice, the lower limit can be influenced by compounds being drawn through by loading solvents and the upper range can be extrapolated up to about a Chrom log D value of 10, using elution with neat acetonitrile. The revised calibration used to generate the Chrom log D values (a modification of CHI log D [39]) has shown comparable incremental changes to those measurements in the traditional octanol-aqueous buffer
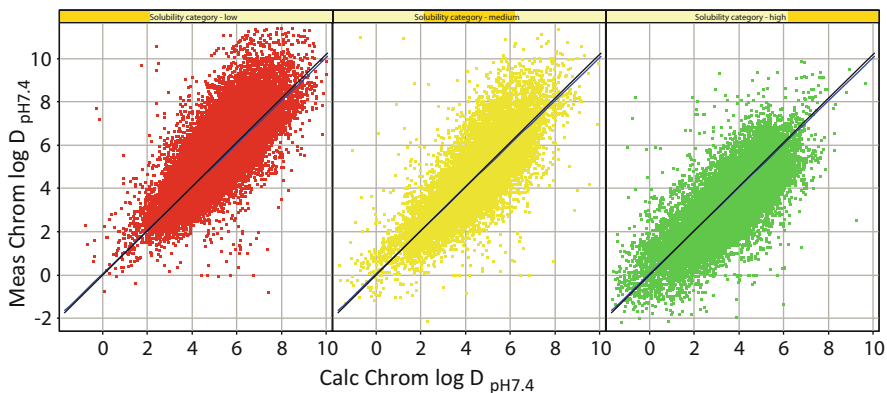
Fig. 9 Trellised plots (by solubility class, coloured as Fig. 8) of measured Chrom log $D_{7.4}$, values versus calculations of that parameter using a GSK in-house model [29], showing the (barely distinguishable) lines of best fit and unity

methodology. In practice, an average offset of approximately 2 log units above OW values is observed with Chrom log D values; this discrepancy has been retained in practice, giving a deliberate means of distinguishing the origins of the data. Accordingly, it has been developed into a convenient, effective and reliable means of measuring lipophilicity, amenable to high-throughput and automated operations, using only a few micrograms of compound from a small volume of a DMSO stock solution.

The relevance and importance of chromatographically derived measurements were established in a wide-ranging analysis of over 100,000 GlaxoSmithKline (GSK) compounds with measured Chrom log D and solubility data [29].

Various key findings which supported the superiority of these measurements when compared to traditional octagonal-buffer (OW) shake flask methods:

1. Normal distribution of values with both calculated and the chromatographically measured values; measured OW values showed a skewed distribution
2. Strong correlation between calculated and the chromatographically measured values – irrespective of solubility

    (a) Correlation with measured OW values only evident for those with >200 μM solubility

3. Few compounds measured in OW had a value >3, even fewer >4, in spite of predictions and distribution of values in chromatographic measurements.

    (a) These lipophilic compounds more likely to be poorly soluble!

4. New predictors trained with the chromatographic data showed clearly better performance than the predictors based on OW measurements (Fig. 9).

Table 4 Changes in log D, at high or low pH, observed for various charged compounds compared to the measured value, X, at pH 7.4

Chrom logD at pH:	Charge characteristic				
	Neutral	Weak acid	Strong acid	Weak base	Strong base
2	X	X	>X	<X	X
7.4	X	X	X	X	X
10.5	X	<X	X	X	>X

if measured value = X, the box is grey, if >X
then black and <X is white

In practice, Chrom log D has been routinely run with aqueous buffers at the three pH values of 2.0, 7.4 and 10.5; these extremes can give an indication of the likely strength (weak or strong) of any acidic or basic centres in a molecule or of zwitterionic character. These charge states are indicated by the difference in the Chrom log D values generated across the three pH values; the actual differences depend on the pK_a values but, as a rule of thumb, the patterns described in Table 4 give a guide. The patterns will be more complex for zwitterionic compounds, although all three values could be the same and the pH 7.4 value could be higher or lower than the other two. The pH variations of distribution values that vary are explained in Sect. 2.1, from where it should also be apparent that the highest measured value(s) in this three pH analysis is at least a very good approximation of chromatographic log P for all but the most acidic or basic of compounds.

In the various recent analyses exploring the link between lipophilicity and poor developability outcomes (Sect. 4.3), it should be noted that, paradoxically, clog P values have often represented a better estimate of the true lipophilicity of a compound rather than a measurement in octanol–buffer. It has been shown that there is a strong correlation between clog P values and chromatographically derived log P measurements, reflective of the quality of the curated data and methodology underpinning the value (see Sect. 2.1.3). Accordingly, many data reviews that explored the impact of lipophilicity will have shown a better correlation with clog P than log $D_{7.4}$ values measured in the traditional octanol–water system. Sometimes calculated log $D_{7.4}$ data can work well too, but often these estimates are subject to error due to poor estimation of the pK_a – and in most work the partitioning behaviour rather than distribution has been used. As a result, due to large numbers in sets and acidic/basic compounds offsetting each other, the outcome has often shown reasonable correlation with clog P due to various inherent compensatory mechanisms. Sect. 4.2 of this chapter reviews the lipophilicity dependence of various developability parameters versus chromatographically measured values; the outcomes indicated enhanced differentiation and resolution of the nature of the dependence (be it driven by intrinsic or effective lipophilicity) as would be expected of more reliable estimates of lipophilicity.

2.2 *Acid/Base Strength: p*K_a

The profound effect that the ionisation state of a molecule has on its behaviour is evident in the other sections of this chapter; the degree of ionisation can make a difference of hundreds- or several thousandfold in the solubility and distribution of a molecule. Charge can be a friend (in improving aqueous solubility and distribution) or a foe [by increasing risk of cardiac liability (lipophilic bases), poor permeation (any charge) or high protein binding (particularly acids)] in the drug optimisation process. The strength of ionisable centres (defined by their pK_a, for example, see Table 5) and the impact this has on the behaviour of the drug is fundamental to understanding the physical attributes of a molecule. In the various body compartments (Sect. 3), there can be significant changes in the pH; this is particularly pertinent to the processes of drug metabolism and pharmaco-kinetics (DMPK) in general and oral absorption in particular. The proximity of a pK_a value of an ionisable centre to the ambient pH will have a very big impact on the degree of ionisation as discussed in the following sections.

Table 5 Typical pK_a value of common motifs in drug discovery (in the absence of modifying substituents)
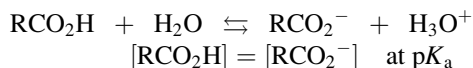
Classes of compounds	Typical pK_a ranges
Pyrimidines (bases)	1–2
Anilines (bases)	1–5
α-amino acids	2–3 (COOH)
Carboxylic acids	3–5
Pyridines (bases)	4–6
Imidazoles (bases)	7
Phenols	8–10
Purines (bases)	8–10
α-Amino acids	9–10.5 (NH_3^+)
Aliphatic and alicyclic amines	9–11
Guanidines	11–14
Azoles (acids)	13–16
Alcohols	13–16

The pK_a refers to the pH at which an ionisable centre, be it an acid or a base, is present with equal proportions of the charged and uncharged forms, i.e. 50% ionised (the term pK_b for bases is no longer used – so pK_a is used to describe acidic or basic relationships). It is derived from the Henderson–Hasselbach equation:
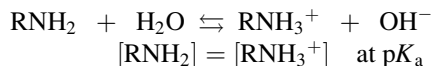
$$pH = pK_a + \log([A^-]/[HA])$$

such that when the species is 50% ionised, then $[A^-] = [HA]$ and the pH = pK_a.

So for an acid:

$$RCO_2H \;+\; H_2O \;\leftrightarrows\; RCO_2^- \;+\; H_3O^+$$
$$[RCO_2H] = [RCO_2^-] \quad \text{at } pK_a$$

For a base:

$$RNH_2 \;+\; H_2O \;\leftrightarrows\; RNH_3^+ \;+\; OH^-$$
$$[RNH_2] = [RNH_3^+] \quad \text{at } pK_a$$

Equation (1), derived from the Henderson–Hasselbach equation, illustrates a useful means for calculating the percentage ionisation of a compound with a given pK_a at a particular pH. This is expanded out to show the level of ionisation at pH values 1, 2 and 3 units away from the pK_a.

$$\% \text{ ionised} = \frac{100}{1 + 10^{[\text{charge}(pH - pK_a)]}} \tag{1}$$

where charge $= +1$ for bases and -1 for acids

pH 1 unit above pK_a	% ionisation $= 100/(1 + 10^1)$, i.e. 9.09% for bases
	% ionisation $= 100/(1 + 10^{-1})$, i.e. 90.9% for acids
pH 2 units above pK_a	% ionisation $= 100/(1 + 10^2)$, i.e. 0.99% for bases
	% ionisation $= 100/(1 + 10^{-2})$, i.e. 99.0% for acids
pH 3 units above pK_a	% ionisation $= 100/(1 + 10^3)$, i.e. 0.099% for bases
	% ionisation $= 100/(1 + 10^{-3})$, i.e. 99.9% for acids
pH 1 unit below pK_a	% ionisation $= 100/(1 + 10^1)$, i.e. 9.09% for acids
	% ionisation $= 100/(1 + 10^{-1})$, i.e. 90.9% for bases
pH 2 units below pK_a	% ionisation $= 100/(1 + 10^2)$, i.e. 0.99% for acids
	% ionisation $= 100/(1 + 10^{-2})$, i.e. 99.0% for bases
pH 3 units below pK_a	% ionisation $= 100/(1 + 10^3)$, i.e. 0.099% for acids
	% ionisation $= 100/(1 + 10^{-3})$, i.e. 99.9% for bases

Even at 2 units away, 99% will be in one form – although the 1% might still have a role in partitioning of an acid above its pK_a and a base below the pK_a as can be seen in Table 1 in the examples used, namely, aspirin and sumatriptan.

2.2.1   Measurement and Calculation of pK_a

There are many methods available to measure pK_a values [40]. Classically, potentiometric titrations [41] have been used, but these require relatively large quantities of material (10 s of mgs) and are quite labour intensive (30 min plus to run each sample). Any physical technique that can differentiate some property of a molecule that is dependent on its ionisation can be used, such as solubility [42], spectrophotometry [43], NMR [44] and capillary electrophoresis [45]. The bases of

Table 6 Inductive and resonance effects of particular substituents, taken from Perrin et al. [44]

+ I (acid weakening)	$-CO_2^-$; $-O^-$; $-NH^-$; -alkyl
+ I (acid strengthening)[a]	$-NH_3^+$; $-NR_3^+$; $-NO_2$; $-SO_2R$; $-CN$, $-F$; $-Cl$; $-Br$; $-I$; $-CF_3$; $-COOH$; $-CONH_2$; $-COOR$; $-CHO$; $-COR$; $-OR$; $-SR$; $-NH_2$; $-Ph$
+ M (acid weakening)	$-F$; $-Cl$; $-Br$; $-I$; $-OH$; $-OR$; $-NH_2$; $-NR_2$; $-NHCOR$; $-O^-$; $-NH^-$; -alkyl
+ M (acid strengthening)	$-NO_2$; $-CN$, $-COOH$; $-COOR$; $-CONH_2$; $-Ph$; $-COR$; $-SO_2R$

[a]Approximately in decreasing order; effects that strengthen acids will weaken bases and vice versa

the measurements are that, as pH varies, the physical behaviour or some spectral characteristic of the molecule will differ, depending on the degree of ionisation. The pK_a of poorly soluble compounds can be hard to measure accurately due to the complex equilibria involved around the dissolution process. One of the more generally used and higher throughput techniques is the spectral gradient analysis (SGA) methodology [46]; this has been applied to an automated 96-well plate format, generating data in about 4 min per sample. SGA does, however, rely on the presence of a good chromophore close enough to the ionisable centre, which produces a spectral shift with changing pH that can be dynamically monitored using spectrophotometry.

The accurate calculation of pK_a values is rather difficult in practice, but useful estimates can be achieved using the classical methods of linear free energy relationships applied to reactions and equilibrium constants, typified by the Hammett equation and subsequent variations [47]. The premise is based on the assumption that within particular classes of acids or bases, substituents produce free energy changes to modulate the pK_a that are additive in a linear fashion. Along these lines, a parent ionophore would be recognised within the structure, which would be assigned a pK_a value derived from a known measurement associated with this motif. This number is modified according to substituents and their distance and connectivity, looking at inductive (I, σ-effects) or mesomeric (M, π- or resonance effects), which can result in donation or withdrawal of electronic charge as a consequence (Table 6). Typically, such calculations would be made using standard tables of collated data based on empirical observations, although this is now usually done computationally and some packages will illustrate how the value has been derived from particular fragmentations of the molecule. The Hammett equation is applicable to aromatic systems and the Taft equation has value in aliphatic systems; these are classically applied to the quantitative impact of substituents on reactivity; the methodology also has great importance in pK_a prediction.

$$\text{Hammett equation}: \quad pK_a = pK_a{}^o - \rho\left(\sum \sigma\right)$$

where $pK_a{}^o$ is the pK_a of the unsubstituted system, σ is a constant applied to a particular substituent and ρ is an empirically derived constant for the particular type of reaction.

Table 7 Illustrative pK_a values of acids and bases as discussed in the text

Acid	CH_3CO_2H	$ClCH_2CO_2H$	Cl_2CHCO_2H	Cl_3CCO_2H
pK_a	4.76	2.85	1.48	0.70
Amine	NH_3	$EtNH_2$	Et_2NH	Et_3NH
pK_a	9.25	10.7	11.2	10.75
Amine	$EtNH_2$	$nPrNH_2$	$HOCH_2CH_2NH_2$	$HO(CH_2)_3NH_2$
pK_a	10.2	10.2	9.5	9.9

$$\text{Taft equation}: \quad pKa = pK_a{}^o - \rho^*\left(\sum \sigma^*\right) + \delta E_s$$

where p$K_a{}^o$ is the pK_a of the unsubstituted system, σ^* is a constant applied to a particular substituent, ρ^* is an empirically derived constant for the particular type of reaction and E_s is a steric term. The σ^* values will be dependent on the number of methylene groups between the substituent and the ionisable centre; each extra methylene reduces the effect by 0.4-fold. Thus a Taft σ^*-value is multiplied by 0.4 for one methylene spacer (X–CH_2–CO_2H or X–CH_2–NH_2), 0.16 for X–CH_2–CH_2 and 0.064 for X–CH_2–CH_2–CH_2.

The strengthening of α-chloroacetic acids illustrates how substituents modulate pK_a, as increased electron withdrawal by each additional alpha chlorine atom increases the acidity by stabilising negative charge on the acid motif through additive σ-inductive effects. Complications can arise, for example, with the stabilisation of protonation states by hydrogen-bonding and accessibility of the charged centre. For example, in Table 7, starting from ammonia, each additional ethylation to ethylamine, diethylamine and triethylamine gives an increased sigma donation from the alkyl chain that stabilises positive charge. However, triethylamine is less basic than diethylamine because the stabilisation of the charged state only has one proton that is able to hydrogen bond with solvating water in the manner $R_3N^+H\cdots OH_2$, compared with two protons with diethyl amine, so the overall stabilising effect of the charged network is reduced. The series expanding from ethanolamine to 3-hydroxypropylamine illustrates how the additional methylenes increase the modulating effect. Steric hindrance can also impact pK_a values, particularly acids, where the larger group has more chance of a steric clash than smaller amino and hydroxyl groups.

A worked example of the Taft equation is illustrated with a calculated pK_a of morpholine [44]:

The molecule is approximated to $CH_3OCH_2CH_2NHCH_2CH_2OCH_3$ for the purposes of the estimate:

$$\mathrm{p}K_a \quad \text{for } R_1R_2NH_2^{+} = 10.59 - 3.23 \times \sum \sigma^{*} + 0.2$$

$10.59 = \mathrm{p}K_a$ for R_2NH; $3.23 = $ experimental ρ^{*} for this system; $0.2 = $ steric constant for a cyclic system

$\Sigma\rho^{*}$ for $2 \times CH_3OCH_2CH_2 -$ equates to $2(0.4 \times 0.4 \times 1.81) = 0.58$
$1.81 = \rho^{*}$ for MeO^{-}

Thus predicted $\mathrm{p}K_a$ is $10.59 - 3.23(0.578) + 0.2 = 8.92$ (the measured value is 8.36).

2.3 Solubility

The solubility of a compound can be described as the propensity of a molecule to act as a solute in a liquid medium to form a solution wherein the solute molecules are dispersed and individually surrounded by solvent molecules. The solute could originate from the gaseous, liquid or solid state – but the end result is the same. Particularly pertinent to drug discovery is the reverse process, whereby around the limit of solubility, a dynamic equilibrium will exist between a solid structure dispersing into the solvent and the reverse process of the molecules coming together and reforming the solid. Life itself is based on chemistry in aqueous solutions so, at some point, no matter what the delivery method, a therapeutic molecule will have to be dissolved in an aqueous environment prior to engagement with its target – thus some level of aqueous solubility is necessary. Often in drug discovery, systemic exposure or transport to a particular location may be required; whilst blood plasma proteins, for example, could aid in this process, a minimum level of solubility will always ultimately be required to give a free aqueous fraction that is at sufficient concentration to ensure a pharmacological action. Solubility is a prerequisite for oral absorption, a chain of events and environments that pose a number of hurdles and challenges to the prototype drug molecule (Sects. 3 and 4). Given the size and nature of drug molecules, they are almost invariably solids in their pure form at room temperature, although in modern practice many are first encountered as amorphous gums. But even an amorphous gum or glass will have a particular solubility and this can be measured; ultimately, the processes of chemical development will identify stable solid forms.

Understanding the "solubility" of drug molecules is a subject that can engender fierce debate in terms of what levels are required and how it should be measured [48]. The different solid forms of compounds (e.g. amorphous versus crystalline polymorphs) will influence the dissolution rate and wettability of the compound [49]. A molecule's solubility changes depending on a number of factors related to the composition of the liquid used to form the solution, such as the solvent/solute composition, pH, and how long the experiment is run. The crystalline and salt forms

of the compound will also have significant effects [50]. An in-depth appraisal of these is beyond the scope of this chapter and is already well documented [48]. In common medicinal chemistry practice, kinetic solubility in pH 7.4 buffer (physiological pH) is a very good benchmark to appraise solubility of experimental molecules and is commonly employed with increasingly high-throughput, such as a multi-well plate assay using chemiluminescent nitrogen detection (CLND) [51].

2.3.1 The General Solubility Equation

By extension of the concepts discussed in Sect. 2.1, aqueous solubility decreases as lipophilicity increases. An additional key concept in solubility is that lattice energy is perturbed as the intermolecular contacts within the solid break down; this can be equated with the breakdown of a solid structure at its melting point. These concepts were developed and empirically substantiated in the derivation of the general solubility equation (GSE) by Yalkowsky and co-workers [52], which provided a useful predictive model of solubility.

The general solubility equation is

$$\log \text{solubility} = 0.5 - 0.01(\text{MP} - 25) - \log P$$

where MP = melting point in Celsius and P is the octanol–water partition coefficient.

In practice, the GSE is also commonly used in terms of log D, whereby

$$\log S_{\text{pH}x} = 0.5 - 0.01(\text{MP} - 25) - \log D_{\text{pH}x}$$

The general solubility equation had been demonstrated to give good prediction of solubility for neutral compounds [52] and acts as a useful guide for ionisable compounds using log D at the pH of the aqueous buffer employed. It is very interesting to consider the implications of the distribution of solubilities in Table 8, which is a matrix of solubility values, calculated using the GSE over a range of log P values and melting points [16]. This is shaded by solubility classifications whereby <30 μM is regarded as poor solubility and >200 μM as good – useful guidelines to follow in drug discovery. The column in the table where log $P = 3$ is particularly notable; it is here that through variation in melting points a compound could lie in each of the three solubility categories. Above this, only low-melting (low lattice energy) solids would be expected to show any suitable levels of solubility; below this only the very highest melting is likely to have issues with solubility. It is probably therefore not a coincidence that the median log P value of a drug is somewhere in the range of 2–3 and melting points are generally low (see Sect. 2.4).

Descriptive terms used by the European Pharmacopoeia [53] for the classification of solubility ranges (Table 9) are consistent with the predicted solubility bands of the GSE; the lowest melting compounds within the log P range of 2–3 (the

Table 8 Computed solubility values based on the general solubility equation across a range of melting points and log P values pertinent to drug discovery, shaded by the predicted solubility. The average clog P value in various analyses of drug molecules usually lies between 2.5 and 3.0

Computed log S values (from GSE) in table		Log P values					
		0	1	2	3	4	5
Melting point values (°C)	50	0.25	-0.75	-1.75	-2.75	-3.75	-4.75
	100	-0.25	-1.25	-2.25	-3.25	-4.25	-5.25
	150	-0.75	-1.75	-2.75	-3.75	-4.75	-5.75
	200	-1.25	-2.25	-3.25	-4.25	-5.25	-6.25
	250	-1.75	-2.75	-3.75	-4.75	-5.75	-6.75
	300	-2.25	-3.25	-4.25	-5.25	-6.25	-7.25

Predicted solubility   <30 μM   30-200 μM   >200 μM

Table 9 European Pharmacopoeia solubility definitions [53]

Descriptive term	Solubility (mg/mL)	Molar solubility MW 400	Log S	Approximate volume (mL) to dissolve 1 g
Very soluble	>1,000	>2.5 M	>0.4	1
Freely soluble	100–1,000	0.25–2.5 M	−0.6 to 0.4	1–10
Soluble	33–100	83–240 mM	−1.1 to −0.6	10–30
Sparingly soluble	10–33	25–83 mM	−1.6 to −1.1	30–100
Slightly soluble	1–10	2.5–25 mM	−2.6 to −1.6	100–1,000
Very slightly soluble	0.1–1	0.25–2.5 mM	−3.6 to −2.6	1,000–10,000
Practically insoluble	<0.1	<250 μM	< − 3.6	>10,000 (or 10 L)

median range for drugs) fall within the generic description of "very slightly soluble". Perhaps it is a misnomer to discuss the solubility of many drug molecules; the majority might be better described as being less insoluble!

In contemporary drug discovery, it is probably pertinent to describe the high proportion of highly lipophilic compounds as insoluble "grease (or fat) balls" which concurs with the findings of Bergström et al. in a study of poorly soluble drugs [54]. Another colloquial term often used for insoluble compounds is "brick dust", which would be a characteristic of the higher melting point compounds. Routine collection of melting points has ceased in high-throughput drug discovery, when only a few milligrams are typically made in early stages, often as lyophilised/amorphous solids or gums. The principles of disruption of planarity and symmetry to improve solubility have been reviewed by Ishikawa and Hashimoto [55]; Fig. 10 is an illustrative example from this paper; this follows the principles of reducing lipophilicity and aromaticity (PFI) and Fsp^3 (Sect. 2.4).

An interesting expansion on the impact of melting point was described by Chu and Yalkowsky [56], where the interesting relationship showing that the melting point and dose are related to the fraction of poorly soluble drugs absorbed was described. The proposed melting point based absorption potential (MPbAP) was
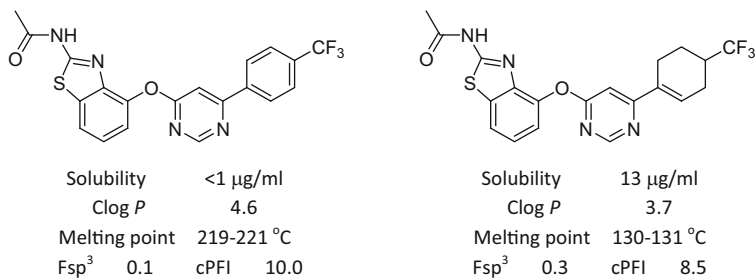
Fig. 10 An illustrative example from [55], where reduction meting point through reducing aromaticity showed a positive benefit (although the solubility is still relatively poor, reflective of the high PFI value, Sect. 2.4)

able to distinguish between high and low absorption (defined as more or less than half the administered dose absorbed) for 90% of the 91 drugs evaluated. As might be expected from the GSE, the more soluble, lower melting point compounds were more likely to be well absorbed.

There was an academic challenge proposed [57] to explore new methodologies for solubility prediction – a summary of the results and many different models proposed has been reviewed [58].

2.3.2 Simulated Gastrointestinal Fluids

Solubility values measured in aqueous buffers are only part of the picture in determining solubility pertinent to drug discovery [48], as the nature of the aqueous environment in the various parts of the body varies greatly in composition and pH. Indeed, the most pertinent measurements or the compositions and pH values of the simulated fluids used to make these measurements are subjects of debate and variation [59]. Table 10 shows some representative compositions and pH for simulated gastric and intestinal fluids in fasted or fed states – hence *Fa*sted or *Fe*d *S*tate *S*imulated (*G*astric or *I*ntestinal) *F*luids. In early phases of drug discovery SGF, FeSSIF and FaSSIF are the most commonly encountered measurements, giving insight on drug dissolution, impact of the changing pH and absorption effects. In later drug development, the more complex interactions with food in the GI tract are modelled by more complex mixtures, giving further insight on absorption in varying dosing regimes. In addition to the individual components and their concentrations, physical variables are controlled including pH, osmotic potential, buffering capacity and surface tension. All of these variables will influence the breakup and dissolution of the drug substance, the form of which can be controlled and modified to expedite dissolution in the stomach.

The pH of the gastric (stomach) preparation is typically 1.5, although this can transiently increase as food enters, before being reduced again by gastric secretion. The pH rapidly increases on exit from the stomach by bile salts to roughly 6.5 in the

Table 10 Typical compositions of simulated gastrointestinal fluids

Fluid	Solvent and pH	Key components	Aqueous acid/base added
FaSSGF	Water 1.6	Sodium taurocholate; lecithin, NaCl	HCl, qs
FeSSGF	1:1 milk/water 5	NaCl, acetic acid, sodium acetate	HCl, NaOH qs
FaSSIF	Water 6.5	Sodium taurocholate; lecithin, maleic acid, NaCl	NaOH qs
FeSSIF	Water (oil?) 6.5 (or lower)	Sodium taurocholate; lecithin, maleic acid, NaCl	NaOH qs

Qs quantum satis – the amount needed

duodenum, where much drug absorption takes place; this may be pertinent to model the oral exposure of a drug taken prior to consumption of food. In the fed state, this pH may be lower and some preparations buffer FeSSIF to pH 5.5 rather than 6.5.

Two key components of the simulated gastrointestinal fluids are taurocholic acid and lecithins. Taurocholic acid is the sulphonic acid amide derivative of cholic acid and taurine (2-aminoethanesulfonic acid); this is a naturally occurring bile acid. Lecithin is a generic term for a mixture of phospholipids, which are the major components of cell membranes; they are used as wetting agents and dispersion enhancers in drug formulations and have a widespread use in the food industry as emulsifiers. An example phosphatidylcholine is shown in Fig. 11. These are likely constituents of the melange in the intestinal fluids and can enhance the solubility of some drugs, as lecithin and taurocholate can form micelles into which drug molecules can partition. Consequently, in the development process, bio-relevant solubility determinations are performed in the simulated fluids described in Table 10; like many aspects of solubility, the exact composition and pH of the fluid is subject of debate; but control of components within, pH, buffering capacity and osmolality are important considerations [59]. In practice, food, excipients and other formulations can enhance solubility [60], sometimes producing supersaturation [61] as the molecules progress through the digestive tract; these can help to deliver compounds with inherently poor physical properties, but they are not a substitute for a molecule with good properties and solubility as they may carry other risks (Sect. 4.3).

The modifiers used in the simulated fluids described above do have an impact on solubility, but it is often the pH–solubility profile that can be very important in drug discovery. As the log D will vary with pH, it follows that solubility will also vary: the higher the ionisation, the more soluble the compound. However, the ionised form is in equilibrium with the unionised, so if the pH of the test solution is at the pK_a of the ionisable centre in the compound, the observed solubility will be limited by the solubility of the unionised form (50% of the material present) that equilibrates with solid. Prototypical pH–solubility profiles for an acid and a base are shown in Figs. 12 and 13.
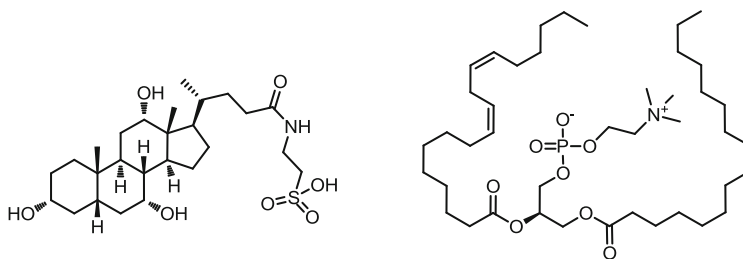
Fig. 11 Taurocholic acid full name 2-{[(3α,5β,7α,12α)-3,7,12-trihydroxy-24-oxocholan-24-yl] amino} ethanesulfonic acid and a typical phosphatidylcholine, the type of phospholipid component found in natural Lecithins
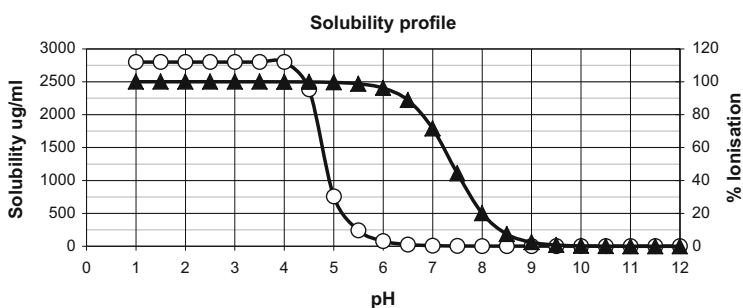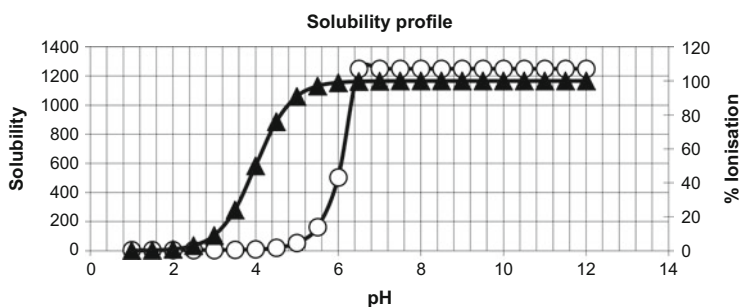

Fig. 12 Solubility (*white circles*) and percentage ionisation (*black triangles*) profiles versus pH for a compound with the following characteristics: single basic pK_a of 7.4, intrinsic solubility of 3 μg/mL and solubility in SGF (pH 1.5) of 2,800 μg/mL. At pH 7.4 the solubility is 6 μg/mL (3 μg/mL is in equilibrium with solids and the freely soluble fraction)

Fig. 13 Solubility (*white circles*) and percentage ionisation (*black triangles*) profiles versus pH for a compound with the following characteristics: single acidic pK_a of 4.0, intrinsic solubility of 3 μg/mL, and 1,250 μg/mL in higher pH buffer. This compound would have poor dissolution in the stomach (as the acid – but a salt would be more soluble); it would have better solubility in the duodenum (pH 6.5) and blood plasma (pH 7.4)

Table 11 Composite table from the data in [62] and [63], showing the decreasing mean aromatic ring count as compounds progressed through the GSK pipeline and the fraction of sp³ centres in molecules in similar phases noted by the Wyeth group

	Candidate selection	First time in human	Phase 1	Phase 2	Proof of concept	Phase 3	Drugs
Mean aromatic ring count (Ritchie, GSK)	3.3	2.9	2.5	2.7	2.3	–	1.6
Fraction of sp³ hybridised atoms (Lovering, Wyeth)	0.36	–	0.38	0.43	–	0.45	0.47

Table 12 Variation in Fsp^3 values within melting point bins from [63]

Melting point	25	75	125	175	225	275	325	375
Fsp^3	0.34	0.33	0.31	0.27	0.24	0.18	0.11	0.10
Count	75	657	1,153	1,253	815	375	93	11

2.4 Escape from Flatland and the Impact of Aromatic Ring Count

As the link between suboptimal physical properties and higher attrition has been established in recent years, further scrutiny indicated similar risks with molecules with high degrees of aromatic character. Ritchie and MacDonald [62] noted the statistically significant correlations of increasing aromatic ring count (#Ar) with poorer outcomes in particular developability assays; furthermore, the average ring counts were noted to reduce with escalating phases of development through to drug molecules (Table 11). This analysis used the Daylight definition of aromatic rings, wherein each aromatic ring (benzenoid or heterocyclic) is counted and fused systems (e.g. indole, naphthalene and purine count as two rings). Almost concurrently, the less favourable prognosis for molecules with excessive aromatic character was also described in a review entitled "escape from flatland" by Lovering et al. [63], wherein the proportion of saturation was expressed as a fraction of sp³ hybridised heavy atoms (Fsp^3), an alternative to #Ar. These authors noted the statistically significant increases in the level of saturation with advances in drug development stages, mirroring the observation of decreased average aromatic ring count (Table 11). Additionally, a correlation with solubility (higher Fsp^3 = more soluble) and an observation that higher melting points correlated with lower Fsp^3 (Table 12), were also noted. These observations were consistent with predictions of the Yalkowsky general solubility equation (Sect. 2.3.1). To a simplistic approximation, with a flatter molecule, there can potentially be more π–π (flat) stacking in the crystal lattice, which contributes to elevated melting points. Additionally, other packing interactions, e.g. edge to face, are also feasible in crystalline structures that can also increase lattice energy. A further impact of aromatic systems is that they have a lesser entropy penalty on a change of physical state or moving from solution into a bound form, due to a lesser change in conformational movement of the rigid system (compare with the greater conformational motion of a saturated system).

Fig. 14 An illustrative
example of the average
distribution of number of
aromatic rings in
representative examples of
drug molecules (*black bars*)
and GSK discovery
examples (*grey bars*)
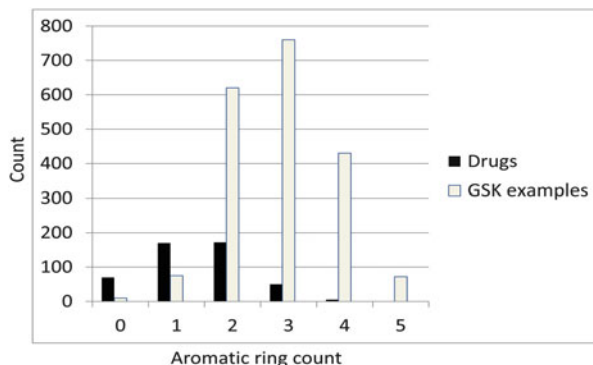
Yang et al. investigated the negative impact of aromaticity as a contributor to molecular topology and its implications on the progress of drug candidate molecules [64]. In a follow-up paper, Lovering further explored Fsp^3 in combination with chiral carbon count (as a measure of complexity), with demonstrable correlations with promiscuity and cytochrome P450 inhibition [65].

The extent to which discovery programmes have been working with overly aromatic compounds is illustrated in Fig. 14. This shows the distribution of aromatic ring counts within a representative set of drug molecules and a cross section of GSK discovery programmes. The high propensity for drugs to possess one or two aromatic rings is immediately apparent; the mean value from Richie and MacDonald's paper is 1.7 [62].

2.4.1 Lipophilicity and Aromaticity Act in Unison

In these kinds of analyses, one evident factor, as noted by Ritchie and MacDonald [60], is that there is a relationship between aromatic ring count and lipophilicity (log P and log D); there is a similar increase in lipophilicity as Fsp^3 decreases within broad sets of data. Similarly, there are often relationships/correlations between lipophilicity and molecular weight, polar surface area, hydrogen bond counts or other such descriptors. One aspect that became clearer within data sets containing more reliable chromatographic hydrophobicity measures [29] was that whilst there was often a relationship between aromaticity and measured Chrom log $D_{7.4}$, there appeared to be an effect of either over and above any correlation between them. This was best visualised by double pie distributions of particular developability parameters (see Sect. 4.2) versus binned Chrom log $D_{7.4}$ values and number of aromatic rings; when either binned Chrom log $D_{7.4}$ values or number of aromatic rings remained constant, then the other parameter also showed worsening outcomes in the developability data (Fig. 15). A notable observation was an apparent diagonal that differentiated between the more soluble and less soluble areas of the plot; this would be expected given the aforementioned trends. The observation was such that a line of Chrom log $D_{7.4}$ + #Ar = 7 described the partitioning well; this led to the
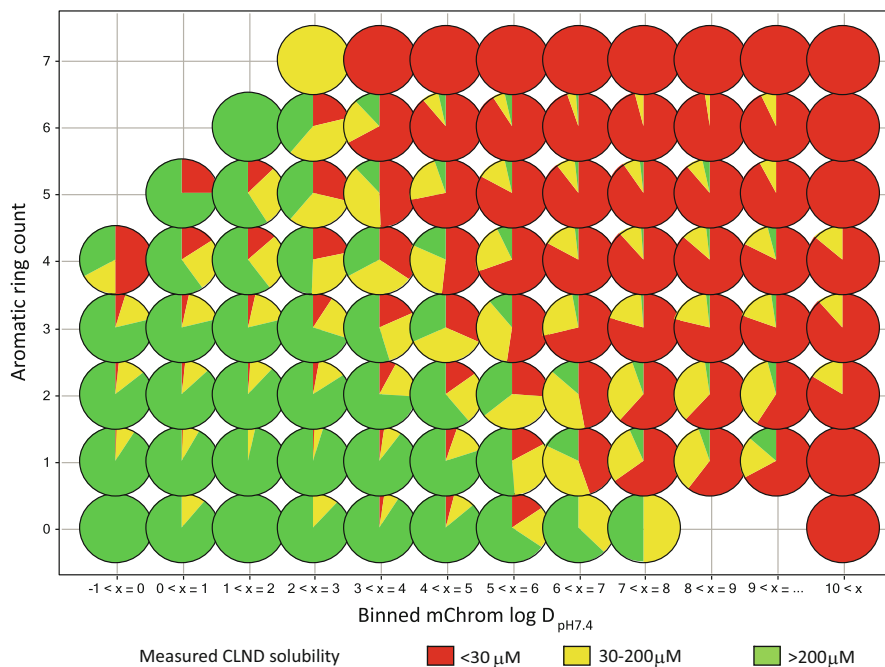
Fig. 15 Pie chart matrix representation of solubility category as a function of binned Chrom log $D_{\text{pH7.4}}$ and aromatic ring count, reproduced with permission from [29]

formulation of the so-called *solubility forecast index*; this showed apparent values for various combinations of lipophilicity estimates/measures, but differentiation was clearly best when the Chrom log $D_{7.4}$ values were used. It may be coincidental, but the value of 7 parallels observations of the general solubility equation (Sect. 2.3.1); if log $D_{7.4} = 3$ is taken as a base line, then conversion to the chromatographic scale adds 2 and the median drug aromatic ring count is 2 – thus $3 + 2 + 2 = 7$. Remarkably, the binned classification in Fig. 16 shows essentially a 1:1:1 ratio of measured solubility values for the Chrom log $D_{7.4}$ + #Ar values between 6 and 7, corresponding to the 1:1:1 ratio in Table 8.

The implications of these observations on other developability parameters are discussed in Sect. 4.2. Evidently, there is something inherent in the flatness of molecules that makes it favourable for such molecules not to reside in aqueous solution and to prefer to be bound – with a favourable energy outcome in the latter process. This not only gives rise to the more favourable binding to the target in question but also influences promiscuity and unfavourable outcomes in other measures of developability [such that the solubility forecast index was renamed property forecast index (PFI) due to the broader impact]. Advances in metal-catalysed cross-coupling methodology and how amenable such reactions have been proven to be for parallel synthesis methodology have contributed to the proportion of compounds with overly aromatic character. Employment of such
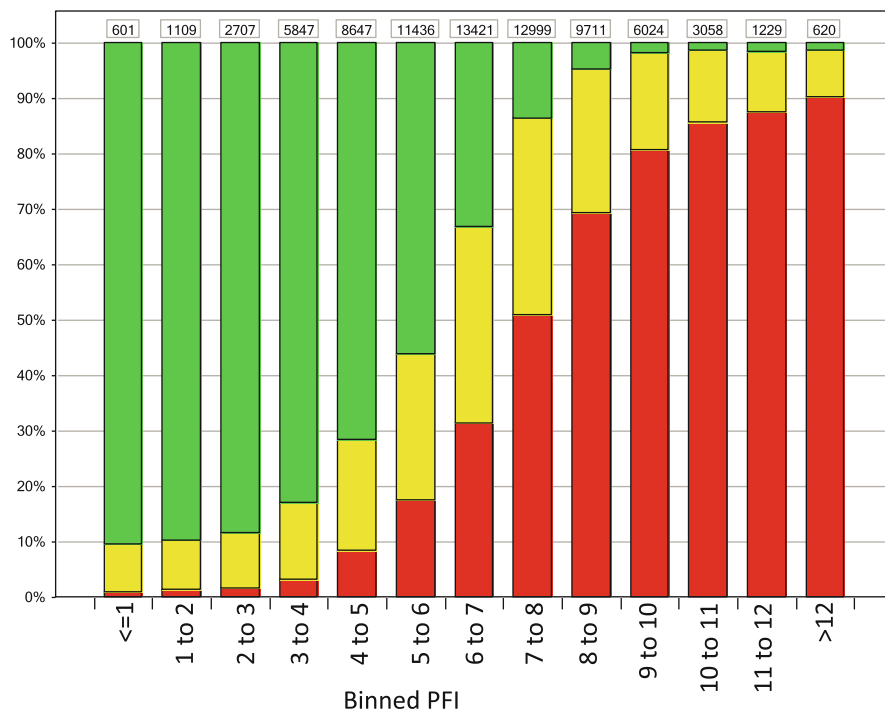
Fig. 16 Distribution of solubility category as a function of measured Chrom log $D_{7.4}$ + #Ar, for 96,000 measured GSK compounds, shaded using the same classifications as Fig. 15, reproduced with permission from [29]

reactions has frequently given benefit to drug discovery programmes, notably in terms of potency, but such chemistry has probably been a significant contributor to the high attrition of the "molecular obesity" era. Heterocyclic aromatic compounds can reduce lipophilicity [66], but it doesn't resolve the inherent issues connected with the flatness of the molecule. Not all heterocycles are equal [66], due to varying lipophilicity, and there may be an effect due to 5-membered rings versus 6; but the underlying observation of the addition Chrom log $D_{7.4}$ + #Ar suggests that the benefits are probably only due to the lipophilicity. These observations have engendered much debate as to the nature of an aromatic ring: for example, if some esoteric variation, twist or tautomer really counts as aromatic; but the simple mantra should be to "escape from flatlands" in addition to minimising lipophilicity. Naturally, there will always be exceptions – the premise of the property forecast index is that it is a probability guideline showing the likelihood of a given outcome; it is necessary to generate empirical data to look for pointers to gain comfort in working in what might be described as sub-optimal space. If PFI is high or Fsp^3 is low and the compound has low solubility and a high melting point, these should be physical warning signs that the compound is probably going to encounter a risk flag in its developability profile somewhere along the line.

3   How the Physical Environment of a Drug Changes from Administration to Target

The journey of a drug molecule from the point of administration to the site of action is a complex one, with many changes of environment. The most common route of administration is via the oral route [67] – which is perhaps the most difficult means of delivery, with a varied range of physical environments [68] and barriers ahead of the path of the molecule. But whichever means of delivery is used, aqueous solubility will be required to ensure that the active molecules can reach their desired targets, be this solubility in the gastrointestinal tract, blood plasma or lung fluid, by way of examples. There may be significant changes in the ambient pH encountered, which may profoundly influence solubility of charged compounds, for example, in moving from the stomach to the duodenum, a molecule sees a rapid rise in local pH from ~1 to ~5 (depending on species and feeding state). Physical properties are also important in the passage of compounds through membranes and their dispersal in various tissues and organs around the body; the metabolism and removal of drugs are also highly dependent on a compound's physical make-up [24]. Together these phenomena make up the science of pharmacokinetics, commonly described by the constituent processes of ADME (ADMET). Drug metabolism and pharmacokinetics are subjects that fill texts on their own [69] – so the focus of this short section will be to consider particular aspects where physical properties have the most profound impact.

3.1   The Passage Through the Body for an Oral Medicine

The stomach is usually the first significant environment for a formulated drug – tablets or capsules are often swallowed whole and there is usually little residence time in the mouth, where the pH will typically range from slightly acidic (pH ~5) to neutral. There is a rapid decrease in pH on entry into the stomach, where the pH of the acidic environment is typically between 1 and 3. This can be a big aid to drug dissolution for basic compounds, as they will have markedly increased solubility due to increased (or complete) ionisation at gastric pH (Sect. 2.1). Conversely, the solubility of compounds containing an acid group may be adversely affected due to the local pH, as the protonated form of the molecule will predominate (even if administered as a salt, as buffering to the low pH will result). Drug formulation or encapsulation may protect the drug from the harsh environment of the stomach. Dissolution in the stomach can also be aided by salt formation, and this can result in supersaturation of compounds, sometimes aided by components in food/formulations; this can maintain a compound in solution as it enters the rapid pH change on entering the duodenum. The various simulated gastrointestinal fluids described in Sect. 2.3.2 are used to try to mimic these varying conditions in order to better understand the likely properties of drugs, their dissolution and the effects of food

Table 13 Representative physical characteristics of pertinent regions of the human gastrointestinal tract [70]; the fluid volumes were determined by magnetic resonance imaging (MRI) under fasting conditions and 1 h after a meal

Region	pH	Fasted volume	Fed volume	Transit time
Oesophagus	~7.0	–	–	10–14 s
Stomach	1–2.5 (up to five fed)	13–72 mL	534–859 mL	Half emptying ~80 min
Duodenum	6.1–7.4	45–319 mL combined	20–156 mL combined	3.2 ($\pm$ 1.6 h) combined
Ileum	6.8–7.8			
Ascending colon	5.3–6.7	1–44 mL combined	2–97 mL combined	Highly variable, dependent on lower movements
Descending colon	5.2–7.0			

Fig. 17 The human digestive track illustrating the regions summarised in Table 13
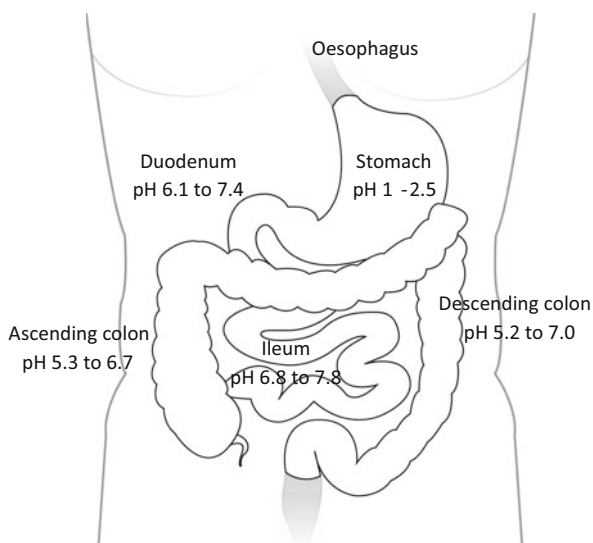
and transit time through the GI tract. Another important consideration is the available volume of aqueous medium (Table 13) to dissolve the drug [70], as this may not be enough to dissolve the dose.

On exiting the stomach, the pH rise encountered by the chyme is rapid, effected by bile and sodium bicarbonate secreted by the pancreas. Other components of bile, such as lipophilic acids and phospholipids, can aid the solubility of molecules through micelle formation (Sect. 2.3.2). Typically, as the absorption process starts, the pH will be at about pH 6.5, so the pK_a and log $D_{6.5}$ are important properties governing the absorption process through the brush border of the small intestine; the pH rises as the passage continues through the ileum (Fig. 17) and absorption can occur along the whole length and into the colon, where the pH can be marginally lower again.

Fig. 18 The developability classification system [73] plotting the predicted human permeability versus the dose vs solubility ratio. The lines at 250 and 500 mL (the units of the ratio) should be compared with the observed available volumes in Table 13
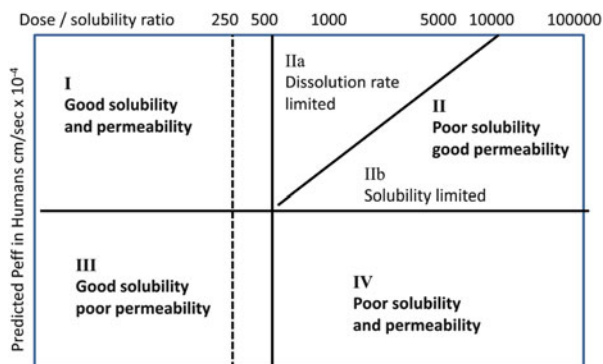
The absorption process has barriers other than physical one; the membrane contains protein transporters that enable both positive transport and efflux, working with and against the absorption process, respectively. Additionally, constituent cells contain many P450 enzymes that are present to oxidise lipophilic molecules as a further line of defence against xenobiotics.

Once absorbed, a compound will enter the blood stream in the hepatic portal vein, where the pH will be 7.4, with the next hurdle to absorption being the liver, where many metabolising enzymes lie in wait in the hepatocytes. Any lipophilic acids will likely be removed via the bile. Any drug that passes through the liver will now enter the systemic circulation; this route will involve further passage of the liver and also the kidneys. The filtration mechanism in the kidneys inherently removes hydrophilic compounds; the oxidation and conjugation mechanisms of metabolism are designed to help facilitate this by means such as sulphation or glucuronidation of xenobiotics or their oxidised metabolites [71].

Many drug molecules will be bound to the plasma proteins in the circulation such as human serum albumin (HSA) (Sect. 4.2.5), the binding to which is physical property dependent. Similarly, the volume of distribution of any compound (Sect. 5.1.1) will also be influenced by the charge and lipophilicity of the particular compound. Another important barrier that some drugs may need to cross is the so-called blood–brain barrier (BBB), which any medication acting on central nervous system or brain targets must cross. The characteristics required to achieve BBB penetration are complex, and the most reliable models are based on multiparameter optimisation methods [72], but lipophilicity charge and size have been established as key components.

3.2 Oral Developability Classification System: Dose, Solubility and Permeation

An important physical classification in oral drug delivery is the developability classification system (DCS) [73] Fig. 18, which classifies oral drugs based on

their solubility and permeation characteristics. This plots the dose solubility ratio versus the effective permeability from the intestine. Solubility and permeability can be measured, whilst the dose is usually estimated (prior to human trials) from the relationship that

$$\mathrm{Dose} \approx (\mathrm{Ceff}_{\mathrm{free}} \times \mathrm{Cl}_{\mathrm{int}})/F\%$$

where $\mathrm{Ceff}_{\mathrm{free}}$ = effective free drug concentration, $\mathrm{Cl}_{\mathrm{int}}$ = intrinsic clearance and $F\%$ = oral bioavailability.

Most oral drugs reside in the ideal quadrant I, with good solubility and good permeability, although some are soluble and poorly permeable. Very few are in quadrant four with low solubility and permeability. The final quadrant, where compounds have poor solubility and good permeability, is often subdivided to differentiate the compounds that have dissolution-limited solubility (this might be addressed using formulation and is most likely to be due to high melting points) or the not so attractive option of the intrinsic solubility being low due to excessive lipophilicity, for example.

An analysis of attrition in 2004 by Kola and Landis [74] highlighted that pharmacokinetic issues had been a major cause of attrition, but improved under-standing and progress in formulation technologies had enhanced the profiles of sub-optimal molecules and this risk had been reduced. Such efforts may have addressed some shortcomings of compounds in DCS class IIa or even IIb; however, in spite of this progress, overall attrition rates remained high and other shortcom-ings were emerging, notably with late-stage, off-target and other cytotoxicity effects. This can be particularly costly, as these issues would surface later in the clinical development process when considerable investment had been made. Many of these effects are likely to have been due to excessive intrinsic lipophilicity (Sect. 4.2.4), whereby the effects are beyond those associated with poor ADMET profiles.

4 Impact of Physical Properties on Developability and DMPK Parameters

4.1 General Aspects of Lipophilicity-Dependent Behaviour

Many aspects of an experimental molecule's behaviour can be correlated with its lipophilicity; this is often the basis of quantitative structure property relationships (QSPR) and a fundamental element of the majority of predictive in silico models used in drug discovery. The majority of the processes involved in the physical chemistry, pharmacokinetics and pharmacology of a drug molecule can be related to the relative propensity of a molecule to dissolve in aqueous or lipid-rich environments and the relative partitioning or distribution between them. There
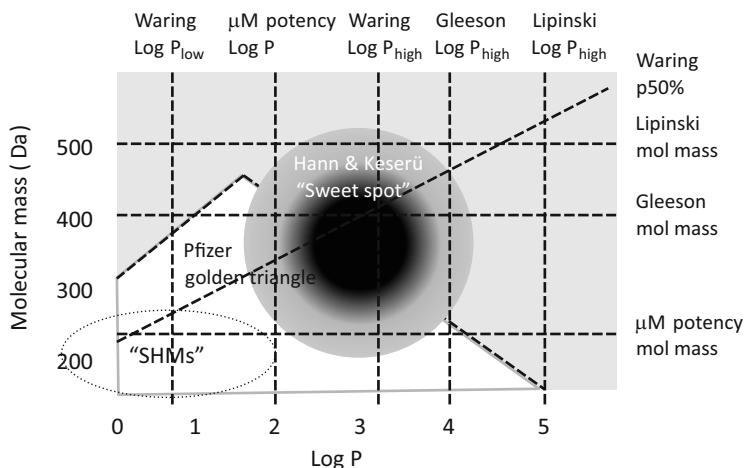
Fig. 19 Adapted version of the drug discovery "sweet spot" defined in the molecular mass–log *P* space proposed by Hann and Keserü [76], based on the concepts described in the text. The *diagonal line* "Waring p50%" is the proposed line for 50% chance of achieving reasonable permeability [77]. Additional regions for "small hydrophilic molecules" (SHMs) with potential paracellular routes of absorption (Sect. 2.1.1) and the Pfizer "golden triangle" (Sect. 4.2.1) have been added. The Pfizer triangle should be shifted right by perhaps a log unit to represent average log *P* values rather than the log *D* data it was based on

have been various reviews on the subject, output from which have been guidelines that suggest target values or optimum ranges for particular physical properties. In a review of lipophilicity, Waring suggested an optimum log *P* band for optimal DMPK outcomes, based on the cumulative suggestions of works looking at various DMPK parameters [75]. Taking this a step further, Hann and Keserü [76] proposed a sweet spot (Fig. 19) defined by log *P* and molecular weight, considering limits suggested by Lipinski [4], Waring [77] and Gleeson [78]. As described in Sect. 2.1.4, chromatographically derived hydrophobicity measurements have been shown to be more reliable and predictable, irrespective of the solubility of the molecule under investigation. Subsequently, in-depth analyses at GSK on the impact of such data in various developability assays clearly showed the relevance and value of the measurements. A reprise of these analyses is presented here, along with perspectives on other reviews and contemporary developments; results within would suggest a more complex picture than those mentioned in Fig. 19, with an impact of not only distribution *and* partition coefficients but also the particular influence of flatness/aromaticity and the facts that molecular weight may not be quite so relevant and that some lipophilicity dependencies are bilinear rather than linear.

Establishing that aromatic ring count has an impact over and above correlation with log *P* or log *D* in many scenarios was a key observation in the development of the PFI [29]; this was especially evident in a double pie plot such as that illustrating solubility (Fig. 15). In the development of predictive models, it is important,
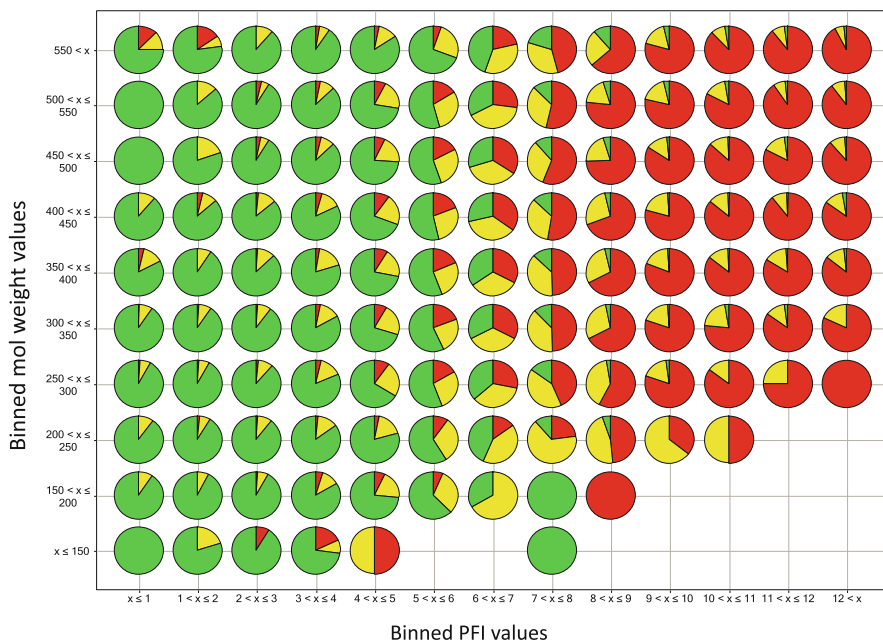
Fig. 20 Categorised multiple pie chart representation of distribution of solubility category as a function of measured Chrom log $D_{pH7.4}$ + #Ar (x-axis) and binned molecular weight (y-axis); coloured as Fig. 15. Reproduced with permission from the supplementary data of [29]

statistically, to use parameters with such orthogonal components in their interrelationships; otherwise parameters that show mutual dependence can aid and abet each other and give artificially inflated correlations [79]. It is interesting to note that a double pie plot of binned molecular weight versus binned PFI (Fig. 20) shows very little variation in solubility in a given PFI binned column, indicating that molecular weight has little impact on the distribution of solubility. In a give molecular weight bin, there is a marked decrease in the proportion of soluble compounds as PFI increases.

4.2   Impact of log P, log D_{pH} and Aromaticity on Particular Parameters

In addition to developability risks associated with poor solubility, the various assays and concepts in the following section have been used as indicators of risk in the drug development process. Each of these has been scrutinised in some detail as numerous predictive models and structure–property relationships with various physical descriptors have been established. Lipophilicity, in some guise, almost invariably plays a central role in all of these analyses; what was clearly perceptible

in the data sets investigated with chromatographic measurements was an enhanced resolution (as would be expected given the observations in Sect. 2.1.4). A clarification of some trends and dependencies were important outcomes, for example, differentiation between partition and distribution effects and establishing bilinear dependencies for permeation and some cytochrome P450 binding. Remarkably, on the chromatographic scale, a PFI value of 7 often gave something of a reflection where below this outcomes were generally good, above it generally poor.

4.2.1 Permeability

The crossing of biological membranes is fundamental to many aspects of drug action; for example, many targets are intracellular and processes of absorption from the intestinal tract and reabsorption in the kidneys rely on permeation of membranes to enable passage through relevant cells. Models to describe these processes are often complex and have been the subject of many research articles [80, 81]; indeed, the nature of the fundamental process of crossing a cell membrane has come under particular scrutiny in recent times, with challenges to the perceived wisdom and models [82, 83]. Permeation of a membrane consists of the partition of a compound from aqueous solution into a membrane and the reverse process as the molecule enters the medium on the other side, Fig. 21 [84]. A major point of contention is whether this process is mediated through the lipid bilayer of the membrane or facilitated by transmembrane proteins, which are liberally dispersed throughout cell membranes. Many of the known transmembrane proteins have been shown to enable passive or active (requiring input of energy via adenosine triphosphate) transport of specific molecules through the membrane and into the cell; others enable efflux mechanisms working in the opposite direction to remove molecules from within the cell. The contention is whether all transmembrane passage occurs via protein or lipid bilayer mediation; Kell et al. [82] argue that it
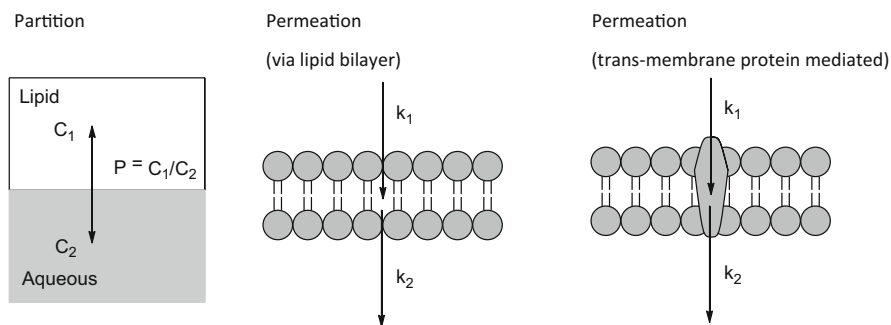
Fig. 21 Comparison between the partitioning (or distribution) and permeation processes, showing the partitioning of a molecule between aqueous and an immiscible solvent such as a lipid/octanol and the two rate constants controlling the passive permeation across a lipid bilayer or membrane, which may include a protein-mediated process (Adapted from [80])
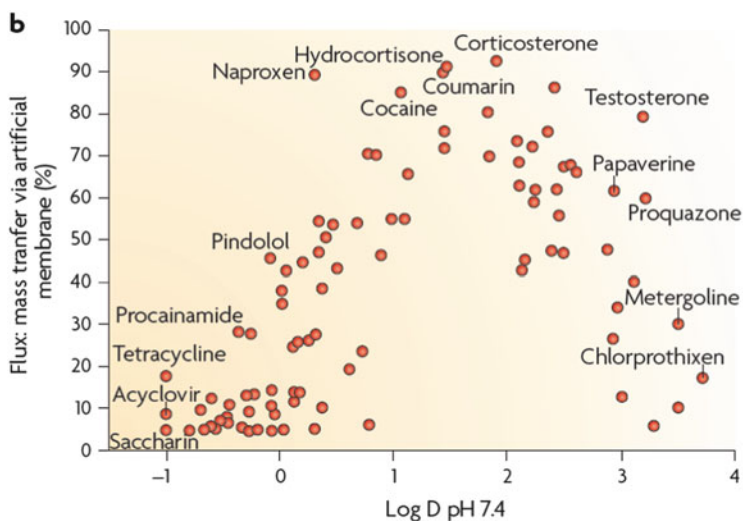
Fig. 22 The 1979 plot, reproduced with permission from [80], of artificial membrane flux versus measured log $D_{7.4}$ for a representative set of drug molecules – illustrating the bilinear response to lipophilicity

is likely that all is mediated by proteins; the review by Sugano et al. describes the coexistence of both mechanisms [82] and others have strongly advocated the role of the lipid bilayer and the relevance of parallel artificial membrane permeation assays (PAMPA) [83]. The definitive answers require more research, but there are fascinating opportunities if the nature, purpose and role of individual transporters and particular specificities of recognition can be understood. One thing that is certain is that the distribution of transporters and their expression levels varies with cell types and the state of those cells; this could offer exciting opportunities for targeting drugs to specific cell types in the future.

It was succinctly demonstrated in 1979 [85] that a set of licensed drug molecules showed a bilinear relationship between measured permeability through an artificial membrane and measured octanol–water log $D_{7.4}$ (Fig. 22). Within this set, it is likely that the OW log $D_{7.4}$ values (measured in the range of approximately -1–3.7) would be reliable given that, as effective drugs, they would be highly likely to be soluble (with such log $D_{7.4}$ values, the GSE or PFI would corroborate this). In subsequent years, as molecular obesity started to take hold, the less reliable measured lipophilicity values clouded the data and more complex models were proposed to try to explain observed trends in permeation in large data sets. The likely shortcomings within both calculated and measured lipophilicity values led to conflicting hypotheses as to its effect on permeation, encompassing extremes represented by each of linear, hyperbolic, sigmoidal, parabolic and bilinear relationships [77].

Chromatographically derived lipophilicity data presented compelling evidence as to the relationship between permeation and this measure of distribution [29]. It is very clear that the permeation rate has a bilinear response, which is consistent with

Table 14 Permeability rules developed by Waring [77] proposing molecular weight and log D cut-offs

Molecular weight	AZlog D
<300	>0.5
300–350	>1.1
350–400	>1.7
400–450	>3.1
450–500	>3.4
>500	>4.5

the 1979 study [84], fitting with the common sense rationale behind this. This would be rationalized by consideration of the two rate constants of importance: k_1, the rate of permeation into the membrane and k_2 the constant for permeation back out again on the other side. Thus, hydrophilic molecules would be expected to have a slow k_1, leading to poorer permeability, but k_1 would increase as lipophilicity increases. Similarly, as lipophilicity increases, then the k_2 parameter will slow; together these phenomena contribute to the bilinear response. This was shown to hold irrespective of the solubility class of the compounds. The patterns observed in artificial membrane data were mirrored in the data produced in MDCK cells, where a similar profile was observed, with the exception of a higher incidence of permeation with more hydrophilic compounds, some of which likely could have traversed the cell preparation by the paracellular route (Sect. 2.1.1). One key observation was that enhanced differentiation was observed when distribution rather than partition coefficients were considered, recognising that permeation rate is a kinetic observation and that the proportion of the more lipophilic uncharged state (the predominant influence on log D, see Sect. 2.1) was shown to be the key factor in determining permeation. The composite PFI value also showed a bilinear response, with optimum permeation between PFI values of 6 and 8.

The size of molecules does have an impact on their permeation; this was described in some detail by Waring [77] who demonstrated that a combination of lipophilicity and molecular weight gave a cut-off whereby there was a 50% chance of achieving a high level of permeability (Table 14). In this analysis, the bigger a molecule, the more lipophilic it needs to be to have a chance of achieving permeability. This is perhaps not very different from methodology used at GSK for a number of years, whereby size has been represented by calculated molar refraction (cmr) and plotted against log $D_{7.4}$ [86, 87]. Revisiting this plot shows a similar trend to that suggested by Waring [77] – and that some big and lipophilic molecules did show good levels of permeability – although in given size bins, the bilinear relationship with log D was apparent (Fig. 23). The findings of this plot are in accordance with the Pfizer "golden triangle" reported by Johnson et al. [88], which looks at oral bioavailability/clearance as opposed to permeation, recognising that big lipophilic molecules would have lower %F due to extensive first-pass metabolism. This fits with the observed distribution of Chrom log $D_{7.4}$ versus cmr values in Fig. 24.

Permeation and oral bioavailability are not to be confused, but the latter cannot be achieved without the former. The various analyses have shown that both of these important factors in drug discovery have complex physical dependencies, amongst
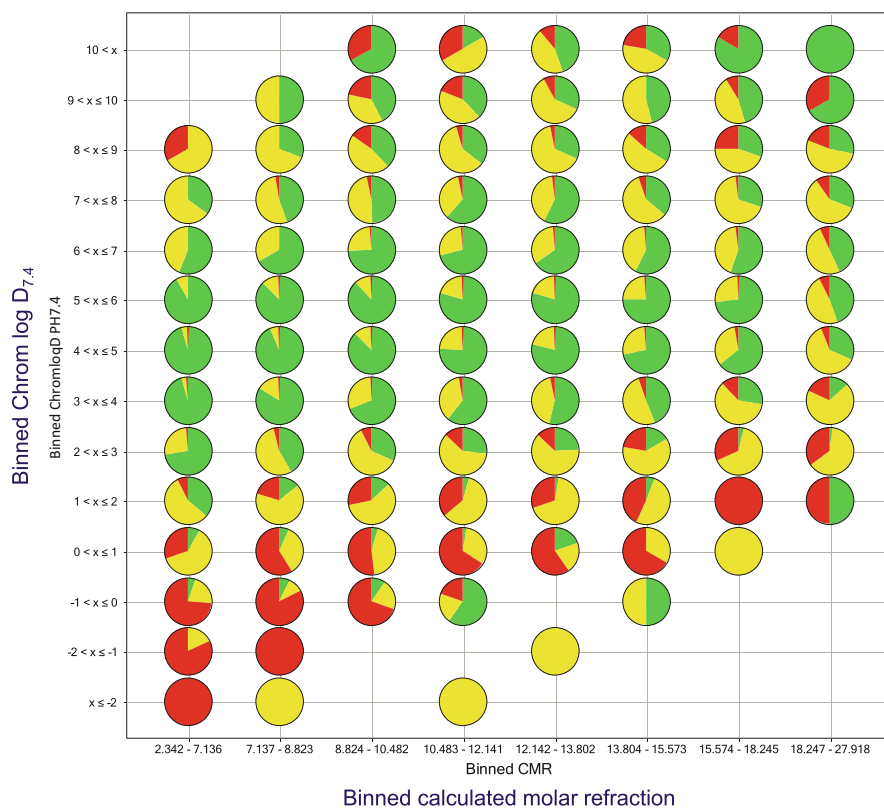
Fig. 23 Binned plot of measured Chrom log $D_{7.4}$ versus calculated molar refraction, with pies shaded by PAMPA permeation classes of compounds within, reproduced with permission from the supplementary data of [29]
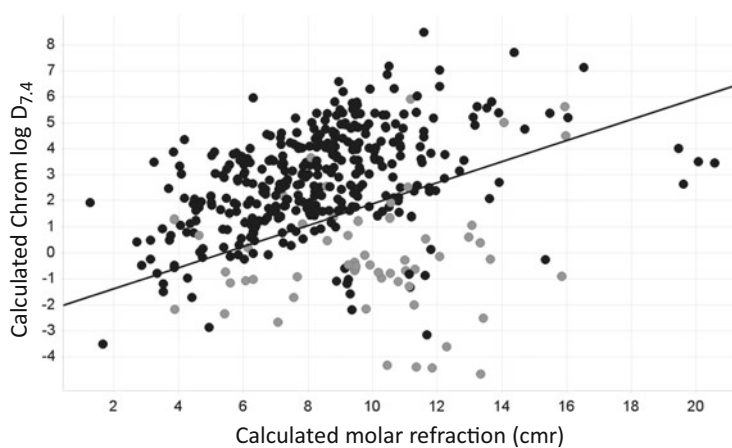
Fig. 24 Plot of calc Chrom log $D_{pH7.4}$ versus calculated molar refraction (cmr), a GSK training set of known drugs, adapted from [87]. *Black spots* have $F > 30\%$ and *grey* $F < 30\%$. The *line* is used as a discriminator of likely good or bad permeation; this is similar to one side of the Pfizer golden triangle [88]

which lipophilicity, aromaticity and size are key components. It should not be noted that the original Lipinski analyses [4] were concerned with predicting solubility and permeability, with molecular weight and log P at the core.

4.2.2 Cytochrome P450s

The propensity for a compound to interfere with metabolic processes mediated by cytochrome P450 enzyme isoforms is another widely used developability benchmark [89]. Lipophilicity is an established risk factor in such interactions [90]. In vitro estimates of P450 modulation are made by monitoring the rates of metabolism of known substrates of particular cytochrome P450 isoforms [91] in the presence of the test substance; high levels of activity in these assays are undesirable, indicative of potential drug–drug interactions with other substrates or inhibitors of particular isoforms. Such assays are run in bactosomes, bacterial membranes containing the human cytochrome P450s co-expressed with human NADPH–cytochrome P450 reductase.

One of the more striking revelations in the study of developability data for the set of compounds with measured Chrom log D values was the relationship between cytochrome P450 interactions and lipophilicity, particularly for the 2D6, 2C9, 2C19 and 3A4 isoforms. As with permeation, clear bilinear responses were observed (Fig. 25). These trends were not so apparent across log P or log $D_{pH7.4}$ values calculated or measured using OW methods. These observations were rationalised by a need to permeate into the bactosome preparation as a prerequisite event ahead of any particular binding to the P450 enzyme contained within. If bactosome permeation is one prerequisite to activity in P450 assays, other physical characteristics such as size, shape, charge and topology are well-established recognition factors for particular isoforms [92], as summarised in Table 15. For example, 1A2, which identifies relatively few active compounds, interacts with smaller, flatter,
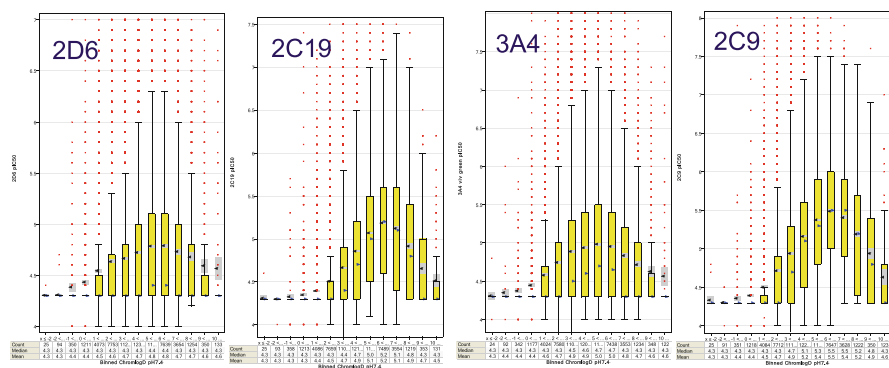
Fig. 25 Box plots of pIC$_{50}$ values versus binned Chrom log $D_{7.4}$ values for the four cytochrome P450 isoforms indicated, showing clear bilinear responses, reproduced with permission from the supplementary data of [29]

Table 15 Influence of descriptors on P450 binding activity: + to +++ represent increasing impact of the parameter. minus signs at either end (e.g. −+++−) are indicative of a bilinear relationship [29]

P450 isoform	Chrom log $D_{pH7.4}$	Size (CMR)	#Ar	Recognition factors
1A2	−	−++−	%Ar not #Ar	Highly aromatic/flat structures Smaller/hydrophobic
2D6	−+++−	−++−	+	Hydrophobic, optimum size, basic
2C9	−+++−	−++−	+++	Hydrophobic, optimum size, aromatic, acidic
2C19	−+++−	−++−	+++	Hydrophobic, optimum size aromatic, basic
3A4	−+++−	+++	+++	Hydrophobic, large, aromatic, basic

molecules (high proportion of aromatic rings rather than #Ar per se). Interestingly, the PFI summation showed a clear differentiation, maintaining the bilinear response for the three isoforms 2C9, 2C19 and 3A4.

4.2.3 hERG Binding

Interactions with the human ether-a-go-go-related gene (hERG) product, causing increases in the cardiac QT interval, are extensively used to identify the potential for cardiovascular risk in drug development [93]. The risk with positively charged lipophilic compounds has been well documented [94], although structure-based design [95], in addition to property-based design [96], has been used to reduce hERG activity. What was clear using chromatographic lipophilicity values was that the intrinsic lipophilicity (*partition*, log P, values) is the key driver of activity but very much driven too by positive charge. Increased aromatic ring count or lower sp^3 proportion has also been shown to increase risk; interestingly, the summation of Chrom log P and #Ar (intrinsic, *i*PFI) showed a significant trend of increased risk, with impact of #Ar over and above any correlation with log P. The value of *i*PFI > 7 showed a marked differentiation, such that when *i*PFI is >7, more than 50% of positively charged compounds exhibit a hERG pIC$_{50}$ > 5.

4.2.4 Promiscuity

Many literature reports and reviews have focussed on the increased risks of toxicity and other side effects due to excessive lipophilicity and aromaticity and other physical parameters (Table 16). These observations have been cited as symptoms of the molecular obesity epidemic that directly contribute to attrition. The underlying reasons for this are particular off-target effects (engaging with other proteins or receptor targets) or promiscuity, which can be defined as the propensity for

Table 16 Promiscuity–property relationships reported by different organizations, taken from Tarcsay and Keserű [97]

Organisation	Number of compounds	Number of assays	Physicochemical properties investigated	Dependency on physicochemical properties	References
Pfizer	75,000 in-house	220 in-house	MW, log P	MW: negative; log P: positive	[98]
Pfizer	1,098 drugs and 430 project compounds	70 Cerep BioPrint panel	log P	log P: positive	[98]
Pfizer	108 clinical candidates	48 Cerep BioPrint panel	log P, TPSA	log P: positive; TPSA: negative	[99]
Organon	138 in-house	Promiscuity was defined as low selectivity	MW, log P	MW: negative; log P: positive	[100]
Novartis	3,138 in house	50–79 in-house	MW, Alog P, HBA, HBD, #o, #N, RotB, #rings, ring assemblies, terminal rotamers, chain assemblies	MW: positive; log P: positive	[101]
AstraZeneca	2,133 drugs and ref compounds	200 Cerep BioPrint panel	MW, clog P, pK_a, #rings MW	MW: positive; log P: positive; pK_a: positive for bases	[5]
Roche	213 in-house	80 Cerep	In silico: MW, HAC, HBD, HBA, RotB, #rings, #Ar rings, #non-Ar, rings, PSA, clog P, amphiphilicity, pK_a, ionisation state, clog D; Measured: solubility, log $D(7.4)$, permeability, pK_a	MW: no; log P: positive	[102]
AstraZeneca	2,267 drugs and ref compounds	200 Cerep BioPrint panel	HAC, clog P and molecular framework (fMF)	HAC: bell shaped; log P: positive	[103]
ChEMBL	40,408 literature	>500 literature	MW, log P, pK_a	MW: positive; log P: positive; pK_a: positive for neutrals and bases	[15]
GSK	>2,500 drugs and in-house	>490 in-house	Chrom log P and #A	log P + #Ar: positive	[29]

Roche	2,413 drugs and drug-like compounds	141 safety-relevant targets	MW, clog P, pK_a, HBA, HBD, #Ar, PSA, RotB	log P: positive pK_a: positive for bases	[104]
Novartis	656 drugs 73	In-house	log P, pK_a	log P: positive pK_a: positive for neutrals and bases	[105]

activity at multiple receptors/targets. Promiscuity has been assessed by screening compounds across multiple biochemical targets, with the output being the proportion of these assays that a given compound engages with at a particular defined affinity. For example, a set of over 2,500 compounds, including exemplars from GSK lead optimisation projects and marketed drugs, was screened in >490 assays incorporating a diverse range of >380 protein targets and phenotypic end points [29]. The output was reviewed using measured chromatographic lipophilicity data, an outcome of which was that the intrinsic lipophilicity, log P, showed a clearer differentiation between the increasing levels of promiscuity as lipophilicity increased. As discussed in Sect. 2.4, higher numbers of targets engaged were observed as the aromatic ring count increased. The observation again held that promiscuity increased with both lipophilicity and #Ar; again, the summation of Chrom log P and #Ar (intrinsic, iPFI) showed utility as a risk indicator (Fig. 26). Fascinatingly, it appeared that values of Chrom log P + #Ar < 7 (or <5 on the OW scale using clog P) are commensurate with the indication of likely low promiscuity; above these values there is >50% chance of inhibiting more than five assays with $pIC_{50} > 5$.

The observations of the GSK exercise are broadly in agreement with the rest of the analyses in Table 16 [97]. However, as with other risk factors, it is unlikely that all variations will be explained by such a simplistic guide as PFI or iPFI, but it does make a good first guideline and expresses a risk for a defined, easily understood and reliably predictable molecular descriptor. The value of performing the analyses with measured lipophilicity data was clearly demonstrable, whereby a differentiation between intrinsic and effective hydrophobicity was evident; however, the quality of the BioByte-derived calculated log P values was also apparent in the analysis [29]. Within Table 16, some analyses, e.g. Leeson et al. [106] and Peters et al. [102], proposed that increased promiscuity was evident for lipophilic bases. However, within the GSK set, a raised overall promiscuity for basic compounds compared with other charge classes was observed, but it appears that log P is the key driver regardless of ionisation class, as within measured or calculated intrinsic lipophilicity bins, little charge-implicated variation was observed [29]. A general observation across data sets is that the average log P values of charged compounds were markedly higher than log P values for neutral compounds, which in itself would be indicative of increased promiscuity. This is borne out in the follow-on study in "Escape from Flatlands II" [65] – wherein the improved overall properties of compounds in the data set were heavily influenced by charge. This could be a risky strategy, as Table 16 would indicate that a good balance of PFI and iPFI is commensurate with minimising risk. One of the clear messages in these analyses is that molecular weight per se is not a promiscuity risk, but may show a signal due to its relationship with lipophilicity, a perception surfaced by Hann and Leach in their revisiting of the concepts of complexity in molecular design [107].
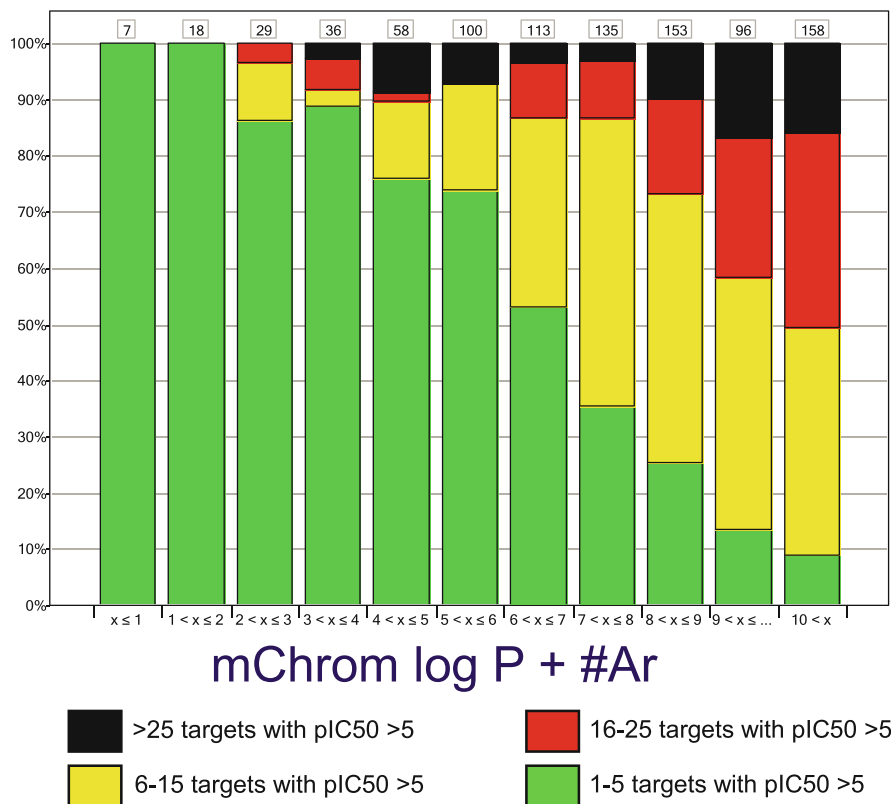
Fig. 26 The distribution of promiscuity scores, based on number of targets engaged with a $pIC_{50} > 5$, by binned iPFI (mChrom log P + #Ar) reproduced with permission from [29]

4.2.5 Human Serum Albumin Binding

The extent of binding to plasma proteins is often cited as a developability concern in drug discovery; whilst it may not be an attrition risk per se, high percentage binding is a characteristic associated with lipophilic compounds and contributes to reductions in both free fraction and drug efficiency (Sect. 5.1) [108]. Human serum albumin is the most abundant plasma protein and usually the benchmark investigated in drug discovery studies; this can be routinely measured by high-throughput HPLC methods or conventional analysis of in vitro or ex vivo samples [109]. The percent bound data should be treated with a little caution [110], as the binding to HSA can be more complex than a simple non-specific binding process, due to specific binding sites that bind some compounds in a stoichiometric manner and a propensity to bind acids to exposed basic residues.

Looking at binding data alongside measured Chrom log P and Chrom log $D_{pH7.4}$ [29], increases in the levels of HSA binding [either represented by % bound or expressed as log K_{HSA}, where K_{HSA} = (% bound/% unbound)] were observed as

both intrinsic lipophilicity and effective lipophilicity increased (Fig. 27). An explanation might be that these observations are reflective of the multiple types of interactions involved with HSA binding, whereby both intrinsic hydrophobicity and effective hydrophobicity can play a role. The PFI summation of Chrom log $D_{pH7.4}$ plus #Ar gave further enhanced resolution of the data, and the impact of #Ar in binding was clearly orthogonal to any impact of lipophilicity. With the chromatographic PFI value of 7, an inflection was again observed, with the box plots suggesting likely binding to above or below 95% at either side of this value.

4.2.6   Intrinsic Clearance

The intrinsic clearance of molecules (Cl_{int}), measured by the rate of disappearance of a test compound in a human liver microsome preparation, is related to hydrophobicity in addition to particular structural liabilities [111]. Interrogating such data using Chrom log D values was supportive of log $D_{7.4}$ and not log P being the more important driver. Clearly, the metabolism of a given compound will depend on a number of steric and structural factors, but, nonetheless, the composite Chrom log $D_{pH7.4}$ + #Ar (PFI) value accentuated a bilinear change in Cl_{int}, consistent with a key role for CYP 450-mediated metabolism. When PFI > 7 more than 50% of compounds showed a Cl_{int} rate of >5 mL/min/kg.

4.3   *The Composite Risks of Poor Physical Profiles*

The preceding sections illustrate the impact of physical properties in various assays used to benchmark the developability of a compound. Taken together, the output of these analyses is presented in Fig. 27 [29], wherein the proportion of compounds meeting certain levels in that assay is presented against binned property forecast index values (or intrinsic PFI for hERG and promiscuity). In addition, the proportion of a representative set of 240 oral drugs (with >30% human bioavailability) is illustrated in the bottom row. The risk of higher PFI/*i*PFI is evident in each row, except for permeation and some of the CYP 450 risks, where the bilinear response has an impact; indeed, the permeation pattern indicates a potential reason for inflated PFI values, given to range where optimum permeation is achieved (in the range 6–8). However, risks for the developability parameters investigated are clearly exacerbated above a PFI/*i*PFI of 7, but a value of <5 would appear to be even more desirable. Eighty-nine percent of the oral drugs have a (calculated) PFI < 7; 65% below five.

   An illustration of this can be seen in Fig. 28 where a composite chance of success (derived from multiplication of the fraction meeting the requirements in any given PFI bin) is plotted against the distribution of drugs in the given bins. This includes the oral set used in Figs. 5 and 24 and the top 100 oral medicines in 2012. Whilst there is a recent movement (reflecting changes noted in [5]) towards compounds in

Assay / target value	PFI = mChrom log $D_{pH7.4}$ + #Ar								
	<3	3-4	4-5	5-6	6-7	7- 8	8-9	9-10	>10
Solubility >200 mM	89	83	72	58	33	13	5	3	2
%HSA <95%	88	80	74	64	50	30	17	8	4
2C9 pIC$_{50}$[a] <5	97	90	83	68	48	32	23	22	38
2C19 pIC$_{50}$ <5	97	95	91	82	67	52	42	42	56
3A4 pIC$_{50}$ <5	92	83	80	75	67	60	58	61	66
Cl$_{int}$ <3 ml/min/kg	79	76	68	61	54	42	41	39	52
Papp >200 nm/s	20	30	46	65	74	77	65	50	33
	iPFI = mChrom log P + #Ar								
hERG pIC$_{50}$ <5 (+1 charge)	86	93	88	70	54	36	29	21	11
[b]Promiscuity <5 hits with pIC$_{50}$ >5	85	78	74	65	49	30	20	13	7
% Oral Drugs with F>30%	35	17	13	10	13	6	5		

Fig. 27 Percentages of compounds achieving defined target values in the various developability assays described in this section, categorised by PFI or *i*PFI bins, adapted from [29].

>67%	34 to 66%	<33%

Shading = % chance of achieving target in that particular bin (or preferred PFI ranges in the drug profiling row)
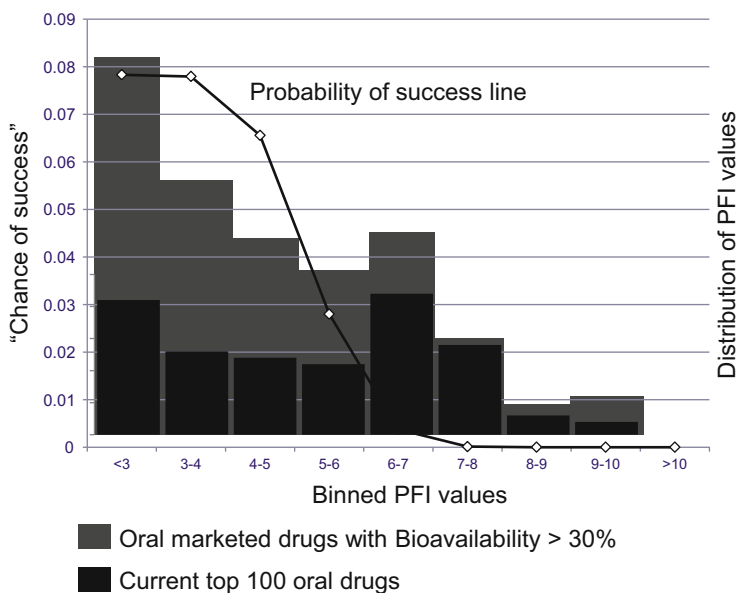
Fig. 28 Plots of the "chance of success", from the multiplication of the risk factors in Fig. 27 versus PFI bins and the distribution of oral drugs within the two sets indicated in the text. On a log scale, the apparent asymptote from PFI > 7 continues downwards as PFI increases

the PFI 6–8 bins, only five molecules in the top 100 oral drugs have a PFI value above eight.

5   Efficiency Metrics and Their Interrelationship with Physical Properties

The preceding sections have focussed on the background to physical properties, the impact of modulating them and the consequences/risks in the drug development process if they are sub-optimal. Consequently, the prototype drug molecule with a better chance of reaching the market will be soluble, modestly lipophilic and have enough structural features to engage effectively with its target without overly interfering with the multitude of other functional molecules and macromolecules that it will encounter. Discussions in the previous sections have been concerned with the passage of the molecule to its site of action at an effective concentration (pharmacokinetics) and off-target effects, rather than on how effectively and efficiently the investigational molecule engages with its target. There are two concepts in contemporary drug discovery that look at these latter properties, namely, drug efficiency and ligand efficiency.

5.1   Drug Efficiency

Drug efficiency is a concept that seeks to quantify the available free fraction of a drug at the site of action. This is derived from the simple definition: $DRUG_{eff}$ (%) = biophase concentration × 100/dose, where the concentration is that measured at the site of action (the free drug hypothesis) [112].

The biophase concentration is defined by the unbound volume of distribution (V_{du}) and can be measured from an administered (intravenous) dose divided by the free concentration (i.e. the product of the total blood concentration and the fraction unbound to plasma proteins). The maximal drug efficiency ($DRUG_{eff\ max}$) for a given compound is thus defined as the reciprocal of the unbound volume of distribution (and can be multiplied by 100 to express as a percentage). Given the multitude of variables influencing the free concentrations and distribution, typical drug efficiencies are usually low (usually single digits or fractions of 1%).

Using the example of Braggio et al. [112], Table 17 illustrates a series of potential development candidates that targeted a peripheral receptor (thus the biophase is unbound fraction in blood plasma). Each of these met standard progression criteria for pharmacokinetics; but the drug efficiency metrics show that the best compounds are likely not to be those with highest potency (C, D and E) but A or particularly B (the latter likely to have the lowest dose based on the combination

Table 17 In vitro and in vivo pharmacokinetic parameters of a series of potential candidates taken from [112]

Parameter	Acceptance criteria	Drug A	Drug B	Drug C	Drug D	Drug E
In vitro pharmacology						
pK_i		6.9	7	7.5	7.4	7.3
In vitro data						
P_{app} (nm/s)	>50	85	90	122	143	115
Fu blood	>1	5.4	2.4	1.5	2.4	1.9
Cli (ng/min/g)	<2	0.5	0.6	0.5	0.7	0.8
In vivo rat DMPK						
Vdss (L/kg)	<5	0.4	0.4	2.3	2.1	1.8
Clearance (mL/min/kg)	<60	3	2	18	42	35
Half-life (h)		1.5	2.3	1.5	0.6	0.6
Bioavailability (%)	>25	60	80	67	50	64
C (%)		1.8	2.7	0.25	0.12	0.23
$DRUG_{eff\ max}$		9.7	6	0.65	1.14	1.06

of DMPK and drug efficiency data). Minimising the dose is the desired outcome of the application of drug efficiency concepts [113].

The concepts of optimisation driven by drug efficiency are summarised in Fig. 29, wherein $DRUG_{eff}$ is plotted versus affinity, illustrating the desirable regions of higher affinity coupled with higher drug efficiency. Lines on the graph also indicate where 50 and 90% receptor occupancy (RO) would be achieved at a dose of 1 mg/kg; RO is related to the free fraction concentration and the pIC_{50}.

In Fig. 29, the upper hashed line is derived from the combination of drug efficiency ($DRUG_{eff}$) and the affinity required reaching the hypothetical PK/PD target of 90% of receptor occupancy at a 1 mg/kg dose. To achieve this, the extreme cases would represent either high affinity and low $DRUG_{eff}$ or low affinity and high $DRUG_{eff}$. The hashed lined represents the case where only 50% receptor occupancy would be required. The ideal outcome is with $DRUG_{eff}$ vs affinity combinations in the upper-right corner region, which will have lower efficacious doses. The directions taken from the lead illustrate the impact of using a "better balanced" approach (driven by properties, arrow A) rather than just chasing potency with a property imbalance (arrow B).

5.1.1 Physical Estimates of Drug Efficiency

The preceding discussion of drug efficiency is largely based on measured in vivo parameters, but useful high-throughput physical methods have been established to produce useful and relevant estimates of drug efficiency. A combination of HPLC columns can be used to estimate protein binding (with immobilised human serum albumin) and volume of distribution (using an immobilised artificial membrane), which have been shown to give reasonable estimates of the actual properties
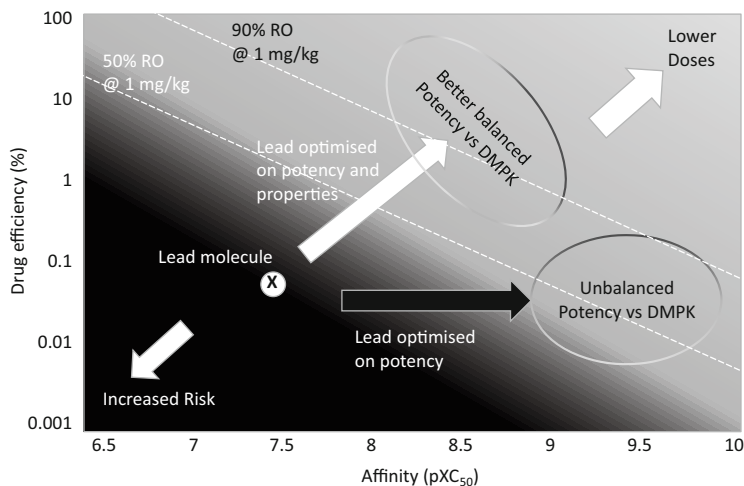
Fig. 29 DRUG_eff vs affinity (pXC_50) plot analysis, adapted from [112], showing impact of pursuing more drug-efficient molecules from hypothetical lead "X", achieved by concurrently optimising both potency and physical properties. *RO* receptor occupancy at the given doses

measured in vivo. These parameters have been measured and the output computed on a set of 70 known drug molecules to generate an estimate of maximal drug efficiency, $DRUG_{eff\ max}$ [114]. This is an empirically derived concept, described by Eq. (2).

$$\log(DRUG_{eff\ max}) = 2 - (0.23 \times \log K_{HSA} + 0.43 \times \log K_{IAM} - 0.72) \quad (2)$$

wherein

$$\log K_{HSA} = \exp(\log K_{HSA})$$
$$\log K_{IAM} = 0.29 \times \exp(\log K_{IAM}) + 0.70$$

The value of the IAM surrogate was established using steady-state human volume of distribution data, which demonstrated a good correlation between the in vivo figures and estimates from the HPLC measurements [115]. Typically, lipophilic and basic compounds will show up with strong binding to IAM, reflective of their generally higher volumes of distribution.

5.2 Ligand Efficiency and Related Measures

The concepts related to ligand efficiency relate to the potency achieved relative to various properties of a given molecule. These have been particularly useful in the rapidly emerging science of fragment-based drug discovery [7–9], but they also

have relevance to any activity-based analysis in hit identification and lead optimi-
sation [10]. Various metrics have been used to describe ligand efficiency in one
form or another [76], but each essentially relates to the activity at a given target
relative to the size or weight of the ligand.

In the most commonly used form, ligand efficiency (LE) is defined as the
binding energy, with units of kcal/mol per heavy atom (heavy – non-hydrogen –
atom count, HAC) in the molecule, derived from dividing the Gibbs free energy of
binding by HAC [11]. The Gibbs free energy of binding (ΔG) is defined as $\Delta G = -RT\ln(K_d)$. In practice, the actual K_d is not always measured directly, but can
usefully be substituted by pK_i or pIC_{50} values (so in strict use, data of the same
origins should be employed). Empirical results have shown that a value of
0.3 kcalmol^{-1}/heavy atom is indicative of a reasonable level of efficiency and
this is a generally accepted target level for use in lead discovery and optimisation,
although the value for many drugs is rather higher [116]:

$$\text{Ligand efficiency (LE)}$$
$$\text{LE} = -\Delta G/\text{HAC}$$
$$\text{LE} \sim -RT\ln(K_d)/\text{HAC}$$

where R is the gas constant, T is temperature and K_d is the kinetic dissociation
constant.

$$\text{LE} = 1.37 \times (pK_i \text{ or } pIC_{50})/\text{HAC}$$

Alternative ligand efficiency measures have been proposed, but increasingly the
simple LE term is gaining wide acceptance and has shown good utility [11]. An
alternative is the binding efficiency index (BEI), derived by dividing an activity
value by molecular weight. Two size-independent ligand efficiency measures that
have been proposed are Fit Quality (FQ) [117] and size independent ligand
efficiency (SILE) [118].

$$\text{BEI} = (pK_i, \text{or } pIC_{50})/(\text{molecular weight})$$
$$\text{FQ} = [pIC_{50} \text{ or } pK_i/\text{HAC}]/\big[0.0715 + (7.5328/\text{HAC}) + \big(25.7079/\text{HAC}^2\big)$$
$$- \big(361.4722/\text{HAC}^3\big)\big]$$
$$\text{SILE} = pIC_{50} \text{ or } pK_i/\text{HAC}^{0.3}$$

The emergence of fragment-based drug design has been a notable feature of the
last decade, often touted as a key element of contemporary practice aimed at
reducing molecular obesity and improving physical property profiles. The essence
of this technique is to identify very small molecules, typically defined as containing
8–18 heavy (non-hydrogen) atoms, which effectively and efficiently bind to a
protein target. Because the molecules are small, this binding is inherently weak –
at roughly the millimolar level – due to the lesser number of productive contacts
and interactions the molecule makes [119]. This requires sensitive biochemical
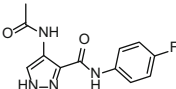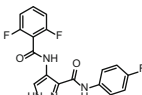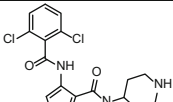
Table 18 Comparing the suggested evolution of a fragment molecule into a candidate whilst maintaining LE and LLE values (taken from [120]), with comparative data for the average drug in the set presented in [87] added in the final row (the target LE/LLE values of 0.3 and the proposed heavy atom counts are perhaps a little too low and too high, respectively)

Compound	Heavy atom count	IC_{50} (μM)	clog P	LE	LLE_{LS}	LLE_{At}
Fragment	12	2,300	1.0	0.3	1.7	0.3
Hit	20	40	1.6	0.3	2.8	0.3
Lead	30	0.25	2.4	0.3	4.2	0.3
Candidate	36	0.01	3.0	0.3	5	0.3
Average drug	24	0.01	2.7	0.46	5.3	0.46

techniques such as surface plasmon resonance, nuclear magnetic resonance, isothermal titration calorimetry or melting point shifts to detect or estimate the kinetics of binding. From this small start, the fragment is then systematically grown, most frequently guided by X-ray crystal structures of the fragment bound into the target protein, into larger molecules with enhanced activity, whilst maintaining effective contacts and ligand efficiency. A guiding principle in all of this is to continually exploit effective interactions, such that each added portion contributes to the binding in an effective manner – this is judged by maintenance or improvement of ligand efficiency. Typically the K_d of a fragment would be in the millimolar to hundreds of micromolar range – which would equate to the value of 0.3–0.4 (kcal/mol) being an effective LE value (Table 18). As the molecule grows, efficient growth would result in at least the maintenance of the LE value as activity and heavy atom count increase proportionately. This is illustrated in Fig. 30 with exemplar compounds in the development of inhibitors for CD38 kinase [121] and P38α kinase [122] by Wyatt, Gill and co-workers at Astex.

5.3 Ligand Lipophilicity Efficiency Measures

As has been outlined in preceding sections, there is an inherent preference for lipophilic molecules to bind in a rather non-specific fashion to protein or other lipophilic environments in biological systems – the basis of, for example, promiscuity, high plasma protein binding and poor efficacy. Increased lipophilicity is often pursued in experimental molecules because it also enhances target affinity – so often the top target value in proposed molecular profiles so commonly used in contemporary drug discovery. A refinement of ligand efficiency has been to consider ligand lipophilicity efficiency (LLE, sometimes LLE_{LS}), which corrects for potency enhanced by the lipophilicity of the molecule, expressed as pK_d minus log P in the most simplistic form [5], although pK_d, pIC_{50}, clog P and log $D_{7.4}$ (defined as lipE [123]) values have all been used in variations on this theme. Typically, a value of greater than five would be regarded as expedient (commensurate with the average values for drugs that indicate a pIC_{50} of around eight and clog $P < 3$). The intrinsic lipophilicity of a molecule is probably the key parameter

CDK2 Ref [121]				
IC$_{50}$ (µM)	185	97	0.85	0.003	0.047
HAC	9	14	19	26	25
LE	0.57	0.39	0.44	0.45	0.40
LLE$_{AT}$	0.43	0.34	0.44	0.43	0.49
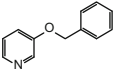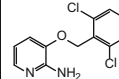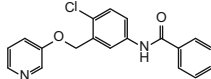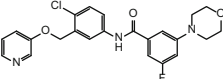
P38α Ref [122]			
IC$_{50}$ (µM)	1000	109	30	0.065
HAC	14	17	24	31
LE	0.29	0.32	0.26	0.32
LLE$_{AT}$	0.13	0.10	0.12	0.25

Fig. 30 Structures, potencies and ligand efficiencies observed during the growth of two fragment hits for kinase targets pursued by Astex [121, 122]

in determining the activity of a compound, so it is perhaps correct to consider log P as the relevant measure, although it is equally valid to consider maximising the gap between potency and effective (log D) lipophilicity. A composite LLE value (LLE$_{AT}$) has been developed by workers as Astex [120], which recognises both the heavy atom and lipophilicity contributions to activity; a correction factor scales LLE$_{AT}$ to reasonably track LE, such that in practice a marked difference between the two would be suggestive of lipophilicity-driven activity.

LLE$_{AT}$ was derived as follows:
A modified free energy term ΔG^* was defined, whereby

$$\Delta G^* = \Delta G - \Delta G_{\text{lipo}}$$
$$\approx RT \ln(\text{IC}_{50}) + RT \ln(P)$$
$$\approx \ln(10)^* RT (\log P - \text{pIC}_{50})$$

From this an idealised profile identified a constant to rescale LLE$_{AT}$ to aim for similar scaling to LE, such that ~0.3 defined activity vs HAC and log P as in Table 18. Thus

$$\text{LLE}_{AT} = 0.111 + [(1.37 \times \text{LLE})/\text{HAC}]$$
[where LLE, as defined above $= \text{pIC}_{50}(\text{or } pK_i) - \text{clog } P$]

The principles of fragment- and/or ligand efficiency-based drug discovery are illustrated in Table 18, showing how hypothetical stages in the discovery process might be described in terms of potency and size, showing activity expectations that would maintain LE and LLE_{AT} at 0.3.

The use of LLE_{AT} in practice is exemplified in the graphs Fig. 31, whereby the LLE_{AT} gives a clear indication that the P38α series [122] was not quite in such a good space as the CDK2 series [121]. The latter showed both good LE and LLE_{AT} (the impact of replacing an aromatic ring and reducing log P is evident in this example too; even though potency and LE dropped, LLE_{AT} improved and the overall properties suggested a molecule that has an increased chance of success). Conversely, whilst the P38α example maintained ligand efficiency, the LLE_{AT} was sub-optimal, suggesting it would be difficult to identify a lower developability risk molecule from this series.

5.4   The Combined Influence of Efficiency Metrics and Physical Property Measures

A rational conclusion to the discussion in this subsection on ligand efficiency measures would be to consider such concepts together with physical properties; it is logical to expect that drug molecules should be ligand efficient and have good solubility and low lipophilicity. This does not contradict the Hansch concept of making molecules "hydrophilic enough, whilst possessing sufficient activity for a pharmacological effect at their target". It should be evident from this chapter that the majority of drugs do indeed posses good physical properties; there is growing evidence in the literature that drugs are also generally highly ligand efficient [116] – and, by implication, should also posses good LLE.

Data are emerging to substantiate this hypothesis. Factor Xa was cited as the most researched target in the industry in the period bridging the millennium, and thus far three licensed drugs have emerged from the research, with several others withdrawn from clinical trials for a number of reasons. In a retrospective, quantitative analysis of what the key drivers of the successful programmes were, plots of ligand efficiency measures against physical measures showed remarkable differentiation between the drugs, clinical candidates and the vast majority of the molecules made [87], for example, the LLE_{AT} vs PFI plot in Fig. 32.

A more extensive analysis by Leeson [116 and personal communication] has indicated that this general principle holds across 46 oral drugs acting at 25 targets, with data extracted from the CHEMBL database. Aside from kinase inhibitors, with very few exceptions, the drugs for any given target are usually amongst those compounds with the best combinations of LE and LLE. Figure 33 illustrates how this observation holds across most target classes (the median value shows that just 2.7% of compounds have both better LE and LLE than the drugs), other than kinases (where 22% are more efficient). Many kinase inhibitors exploit bulk to achieve selectivity over related kinase enzymes [124]; thus,
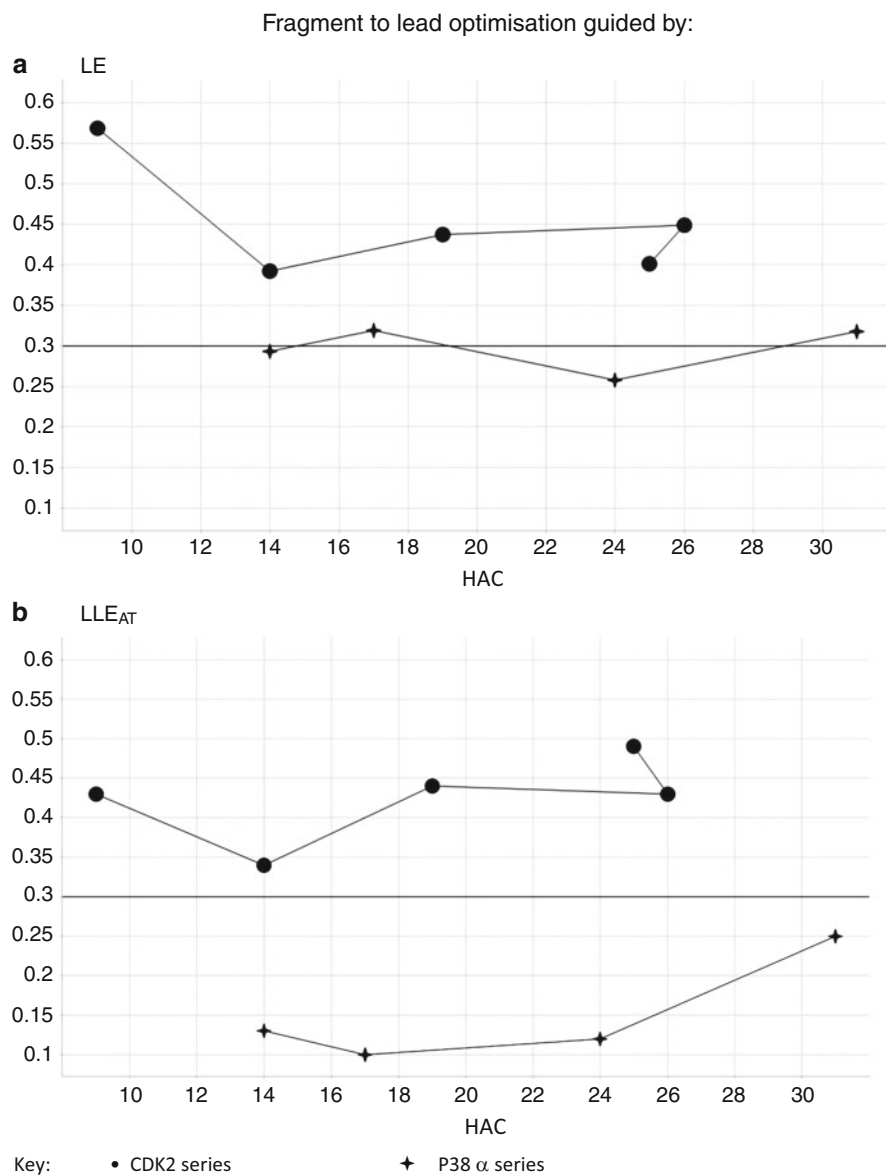
Fig. 31 The plots of fragment to lead tracking of (**a**) LE and (**b**) LLE$_{AT}$ versus heavy atom count (HAC) for the Astex kinase examples discussed in the text and Fig. 30
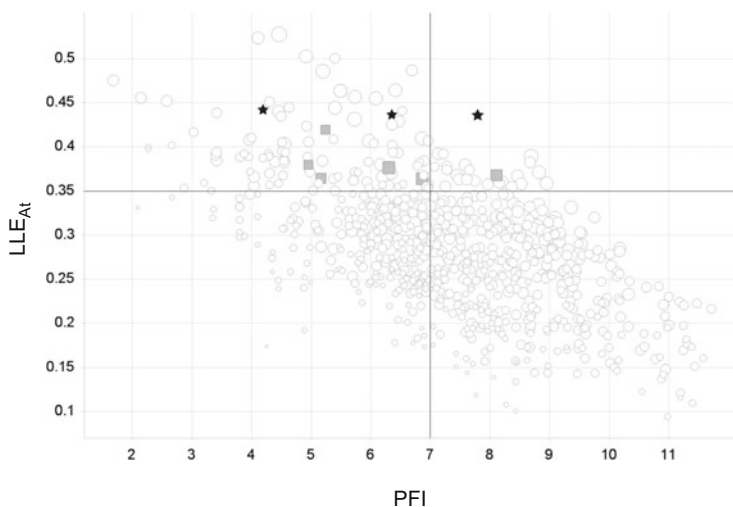
Fig. 32 Plot of LLE$_{AT}$ versus PFI for a representative set of 773 Factor Xa inhibitors, adapted from the review by Young [87]. The three drugs are *black stars*, other clinical candidates exemplified in the review are *grey squares* and other compounds in the set are *white circles* (all sized by pK_i)
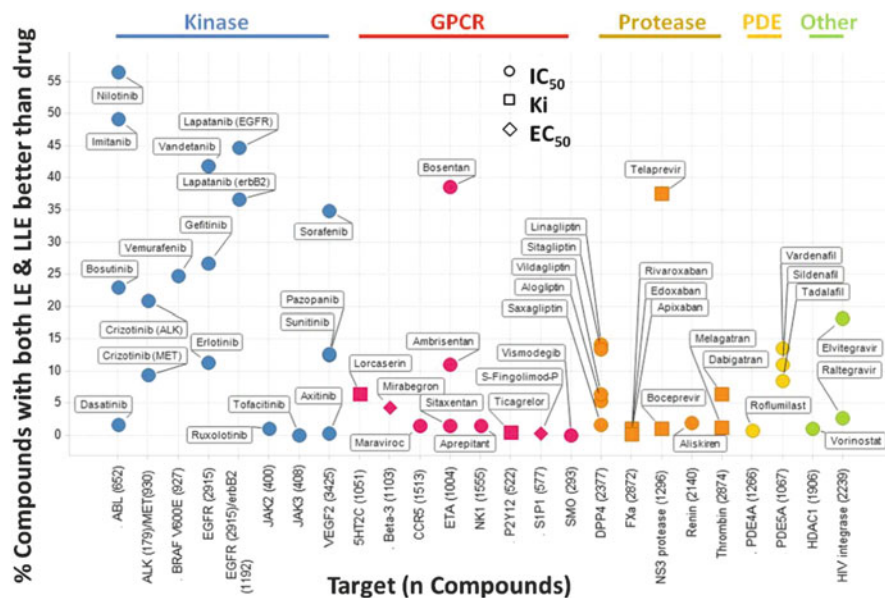
Fig. 33 Leeson's analysis of the relative ligand efficiencies of 46 oral drugs acting at 25 targets based on in vitro affinity data collected from CHEMBL and used with permission from [116]
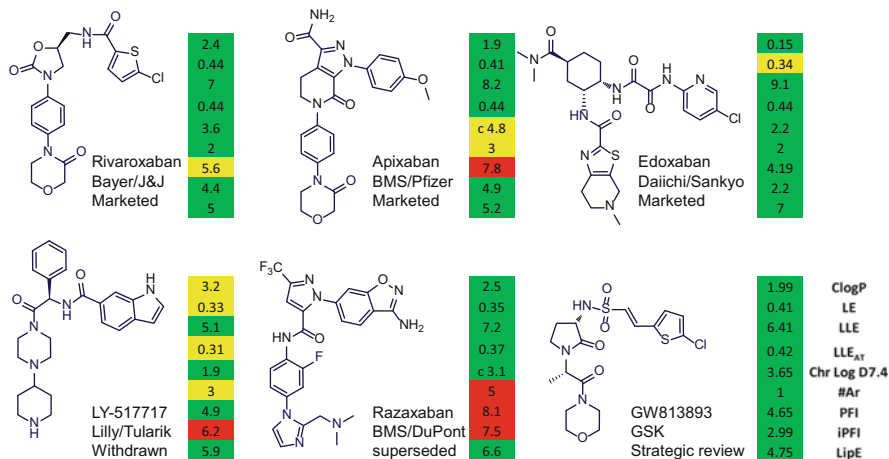
Fig. 34 Structures of representative Factor Xa inhibitors [87] illustrating the combinations of favourable property and efficiency measures and outcomes with the molecules. Colouration reflects efficiency and PFI values based on measure Chrom log $D_{7.4}$ and clog P data. It should be noted that apixaban is a highly optimised and potent molecule, but has a relatively poor translation of its intrinsic affinity into anticoagulant activity (the high PFI is indicative of this)

many anticancer drugs inhibiting kinases show non-optimal LE and LLE values for their target. However, it has been shown that kinase inhibitors produced by fragment-based methods at Astex have markedly better properties than compounds at competitor companies for the same targets; furthermore, a recent study of VEGF kinase inhibitors has indicated that the clinical outcome can be equated with LLE [125].

The pertinent question in optimisation might thus be: which would be the particular value(s) to focus on? The answer may be a combination of them all. Ligand-efficient molecules might be potent due to lipophilicity – although LLE or LipE may compensate for this. log $D_{7.4}$ or PFI can be engineered into favourable space using polar/charged groups, but this could still be a lipophilic base with a hERG risk. Analysis of factor Xa compounds indicated that the drugs are those that possess favourable properties across all of the predictors (Fig. 34). Many of these factors are indicative of risk, so minimising risk across all of them would appear to be a good strategy; although the validity of all in the plethora of metrics, indices and measures summarised above, has been questioned by Schultz [126]. The conclusion of Schultz's analysis, suggesting that LipE is the most important parameter, was really no different to the Hansch mantra [2] of making a molecule "as hydrophilic as feasible without loss of activity".
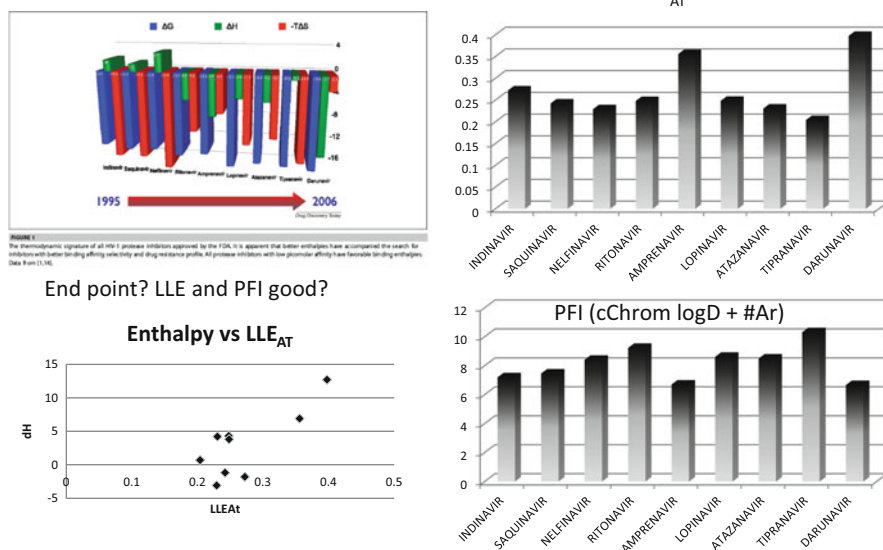
Fig. 35 Friere's thermodynamic signature for marketed HIV protease inhibitors, used with permission from [127] and an analysis of the same compounds using ligand efficiency and PFI metrics

5.5 *The Thermodynamics of Efficient Binding*

One physical aspect of effective drug binding that is now coming to the fore is increased attention to the thermodynamics of the interactions. The free energy of binding, ΔG, is composed of the enthalpic (ΔH) and entropic (ΔS) components, where $\Delta G = \Delta H - T\Delta S$ and the components can be measured by isothermal titration calorimetry (ITC). Data and analyses from these measurements are emerging; whilst it may be too early to draw too many firm conclusions [76], there is growing evidence that efficient binders have a strong entropic component to their binding. Early studies by Freire demonstrated that the enthalpic signature of HIV protease inhibitors and stains improved with time (with a minor blip to address resistance issue with the former series) (Fig. 35) [127]. But others have shown in wider sets there is no general correlation with ΔH and any particular metric and it has been observed that as potency increases in series the entopic component appears to increase [76]. An unpublished analysis of Astex data by Glyn Williams (see [76] and personal communication) has shown that, as might be expected, their fragment-derived series have strong enthalpic signatures, but that there is also an increase in entropic contribution for the more potent molecules. Looking at the impact of efficiency metrics and PFI on the HIV-1 protease set of Freire showed that those with the highest enthalpic signature also had the most favourable LLE_{AT} and PFI values, but a correlation between them is hard to substantiate (Fig. 35). Schultz has

Table 19 Summary of interactions and their thermodynamic consequences

Enthalpic interactions – driven by specific interactions; match size shape and electrostatics
+ Hydrogen bonds
+ Salt bridges
+ "Good fit" van der Waals contacts
− Large enthalpic cost from desolvating polar functionality
Entropy-driven binding – liberation of molecules versus confinement/constraint
+ Loss of water from binding site/order around solvated molecule
− Change in translational/rotational/conformational freedom
+ If little conformational change free vs bound

suggested a similar trend [128] based on LipE values in this series and in the same manuscript demonstrates the potential impact of using enthalpy-driven metrics to guide lead optimisation.

Moving forward, in the process of making ligand-efficient molecules with enthalpy-driven activity, it would appear sensible to exploit the nature of particular interactions that maximise the chances of achieving the desired outcome; these are summarised in Table 19. If the working hypothesis is that "better" molecules have enthalpy-driven activity, then what might the drivers of this be? Enthalpically driven interactions are characterised by optimal contacts and specific hydrogen bonds; these are, by their nature, likely to drive *specificity*, with fewer off-target effects (due to poor active site fit and likely polar contact mismatches at binding sites other than that targeted). Conversely, entropically driven activity is characterised by suboptimal, non-specific, interactions, driven by lipophilicity and/or aromatic contacts. Such molecules would, as discussed in Sect. 4.2.4, have increased chances of binding to other proteins in a *promiscuous* fashion due to lesser specificity. However, more work and analysis is required in this area.

6 Conclusions

The various analyses investigating the impact of physical properties have demonstrated that drug molecules have better physical properties than the majority of experimental molecules. Some have questioned the statistical methods and the strength of correlations [79], but the link between chance of success and physical properties appears inexorable. The complex requirements for the discovery of an efficacious drug molecule mean that it is necessary to maintain activity during the optimisation of pharmacokinetics, pharmacodynamics and toxicology; these are all multi-factorial processes. It is thus perhaps unlikely that a simple correlation between properties might be established; good properties alone are not a guarantee of success and some effective drugs have what might be described as sub-optimal properties. However, it is clear that the chances of success are much greater with better physical properties (solubility, shape and lower lipophilicity). These
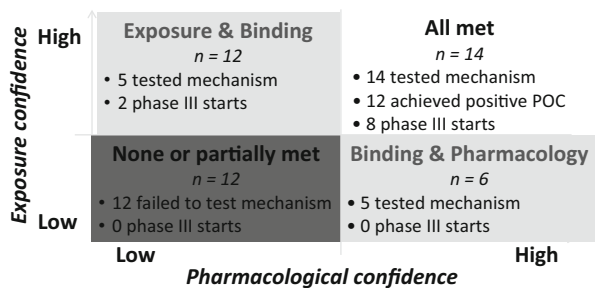
Fig. 36 A representation of the Pfizer three pillar principles and the clinical outcomes presented in the analysis by Morgan et al. [129] based on the confidence levels in exposure and pharmacology for 44 clinical assets. Due to low confidence in the exposure of 18 molecules, the authors concluded that they "could not conclude that the mechanism was tested adequately in 43% of cases"

principles are evident in both the broader analyses with attrition/progression as a marker and also in the particular risk/activity values in various developability screens.

What is required in any molecule is enough aqueous solubility to enable engagement with the target following delivery by whatever method of administration and transport to the compartment/site of action at a suitable concentration. The developability risks discussed in the various sections of this chapter clearly indicate that as lipophilicity and aromaticity increase, the chances of achieving requisite levels of solubility and pharmacokinetics commensurate with delivery to the site of action are diminished. Likewise as these parameters are inflated, the chances of encountering off-target effects clearly increase as well.

The growing consensus in the industry conforms to the hypothesis that development portfolios with lower physicochemical risks have a higher chance of producing efficacious medicines [1], yet there are inefficient drugs in poor physical space, but these have been produced at great cost during a period of high attrition. The evidence currently supports the counterargument that poorer physical properties have produced higher attrition – it now remains to see how the shapelier and leaner molecules of the next generation perform against validated targets. An insightful analysis by Morgan et al. at Pfizer [129] on the progress of their clinical assets identified three pillars on which to base levels of confidence in drug discovery: (1) pharmacokinetic exposure, (2) engagement at the site of action and (3) confidence in the validity of the target and the pharmacological outcome. Confidence in the exposure and target engagement can be achieved with good medicinal chemistry practice; this alone would increase progression even without high confidence in the pharmacological outcome (Fig. 36). Without confidence in the quality of the molecule, it is a costly exercise to conduct a trial knowing that the outcome could be influenced by shortcomings that could have been controlled. Given the low success rate, it would only require, as suggested by Leeson and Springthorpe [5], a 5% improvement in attrition rates to double the output of new medicines.

References

1. Meanwell NA (2011) Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. Chem Res Toxicol 24:1420–1456
2. Hansch C, Bjorkroth J, Leo A (1987) Hydrophobicity and central nervous system agents: on the principle of minimal hydrophobicity. J Pharm Sci 76:663–687
3. Hann MM (1994) Considerations for the use of computational chemistry techniques by medicinal chemists. In: King FD (ed) Medicinal chemistry, principles and practice. RSC, Cambridge, pp 130–142
4. Lipinski CA et al (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26
5. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 6:881–890
6. Hann MM (2011) Molecular obesity, potency and other addictions in drug discovery. Med Chem Comm 2:349–355
7. Hubbard RE, Murray JB (2011) Experiences in fragment-based lead discovery. Methods Enzymol 493:509–531
8. Congreve M et al (2008) Recent developments in fragment-based drug discovery. J Med Chem 51:3661–3680
9. Erlanson DA et al (2004) Fragment-based drug discovery. J Med Chem 47:3463–3482
10. Muresan S, Sadowski J (2008) Properties guiding drug- and lead-likeness. In: Mannhold R (ed) Molecular drug properties – measurement and prediction. Wiley-VCH, Weinheim, pp 439–461
11. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: a useful metric for lead selection. Drug Discov Today 9:430–431
12. Leeson PD, Empfield JR (2010) Reducing the risk of drug attrition associated with physico-chemical properties. Annu Rep Med Chem 45:393–407
13. Van De Waterbeemd H et al (2001) Property-based design: optimization of drug absorption and pharmacokinetics. J Med Chem 44:1313–1333
14. Tarcsay Á, Nyíri K, Keserű GM (2012) Impact of lipophilic efficiency on compound quality. J Med Chem 55:1252–1260
15. Gleeson MP et al (2011) Probing the links between in vitro potency, ADMET and physicochemical parameters. Nat Rev Drug Discov 10:197–208
16. Hill AP, Young RJ (2010) Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity. Drug Discov Today 15:648–655
17. Roden DM, George AL Jr (2002) The genetic basis of variability in drug responses. Nat Rev Drug Discov 1:37–44
18. Faller B et al (2011) Evolution of the physicochemical properties of marketed drugs: can history foretell the future? Drug Discov Today 16:976–984
19. Zhao H (2010) Lead optimization in the nondrug-like space. Drug Discov Today 16:158–163
20. Van de Waterbeemd H, Gifford E (2003) Admet in silico modelling: towards prediction paradise? Nat Rev Drug Discov 2:192–204

21. Bickerton GR et al (2012) Quantifying the chemical beauty of drugs. Nat Chem 4:90–98
22. Kerns EH, Di L (2004) Physicochemical profiling: overview of the screens. Drug Discov Today Technol 1:343–348
23. Kerns EH, Di L (2004) Drug-like properties: concepts, structure design and methods: from ADME to toxicity optimization. Academic, Amsterdam, Boston
24. Wan H, Holmen AG (2009) High throughput screening of physicochemical properties and in vitro ADME profiling in drug discovery. Comb Chem High Throughput Screen 12:315–329
25. Smith RN et al (1975) Selection of a reference partitioning system for drug design work. J Pharm Sci 64:599–606
26. Tute MS (1996) Lipophilicity: a history. In: Mannhold R et al (eds) Methods and principles in medicinal chemistry. Wiley, New York, pp 7–26
27. Fujita T, Iwasha J, Hansch C (1964) A new substituent constant, π, derived from partition coefficients. J Am Chem Soc 86:5175–5180
28. Leo A, Hansch C, Elkins D (1971) Partition coefficients and their uses. Chem Rev 71:525–616
29. Young RJ et al (2011) Getting physical in drug discovery II: the impact of chromatographic hydrophobicity measurements and aromaticity. Drug Discov Today 16:822–830
30. He Y-L et al (1998) Species differences in size discrimination in the paracellular pathway reflected by oral bioavailability of polyethylene glycol and D-peptides. J Pharm Sci 87:626–633
31. Kirch W, Görg KG (1982) Clinical pharmacokinetics of atenolol. Eur J Drug Metab Pharmacokinet 7:81–91
32. Bunnage ME et al (2007) Discovery of potent & selective inhibitors of activated thrombin-activatable fibrinolysis inhibitor for the treatment of thrombosis. J Med Chem 50:6095–6103
33. Hansch C, Leo A (1979) Substituent constants for correlation analysis in chemistry and biology. Wiley, New York
34. Leo AJ (1993) Calculating log P_{Oct} from structures. Chem Rev 93:1281–1306
35. Rekker RE et al (1993) On the reliability of calculated log P-values: Rekker, Hansch-Leo and Suzuki approach. Quant Struct Act Relat 12:152–157
36. Ribeiro MMB et al (2010) Drug–lipid interaction evaluation: why a 19th century solution? Trends Pharmacol Sci 31:449–454
37. Wenlock MC, Barton P, Luker T (2011) Lipophilicity of acidic compounds: impact of ion pair partitioning on drug design. Bioorg Med Chem Lett 21:3550–3556
38. Valkó K (2004) Application of high-performance liquid chromatography based measurements of lipophilicity to model biological distribution. J Chromatogr A 1037:299–310
39. Valkó K et al (1997) Chromatographic hydrophobicity index by fast-gradient RP-HPLC: a high-throughput alternative to log P/log D. Anal Chem 69:2022–2029
40. Albert A, Serjeant EP (1984) The determination of ionization constants, 3rd edn. Chapman and Hall, New York
41. Avdeef A, Bucher JJ (1978) Accurate measurements of the concentration of hydrogen ions with a glass electrode. Anal Chem 50:2137–2142
42. Avdeef A et al (2000) pH-metric solubility: correlation between the acid–base titration and the saturation shake-flask solubility-pH methods. Pharm Res 17:85–89
43. Allen RI et al (1998) Multiwavelength spectrophotometric determination of acid dissociation constants of ionisable drugs. J Pharm Biomed Anal 17:699–712
44. Gift AD et al (2012) Experimental determination of pK_a values by use of NMR chemical shifts, revisited. J Chem Educ 89:1458–1460
45. Cleveland JA et al (1993) Automated pK_a determination at low solute concentrations by capillary electrophoresis. J Chromatogr A 652:301–308
46. Box K et al (2003) High throughput measurement of pK_a values in a mixed-buffer linear pH gradient system. Anal Chem 75:883–892

47. Perrin DD, Dempsey B, Serjeant EP (1981) pK_a prediction for organic acids and bases. Chapman and Hall, London
48. Di L, Fish PV, Mano T (2012) Bridging solubility between drug discovery and development. Drug Discov Today 17:486–495
49. Sugano K et al (2007) Solubility and dissolution profile assessment in drug discovery. Drug Metab Pharmacokinet 22:225–254
50. Huang LF, Tong WQ (2004) Impact of solid state properties on developability assessment of drug candidates. Adv Drug Deliv Rev 56:321–334
51. Bhattachar SN et al (2006) Evaluation of the chemiluminescent nitrogen detector for solubility determinations to support drug discovery. J Pharm Biomed Anal 41:152–157
52. Jain N, Yalkowsky SH (2001) Estimation of the aqueous solubility I: application to organic non-electrolytes. J Pharm Sci 90:234–252
53. European pharmacopeia. http://pharmeuropa.edqm.eu/home/
54. Bergström CA et al (2007) Poorly soluble marketed drugs display solvation limited solubility. J Med Chem 50:5858–5862
55. Ishikawa, Hashimoto (2011) Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. J Med Chem 54:1539–1554
56. Chu KA, Yalkowsky SH (2009) An interesting relationship between drug absorption and melting point. Int J Pharm 373:24–40
57. Llinàs A, Glen RC, Goodman JM (2008) Can you predict solubilities of thirty-two molecules using a database of one hundred reliable measurements? J Chem Inf Model 48:1289–1303
58. Elder D, Holm R (2013) Aqueous solubility: simple predictive methods (in silico, in vitro and bio-relevant approaches). Int J Pharm 453:3–11
59. Jantratid E et al (2008) Dissolution media simulating conditions in the proximal human gastrointestinal tract: an update. Pharm Res 25:1663–1676
60. Holm R et al (2013) Bile salts and their importance for drug absorption. Int J Pharm 453:44–45
61. Bevernage J et al (2013) Evaluation of gastrointestinal drug supersaturation and precipitation: strategies and issues. Int J Pharm 453:25–35
62. Ritchie TJ, Macdonald SJF (2009) The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? Drug Discov Today 14:1011–1020
63. Lovering F, Bikker J, Humblet C (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52:6752–6756
64. Yang Y et al (2012) Beyond size, ionization state, and lipophilicity: influence of molecular topology on absorption, distribution, metabolism, excretion, and toxicity for drug-like compounds. J Med Chem 55:3667–3677
65. Lovering F (2013) Escape from Flatland 2: complexity and promiscuity. Med Chem Comm 4:515
66. Ritchie TJ et al (2011) The impact of aromatic ring count on compound developability – further insights by examining carbo- and hetero- aromatic and aliphatic ring types. Drug Discov Today 16:164–171
67. Kimura T, Higaki K (2002) Gastrointestinal transit and drug absorption. Biol Pharm Bull 25:149–164
68. Davies B, Morris T (1993) Physiological parameters in laboratory animals and humans. Pharm Res 10:1093–1095
69. Riley RJ et al (2002) The influence of DMPK as an integrated partner in modern drug discovery. Curr Drug Metab 3:527–550
70. Schiller C et al (2005) Intestinal fluid volumes and transit of dosage forms as assessed by magnetic resonance imaging. Aliment Pharmacol Ther 22:971–979
71. Uetrecht JP, Trager W (2007) Conjugation pathways. Drug metabolism, chemical and enzymatic aspects. Informa Healthcare, New York, pp 130–144

72. Wager TT et al (2010) Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of drug-like properties. ACS Chem Neurosci 1:435–449

73. Butler JM, Dressman JB (2010) The developability classification system: application of biopharmaceutics concepts to formulation development. J Pharm Sci 99:4940–4954

74. Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–715

75. Waring MJ (2010) Lipophilicity in drug discovery. Expert Opin Drug Discov 5:235–248

76. Hann MM, Keserű GM (2012) Finding the sweet spot – the role of nature and nurture in medicinal chemistry. Nat Rev Drug Discov 11:355–365

77. Waring MJ (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates – molecular weight dependent lower logD limits based on permeability. Bioorg Med Chem Lett 19:2844–2851

78. Gleeson MP (2008) Generation of a set of simple, interpretable ADMET rules of thumb. J Med Chem 51:817–834

79. Kenny PW, Montanari CA (2013) Inflation of correlation in the pursuit of drug-likeness. J Comput Aided Mol Des 27:1–13

80. Sugano K et al (2010) Coexistence of passive and carrier-mediated processes in drug transport. Nat Rev Drug Discov 9:597–614

81. Gleeson MP, Hersey A, Hannongbua S (2011) In-silico ADME models: a general assessment of their utility in drug discovery applications. Curr Top Med Chem 11:358–381

82. Kell DB, Dobson PD, Oliver SG (2011) Pharmaceutical drug transport: the issues and the implications that it is essentially carrier-mediated only. Drug Discov Today 16:704–714

83. Di L et al (2012) Evidence-based approach to assess passive diffusion and carrier-mediated drug transport. Drug Discov Today 17:905–912

84. Kubinyi H (1978) Drug partitioning: relationships between forward and reverse rate constants and partition coefficient. J Pharm Sci 67:262–263

85. Kubinyi H (1979) Lipophilicity and drug activity. Prog Drug Res 23:97–198

86. Glen RC et al (1995) Computer-aided design and synthesis of 5-substituted tryptamines and their pharmacology at the 5-HT1D receptor: discovery of compounds with potential anti-migraine properties. J Med Chem 38:3566–3580

87. Young RJ (2011) The successful quest for oral factor Xa inhibitors; learnings for all of medicinal chemistry? Bioorg Med Chem Lett 21:6228–6235

88. Johnson TW et al (2009) Using the Golden Triangle to optimize clearance and oral absorption. Bioorg Med Chem Lett 19:5560–5564

89. Obach RL et al (2005) In vitro cytochrome P450 inhibition data and the prediction of drug–drug interactions: qualitative relationships, quantitative predictions, and the rank-order approach. Clin Pharmacol Ther 78:582–592

90. Lewis DFV et al (2007) Quantitative structure-activity relationships (QSARs) in inhibitors of various cytochromes P450: the importance of compound lipophilicity. J Enzyme Inhib Med Chem 22:1–6

91. Lewis DFV et al (2004) Compound lipophilicity for substrate binding to human P450s in drug metabolism. Drug Discov Today 9:530–537

92. Lewis DFV, Dickins M (2002) Substrate SAR in human p450s. Drug Discov Today 7:918–925

93. Jamieson C et al (2006) Medicinal chemistry of hERG optimizations: highlights and hang-ups. J Med Chem 49:5029–5046

94. Waring MJ, Johnstone C (2007) A quantitative assessment of hERG liability as a function of lipophilicity. Bioorg Med Chem Lett 17:1759–1764

95. Wood A, Armour D (2005) The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. Prog Med Chem 43:239–271

96. Diller DJ (2009) In silico hERG modelling: challenges and progress. Curr Comput Aided Drug Des 5:106–121

97. Tarcsay Á, Keserű GM (2013) Contributions of molecular properties to drug promiscuity. J Med Chem 56:1789–1795
98. Hopkins AL, Mason JS, Overington JP (2006) Can we rationally design promiscuous drugs? Curr Opin Struct Biol 16:127–136
99. Hughes JD et al (2008) Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett 18:4872–4875
100. Morphy R, Rankovic Z (2007) Fragments, network biology and designing multiple ligands. Drug Discov Today 12:156–160
101. Azzaoui K et al (2007) Modeling promiscuity based on in vitro safety pharmacology profiling data. ChemMedChem 2007(2):874–880
102. Peters J-U et al (2009) Pharmacological promiscuity: dependence on compound properties and target specificity in a set of recent Roche compounds. ChemMedChem 4:680–686
103. Yang Y et al (2010) Investigation of the relationship between topology and selectivity for drug-like molecules. J Med Chem 53:7709–7714
104. Peters J-U et al (2012) Can we discover pharmacological promiscuity early in the drug discovery process? Drug Discov Today 17:325–335
105. Lounkine E et al (2012) Large-scale prediction and testing of drug activity on side-effect targets. Nature 486:361–367
106. Leeson PD et al (2011) Impact of ion class and time on oral drug molecular properties. Med Chem Comm 2:91–105
107. Leach AR, Hann MM (2011) Molecular complexity and fragment-based drug discovery: ten years on. Curr Opin Chem Biol 15:489–496
108. Trainor GL (2007) The importance of plasma protein binding in drug discovery. Expert Opin Drug Discov 2:51–64
109. Valkó K et al (2003) Fast gradient HPLC method to determine compounds binding to human serum albumin: relationships with octanol/water and immobilized artificial membrane lipophilicity. J Pharm Sci 92:2236–2248
110. Smith DA et al (2010) The effect of plasma protein binding on in vivo efficacy: misconceptions in drug discovery. Nat Rev Drug Discov 9:929–939
111. Riley RJ et al (2005) A unified model for predicting human hepatic metabolic clearance from in vitro intrinsic clearance data in hepatocytes and microsomes. Drug Metab Dispos 33:1304–1311
112. Braggio et al (2010) Drug efficiency: a new concept to guide lead optimization programs towards the selection of better clinical candidates. Expert Opin Drug Discov 5:609–618
113. Montanari D et al (2011) Application of drug efficiency index in drug discovery: a strategy towards low therapeutic dose. Expert Opin Drug Discov 6:913–920
114. Valkó K et al (2012) In vitro measurement of drug efficiency index to aid early lead optimization. J Pharm Sci 101:4155–4169
115. Valkó K, Nunhuck SB, Hill AP (2011) Estimating unbound volume of distribution and tissue binding by in vitro HPLC-based human serum albumin and immobilized artificial membrane-binding measurements. J Pharm Sci 100:849–862
116. Hopkins AL et al (2014) The role of ligand efficiency measures in drug discovery. Nat Rev Drug Discov 13:105–121
117. Reynolds CH et al (2008) Ligand binding efficiency: trends, physical basis, and implications. J Med Chem 51:2432–2438
118. Nissink JWM (2009) Simple size-independent measure of ligand efficiency. J Chem Inf Model 49:1617–1622
119. Hann MM, Leach AR, 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
120. Mortenson PN, Murray CW (2011) Assessing the lipophilicity of fragments and early hits. J Comput Aided Mol Des 25:663–667
121. Wyatt PG et al (2008) Identification of N-(4-Piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using

fragment-based X-ray crystallography and structure based drug design. J Med Chem 51:4986–4999

122. Gill AL et al (2005) Identification of novel p38alpha MAP kinase inhibitors using fragment-based lead generation. J Med Chem 48:414–426
123. Freeman-Cook KD, Hoffman RL, Johnson TW (2013) Lipophilic efficiency: the most important efficiency metric in medicinal chemistry. Future Med Chem 5:113–115
124. Gill AL et al (2007) A comparison of physicochemical property profiles of marketed oral drugs and orally bioavailable anti-cancer protein kinase inhibitors in clinical development. Curr Top Med Chem 7:1408–1422
125. McTigue M et al (2012) Molecular conformations, interactions, and properties associated with drug efficiency and clinical performance among VEGFR TK inhibitors. Proc Natl Acad Sci U S A 109:18281–18289
126. Shultz MD (2013) Setting expectations in molecular optimizations: strengths and limitations of commonly used composite parameters. Bioorg Med Chem Lett 23:5980–5991
127. Freire E (2008) Do enthalpy and entropy distinguish first in class from best in class? Drug Discov Today 13:869–874
128. Shultz MD (2013) The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations. Bioorg Med Chem Lett 23:5992–6000
129. Morgan P et al (2012) Can the flow of medicines be improved? fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. Drug Discov Today 17:419–424

Improving Solubility via Structural Modification

Michael A. Walker

Abstract The examples and discussion presented in this review are intended to serve as resource for medicinal chemists engaged in the task of optimizing drug physical properties. A discussion of the factors governing aqueous solubility is presented followed by specific examples drawn from the recent literature. According to the general solubility equation (GSE), the factors involved in the solubility of a compound are represented by $\log P$ and melting point. Improved solubility can be accomplished by reducing $\log P$ or melting point by increasing polarity or disrupting intermolecular interactions in the solid state. Tactics for increasing polarity include introducing a solubilizing appendage onto the drug or modifying the template or attached substituents. The melting point of a compound can be lowered by disrupting specific intermolecular interactions or changing the topology or shape of the molecule.

Keywords Crystal lattice stability, Enthalpy of solvation, Entropy of solvation, Fluorine, General solubility equation, Hydrogen bonding, Hydrophobicity, Matched molecular pairs, Melting point, Molecular planarity, Molecular symmetry, Packing efficiency, Solubilizing appendage, Solvation, X-Ray crystallography

Contents

M.A. Walker (✉)
Department of Medicinal Chemistry, Bristol Myers-Squibb, 5 Research Pkwy, Wallingford, CT 06492
e-mail: michael.a.walker@bms.com

Abbreviations

C_k	The Kitaigorodski packing coefficient
CSD	Cambridge Structural Database
DGAT1	Diacylglycerol *O*-acyltransferase 1
GSE	General solubility equation
ICAM 1	Intercellular adhesion molecule 1
MMP	Matched molecular pair
mp	Melting point
PDE	Phosphodiesterase
PEG	Polyethylene glycol
S1P1	Sphingosine 1-phosphate receptor subtype 1
sol	Solubility
SPT	Scaled particle theory
T_m	Melting temperature

1   Introduction

The increasing demand for higher potency and selectivity has driven the search for new drugs into chemical space occupied by large hydrophobic molecules. As a consequence, poor aqueous solubility is encountered more often today than in years past. To illustrate this phenomenon, a search of PubMed using the search term "inhibitor" shows a practically linear increase in the number of publications over the last 40 years (Fig. 1). However, over the same period of time, documents containing the term "poor solubility" have grown at an almost exponential rate.

Aqueous solubility is recognized as one of two parameters limiting the oral bioavailability of a drug, the other being the rate of absorption [1]. Not surprisingly, a number of reviews have recently appeared in the literature which highlight the growing issue of poor physical properties, including solubility, and the impact they can have on pharmacokinetics (PK) and pharmacology [2–10].

The factors controlling the aqueous solubility of organic molecules are complex and in some cases still a matter of debate. Because of this, drug solubility issues are usually solved by a combination of empirical and rational drug design strategies. In order to support an empirical approach, the literature serves as a rich source of examples of solubility-enhancing structural transformations that could potentially be applied in a tactical manner. Therefore, this review draws upon the current literature in order to capture a broad array of different methods that have been employed to successfully overcome poor drug solubility.

Fig. 1 Graphical representation of the number of citations per year in PubMed using "inhibitor" or "poor solubility" as the search terms. The number of citations for the search term "poor solubility" are multiplied by 10
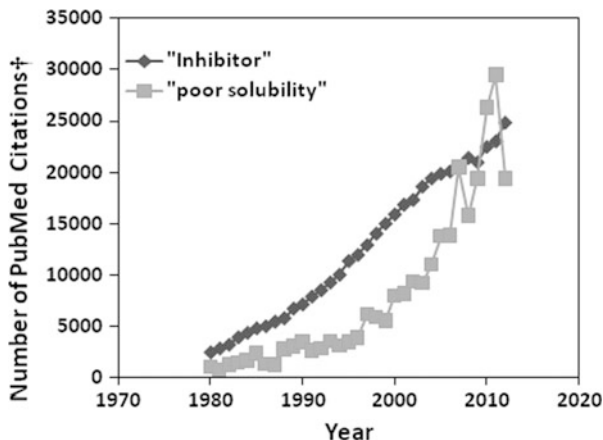
2 Description of Aqueous Solubility

2.1 Definition of Solubility

Solubility is defined as the concentration of a substance in solution which is in equilibrium with undissolved solid. This is sometimes referred to as equilibrium or thermodynamic solubility. Small molecule drugs can exist in several different solid forms, each having a different solubility. Therefore, solubility is dependent on both the solution-state and solid-state properties of a compound.

2.2 General Solubility Equation

The general solubility equation (GSE, Eq. (1)) developed by Jain and Yalkowsky [11, 12] is useful for providing insight into the solution- and solid-state factors governing aqueous solubility. The equation correlates solubility to two easily measured properties, $\log P$ and melting point (mp). For example, the solubility of a compound in water ($\log S_w$), ignoring the solid state, can be approximated by $\log P$. Furthermore, the "ideal solubility" of a crystalline solid in an "ideal solvent" is simply a function of its melting point, assuming that the entropy of melting is 13.5 cal K^{-1} mol^{-1}. Combining these two approximations yields the GSE. Using the GSE one can determine the relative contributions of hydrophobicity ($\log P$) and solid-state stability (mp) towards the solubility of a given compound.

It can also be used to predict the improvement in solubility accompanying a change in log$P$ or mp. All other things being equal, a change in mp of $100°C$ or log$P$ of 1 will result in a 10-fold increase/decrease in solubility.

$$\log S_w = 0.5 - \log P - 0.01(\text{MP} - 25) \tag{1}$$

2.3 Solvation of Organic Molecules

Solvation is defined as the process of transferring a molecule from the gas phase into solution. The underlying reason for the limited solubility of organic compounds in water can be understood by examining the corresponding change in free energy (ΔG_{solv}) [13]. The enthalpy of solvation (ΔH_{solv}) corresponds to the net change in water–water and water–solute intermolecular interaction energies. For organic compounds, ΔH_{solv} is usually negative, meaning that it is an enthalpically favored process. On the other hand, the entropy of solvation (ΔS_{solv}), which represents the change in the level of order in the system, is negative meaning that solvation is an unfavorable process. Therefore, organic molecules display limited solubility in water due to a loss in entropy.

There are currently two theories which explain the negative entropy of solvation. According to the first hypothesis, an ordered shell of water molecules held together by hydrogen bonding encapsulates the solute [14]. The increased order of the water molecules which make up the solvation shell results in a decrease in entropy. A more recent model attributes the loss in entropy to the formation of the cavity in the solvent large enough to accommodate the solute [15, 16]. The entropy cost on cavity formation can be derived using scaled particle theory (SPT), which treats the solvent and solute as hard spheres without the need to invoke the formation of a H-bonded shell of water molecules [17]. Both models predict that the loss in entropy is related to the size of the molecule. All things being equal, solubility decreases as the size of the molecule increases due to a corresponding loss in entropy.

2.3.1 Effect of Molecular Size and Shape on Solvation

The effect of size can be appreciated by examining the influence that molecular surface area has on solubility [18]. Figure 2 shows that for a series of C_5–C_{15} n-alkane homologues there is a correlation between logS_w and calculated molecular surface area [19]. This correlation provides insight into the conformational behavior of n-alkanes in aqueous solution. Specifically, the change in solubility remains constant over the entire series of n-alkanes, indicating that these molecules all adopt an extended linear configuration [20]. This is interesting since one might have expected the longer n-alkanes to fold back on themselves in order to minimize solvent-exposed surface area and the loss in entropy.

Fig. 2 Solubility ($\log S_w$) versus the calculated surface area for the C_5–C_{15} *n*-alkanes congeners
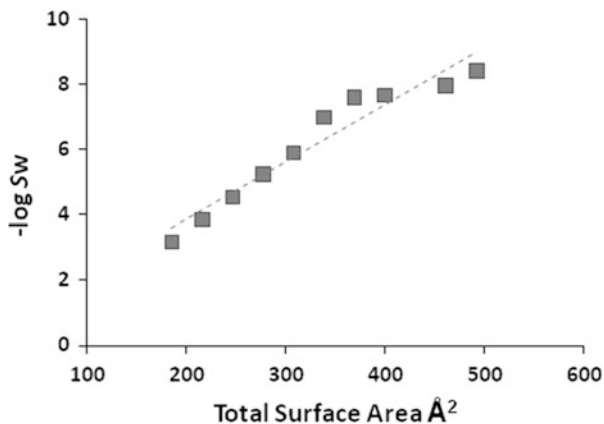
2.3.2 Water-Solute Hydrogen Bonding

Organic compounds containing heteroatoms such as O and N can form stabilizing hydrogen bonds to water. This can lead to an improvement in solubility by lowering the solvation energy [21]. There are several scales available which can be used to rank the relative H-bonding donor and acceptor ability of different heteroatom-containing substituents [22–25]. The extensive amount of data available on this topic make it impossible to cover in the current review. However, some interesting observations from the recent literature are highlighted below.

The H-bond basicities (pK_{BHX}) of the heterocycles shown below reveal an interesting dichotomy between 5- and 6-membered ring compounds [26]. For the 5-membered ring heterocycles **1**, **2**, and **3**, pK_{HBX} increases as the number of ring nitrogen atoms increases. In contrast pK_{HBX} decreases for 6-membered ring compounds (**6**, **7**, **8**, **9**, and **10**). With respect to regiochemistry, the 1,3-arrangement (**2** and **4**) is preferred over the 1,2-arrangement (**3** and **5**) for the 5-membered heterocycles while for the 6-membered series the relative preference is 1,2 (**7**) > 1,3 (**8**) > 1,4 (**9**).

	1	2	3	4	5
pK_{BHX}	0.23	1.86	1.84	1.30	0.81

	6	7	8	9	10
pK_{BHX}	1.86	1.65	1.07	0.92	0.32
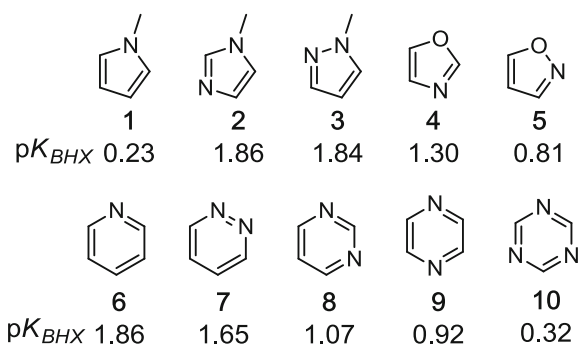
With respect to non-heterocyclic functional groups, it is interesting to note that the amide and urea analogues **12–14** exhibit higher H-bond acceptor strength (pK_{BHX}) than the basic heterocycles despite being neutral in solution. An interesting aspect of the amide and urea functionalities is that inclusion of the group in a ring, as in **12** and **14**, increases the H-bond strength compared to the acyclic analogues **11** and **13**, respectively. This might be due to an increase in conjugation of N to the attached carbonyl by reinforcement of planarity. This would be expected to increase the electron density at the carbonyl oxygen and thereby increase its affinity for an available hydrogen atom.

	11	**12**	**13**	**14**
pK_{BHX}	2.44	2.60	2.44	2.79
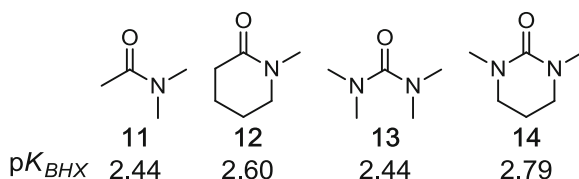
2.4 Crystal Lattice Stability and Packing Efficiency

The contribution of crystal lattice stability to solubility is an important factor to consider since drugs are commonly delivered via a solid dosage form. The stability of the crystal lattice is generally determined from its melting point. By definition, at the melting temperature (T_m), solid and liquid are in equilibrium and the change in free energy is zero ($\Delta G_m = 0$). Rearrangement of the Gibbs free energy equation reveals that at the melting point, T_m is equivalent to the enthalpy of melting (ΔH_m) divided by the entropy of melting (ΔS_m) (Eq. (2)). ΔH_m can be thought of as the energy required to overcome the intermolecular forces holding the molecules in place while ΔS_m represents the increase in translational, rotational, and conformational degrees of freedom upon melting. The intermolecular forces restraining molecules in the crystal lattice are easy to appreciate since they consist of standard intermolecular interactions. In contrast, with the exception of molecules with low or no conformational mobility, ΔS_m is more difficult to assess even though it can have a significant impact on solubility [27–29].

Crystalline stability
T_m (melting temperature)

$$^1T_m = \frac{\Delta H_m}{\Delta S_m} \quad (2)$$

Packing efficiency
C_k (packing coefficient)

$$^2C_k = \frac{Z \times V_{mol}}{V_{cell}} \quad (3)$$

---

[1] ΔH_m, melting (fusion) enthalpy; ΔS_m, melting (fusion) entropy.
[2] Z, number of molecules per unit cell; V_{mol}, molecular volume ($Å^3$); V_{cell}, unit cell volume ($Å^3$).

Another consideration is the efficiency with which the molecules pack. The Kitaigorodski packing coefficient (C_k, Eq. (3)) which represents the fraction of the cell volume occupied by the compound is useful in this respect. Regardless of structure, organic molecules tend to pack closely to one another with efficiencies in the range of $C_k = 0.65$–0.80. This is comparable to the packing efficiency calculated for the closest packing arrangement of identical spheres ($C_k = 0.74$).

2.4.1   Effect of Molecular Shape on Packing Efficiency

Shape can play a role in packing efficiency and lattice stability. It is well known that "flat," non-flexible molecules pack efficiently. However, compounds with "rodlike" or "globular" shapes also pack well [30]. Furthermore, according to the "close-packing principle," even compounds with complex structures can pack efficiently, especially if the overall shape of the molecule is self-complementary. As an example of self-complementary packing, the X-ray crystal structure of compound **15** (Fig. 3) [31] shows that even though it has an awkward shape, the diketopiperazine packs well due to interdigitation of the bridging alkenyl rings. Since molecules will always seek the most efficient packing arrangement, it is difficult to predict what effect modifications to the shape of a molecule will have on crystal lattice stability.

2.4.2   Intermolecular Interactions in the Solid State

Organic crystalline solids are held together through a combination of weak intermolecular interactions which include hydrogen bonding, halogen bonding, electrostatic attraction, and van der Waals forces [32]. Examples of these are demonstrated by **16–20**. H bonding is usually among the strongest interactions occurring in the solid state and, as such, can have a large influence on lattice architecture. For example, amides and carboxylic acids tend to self-associate via a select set of intermolecular configurations, such as that illustrated by **16** and **17** [33, 34], in order to maximize bonding and reduce repulsion [35]. As a consequence, this can dictate the overall packing configuration of the crystal. In the solid state, hydrogen bonding is not limited to conventional functional groups such as O-H and N-H but, as illustrated by **18**, can also include C-H, which can act as a H-bond donor [36]. Halogen atoms are able to form diverse intermolecular interactions in the solid state [37]. They can act as H-bond acceptors (cf. **19**) [38] or form intermolecular halogen-halogen bonds (cf. **20**) [39]. Numerous other intermolecular interactions also exist but are not covered in the current review [40].

Fig. 3 Structure
(**a**) and crystal
packing configuration
(**b**) of compound **15**
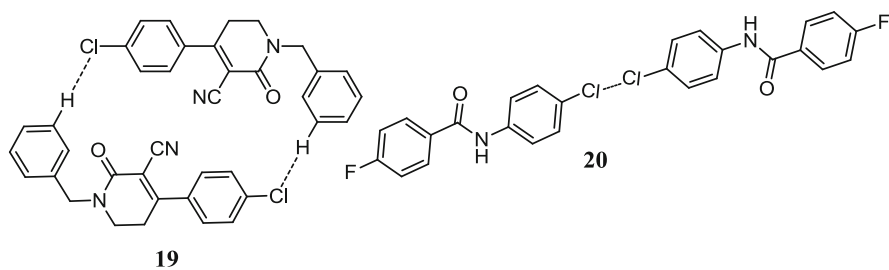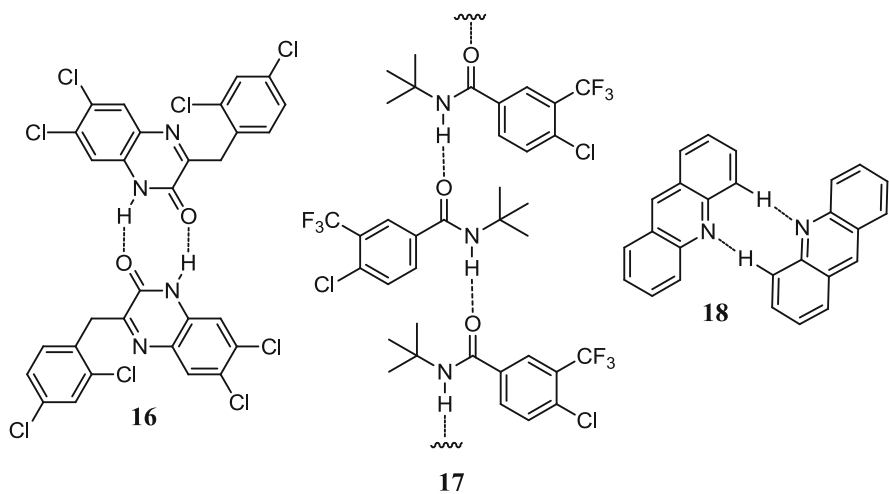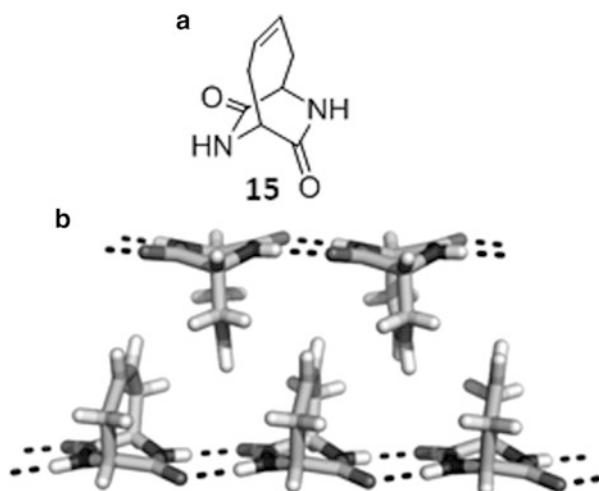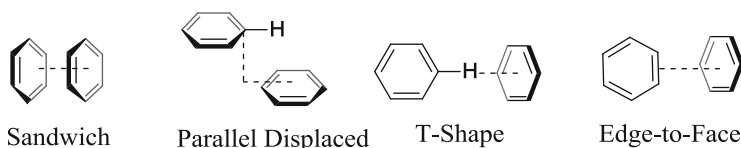
Fig. 4 Configurations of aryl–aryl interactions

Aryl–aryl interactions deserve special attention since drug-like molecules tend to be built from aromatic and heteroaromatic templates. Aromatic–aromatic interactions can be classified according to the relative geometry of the two rings, as illustrated in Fig. 4. The π-stacking interaction, also referred to as the "sandwich" arrangement, involves the parallel alignment of the rings lying directly on top of one another [41, 42]. When the centers of the rings are not aligned, the assembly is referred to as "parallel displaced." Likewise, there are two configurations where the planes of the two rings are not parallel to one another, the "T-shaped" and the "edge-to-face" motifs [43, 44]. These interactions can play a significant role in the packing and stability of the crystal lattice of small molecule drugs.

2.4.3 The Effect of Molecular Symmetry on Crystal Packing

Compounds with high molecular symmetry tend to form more stable, higher melting, solids than compounds with lower symmetry [28, 45–47]. As a typical example, *para*-disubstituted benzenes usually exhibit higher melting points than the corresponding *ortho*- or *meta*-substituted isomers regardless of the identity of the substituents [48, 49]. However, the relationship between melting point and symmetry can be more subtle, as illustrated by the homologous dibasic carboxylic acids in Fig. 5 [50]. For this series of compounds, the melting point decreases as the number of carbon atoms increases. Interestingly, the melting points alternate depending on whether the number of carbon atoms in the molecule is even or odd. The even-numbered analogues tend to have higher melting points than the odd-numbered compounds. The reason for this is only evident from the X-ray crystal structures of these compounds, as exemplified by adipic (**21**) and glutaric acid (**22**). These compounds display different conformations with respect to the carboxylic acid termini. For adipic acid, which has an even number of carbon atoms, the two carboxylic acids are coplanar making it more symmetrical than glutaric acid where the terminal acids are skewed [51]. This suggests that large changes in crystal lattice stability can be achieved through seemingly minor structural modifications such as changing the number of carbon atoms from even to odd.

Fig. 5 Relationship
between chain length,
solubility, and melting
point for dibasic acids,
$n = 0–7$. Single-crystal
X-ray structures depicting
the conformations of adipic
acid (**21**) and glutaric acid
(**22**) in the solid state
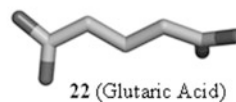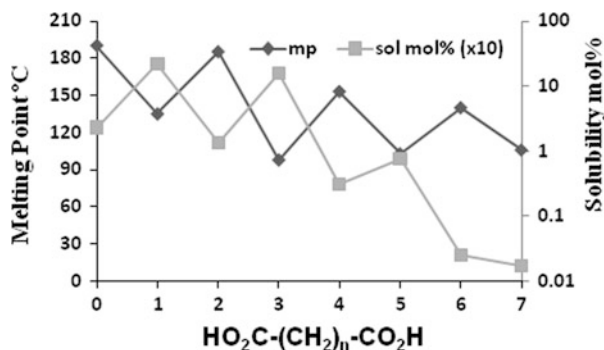
3 Tactics for Improving Solubility

As suggested by the GSE, one can potentially increase the aqueous solubility
of a molecule by reducing either the log$P$ or the melting point. This can be
accomplished by adding or removing substituents in order to increase polarity or
to destabilize crystalline stability. Based on the desired outcome, lowering log$P$
or melting point, there are a number of tactics at one's disposal. In the sections
below, these have been divided into two basic categories according to whether
the goal is to reduce hydrophobicity or lower solid-state stability. These two
categories have been further divided based on the means by which log$P$ or melting
point is lowered. These classifications are arbitrary but are intended to convey
some sense of how the structural modification might impact other properties of
the molecule such as potency and pharmacokinetics.

3.1 Reducing LogP

3.1.1 Introducing a Solubilizing Appendage

A solubilizing appendage can be a basic, acidic, or neutral group tethered to the drug
via a linker. Its primary function is to enhance solubility by adding additional polarity
to the molecule in a fashion that does not directly interact with the target. Protonatable
amines are generally employed as basic solubilizing groups while carboxylic acids
serve as acidic appendages. Neutral groups are polar substituents which can
hydrogen-bond to water. Since the goal is to increase solubility without interfering
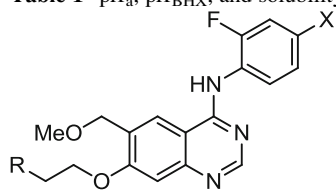
with the activity of the drug, the site of attachment is chosen such that the appendage is directed away from the target-binding domain when the drug is in residence. A structure of the drug bound to the target is useful for determining the appropriate placement of the appendage as well as the length of the linker. This is a very effective tactic for improving aqueous solubility and even large hydrophobic molecules such as C_{60} have been solubilized by introducing an appropriate appendage [52].

Introducing Basic Appendages

There are numerous examples in the literature where a basic appendage has been introduced to improve aqueous solubility. The solubility-enhancing effect of an amine group is derived from its ability to make a strong hydrogen bond to water and/or to form a charged species at a pH $\leq$ pKa of the conjugate acid. In general, the pK_a of the amino group will determine the expected increase in solubility. However, as illustrated in the examples below, H-bond basicity might be a better parameter to use. Both H-bonding basicity and pK_a can be modulated by electronic and steric effects. Therefore, one advantage of this class of appendages is that there are a large number of diverse amines to choose from that encompass a broad range of H-bonding ability and pK_a. This allows for precise tuning of the molecule in order to achieve the correct balance between solubility and pharmacological properties. This capability is often necessary since compounds containing amine groups have an increased potential to cause phospholipidosis or QT prolongation [6, 53].

The analogues shown in Table 1 demonstrate the effect that modulating basicity (pK_a) and H-bond basicity (pK_{BHX}) can have on solubility [54]. For this series of compounds, solubility can be correlated with either the pK_{BHX} or pK_a of the attached amine. However, when the solubilities of **23** and **26** were measured at pH 12.7, where protonation does not participate in solubilizing the compounds, the relative solubilities of the two molecules were retained. Compound **26** was found to be 17-fold more soluble than the triazole **23** at this pH. This suggests that, at least for the non-protonated species, H bonding to water makes a significant contribution to solubility.

The importance of H bonding to water appears to be more significant for the compounds in Table 2. Assuming that the relative basicities of the attached amines are **27** ~ **28** > **29** > **30**, there is an inverse correlation between solubility and amine pK_a for this chemotype [55]. In the case of **29** and **30**, the higher solubilities relative to the more basic amines might be due to additional H bonding to water from the adjacent OEt and OH groups. The origin of the moderately elevated solubility of **28** compared to **27** is not obvious since one might expect the *tert*-butyl group to sterically interfere with H bonding to water. In this case, the compact size of the *tert*-butyl group versus azepine might be responsible for reducing its solvation energy.

Table 1 pK_a, pK_{BHX}, and solubility of compounds **23–26**
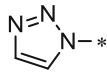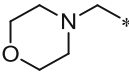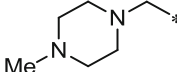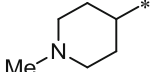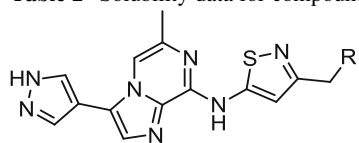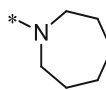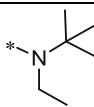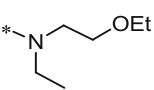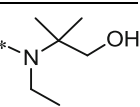
R	$\underset{N}{\overset{N=N}{\bigvee}}N-*$	$\overset{O}{\underset{}{\bigcirc}}N\diagup*$	$\underset{Me}{\overset{}{N}}\diagdown N\diagup*$	$\underset{Me}{\overset{}{N}}\diagdown\overset{*}{\bigcirc}$
X	Br	Cl	Cl	Br
Compound	**23**	**24**	**25**	**26**
pK_a	5.3	7.2–5.2	7.9–5.4	9.4–5.3
pK_{BHX}	–	1.78	1.88	2.11
Solubility (µM)	0.7	20	60	330

Table 2 Solubility data for compounds **27–30**

R=	$*-N\bigcirc$	$*-\overset{}{N}\diagup\diagdown$	$*-\overset{}{N}\diagdown\diagup OEt$	$*-\overset{}{N}\diagdown\diagup OH$
	27	**28**	**29**	**30**
Solubility at pH 6.5 (µg/mL)	10	130	5,100	5,700

Introducing Acidic Appendages

Carboxylic acid appendages having a pK_a of 4–6 can potentially yield improved solubility through deprotonation at physiological pH (≥ 7.0). However, a disadvantage of this is that formation of the negatively charged species in the intestine might also hamper absorption. Another potential issue is that carboxylic acids are subject to glucuronidation which can negatively affect the toxicity profile of a drug. This might explain the smaller number of examples appearing in the literature of acidic appendages being utilized as solubility-enhancing groups compared to basic appendages.

An example of a solubility-enhancing acidic appendage comes from the phosphodiesterase (PDE) 4 inhibitor area (Table 3) where the lead compound **31** suffered from poor solubility (0.3 µg/mL) due to high hydrophobicity (clog$P = 5.20$) [56]. After an attachment point for a solubilizing appendage was identified using X-ray structural data, a series of linear carboxylic acids ($-(CH_2)_nCO_2H$) was synthesized. A dramatic improvement in solubility was observed for all of the analogues. Solubility increased from 500 to 1,400 µg/mL as the linker chain was lengthened

Table 3 Activity and solubility of PDE 4 inhibitors **31–36**
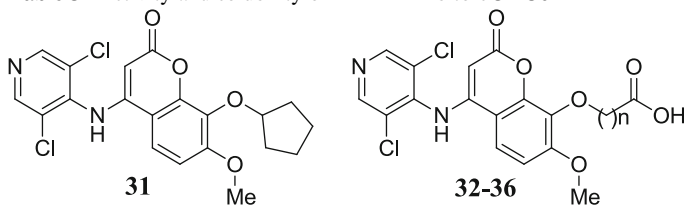
Compound	n	IC$_{50}$ (nM)	Solubility at pH $= 7.4$ (µg/mL)
31	–	5	0.3
32	4	43	500
33	5	17	900
34	6	2	1,400
35	7	2	300
36	8	1	300
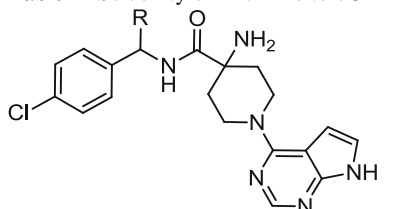
from $n = 4$ to 6 after which a modest drop off to 300 µg/mL was observed. The slight attenuation in solubility for the longer-chain derivatives is probably due to increased hydrophobicity which offsets the solubilizing effect of the CO$_2$H group. Interestingly, the in vitro potency was also sensitive to the length of the linker. A minimal length of $n = 6$ was required in order to achieve the same activity as the parent, suggesting that the carboxylate needed to be positioned far enough away from the enzyme in order to avoid interfering with drug binding.

Introduction of Neutral Appendages

Neutral appendages provide improved solubility by means of hydrogen bonding to water. This can be achieved by using a group with high H-bonding basicity such as a heterocycle, amide, or urea. However, in certain cases, strong H-bonding functionality can lead to poor membrane permeability. As an alternative, weaker H-bond acceptors, such as ethers, have been examined. However, weak H-bonding groups have been utilized less often since, in order to achieve sufficient solubility, multiple acceptors need to be present which can lead to excessively large structures.

An article describing the discovery of novel Akt kinase inhibitors provides an illustrative example of improved solubility achieved by means of a strong H-bond acceptor (Table 4) [57]. Amide groups R$=$CONH$_2$ (**38**) and CH$_2$CON (Me)$_2$ (**39**) attached to the α-position of the benzylamide portion of the template provided a 2–3-fold improvement in solubility over the unsubstituted parent compound **37**, while the reverse amide **40** resulted in a ≥10-fold increase. Interestingly, **40** exhibited solubility comparable to analogues containing basic amine solubilizing appendages, such as **41**.

The attachment of large polyethylene glycol (PEG) groups (MW ≥ 1,000) is often employed in the development of protein therapeutics but is rarely pursued as a means for improving the solubility of small molecules [58–60].

Table 4 Solubility of Akt inhibitors **37–41**
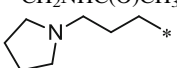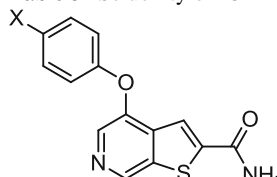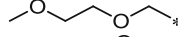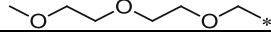
Compound	R	Sol (µM)
37	H	150
38	−CONH₂	460
39	−CH₂CON(Me)₂	400
40	−CH₂NHC(O)CH₃	>1,200
41	(pyrrolidinyl-propyl)*	>1,500

Table 5 Solubility of ICAM 1 inhibitors **42–45**

Compound	X	Solubility (µg/mL)
42	HO~~~*	6.9
43	～O～O～*	6.2
44	～O～～O～*	25.7[a]
45	～O～～O～～O～*	160

[a] Methyl amide

The reason for this is that PEG groups drastically increase the size of the molecule and can interfere with binding. More often shortened PEG groups are examined as solubility-augmenting substituents for small molecule drugs. For example, ethylene glycol-based functionality was able to significantly improve the solubility of the intercellular adhesion molecule 1 (ICAM 1) inhibitor **42** (Table 5) [61]. The increase in solubility is associated with the length of the glycol chain indicating that multiple H-bond acceptor groups are required in order to realize a noticeable effect (compare **44** and **45**).

3.1.2 Template and Substituent Modifications

The solubilizing appendage approach can be viewed as somewhat inefficient since it involves attaching a moderately large structural moiety whose only function is to

increase solubility. A more economical method would be to modify the drug itself. Dissecting a drug into its primary components, the template consists of the core scaffold of the structure and in most cases defines the chemotype from which analogues are produced. The substituents are the groups which are appended to the template and lie on the periphery of the drug, often in direct contact with the target. Changes made to either the template or ligands can affect both solubility and drug affinity. The major challenge in this method is to improve solubility without reducing biological activity.

Substituent Modifications

The primary process of drug discovery involves the synthesis of structurally similar analogues by varying the substituents ligands attached to a template. The ligand can be (1) a single atom such as halogen; (2) a small functional substituent, for example, CO_2H, NH_2, or OH; or (3) a large group such as a heterocycle. The influence that most substituents have on solubility is generally dictated by their intrinsic hydrophobic and/or acidic or basic properties. Since understanding substituent properties is a necessary tool for drug design, a general overview of solubility SAR is presented in the following sections. A thorough analysis of this subject is too broad to be covered in a single review; therefore, the following sections highlight topics of special interest as well as provide a small number of examples from the recent literature.

Analysis of Substituent Effects on Solubility

Matched molecular pair (MMP) analysis is a useful tool to determine the effect a substituent can have on physicochemical and pharmacological properties [62]. Selected results from three recent studies by Leach et al. [63], Zhang et al. [64], and Gleeson et al. [65] which examine the effect of replacing hydrogen with commonly encountered functional groups are summarized in Table 6. The data include the mean change in solubility accompanying the introduction of the substituent. This indicates whether a substituent results in an increase or decrease in solubility and to what degree. In addition, the analysis also includes the percentage of analogues where the substituent was associated with an increase in solubility.

As might be anticipated, the hydrophobic substituents F, Cl, CF_3, and Me displayed a moderate to large mean reduction in solubility across all three studies while the polar groups OH and NH_2 consistently resulted in an increase. There is less agreement among the studies concerning the change in solubility brought about by SO_2Me and CO_2H. In a number of instances, these substituents led to a reduction in solubility (data not shown). This is likely due to increased crystal lattice stability since SO_2Me and CO_2H are known to form strong intermolecular interactions in the solid state. The apparent disparity between the studies might be due to differences in the experimental conditions used to measure solubility or the chemotypes being examined.

Table 6 MMP analysis of common substituents

ΔLogS (% compounds with increased sol)			
	Leach et al. [63]	Zhang et al. [64]	Gleeson et al. [65]
F	−0.22 (34)	−0.45 (22)	−0.1 (9)[Ar]
Cl	−0.67 (14)	−1.45 (4)	−0.35 (5)[Ar]
CF$_3$	−0.81 (17)	−0.77 (25)	−0.54 (3)
Me	−0.21 (33)	−0.50 (26)	−0.11 (11)
OMe	−0.11 (42)	−0.24 (43)	−0.03 (19)
CN	−0.26 (36)	–	−0.14 (9)
OH	0.07 (56)	0.97 (85)	0.31 (48)[Ali]
NH$_2$	–	0.76 (61)	0.37 (54)[Ali]
SO$_2$Me	0.26 (71)	−0.38 (0)	0.01 (27)
CO$_2$H	–	−0.05 (45)	0.57 (56)

An interesting trend revealed by these studies is that even though the hydrophobic groups F, Cl, CF$_3$, and Me are associated with a mean reduction in solubility, there were a significant number of examples where these groups yielded an improvement in solubility. The causes for the solubility-augmenting ability of these substituents have not been comprehensively examined; however, the examples presented in the following sections might shed some light on this phenomenon.

The Dual Nature of Fluorine

Interest in fluorine as a substituent in drug design has been growing due to its favorable effect on potency and metabolism [66]. Moreover, new synthetic methods have made fluorinated compounds more readily accessible, meaning that the number of compounds containing fluorine is likely to increase in the upcoming years. Because of this, it is worthwhile to look beyond the pharmacology of organic fluorides and consider the physiochemical properties of these compounds, particularly their aqueous solubility. Although F is generally considered to be a hydrophobic group, leading to reduced solubility, this is not always the case. The MMP analysis presented in Table 6 indicated that although F substitution resulted in a 1.3–3-fold mean loss in solubility, a significant portion of the compounds examined (up to 34%) showed an increase. A brief review of the literature suggests that these results are not anomalous and that there are particular structural arrangements that cause fluorine to act as a solubility enhancer.

The recognition of fluorine's potential to increase solubility is not new and today is built into calculations of clog$P$. The circumstances under which F can lead to a reduction in log$P$ were reported by Hansch based on an investigation of halogen atom hydrophobicity. Table 7 shows the hydrophobicity constants (π_x) [67] determined for F, Cl, and Br using either benzene (X=H) or pentane (X=H) as reference templates. As expected, when H is substituted with F, Cl, or Br on benzene, there is a net increase in hydrophobicity ($\pi_x > 0$). However, when the reference template is pentane, there is a shift towards lower π_x and F

Table 7 Hansch hydrophobicity constants $(\pi_x)^a$ for F, Cl, and Br

[b]Ref		$CH_3CH_2CH_2CH_2CH_2X$
x	π_x	π_x
F	0.14	−0.17
Cl	0.71	0.39
Br	0.86	0.60

[a]$\pi_x = \log P_{R\text{-}X} - \log P_{R\text{-}H}$ where x = F, Cl, or Br and R = benzene or pentane
[b]Reference template
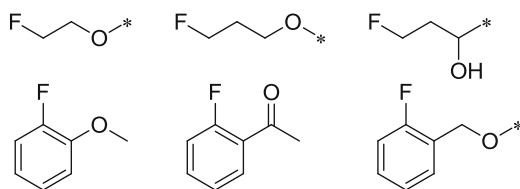
Fig. 6 Fluorine-containing fragments resulting in a decrease in $\log D$

actually results in a net reduction in hydrophobicity ($\pi_x < 0$) [68]. It appears that fluorination has an opposite effect on hydrophobicity depending on whether the attachment site is an aromatic ring or an alkyl group.

More recent studies indicate that fluorine's effect on aryl and alkyl groups is more complex and that there are other structural elements which can interact with fluorine to produce a decrease in hydrophobicity. For example, analysis of compounds from the Roche in-house collection found that introducing fluorine 2–3 carbon atoms away from an oxygen atom (Fig. 1), regardless of whether the site of substitution was aromatic or alkyl, led to a reduction in $\log D$ (Fig. 6) [69].

There are examples in the literature where fluorination has been rationally pursued in order to increase drug solubility. For example, the antibiotic analogues **46–49** in Fig. 7 show a similar reduction in hydrophobicity (partitioning between CHCl$_3$ and pH 7.4 buffer) when fluorine is introduced onto the cyclopropyl ring [70]. Regardless of the starting template, **48** or **46**, a comparable reduction in hydrophobicity ($\Delta\log K_{CHCl3} < 0$) is achieved upon fluorination.

The compounds in Table 8 are taken from a report describing diacylglycerol O-acyltransferase 1 (DGAT1) inhibitors [71]. Variation of the substitution at the terminal phenyl ring led to a dramatic improvement in solubility for the bis-fluorinated analogue **54** compared to the others (**50–53**). Interestingly, the hydrophobicity of **54** ($\log D_{7.4} = 2.5$) is similar to the parent compound **50** ($\log D_{7.4} = 2.5$), indicating that in this case the fluorines do not alter hydrophobicity. This implies that the crystal lattice has been weakened. Reduced cohesive interaction energy probably plays a role since the introduction of Me and Cl has no effect on solubility (cf. **51**, **52**, and **53**).

$$\Delta \log K_{CHCl3}$$

-0.75

-0.45
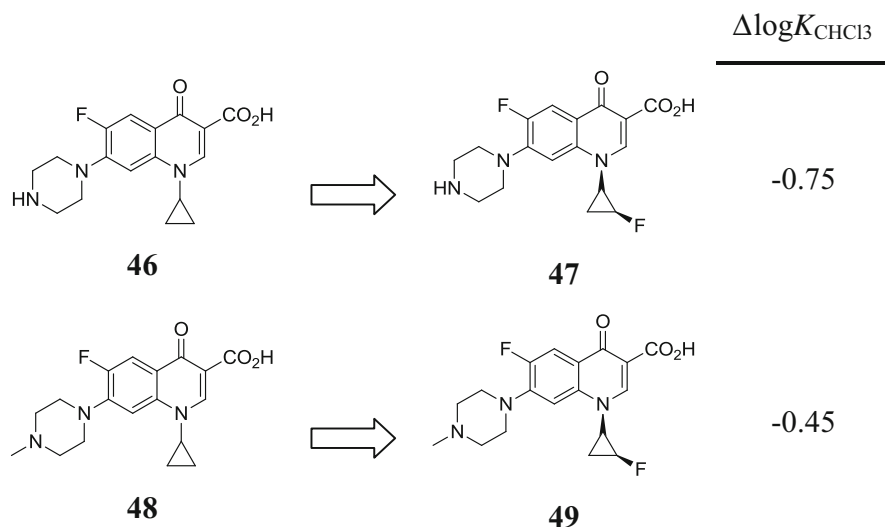
Fig. 7 Antibiotics showing lower hydrophobicity upon fluorination

Table 8 $\log D_{7.4}$ and solubility of DGAT1 of **50–54**

Compound	R^1	R^2	$\log D_{7.4}$	Solubility (μM)
50	H	H	2.4	3
51	Cl	H	2.9	5
52	Me	H	2.7	6
53	Cl	Me	3.4	1
54	F	F	2.5	140

Adamantane carboxylic acid was used by Jasys et al. to study the effect fluorine substitution had on both hydrophobicity and crystal lattice stability (Table 9) [72]. In line with Hansch's observation concerning the fluorination of alkyl templates, the measured $\log P$ decreased incrementally as the number of fluorines increased from 0 to 2 (cf. **55–57**), with the exception of the trifluoro-derivative **58**. The introduction of the third fluorine atom resulted in a small shift back towards increased hydrophobicity ($\log P = 2.50$) compared to the bis-substituted analogue **57** ($\log P = 2.16$). However, this last value should be viewed with caution since it is known that carboxylic acids occasionally yield an artificially elevated $\log P$

Table 9 Log$P$, clog$P$, and melting point data for fluoro-adamantane carboxylate derivatives **55–58**
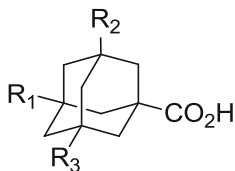
Compound	R^1, R^2, R^3	Log$P$	clog$P$	mp (°C)
55	**H, H, H**	3.26	2.47	173
56	**F, H, H**	2.55	2.01	155
57	**F, F, H**	2.16	1.55	163
58	**F, F, F**	2.50	1.09	202

due to aggregation [73]. The calculated clog$P$ indicates that **58** should be less hydrophobic (clog$P$ = 1.09) than **57** (clog$P$ = 1.55). In contrast to hydrophobicity, the outcome of fluorination on melting point is not constant. This is likely due to two opposing factors, intermolecular attraction versus molecular symmetry, making the results difficult to interpret.

Intramolecular Interactions

Intramolecular interactions between functional groups can change the overall shape of a molecule in solution and thereby alter hydrophobicity. An example of this type of interaction is demonstrated by the ability of phenyl rings to associate with appropriately positioned substituents (Table 10) [68]. This can be seen by comparing the hydrophobicity constants (π_x) with those derived from 3-phenyl-propane and using pentane as the reference template. For the 3-phenyl-propyl analogues, the positioning of the phenyl ring three carbons away from the substituent (X) allows it to assume a conformation where the two groups are aligned. In water, this conformation is reinforced by burying hydrophobic surface and a potentially stabilizing interaction between phenyl and the substituent. In agreement with this mechanism, the substituent hydrophobicity constants determined using the 3-phenyl-propyl template are lower than the pentyl. Moreover, the shift in hydrophobicity ($\Delta\pi_x$) between the two templates is nearly the same for all the substituents ($\Delta\pi_x$ = -0.51 to -0.65), suggesting a common mechanism. When toluene was used as the reference template (data not shown), the effect on π_x was lower and more variable.

Examples of Substituent Modifications

The examples reviewed in this section come from the recent literature with special emphasis being placed on the large effect that minor variations in structure can have on solubility.

Table 10 Hydrophobicity constants using pentyl and 3-phenyl-propyl templates

conformation A			conformation B (in water)		
X=	OH	F	Cl	CN	OMe	CONH$_2$
π_x (pentyl)	−1.16	−0.17	0.39	−0.84	−0.47	−1.71
π_x (3-phenyl-propyl)	−1.80	−0.73	−0.13	−1.47	−0.98	−2.24
[a]$\Delta\pi_x$	−0.64	−0.56	−0.52	−0.63	−0.51	−0.53

[a]$\Delta\pi_x = \pi_x$ (pentyl) $- \pi_x$ (3-phenyl-propyl)

The poor solubility of the tetrahydropyrazolo[1,5-*a*]pyrimidine-based antituber-culosis agent **59** (solubility at pH 6.8 < 4 μM) is likely caused by its relatively high hydrophobicity (log*P* 6.3) [74]. However, reducing hydrophobicity by exchanging the phenyl ring for pyridyl (**60**) produced only a modest increase in solubility (solubility at pH 6.8 = 9 μM). Unexpectedly, a larger improvement in solubility was achieved by simply removing one of the fluorine atoms from the trifluoromethyl substituent (**61**, solubility at pH 6.8 = 212 μM). The difference in log*P* between the difluoromethylated analogue **61** and the trifluoromethylated derivative **60** is small (log*P* = 4.6, 4.4 resp.) and consistent with the SAR presented above. The improve-ment in solubility accompanying the slight increase in hydrophobicity suggests that crystal packing has been affected by removing the fluorine atom.

59
log *P* 6.3
pH 6.8 sol < 4 μM

60
log *P* 4.4
pH 6.8 sol = 9 μM

61
log *P* 4.6
pH 6.8 sol = 212 μM
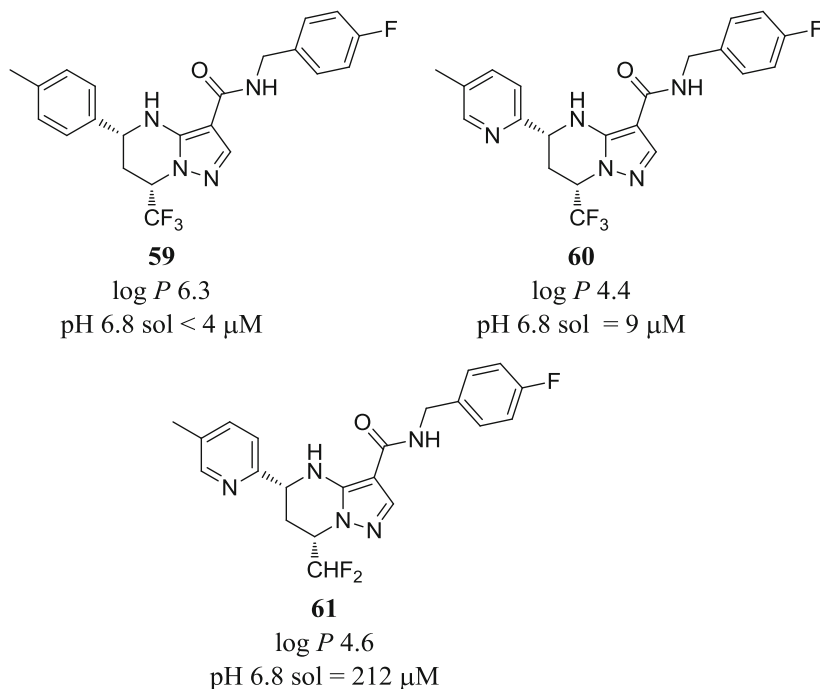
Solubility can sometimes be improved by simply rearranging the atoms of a heterocyclic ring. For example, a wide variation in solubility was observed among the quinoline regio-isomers **63–69** compared to the parent naphthalene derivative **62** [75]. Analogues where the heterocyclic N atom was adjacent to the linker NH (**63** and **64**) were considerably less soluble than those where it was farther away. In contrast, quinoline derivatives **70–72** display the opposite behavior [76]. For this template, isomers **71** and **72** having the ring nitrogen adjacent to the linking N atom were more soluble.

62–69

62
0.1 μM

63
0.7 μM

64
18 μM

65
30.3 μM

66
60.5 μM

67
84.9 μM

68
101.3 μM

69
138.5 μM

70–72

70
3 μM

71
590 μM

72
1100 μM
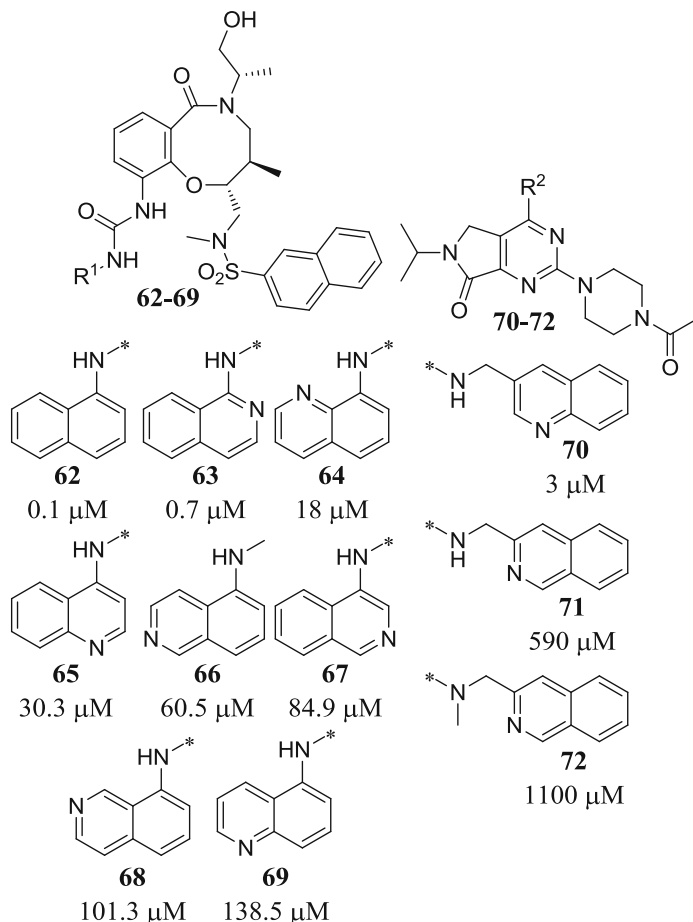
Seemingly minor changes in substituent stereochemistry can also alter the physiochemical properties of a molecule. For example, the relative stereochemistry of the [3.1.0]bicyclohexane ring connected to N-terminal carbamate proved to be a critical parameter in determining the solubility of **73** and **74** [77]. Compound **74**

containing the *cis*-stereochemistry was approximately 7 times more soluble than the related *trans*-isomer **73**. How such a large change in solubility can be caused by simply altering the stereochemistry of a group lying on a peripheral site of the molecule is not immediately apparent. However, it is known that intramolecular H bonding between terminal groups in peptidomimetics can decrease aqueous solubility by reducing the number of H-bonding interactions with water [78]. Therefore, it is possible that the *trans*-isomer **74** is less soluble due to intra-molecular H bonding which is hindered in the *cis*-analogue.
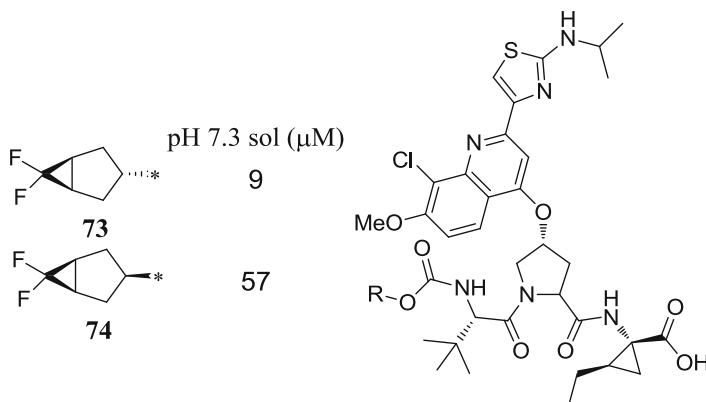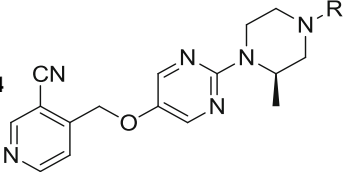
Template Modifications
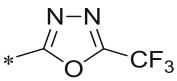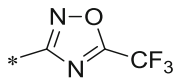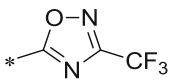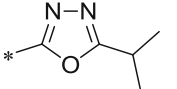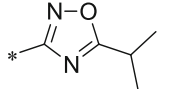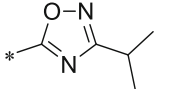
One can view a template as being the central body of a molecule which remains unchanged during analogue synthesis. In most cases, modification of this portion of the molecule is only carried out if it is believed that the desired pharmacological or physiochemical properties cannot be achieved through substituent modification. With respect to improving solubility, the structural changes can be subtle such as, replacing C with N/O or more dramatic such as altering the shape and topology of the molecule. The following examples capture a few of these tactics.

One of the most straightforward alterations to reduce hydrophobicity is to replace phenyl with a polar heterocycle. Pyridine is often chosen since it is structurally similar to phenyl and preserves the overall shape of the molecule. A demonstration of the potential magnitude in solubility improvement which can accompany this change is seen by comparing the solubility of the phenyl analogues **75** and **77** (solubility at pH 6.8 = <1 μM) with the corresponding pyridyl compounds **76** and **78** (solubility at pH 6.8 = 31 and 98 μM, respectively) [79]. It is interesting to note that this increase in solubility is brought about by changing a single carbon atom to nitrogen.

Table 11 Solubility (pH 7.4) and calculated H-bond basicities ($\log K_\beta$) of oxadiazole-analogues

	1,3,4-Oxadiazole	1,2,4-Oxadiazole	1,2,4-Oxadiazole
Compound	**79**	**80**	**81**
Solubility at pH 7.4 μM	118	3.1	4.8
calc $\log K_\beta$	1.4, 1.0	0.4, −0.8	0.2, 0.4
Compound	**82**	**83**	**84**
Solubility at pH 7.4 μM	150	5.9	12.6
calc $\log K_\beta$	2.9, 2.6	1.8, 0.8	1.6, 1.6
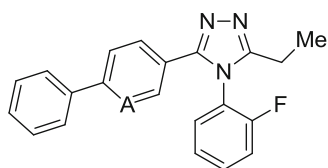
75 (CH); pH 6.8 sol < 1 μM
76 (N); pH 6.8 sol = 31 μM
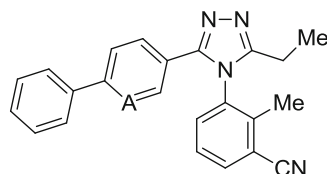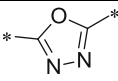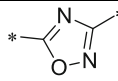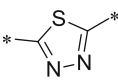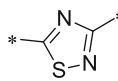
77 (CH); pH 6.8 sol < 1 μM
78 (N); pH 6.8 sol = 98 μM

Rearrangement of the heteroatoms in the template can have significant effects on solubility. For example, analysis of a large set of heterocyclic derivatives conducted by Ritchie et al. showed that compounds having a 1,3,4-oxadiazole ring were in general more soluble than those with the 1,2,4-regiochemistry as well as a number of other heterocycles [80]. This trend has also been observed by others. Goldberg et al. demonstrated that compounds **79** and **82** containing the 1,3,4-oxadiazole-based template were more soluble than **80**, **81**, **83**, and **84** which are derived from the 1,2,4-oxadiazole regardless of the regiochemical disposition (Table 11) [81]. In order to rationalize the differences in solubility, the H-bonding basicity ($\log K_\beta$) [82] of the ring nitrogen atoms was calculated for the parent heterocycles using the method of Kenny [83]. The outcome revealed that the higher solubility of the 1,3,4-isomers was associated with correspondingly higher overall H-bond basicity.

A similar result was reported for the sphingosine 1-phosphate receptor subtype 1 (S1P1) inhibitors compiled in Table 12 [84]. In agreement with the Goldberg study, the 1,3,4-oxadiazole **85** (solubility = 187 μg/mL) was found to be more soluble than the corresponding 1,2,4-isomer, **88** (solubility = 46 μg/mL). Furthermore, this relationship

Table 12 Solubility of compounds **85–88**

	1,3,4-Oxadiazole	1,3,4-Thiadiazole
Compound	**85**	**86**
Solubility (µg/mL)	187	46

	1,3,4-Thiadiazole	1,2,4-Thiadiazole
Compound	**87**	**88**
Solubility (µg/mL)	256	160

also appears to apply to other 5-membered ring heterocycles such as thiadiazole. Like the oxadiazole templates, 1,3,4-thiadiazole **87** (solubility = 256 µg/mL) was found to be more soluble than the 1,2,4-analogue **88** (solubility = 160 µg/mL).

Reducing the degree of aromaticity of a molecule by transforming aryl rings into their corresponding saturated forms is a useful method for increasing solubility [4]. This is nicely illustrated in a program targeting compounds that downregulate the androgen receptor [85]. Addition of two hydrogen atoms to the poorly soluble analogue **89** (solubility at pH 7.4 = 5.4 µM) to afford **90** resulted in a ~100-fold increase in solubility (solubility pH 7.4 = 577 µM). It is of interest to note that this modification results in a lowering of the hydrophobicity ($\log D_{7.4}$) which is probably responsible for the improved solubility.

89
$\log D_{7.4} = 3.2$
sol pH 7.4 = 5.4 µM

90
$\log D_{7.4} = 2.5$
sol pH 7.4 = 577 µM
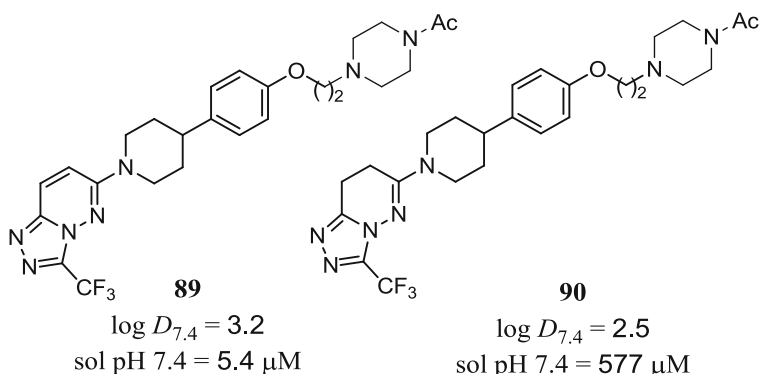
A different approach to reducing aromaticity is to replace the aromatic ring with an isosteric saturated ring. In a report by Stepan et al., the central phenyl ring of the γ-secretase inhibitor **91** was exchanged for a bicyclo[1.1.1]pentane

to yield **92** [86]. The benefit of using this bicyclic ring is that it can mimic the *para*-substituted phenyl by positioning the two substituents at 180° from each other. This resulted in a 12- and 33-fold improvement in aqueous solubility for **92** at pH 5.5 and 7.4, respectively, compared to **91**. As with the previous example, the increase in solubility is most likely brought about by a reduction in hydrophobicity as evidenced by the lower Log*D*.
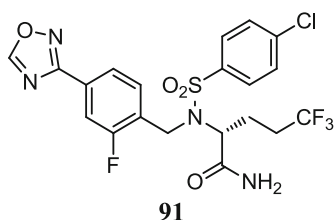
91	**92**
log *D* = 4.7	log *D* = 3.8
sol pH 6.5 = 1.7 μM	sol pH 6.5 = 19.7 μM
sol pH 7.4 = 0.9 μM	sol pH 7.4 = 29.4 μM
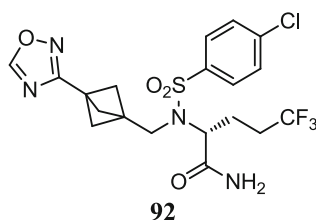
3.2   Disrupting Crystal Lattice Stability

Based on the GSE, Wassvik et al. estimated that solid-state stability plays a major role in limiting the solubility for compounds having log$P \leq 2$ and mp $\geq$250°C [87]. Compounds which meet this profile are prime candidates for solubility optimization by means of disrupting crystal packing energy. There are two general approaches for reducing crystal lattice stability. The first is to directly target specific intermolecular interactions which contribute to cohesive binding in the solid state. The second is to introduce modifications which alter the symmetry or shape of the molecule.

3.2.1   Disruption of Intermolecular H-Bonding Interactions

Without the aid of a single-crystal X-ray structure, it is difficult to determine which portion of the molecule to modify in order to disrupt crystal packing. Fortunately, it is known that if a molecule possesses H-bond donating and accepting functionality, these groups will most likely be involved in H-bonding interactions in the solid state [88]. Therefore, targeting these groups for modification is usually the most reliable and straightforward method for disrupting crystal lattice stability.

   H bonding can be impeded in two different ways. In the first, removing the acceptor or capping the donor group will completely eliminate the H bond. In contrast, in the second approach, the H bond is retained but weakened by introducing functionality which alters the steric or electronic environment around donor and/or acceptor groups.

Table 13 Solubility, log$P$, and melting point data for compounds **93–96**
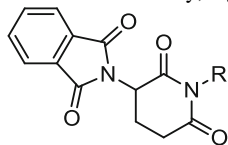
Compound	R	Solubility at pH 6.4 (µg/L)	Log$P$	mp °C
93	H	52	0.5	275
94	Methyl	276	1.1	159
95	*n*-Propyl	57	2.1	136
96	*n*-Pentyl	7	3	105
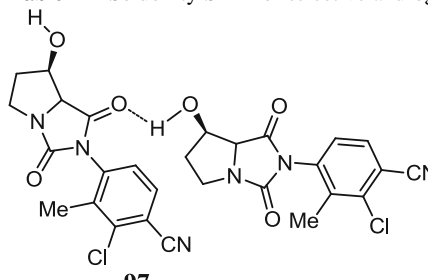
Removing Hydrogen-Bonding Functionality

Thalidomide (**93**, Table 13) is a poorly soluble, high melting compound (mp = 275 °C) with low hydrophobicity (log$P$ = 0.5). The high melting point is likely due to the strong H-bond forming imide functionality which leads to strong inter-molecular cohesion in the solid state [89]. Removing the imide N–H donor group by capping the nitrogen with a methyl moiety (**94**) eliminates the imide-imide H-bonding interaction [90]. Although there is an increase in log$P$, the net result is an improvement in solubility due to the lower melting point. The melting point can be further reduced by increasing the length of the N-alkyl group, as with **95** and **96**. However, for the larger N-alkyl groups, the reduction in solid-state stability is offset by increased hydrophobicity which leads to reduced solubility.

The androgen receptor modulator **97** (Table 14) displays solid-state limited solubility (solubility = 19 µg/mL, log$P$ = 1.3, mp = 255–257°C) [91]. A single-crystal X-ray structure suggested that H bonding between the imide carbonyl oxygen and the C7-hydroxyl group contributed to crystal lattice stability; therefore, the imide carbonyl was targeted for removal. In the absence of the carbonyl, a large increase in solubility was observed (e.g., **98**, solubility = 245 µg/mL). Alkylation at the former carbonyl site and modification of the C7 stereochemistry further reduced melting point and, as a consequence, increased aqueous solubility (cf. **99–101**). Interestingly, with the exception of compound **100**, there is a direct correlation between solubility and melting point for this series. The reason that **100** does not conform to this trend can be conjectured based on its elemental analysis which shows that it crystallizes as a hydrate.

Weakening Intermolecular Hydrogen Bonding

Intermolecular hydrogen bonding can be weakened by introducing substituents in the vicinity of the bond. An example of this can be seen with the metabolites of the leflunomide analogue shown in Table 15 [92]. Two intermolecular H-bonding motifs, designated **A** and **B**, are possible, as indicated by X-ray crystal-lography. The phenyl ring substitution pattern appears to dictate which is formed.

Table 14 Solubility SAR for selective androgen receptor modulators **97–101**

Compound		Solubility (µg/mL)	mp (°C)	LogP
97		19	255–257	1.3
98		245	197–198.5	1.3
99		613	175–177	2.2
100		242	125–127	2.7
101		1,140	99–101	2.3

Two modifications were observed to be critical: substitution at R_1 and/or introducing fluorine anywhere on the ring. In general, motif **A** is only formed when $R_1 = H$ (cf. **102–105**) while the presence of fluorine anywhere on the phenyl ring hampered the formation of motif **B** (cf. **104**, **105**, **108**, **109**). This appears to be the result of an inductive or through space interaction since all of the compounds exhibit the same flat conformation where R_1 is aligned with the amide N–H.

The B-Raf^{V600E} kinase inhibitor **110** displayed solubility-limited absorption (solubility at pH 7.4 = 9 µg/mL) which hampered oral exposure [93]. The relatively low hydrophobicity (clog$P$ = 1.61) coupled with an elevated melting point (mp = 229°C) indicated that crystalline stability was a significant contributor t the poor solubility. A single-crystal X-ray structure of **110**-H_2O showed extensive H bonding between neighboring molecules, as well as water, in the crystal lattice

Table 15 Intermolecular H bonding of leflunomide analogue metabolites

H bonding	Compound	R^1	R^2	R^3	R^4
A, B	**102**	H	Cl	H	Cl
A, B	**103**			Cl	H
A	**104**	H	F	H	F
A	**105**			H	H
B	**106**	Cl	H	Cl	H
B	**107**	Br		Br	
None	**108**	F	H	H	H
None	**109**				F

(Fig. 8). The structure features four hydrogen bonds involving the sulfonamide and the azaindole which hold two molecules together in a head-to-tail arrangement. Replacing the C3-methoxy group with cyclopropyl and one of the *ortho*-fluorine atoms located on the phenyl ring with Cl yielded **111** which exhibits a significant reduction in melting point and an increase in solubility. An X-ray crystal structure of **111** shows that the intermolecular H-bonding arrangement has been significantly altered compared to **110**. Most notably, the pyridine nitrogen and sulfonamide oxygen atoms no longer participate in H bonding, leaving the four remaining H bonds to be shared among three molecules rather than two.

3.2.2 Disrupting Planarity and Symmetry

Molecular planarity and symmetry play an important role in crystal lattice stability and therefore provide valuable targets for structural modifications. There are a number of reports in the literature which show solubility improvement via introducing substituents which alter the planarity or symmetry. A review summarizing this approach has been recently published [94]. Therefore, in order to avoid duplication, the present review will focus on examining the effect these modifications have on the packing arrangement of the molecules in the crystal. It has been hypothesized that the lower crystal lattice stability brought about by alterations in

110
mp = 229 °C
clog P = 1.61
sol pH 7.4 = 9 µg/mL

111
mp = 164 °C
clog P = 2.19
sol pH 7.4 = 127 µg/mL
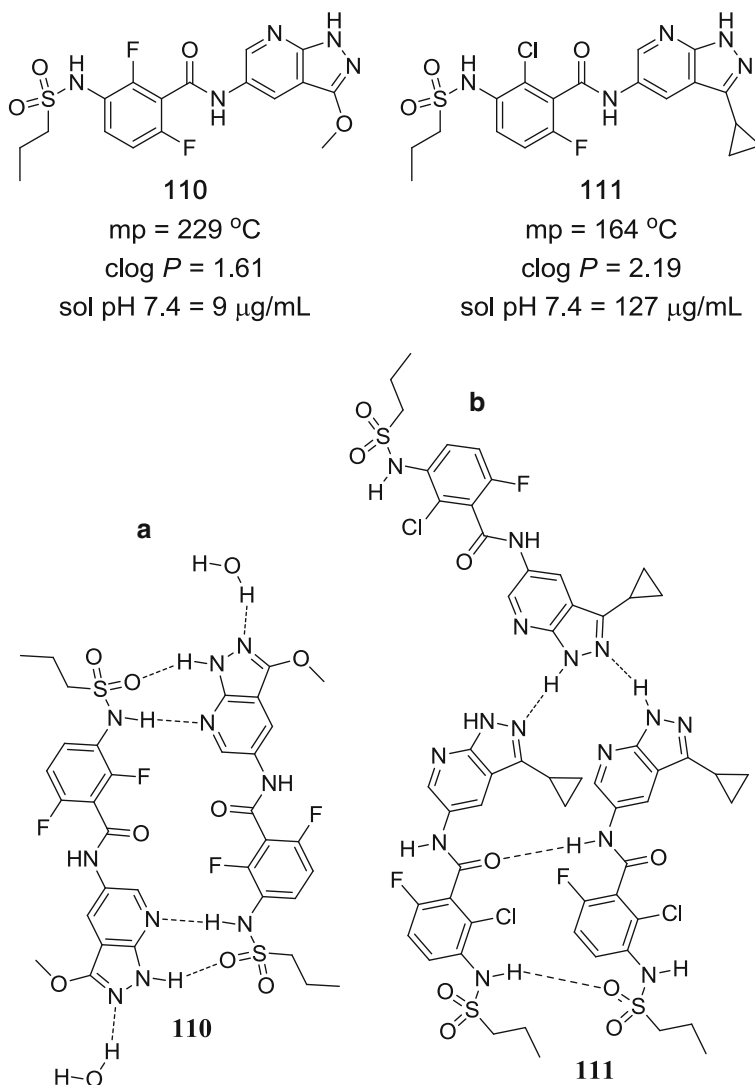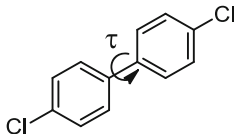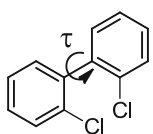
Fig. 8 Crystal packing arrangements of **110** (**a**) and **111** (**b**)

symmetry or planarity is due to a reduction in packing efficiency and the loss of π-stacking interactions. However, as seen in the following, there is often only a minor change in packing efficiency accompanying these modifications compared to the more significant changes in intermolecular interactions. Moreover, π-stacking is only one of a number of other interactions that are altered.

Disrupting the planarity of biphenyl derivatives or aryl amides via introducing groups which widen the torsional bond angle at the phenyl-phenyl or amide NH-aryl junctions often results in a reduction in melting point and increase in

Table 16 Properties of 4,4-
and 1,1′-dichlorobiphenyl

	112 4,4′-Dichlorobiphenyl	**113** 2,2′-Dichlorobiphenyl
Log$S$	−3.567	−2.575
mp (°C)	144–145	61–62
Log$P$	5.02	5.02
τ	38.3°	66.5°
C_k	0.695	0.678

solubility. The Cambridge Structural Database (CSD) provides a few sets of compounds which can be examined as models for evaluating the changes in crystal packing associated with these modifications.

As shown in Table 16, 4,4′-dichlorobiphenyl (**112**) is more symmetrical than the corresponding 2,2′-isomer **113** and in the solid state displays a slightly more planar conformation around the bond connecting the two phenyl rings ($\tau = 38.3°$ vs. 66.5°, for **112** and **113**, respectively). Examination of log$P$ and melting point data indicates that the lower solubility of 4,4′-dichlorobiphenyl is most likely the result of higher crystal lattice stability [95]. Although the 4,4′-analogue packs slightly more efficiently than the 2,2′-analogue ($C_k = 0.695$ vs. 0.678), the difference in stability is more ostensibly manifested through alterations in the type and number of intermolecular interactions in the solid state. Figure 9 shows that 4,4′-dichlorophenyl forms a linear chain of molecules held together by Cl···Cl bonding. These are aligned in a parallel fashion which appears to allow edge-to-face aryl–aryl binding between parallel strands. In addition, the Cl atoms are close enough to the phenyl ring of the adjacent strand to potentially to be involved in Cl···H or Cl··· π-bonding. In contrast, 2,2′-dichlorobiphenyl does not have the Cl···Cl interaction but instead forms a network of C–H···Cl bonds. However, like the 4,4′-isomer, edge-to-face aryl–aryl interactions are also present. This cursory analysis implies that the differences in stability between **112** and **113** might be due to the differences in the intermolecular interactions of the Cl group.

An example of the potential effect *ortho*-substitution can have on amide bond geometry and crystal packing is illustrated by compounds **114** and **115** in Fig. 10. The *ortho, ortho*-dichloro analogue **115** is significantly less planar ($\tau = 61.3°$) than the corresponding *meta, meta*-analogue ($\tau = 0.05°$) and has a correspondingly lower melting point (mp = 178 vs. 189°C). The X-ray crystal structures of the two compounds indicate that the amide groups and aryl rings of both compounds are involved in similar H-bonding and π-stacking contacts. It appears that there is no apparent difference in the type and number of intermolecular interactions between the two isomers. This suggests the possibility of an additional interaction contributing to the stability of **114** or that the H-bonding and/or pi-stacking associations for this compound are stronger compared to **115**.
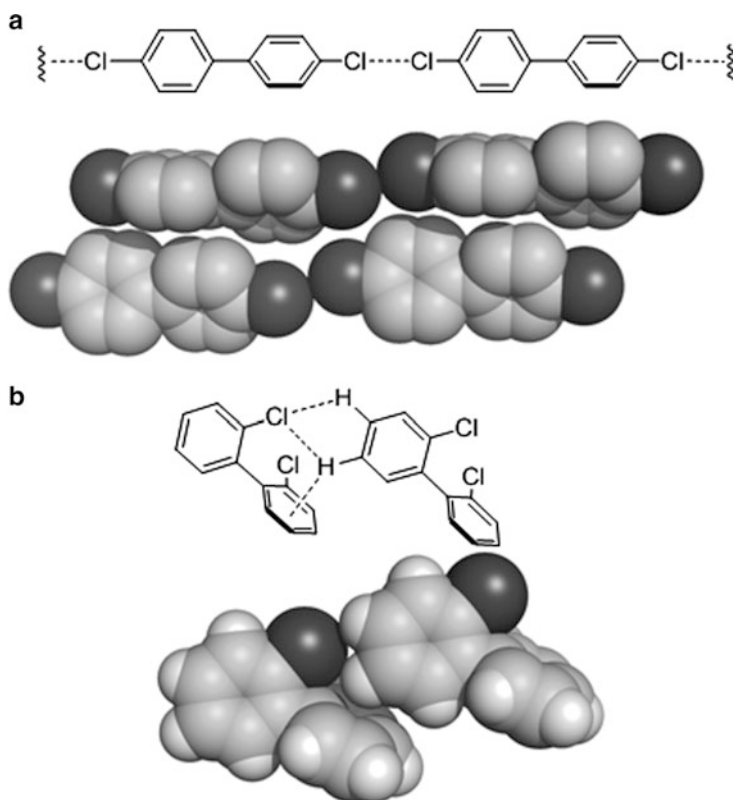
Fig. 9 Crystal packing 4,4′-dichlorophenyl (**a**) and 2,2-dichlorobiphenyl (**b**)

The ATR protein kinase inhibitor **116** (Table 17) exhibits solid-state limited solubility, as revealed by its modest hydrophobicity ($\log D_{7.4} = 2.2$) and high melting point (mp = 221°C) [96]. A single-crystal X-ray structure of **116** (Fig. 11) shows that in addition to the π-stacking of the pyrimidine and indole rings (**a**), the sulfone group is involved in strong H-bonding interactions with the indole-NH (**b**) and the methylsulfone-CH (**c**). In an attempt to reduce crystal lattice stability, a methyl group was introduced onto the morpholine group to yield **117**, which displays an increase in solubility proportional to its lower melting point. As observed in the X-ray crystal structure, methylation forces the adjacent indole group to rotate ($\tau = 22.3°$) so that it is no longer in the same plane as the pyrimidine group. The result is that the molecules must realign themselves in order to maintain a π-stacking interaction (Fig. 11**d**). Methylation also affects the lattice H-bonding network since the indole-NH-sulfone H bond is missing in **117**. Of the two interactions which are affected, one could assume that the loss of this H bond plays a larger role in reducing the melting point compared to a presumed difference in energy of the corresponding π-stacking arrangements.
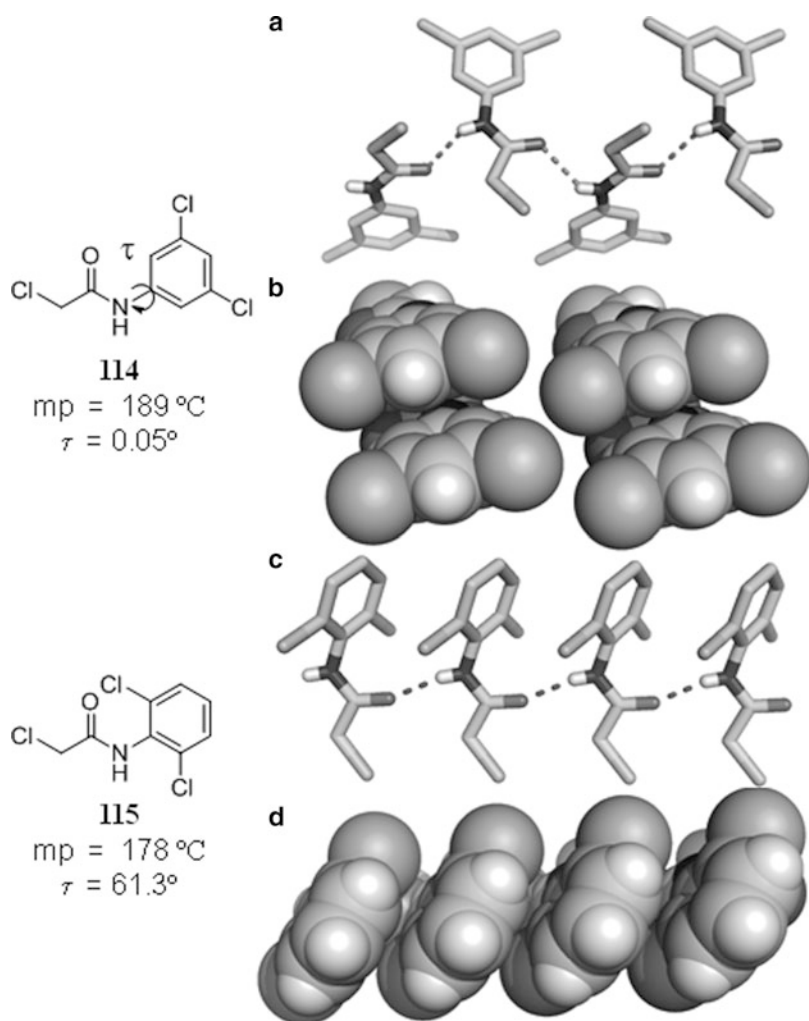
Fig. 10 Solubility, melting point, and crystal packing arrangement for **114** and **115**

Table 17 Melting point, log$D_{7.4}$, and solubility data for **116** and **117**
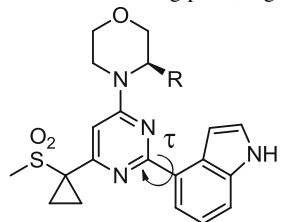
Compound	R	mp (°C)	τ	Log$D_{7.4}$	Solubility (µM)
116	H	221	3.7°	2.2	4.4
117	Me	204	22.3°	2.7	10

Fig. 11 Crystal packing arrangement of **116** (**a–c**) and **117** (**d**)
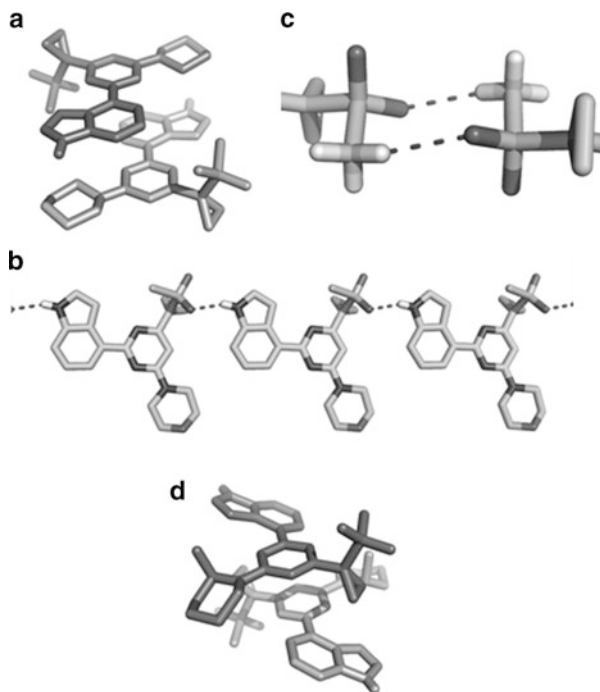
4 Summary and Conclusion

Crystalline solubility is dependent on the ability of a compound to escape the forces holding it in the lattice (crystal lattice energy) and on its potential to interact with water when it is in solution (solvation energy). Often the structural modifications which lead to a change in one property (solid-state stability or solvation energy) will affect the other in the opposite direction. Therefore, the challenge is to find an appropriate balance between the changes in crystal lattice and solvation energies such that the net effect is an increase in solubility.

The GSE relates the crystal lattice energy of a compound to its melting point and its solvation energy to log*P*. Thus, structural modifications which lead to a reduction in melting point or log*P* can be used to increase solubility. Since poor solubility is usually the result of high hydrophobicity, methods for lowering log*P* by increasing polarity are more often found in the literature. The tactics for accomplishing this can be categorized according to the site on the molecule which is modified, relative to how it binds to its target. On the other hand, compounds which are poorly soluble but not highly hydrophobic (log*P* ≤ 2) probably suffer from solid-state limited solubility. A high melting point ($\geq 250°$C) is indicative of this. For these compounds, disrupting intermolecular interactions which exist in the solid state will result in a lower melting point and improved solubility.

While the effect a given group has on solubility can usually be predicted based on its chemical properties, there are cases where it is not. In particular, there are numerous examples where intrinsically hydrophobic substituents have resulted in an increase in solubility rather than a decrease. One goal of this review was to highlight these instances in order to uncover the underlying cause for this apparent contradictory behavior. An examination of these examples has identified a few consistent structural contexts under which solubility can be improved using nominally hydrophobic substituents. The incorporation of these structural features in drug design expands the methods available for optimizing solubility and allows for the fine-tuning of molecular properties.

The optimization of physical properties in parallel with biological activity and pharmacology is recognized as an essential component of successful drug discovery. This is especially true in the modern era of medicinal chemistry due to the higher demands placed on potency and selectivity since this often has the effect of increasing the size and hydrophobicity of the molecules being examined as potential drugs. While certain remedies have been institutionalized by most companies in order to tackle this issue, there is no panacea to avoid this trend completely. As such, it is important that medicinal chemists be equipped with the appropriate knowledge to solve problems such as poor solubility as they arise. Unfortunately, the factors which control solubility are not completely understood, meaning that solubility issues can only be solved through a combination of rational and empirical molecular design. This requires a fundamental understanding of the principles of solvation and crystalline stability as well as a body of literature examples as a source of ideas.

References

1. Amidon GL, Lennernäs HL, Shah VP, Crison JR (1995) A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm Res 12:413–420
2. Lipinski CA (2000) Drug-like properties and the causes of poor solubility and poor permeability. J Pharm Tox Meth 44:235–249
3. Lovering F, Bikker J, Humblet C (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52:6752–6756
4. Ritchie TJ, Macdonald SJF (2009) The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? Drug Disc Today 14:1011–1020
5. Hill AP, Young RJ (2010) Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity. Drug Disc Today 15:649–655
6. Yang Y, Engkvist O, Llinàs A, Chen H (2012) Beyond size, ionization state, and lipophilicity: influence of molecular topology on absorption, distribution, metabolism, excretion, and toxicity for drug like compounds. J Med Chem 55:3667–3677
7. Hann MM (2011) Molecular obesity, potency and other addictions in drug discovery. Med Chem Comm 2:349–355
8. Faller B, Ottaviani G, Ertl P et al (2011) Evolution of the physicochemical properties of marketed drugs: can history foretell the future? Drug Discov Today 16:976–984

9. Di L, Fish PV, Mano T (2012) Bridging solubility between drug discovery and development. Drug Discov Today 17:486–495
10. Meanwell NA (2011) Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. Chem Res Toxicol 24: 1420–1456
11. Ran Y, Yalkowsky SH (2001) Prediction of drug solubility by the general solubility equation (GSE). J Chem Inf Comput Sci 41:354–357
12. Ran Y, Jain N, Yalkowsky SH (2001) Prediction of aqueous solubility of organic compounds by the general solubility equation (GSE). J Chem Inf Comput Sci 41:1208–1217
13. Ben-Naim A (2001) On the evolution of the concept of solvation thermodynamics. J Sol Chem 30:475–487
14. Frank HS, Evans MW (1945) Free volume and entropy in condensed systems. III. Entropy in binary liquid mixtures; partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes. J Chem Phys 13:507–532
15. Ben-Naim A, Tenne R (1977) Application of the scaled particle theory to the problem of hydrophobic interaction. J Chem Phys 67:627–635
16. Lee B, Graziano G (1996) A two-state model of hydrophobic hydration that produces compensating enthalpy and entropy changes. J Am Chem Soc 118:5163–5168
17. Reiss H, Frisch HL, Lebowitz JL (1959) Statistical mechanics of rigid spheres. J Chem Phys 31:369–380
18. Tolls J, van Dijk J, Verbruggen EJM et al (2002) Aqueous solubility-molecular size relationships: a mechanistic case study using C10- to C19-alkanes. J Phys Chem A 106:2760–2765
19. Khadikar RV, Mandloi D, Bajaj AV, Joshi S (2003) QSAR Study on solubility of alkanes in water and their partition coefficients in different solvent systems using PI index. Bioorg Med Chem Lett 13:419–422
20. Ferguson AL, Debenedetti PG, Panagiotopoulos AZ (2009) Solubility and molecular conformations of n-alkane chains in water. J Phys Chem B 113:6405–6414
21. Abraham MH, Ibrahim A, Zissimos AM et al (2002) Application of hydrogen bonding calculations in property based drug design. Drug Discov Today 7:1056–1063
22. Raevsky O (2008) H-Bonding parameterization in quantitative structure–activity relationships and drug design. In: Mannhold R (ed) Molecular drug properties. Wiley-VCH, Weinheim
23. Gurka D, Taft RW (1969) Studies of hydrogen-bonded complex formation with p-fluorophenol. IV. Fluorine nuclear magnetic resonance method. J Am Chem Soc 91:4794–4801
24. Arnett EM, Joris L, Mitchell E et al (1970) Hydrogen-bonded complex formation. III. Thermodynamics of complexing by infrared spectroscopy and calorimetry. J Am Chem Soc 92:2365–2377
25. Berthelot M, Laurence C, Safar M, Besseau F (1998) Hydrogen-bond basicity pK_{HB} scale of six-membered aromatic N-heterocycles. J Chem Soc Perkin Trans 2:283–290
26. Laurence C, Brameld KA, Graton J et al (2009) The pK(BHX) database: toward a better understanding of hydrogen-bond basicity for medicinal chemists. J Med Chem 52:4073–4086
27. Yalkowsky SH, Valvani SC (1980) Solubility and partitioning I: solubility of nonelectrolytes in water. J Pharm Sci 69:912–922
28. Dannenfelser RM, Yalkowsky SH (1996) Estimation of entropy of melting from molecular structure: a non-group contribution method. Ind Eng Chem Res 35:1483–1486
29. Chickos JS, Braton CM, Hesse DG, Liebman JF (1991) Estimating entropies and enthalpies of fusion of organic compounds. J Org Chem 56:927–938
30. Dunitz JD, Filippini G, Gavezzotti A (2000) Molecular shape and crystal packing: a study of $C_{12}H_{12}$ isomers, real and imaginary. Helv Chim Acta 83:2317–2335
31. Du Y, Creighton CJ, Tounge BA, Reitz AP (2004) Noncovalent self-assembly of bicyclo [4.2.2]diketopiperazines: influence of saturation in the bridging carbacyclic ring. Org Lett 6: 306–312
32. Dunitz JD, Gavezzotti A (2009) How molecules stick together in organic crystals: weak intermolecular interactions. Chem Soc Rev 38:2622–2633

33. Zhang J, Wang Y, Wang Q, Xu L (2012) 6,7-Dichloro-3-(2,4-dichlorobenzyl)- quinoxalin-2 (1H)-one. Acta Crystallogr Sect E 68:2481

34. Zhou Y, Ren L, Lu Y, Zhang F, Chen G (2012) N-[4-Chloro-3-(trifluoromethyl)phenyl]-2,2-dimethylpropanamide. Acta Cryst Sect E 68:2534

35. Gavezzotti A, Filippini G (1994) Geometry of the intermolecular X-H-Y (X, Y = N, O) hydrogen-bond and the calibration of empirical hydrogen-bond potentials. J Phys Chem 98:4831–4837

36. Mei X, Wolf C (2004) Formation of new polymorphs of acridine using dicarboxylic acids as crystallization templates in solution. Cryst Growth Des 4:1099–1103

37. Metrangolo P, Resnati G (2012) Halogen bonding: where we are and where we are going. Cryst Growth Des 12:5835–5838

38. Pathak S, Kundu A, Pramanik A (2012) Regioselective synthesis of two types of highly substituted 2-pyridones through similar multicomponent reactions. Tet Lett 53:3030–3034

39. Cortez-Maya S, Cortes EC, Hernández-Ortega S et al (2012) Synthesis of 2-aminobenzophenone derivatives and their anticancer activity. Syn Comm 42:46–54

40. Allen FH (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. Acta Cryst B 58:380–388

41. Gavezzotto A (1989) On the preferred mutual orientation of aromatic groups in organic condensed media. Chem Phys Lett 161:67–72

42. Główka ML, Martynowski D, Kozłowska K (1999) Stacking of six-membered aromatic rings in crystals. J Mol Struct 474:81–89

43. Martinez CR, Iverson BL (2012) Rethinking the term "pi-stacking". Chem Sci 3:2191–2201

44. Wheeler SE, Houk KN (2008) Substituent effects in the benzene dimer are due to direct interactions of the substituents with the unsubstituted benzene. J Am Chem Soc 130:10854–10855

45. Pinal R (2004) Effect of molecular symmetry on melting temperature and solubility. Org Biomol Chem 2:2692–2699

46. Wei J (1999) Molecular symmetry, rotational entropy and elevated melting points. Ind Eng Chem Res 38:5019–5027

47. Brown RJC, Brown RFC (2000) Melting point and molecular symmetry. J Chem Edu 77:724–731

48. Holler AC (1948) An observation on the relationship between the melting points of the disubstituted isomers of benzene and their chemical constitution. J Org Chem 13:70–74

49. Gavezzotti A (1995) Molecular symmetry, melting temperatures and melting enthalpies of substituted benzenes and naphthalenes. J Chem Soc Perkin Trans 2:1399–1404

50. Thalladi VR, Nüsse M, Boese R (2000) The melting point alternation in α, ω-alkanedicarboxylic acids. J Am Chem Soc 122:9227–9236

51. Gopalan RS, Kumaradhas P, Kulkami GU, Rao CNR (2000) An experimental charge density study of aliphatic dicarboxylic acids. J Mol Struct 521:97–106

52. Troshina OA, Troshin PA, Peregudov AS et al (2006) Photoaddition of N-substituted piperazines to C60: an efficient approach to the synthesis of water-soluble fullerene derivatives. Chemistry 12:5569–5577

53. Fischer H, Atzpodien EA, Csato M et al (2011) In silico assay for assessing phospholipidosis potential of small druglike molecules: training, validation, and refinement using several data sets. J Med Chem 55:126–139

54. Hennequin LF, Stokes ES, Thomas AP et al (2002) Novel 4-anilinoquinazolines with C-7 basic side chains: design and structure activity relationship of a series of potent, orally active, VEGF receptor tyrosine kinase inhibitors. J Med Chem 45:1300–1312

55. Yu T, Tagat JR, Kerekes AD et al (2010) Discovery of a potent, injectable inhibitor of aurora kinases based on the imidazo-[1,2-a]-pyrazine core. ACS Med Chem Lett 1:214–218

56. Govek SP, Oshiro G, Anzola JV et al (2010) Water-soluble PDE4 inhibitors for the treatment of dry eye. Bioorg Med Chem Lett 20:2928–2932

57. Addie M, Ballard P, Buttar D et al (2013) Discovery of 4-amino-N-[(1S)-1-(4-chlorophenyl)-3-hydroxypropyl]-1-(7H-pyrrolo[2,3-d]pyrimidin-4-yl)piperidine-4-carboxamide (AZD5363), an orally bioavailable, potent inhibitor of Akt kinases. J Med Chem 56:2059–2073

58. Greenwald RB, Choe YH, McGuire J, Conover CD (2003) Effective drug delivery by PEGylated drug conjugates. Adv Drug Del Rev 55:217–250

59. Zalipsky S, Gilon C, Zilkha A (1983) Attachment of drugs to polyethylene glycols. Eur Polym J 19:1177–1183

60. Smith JL, Rossiter KI, Semko CM et al (2013) PEG conjugates of potent α4 integrin inhibitors, maintaining sustained levels and bioactivity in vivo, following subcutaneous administration. Bioorg Med Chem Lett 23:4117–4119

61. Zhu GD, Arendsen DL, Gunawardana IW et al (2001) Selective inhibition of ICAM-1 and E-selectin expression in human endothelial cells. 2. Aryl modifications of 4-(aryloxy)thieno[2,3-c]pyridines with fine-tuning at C-2 carbamides. J Med Chem 44:3469–3487

62. Dossetter AG, Griffen EJ, Leach AG (2013) Matched molecular pair analysis in drug discovery. Drug Discov Today 18:724–731

63. Leach AG, Jones HD, Cosgrove DA et al (2006) Matched molecular pairs as a guide in the optimization of pharmaceutical properties; a study of aqueous solubility, plasma protein binding and oral exposure. J Med Chem 49:6672–6682

64. Zhang L, Zhu H, Mathiowetz A, Gao H et al (2011) Deep understanding of structure–solubility relationship for a diverse set of organic compounds using matched molecular pairs. Bioorg Med Chem Lett 19:5763–5770

65. Gleeson P, Bravi G, Modi S, Lowe D et al (2009) ADMET rules of thumb II: a comparison of the effects of common substituents on a range of ADMET parameters. Bioorg Med Chem 17:5906–5919

66. Hagmann WK (2008) The many roles for fluorine in medicinal chemistry. J Med Chem 51:4359–4369

67. Fujita T, Iwasa J, Hansch C (1964) A new substituent constant, π, derived from partition coefficients. J Am Chem Soc 86:5175–5180

68. Hansch C, Anderson SM (1967) The effect of intramolecular hydrophobic bonding on partition coefficients. J Org Chem 32:2583–2586

69. Böhm HJ, Banner D, Bendels S et al (2004) Fluorine in medicinal chemistry. ChemBioChem 5:637–643

70. Atarashi S, Imamura M, Kimura Y (1993) Fluorocyclopropyl quinolones. 1. Synthesis and structure-activity relationships of 1-(2-fluorocyclopropyl)-3-pyridonecarboxylic acid antibacterial agents. J Med Chem 22:3444–3448

71. Goldberg FW, Birch AM, Leach AG et al (2013) Discovery and optimization of efficacious neutral 4-amino-6-biphenyl-7,8-dihydropyrimido[5,4-f][1, 4] oxazepin-5-one diacylglycerol acyl transferase-1 (DGAT1) inhibitors. Med Chem Comm 4:165–174

72. Jasys VJ, Lombardo F, Appleton TA et al (2000) Preparation of fluoroadamantane acids and amines: Impact of bridgehead fluorine substitution on the solution- and solid-state properties of functionalized adamantanes. J Am Chem Soc 122:466–473

73. Wenlock MC, Barton P, Luker T (2011) Lipophilicity of acidic compounds: impact of ion pair partitioning on drug design. Bioorg Med Chem Lett 21:3550–3556

74. Yokoawa F, Wang G, Chan WL et al (2013) Discovery of tetrahydropyrazolopyrimidine carboxamide derivatives as potent and orally active antitubercular agents. ACS Med Chem Lett 4:451–455

75. Scully SC, Tang AJ, Lundh M et al (2013) Small-molecule inhibitors of cytokine-mediated STAT1 signal transduction in β-cells with improved aqueous solubility. J Med Chem 56:4125–4129

76. Cantin LD, Bayrakdarian M, Buon C et al (2012) Discovery of P2X3 selective antagonists for the treatment of chronic pain. Bioorg Med Chem Lett 22:2656–2671

77. Sheng XC, Appleby T, Butler T et al (2012) Discovery of GS-9451: an acid inhibitor of the hepatitis C virus NS3/4A protease. Bioorg Med Chem Lett 22:2629–2634

78. Kuhn B, Mohr P, Stahl M (2010) Intramolecular hydrogen-bonding in medicinal chemistry. J Med Chem 53:2601–2611

79. Sugane T, Tobe T, Hamaguchi W et al (2013) Atropisomeric 4-phenyl-4*H*-1,2,4-triazoles as selective glycine transporter 1 inhibitors. J Med Chem 56:5744–5756

80. Ritchie TJ, Macdonal SJF, Peace S et al (2012) The developability of heteroaromatic and heteroaliphatic rings – do some have a better pedigree as potential drug molecules than others? Med Chem Comm 3:1062–1069

81. Goldberg K, Groombridge S, Hudson J et al (2012) Oxadiazole isomers: all bioisosteres are not created equal. Med Chem Commun 3:600–604

82. Abraham MH, Duce PD, Prior DV et al (1989) Hydrogen bonding. Part 9. Solute proton donor and proton acceptor scales for use in drug design. J Chem Soc Perkin Trans II 1355–1375

83. Toulmin A, Wood JM, Kenny PW (2008) Toward prediction of alkane/water partition coefficients. J Med Chem 51:3720–3730

84. Ren F, Deng G, Wang H (2012) Discovery of novel 1,2,4-thiadiazole derivatives as potent, orally active agonists of sphingosine 1-phosphate receptor subtype 1 (S1P(1)). J Med Chem 55:4286–4296

85. Bradbury RH, Acton DG, Broadbent NL et al (2013) Discovery of AZD3514, a small-molecule androgen receptor downregulator for treatment of advanced prostate cancer. Bioorg Med Chem Lett 23:1945–1948

86. Stepan AF, Subramanyam C, Efremov IV et al (2012) Application of the bicyclo[1.1.1]pentane motif as a nonclassical phenyl ring bioisostere in the design of a potent and orally active γ-secretase inhibitor. J Med Chem 55:3414–3424

87. Wassvik CM, Holmén AG, Draheim RI et al (2008) Molecular characteristics for solid-state limited solubility. J Med Chem 51:3035–3039

88. Etter MC (1990) Encoding and decoding hydrogen-bond patterns of organic compounds. Acc Chem Res 23:120–126

89. Abraham MH (2004) Hydrogen bond and other descriptors for thalidomide and its N-alkyl analogs; prediction of physicochemical and biological properties. Eur J Pharm Sci 21:465–469

90. Goosen C, Laing TJ, du PJ, Goosen TC, Flynn GL (2002) Physicochemical characterization and solubility analysis of thalidomide and its N-alkyl analogs. Pharm Res 19:13–19

91. Li JJ, Sutton JC, Nirschl A et al (2007) Discovery of potent and muscle selective androgen receptor modulators through scaffold modifications. J Med Chem 50:3015–3025

92. Venkatachalam TK, Zheng Y, Ghosh S, Uckun FM (2005) Structural influence on the intermolecular/intramolecular hydrogen bonding in solid state of substituted leflunomides: evidence by X-ray crystal structure. J Mol Struct 753:103–115

93. Wenglowsky S, Moreno D, Rudolph J et al (2012) Pyrazolopyridine inhibitors of B-Raf[V600E]. Part 3: An increase in aqueous solubility via the disruption of crystal packing. Bioorg Med Chem Lett 22:912–915

94. Ishikawa M, Hashimoto Y (2011) Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. J Med Chem 54:1539–1554

95. Paasivirta J, Sinkkonen SI (2009) Environmentally relevant properties of all 209 polychlorinated biphenyl congeners for modeling their fate in different natural and climatic conditions. J Chem Eng Data 54:1189–1213

96. Foote KM, Blades K, Cronin A et al (2013) Discovery of 4-{4-[(3*R*)-3-methylmorpholin-4-yl]-6-[1-(methylsulfonyl)cyclopropyl]pyrimidin-2-yl}-1*H*-indole (AZ20): a potent and selective inhibitor of ATR protein kinase with monotherapy in vivo antitumor activity. J Med Chem 56:2125–2138

Tactics to Avoid Inhibition of Cytochrome P450s

Andrew G. Leach

Abstract This chapter will focus on avoiding inhibition of cytochrome P450s. This problem leads to drug–drug interactions that can severely restrict the patients who can receive a particular drug. The first sections [(1) background and motivation, (2) the structure of the cytochrome P450s, and (3) overview of the mechanism of cytochrome P450s] provide essential background to understanding both the problem and some of the solutions. The subsequent three sections [(4) trends for inhibition to be modulated by bulk properties (database analysis), (5) metal chelation, and (6) reactive metabolites and mechanism-based inhibition] will outline some general methods for tackling the problem, and each section has a summary in which the strategic approaches that follow from the information are presented. In the final section [(7) examples of successful reduction of P450 liability], how these strategies can be reduced to tactical application is illustrated with published examples.

Keywords Cytochrome P450, Inhibition, Lead optimization, Medicinal chemistry

Contents

A.G. Leach (✉)
Liverpool John Moores University, James Parsons Building, Byrom Street,
Liverpool L3 3AF, UK
e-mail: a.g.leach@ljmu.ac.uk

1   Background and Motivation

The cytochrome P450s (CYPs) are a class of iron containing enzymes that catalyze a range of redox reactions in biological contexts. In the human system, they play a key role in the metabolism of many substrates both natural and unnatural. It is this latter effect that is most directly relevant to medicinal chemists. The co-administration of two compounds one of which relies upon a particular CYP (or any other mechanism for that matter) for its metabolism may have its exposure significantly altered (usually increased) if the other compound is an inhibitor of that CYP (or other mechanism) [1]. This is a so-called drug–drug interaction and is a key cause of concern when a compound is found to be an inhibitor of a CYP [1–4].

A relatively small fraction of the large family of CYP isoforms is responsible for the metabolism of most drugs [2]. These are the isoforms that are able to bind (and in some cases metabolize) a large range of compounds with properties like those of drugs. The most significant of these isoforms are in order: 3A4, 2C9, 1A2, 2E1, 2D6, 2C19, 2C8, and 2A6 [2]. The reliance upon any single one of these enzymes for the elimination of a compound is also a liability that will not be explored further here. The inhibition of any of this small set of isoforms is likely to lead to a limitation of the other drugs that a compound can be co-dosed with. This can be a more severe problem in certain classes of disease targets; any affliction of the elderly or other groups likely to be on complex regimes of medication will require treatments that do not cause any of these concerns.

A stark example of the drug–drug interaction problem is provided by the withdrawal of mibefradil from the market in light of interactions with a range of other drugs [5]. Mibefradil is an inhibitor of the 2D6 and 3A4 isoforms of human CYPs. These isoforms are responsible for the metabolism of many drugs including antibiotics, antihistamines, and statins. In the latter case, these compounds which rely upon the 3A4 isoform for their elimination reached elevated levels when co-dosed with mibefradil and caused rhabdomyolysis.

This kind of drug–drug interaction can also be turned to advantage. Compounds that are not able to achieve optimal exposure due to high clearance by one of the CYPs might be improved by co-administrating with a CYP inhibitor. This general principle has been described in patents and has focused on the use of ritonavir, a 3A4 inhibitor [6]. It has been found in clinical practice that the HIV protease inhibitors saquinavir and indinavir achieve better plasma levels and efficacy when co-dosed with low doses of ritonavir [7].

By contrast, CYPs are also responsible for some metabolic transforms of natural substrates; some medicinal chemistry programs involve inhibiting these in order to modify physiological levels of substrates or product. For instance, the drug anastrozole inhibits CYP19 that is otherwise known as aromatase and catalyzes the transformation of androgens to estrogens. Some forms of breast cancer are

driven by estrogens such that reducing their production leads to an effective treatment for that condition. Optimization of compounds such as anastrozole that target one CYP is a significant challenge due to the large family of related enzymes which it is impractical to test against.

One mechanism by which compounds can affect the levels of CYP-mediated metabolism is to modulate the expressed levels of the CYPs or their associated cofactors. For instance, a number of compounds are known to be inducers of CYPs, including the herbal treatment St John's wort [8]. By contrast, some compounds, including many that are involved in the inflammatory cascade, are thought to reduce the levels of CYPs, and so levels of metabolism can be decreased in this state [9]. Similarly, there are significant differences in the levels and activity of many of the CYPs between the two genders, between different ethnic groups and between people with different disease states [2].

Publications in the field of medicinal chemistry have recently highlighted the increased likelihood of compounds safely traversing clinical trials if they have lower molecular weight and lipophilicity [10–13]. One simple way of reducing both of these properties without changing the shape of the molecule (and hence its likelihood of fitting in its target active site) is to transform aromatic carbon atoms into aza nitrogen atoms. This is generally beneficial but these aza nitrogen atoms can have a propensity to coordinate to metals in general and iron in particular and as such are more likely to be inhibitors of CYPs [14–16]. Curiously, reduction of lipophilicity is also an oft-cited means for reducing the CYP inhibition potential of compounds.

This chapter will focus on the approaches that have been adopted to avoid CYP inhibition without sacrificing the other good properties of a molecule. These will be compounds that are direct enzyme inhibitors and not modulators of the amount of enzyme that is present. The chapter will encompass reversible and irreversible inhibitors and compounds that bind to the iron of CYPs (type II) and those that do not (type I).

2   The Structure of the Cytochrome P450s

The structures of a number of the most important human isoforms of the CYPs have been solved in either apo- or holo-form. The structure of 3A4 illustrates many of the key features of this class of enzymes [17–19]. The complexes of the protein with two different inhibitors are shown in Fig. 1. The first is the complex with erythromycin, taken from the protein databank file with code 2J0D [17] This macrolide antibiotic can be accommodated in what appears to be a large and open active site. The first view in panel a shows that the added compound as well as the heme (carbons colored light blue) can be seen from the exterior, and a wide channel is the likely route taken by the erythromycin during complexation. The protein is color coded with lipophilic regions green and hydrophilic regions purple. The close-up of the region immediately above the heme in panel b shows that this region has significant hydrophilic character with a more lipophilic zone a little further back. The second complex (from the structure with code 2V0M) is that with ketoconazole [17]. This compound is known to show unusual kinetics of inhibition
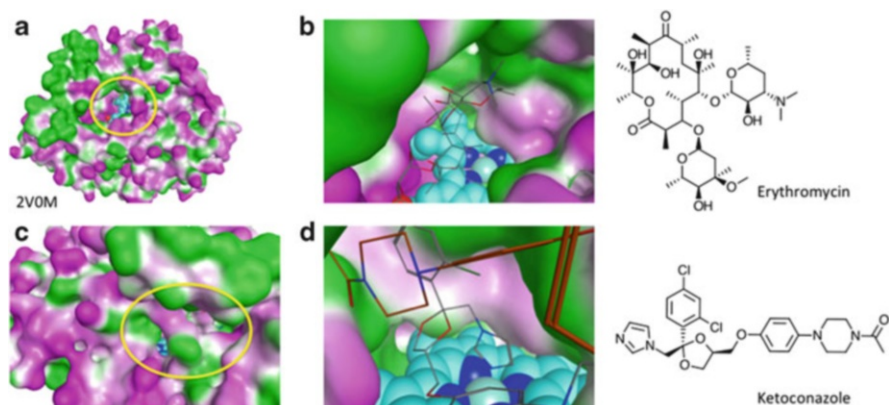
Fig. 1 Two views of the crystal structure of the complex between erythromycin (**a** and **b**) and CYP 3A4, and ketoconazole (**c** and **d**) and CYP 3A4. Panels **a** and **c** show a view from a distance from the protein with glimpses of the active site heme in blue. Panels **b** and **d** show a close-up of the binding site including the heme

of 3A4, and this may relate to the observation that two molecules are bound to the protein in the crystallographic structure. Panel c shows that the channel linking the binding site to the solvent is largely closed over in this complex. It is a key challenge to using the structures of these proteins in drug design that they are highly mobile. The 3A4 isoform may represent the most flexible CYP, and this is a likely contributor to this isoform being the most promiscuous in terms of having drugs as substrates and inhibitors. The zoomed in view of the binding site in panel d shows the two molecules bound, one remote from the heme with carbons colored brown and one close to the heme with carbons colored gray. This latter molecule presents an imidazole nitrogen to the iron atom and forms a covalent bond. This type of interaction is an important one for many inhibitors and will be discussed below. The effects of each molecule on the binding of the other within this complex have been explored with simulations that reveal that the increased shape complementarity of the binding site to the second molecule once the first is already bound leads to positive cooperative binding [20].

For the other most important isoforms, except for 2C19, there are structures available. The relevant PDB codes are given below:

2C9: 1R9O, 1OG2, and 1OG5 [21, 22]
1A2: 2HI4 [23]
2E1: 3T3Z, 3E6I, 3KOH, 3GPH, 3E4E, and 3LC4 [24–26]
2D6: 2F9Q, 3TDA, 3TBG, and 3QM4 [27, 28]
2C8: 1PQ2, 2NNI, 2NNJ, 2VN0, and 2NNH [29, 30]
2A6: 4EJJ, 3T3Q, 3T3R, 3EBS, 2PG5, 2PG6, 2PG7, 2FDU, 2FDV, 2FDW, 2FDY, 1Z10, and 1Z11 [19, 26, 31–33]

The heme present in the CYPs is a B type heme with structure shown in Fig. 2. The fifth ligand is a deprotonated active site cysteine. In the resting state of the
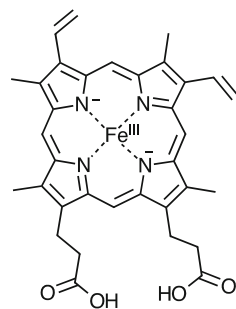
enzyme, the sixth site in the octahedral coordination sphere is occupied by water. The porphyrin has a significant effect on the types of ligands that are able to form direct interactions with the iron atom. The four nitrogens that interact with the iron are held by the covalent framework of the porphyrin in such a way that they are closer to the metal than they would be if they were not so linked. The consequence of this is that the coordination sphere of the iron is more sterically crowded than it would otherwise be. Hence, a number of ligand types that have the potential to bind to iron may not be able to do so in this particular environment [34]. Examples are discussed in the section on metal binding.

3   Overview of the Mechanism of Cytochrome P450s

A compound will be an inhibitor of CYPs if it disrupts any of the steps of the mechanism by which they operate. This mechanism is complex and remains the subject of research and much heated debate. The structures of several CYPs have been elucidated and reveal that the active site contains a porphyrin complex of iron, the so-called heme, that becomes the functional oxidant. The remainder of the active site is highly variable both in shape and in the polarity, hydrophobicity and charge type of the residues lining them. Hence, a broad range of substrates and inhibitors are known for this class of enzymes.

The most likely mechanism by which CYP-based metabolism of drugs takes place is shown in Fig. 3. The resting state of the enzyme **1** has iron in its Fe^{III} state. Initial binding of a substrate to give **2** prompts the first redox step leading to **3**; the details of the timing of these two events are not clear and are likely variable. A molecule of NADPH provides an electron to **2** via the cytochrome P450 reductase enzyme (other enzymes may also perform this feat) that mediates the electron transfer from NADPH to first an FAD cofactor and then an FMN cofactor which transfers the electron to the heme system [35, 36]. The two enzymes interact with one another in the context of membranes, usually in the endoplasmic reticulum and, hence, the structure of the interaction complex is challenging to establish. How such a complex might achieve its chemical outcome is therefore unclear.
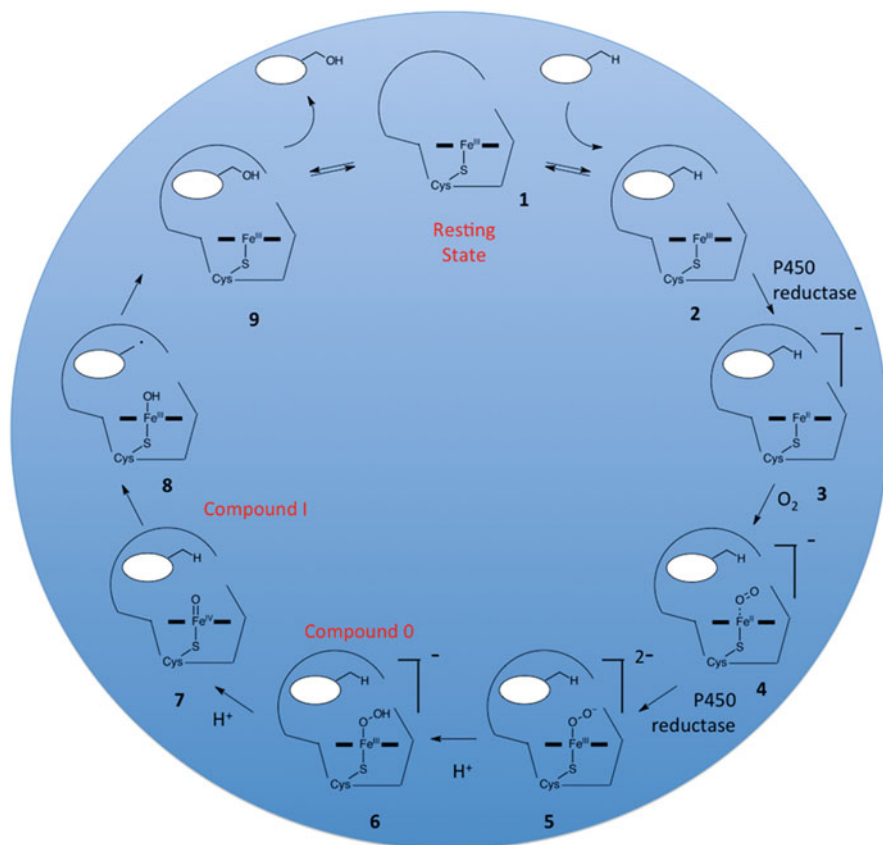
Fig. 3 The mechanism by which CYPs oxidize aliphatic C–H bonds

The reduced form of the enzyme, **3**, has a high affinity for oxygen which binds to create complex **4**. At this point, a second electron is transferred from NADPH through the intermediacy of the cytochrome P450 reductase which generates the superoxide complex **5**. Initial protonation of this complex creates the peroxide containing Compound 0 (**6**) which has been implicated as the active oxidant in some enzymatic transformations. Addition of a further proton prompts the concerted cleavage of the O–O bond generating water and Compound I (**7**) which is the oxidant most usually cited as being relevant to the observed transformation. The structure of this complex has evaded observation until recently [37, 38].

The mechanism by which Compound I (or any of the alternative postulated species) brings about the oxidation of various compounds is a subject of some debate and some ingenious experimentation; one of the more commonly accepted ones is illustrated in Fig. 3 [39, 40]. The oxygen atom of Compound I can abstract a hydrogen atom. The radical thus generated (**8**) rebounds and traps back onto the

oxygen, cleaving the Fe–O bond and generating an alcohol (**9**) in which the oxygen atom has been inserted into a C–H bond. In substrates where the oxidized position is adjacent to a heteroatom, the species generated (acetal, aminal, etc.) is unstable and can lead to cleavage of the carbon–heteroatom bond. Aryl or alkenyl oxidations tend to proceed through epoxidation and subsequent rearrangements [39]. The apparently unpredictable nature of the reaction that might be undergone by any given substrate can be rationalized in part by the observation that multiple spin states of the iron may be accessible at each stage of the catalytic cycle [41, 42]. The interplay of chemical reactivity with the orientation of the substrate in the binding site which will be governed by a complex set of molecular recognition elements makes predicting the site at which any given substrate will be metabolized and which reaction it will undergo extremely challenging [43].

Many assessments of inhibition of the CYPs are made by monitoring the decrease in the turnover of a standard substrate, usually selected to be easily detectable. Such measurements are not able to distinguish if a compound inhibits that reaction by being an inhibitor or by being a competitive substrate. Such distinction requires a more detailed study that is rarely performed; some compounds may be incorrectly classified as inhibitors although their ability to cause drug–drug interactions remains an issue.

4 Trends for Inhibition to be Modulated by Bulk Properties (Database Analysis)

Many reports describe attempts to model the inhibition of various CYP isoforms using computed or measured molecular descriptors. Many of these have been summarized by Refsgaard et al. who showed that the various studies are not consistent in terms of the descriptors that are found to be significant [44]. More recently, Gleeson has analyzed the inhibition of five of the CYP isoforms that are routinely measured within GSK (1A2, 2C9, 2C19, 2D6, and 3A4) in order to arrive at some "rules of thumb." This analysis appeared to show that molecular weight (MW), charge type and lipophilicity are key determinants of activity against some or all of these isoforms. This contradicts experience at AstraZeneca, and to probe some of this uncertainty, the full set of data concerning the inhibition of human cytochrome P450s curated by the EBI in the Chembl database has been analyzed.[1] Each compound in Chembl is annotated with its molecular weight and three different computed lipophilicity measures: AlogP, ACDlogP, and ACDlogD. This permits a comparison of the data in Chembl with that analyzed by Gleeson and enables some of the conclusions to be challenged.

---

[1] Data that are categorized as either IC_{50}, AC_{50} or potency and have molar units have been included. Properties are calculated on all compounds with molecular weight <1,000 Da.

4.1 Molecular Weight Dependence

In Fig. 4, the dependence of P450 inhibition upon molecular weight is shown in two ways. In the left hand column, the direct plot of pXC_{50} (all measurements transformed to a linear scale) against molecular weight shows that there is only a very weak link between these two parameters with the best correlation being for 3A4, which has an R^2 value of 0.03, which is to say that this linear dependence accounts for less than 5% of the variation in all cases. The coefficient linking molecular weight and pXC_{50} is always less than 0.0012, which is to say that variation of 1,000 in MW is required to bring about an order of magnitude variation in CYP inhibition. The same data is replotted in the right hand column in the format employed by Gleeson. The binning procedure into molecular weight sets for <300, 300–500, 500–700, and >700 Da shows that the data in Chembl apparently recapitulates the findings of Gleeson: 1A2 inhibition decreases as MW increases, 2C9 shows a parabolic dependence upon MW; 2C19 has little dependence upon MW; 2D6 has a weak but "clearly evident" parabolic dependence upon MW; and 3A4 inhibition potency increases with the MW of the inhibitor. The dataset suggests that such correlations are driven by the arbitrary binning of a continuous variable and by the different number of compounds populating each part of the molecular weight range. This is not to say that the conclusions are incorrect but that they are overstated by Gleeson. Based even on the presentation by Gleeson, a change in molecular weight of over 400 Da (i.e., from <300 to >700) would bring about a change of 0.8 log units in inhibition of 1A2, 0 for 2C9, 0.1 log units for 2C19 and 2D6, and 0.82 log units for 3A4. This magnitude of change in molecular weight is unlikely to be achieved without significantly altering the other properties of a molecule.

4.2 Charge State and Lipophilicity Dependence

The Chembl dataset can also be analyzed in such a way as to probe the variation with charge state and lipophilicity. In Fig. 5, the pXC_{50} values for inhibition of the same five isoforms are plotted against values of AlogP (almost identical plots arise for ACDlogP and ACDlogD also) [45]. For each isoform, the plots are shown for compounds identified as acids, bases, and neutral molecules. The zwitterion set was excluded because there are many less molecules in that group. As with the molecular weight analysis, the line of best fit is shown for each plot along with the R^2 value for that correlation. These show that the strongest linear dependence upon lipophilicity is for bases inhibiting 2D6 which has R^2 of 0.05, showing that like molecular weight, lipophilicity accounts for less than 5% of the variation in inhibition that is observed. The correlations are similar to those found by Leach and Kidley when studying the AstraZeneca dataset for the same five isoforms [34].
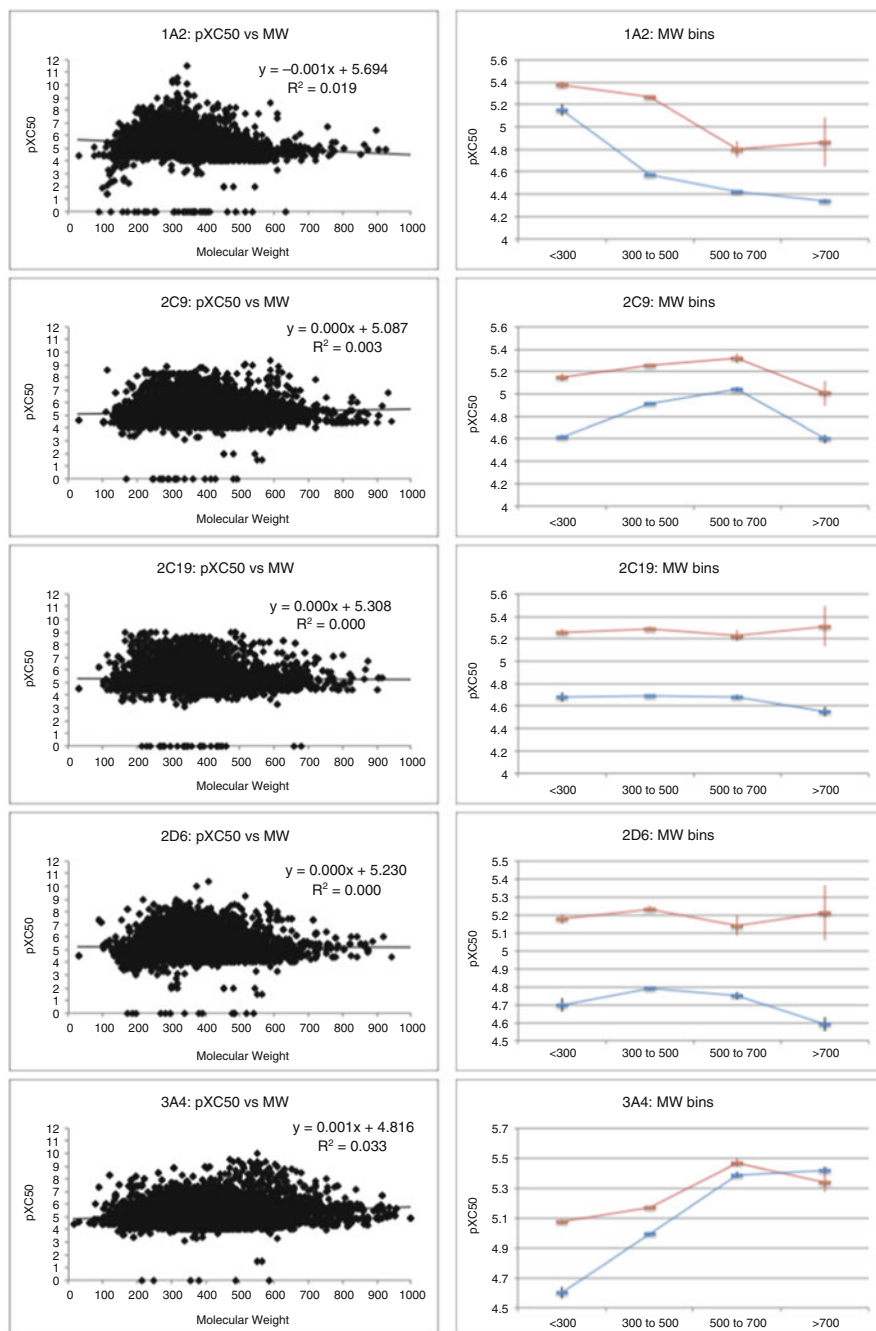
Fig. 4 The dependence of CYP inhibition upon molecular weight
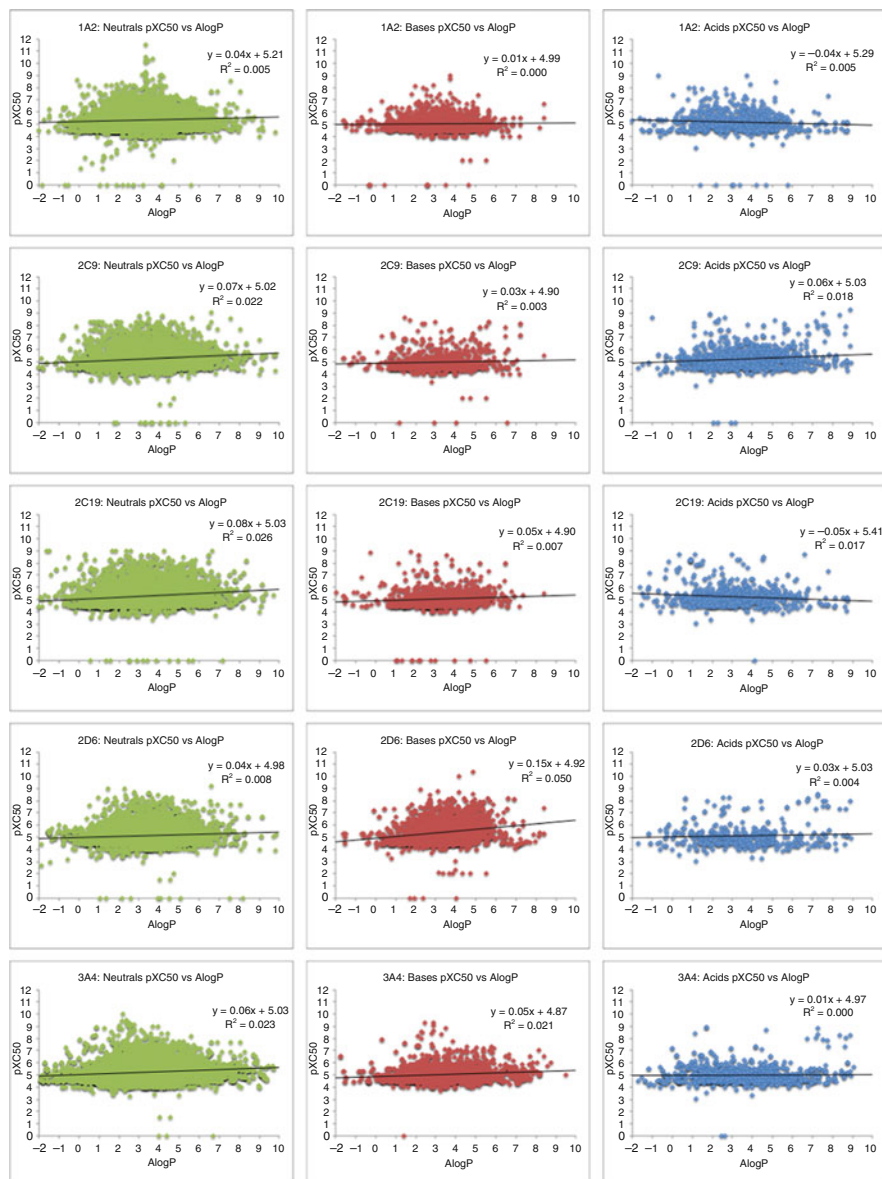
Fig. 5 Dependence of CYP inhibition upon lipophilicity for neutral molecules (*green*, left hand column), bases (*red*, central column), and acids (*blue*, right hand column)

The lines of best fit lead to the conclusion that when the lipophilicity increases by 4 log units going from AlogP = 1 to AlogP = 5, the average inhibition changes as indicated:

1A2: neutral molecules +0.2; bases 0.0; acids −0.1
2C9: neutral molecules +0.3; bases +0.1; acids +0.2
2C19: neutral molecules +0.3; bases +0.2; acids −0.3
2D6: neutral molecules +0.2; bases +0.6; acids +0.1
3A4: neutral molecules +0.2; bases +0.2; acids 0.0

The coefficients of the AlogP term for the line of best fit are generally less than 0.08 which implies that a change in lipophilicity of over 12 log units is on average required to change the pXC_{50} by an order of magnitude. It is clear that even large changes in lipophilicity only have a weak effect on inhibition of any of these isoforms; changes of the order of 4 log units are unlikely to leave other properties of a molecule unaffected.

There are a number of other isoforms for which data is provided in the Chembl database, and among these are five which do show correlations with R^2 higher than the highest found for the five isoforms described in detail above for at least one charge type and where there are at least 20 compounds contributing to the correlation. These are shown in Fig. 6. For four of the isoforms, it is neutral molecules that show a correlation with lipophilicity (1A1, 3A5, 4A11, and 4 F2) while one isoform shows a correlation for acids (19A1). Even among this set which show higher correlations, the highest only has R^2 of 0.27 (19A1 acids), and this statistic is clearly driven by clumping of the data as well as a genuine dependence upon lipophilicity. For three of the five sets, inhibition decreases as lipophilicity increases.

Once again, the data for the five isoforms selected by Gleeson can be replotted in the format employed by Gleeson, and this is shown in Fig. 7. Molecules are binned into groups according to their lipophilicity using the ranges <3 (green), 3 to 5 (orange), and >5 (red). The data from Chembl again resembles that from Gleeson when plotted this way. This suggests that the underlying datasets would have similar properties. The conclusions Gleeson drew are as follows: 1A2 shows little dependence on charge state but with neutral molecules most potent and the effect of lipophilicity varies with charge state; 2C9 shows neutral and acidic molecules are more potent and increasing lipophilicity increases potency; 2C19 shows little variation with charge state and potency increases with lipophilicity (the Chembl dataset differs on this last point for acids only); 2D6 is inhibited most strongly by bases and all charge states apart from acids show increasing inhibition with increasing lipophilicity; and 3A4 is most potently inhibited by neutral molecules, acids are weaker inhibitors, and increasing lipophilicity leads to increasing potency. Once again, the underlying data reveals that the binning of the data into groups exaggerates what is actually a rather weak link between potency of inhibition and charge state or lipophilicity. Further work from Ritchie and coworkers who, like Gleeson, studied the GSK dataset confirmed a weak dependence upon logD [46].

Focusing on the compounds with the most oral-drug-like properties (AlogP or clogP < 3) in Fig. 7 shows that within this group the effect of changing charge type is also weak. The position of acids differs between the Chembl dataset and the GSK dataset, but general conclusions are that:
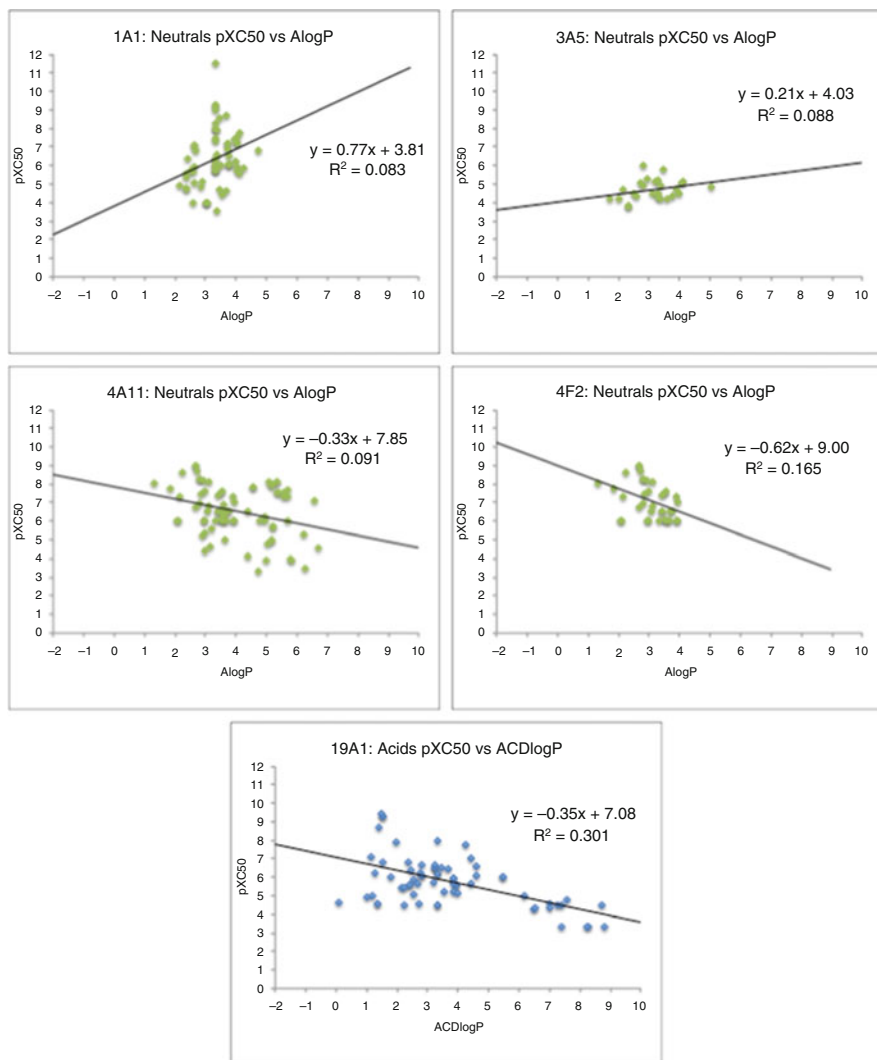
Fig. 6 Isoform and charge type combinations showing a stronger linear correlation than the examples in Fig. 5

1. for 1A2 and 3A4, neutral molecules are more potent than bases and acids;
2. for 2C9 and 2C19, neutral molecules are more potent than bases;
3. for 2D6, bases are more potent than acids or neutrals.

In all cases, the differences between charge types are less than 0.2 log units within this less lipophilic group. Even in the high lipophilicity group, the largest difference between charge types is between acids and neutrals in 3A4 where according to Gleeson there is a 0.8 log unit difference on average but only a 0.4 log unit difference according to the Chembl dataset.
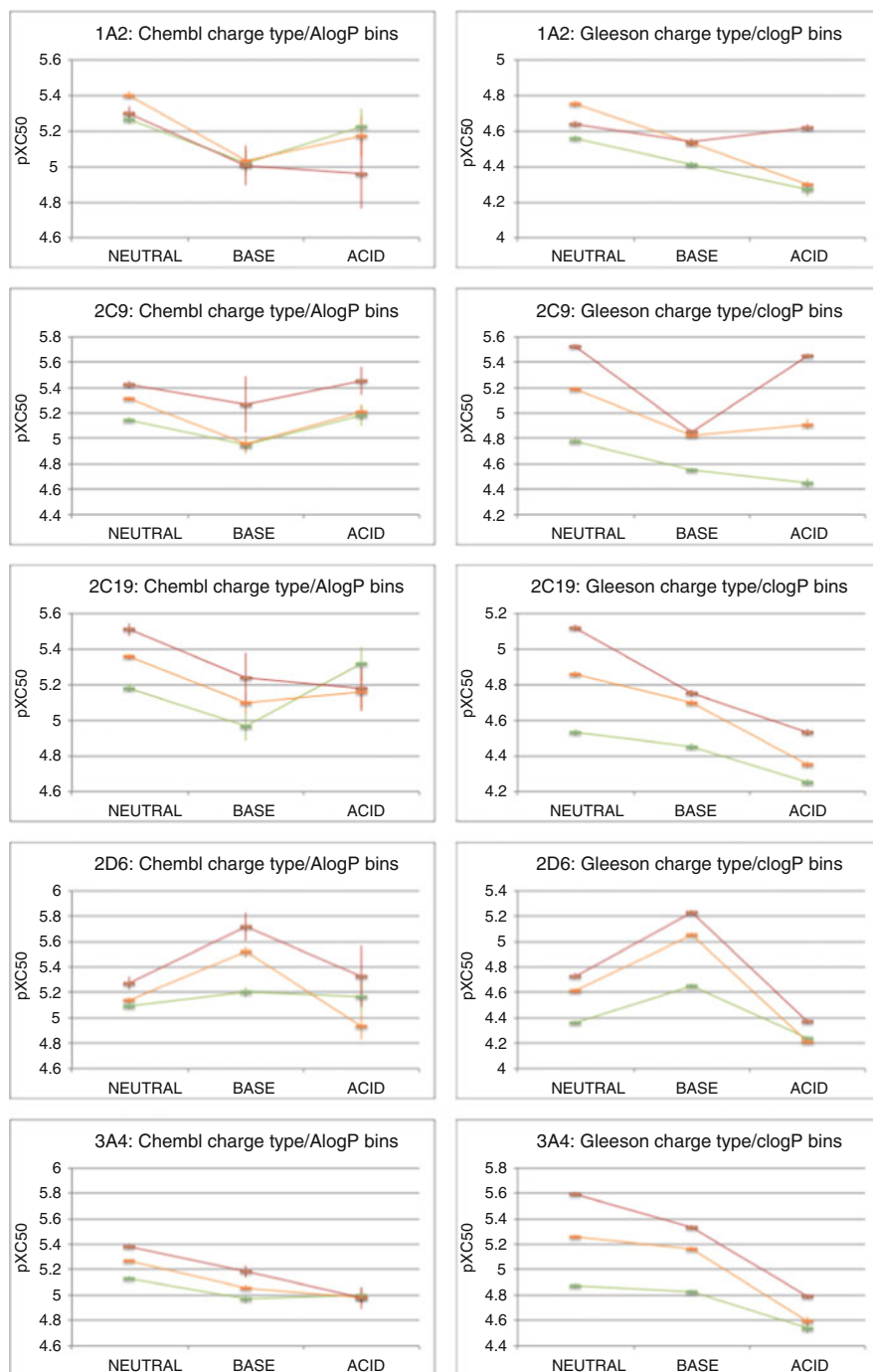
Fig. 7 Plots analogous to those plotted by Gleeson for Chembl data (*left hand column*) and Gleeson's own data (*right hand column*) [12]

4.3   Other Properties: Shape

The weak dependence upon any of molecular weight, lipophilicity, or charge state leaves the question of what property-based design can be undertaken. Some hints are provided by some other recent publications. Ritchie et al. who, like Gleeson [12], studied the GSK dataset found that inhibition of 3A4 increases with the number of aromatic rings present in the molecule [47]. This was developed further when the aromatic rings were subset by whether they are carbo- or heteroaromatic [48]. This subsetting showed that for 2C9, 2C19, 2D6, and 3A4, carboaromatic rings are more of a liability than heteroaromatic rings. For 1A2, the reverse trend was found with heteroaromatic rings more of a liability. Aliphatic ring types generally have a weaker influence, but heteroaliphatic rings are found to be beneficial in reducing inhibition of 1A2, 2C9, and 2C19. The benefit endowed by heteroaliphatic rings resonates with the observations of Lovering et al. concerning the shape of molecules. They showed that increasing the proportion of carbons in a molecule that are sp^3 or chiral can have a benefit [49].

In addition to reconfirming enhanced inhibition of 3A4 for bases and neutral molecules compared to acids, Chen et al. found the same dependence on lipophilicity as Gleeson but also showed that there is an additional variation caused by the fraction of the molecule that is included in the "molecular framework": the fraction of the heavy atoms that are included inside the part of the molecule bounded by rings. The higher this fraction, the more potent the inhibitors. This suggests that having more of the molecule as side chains rather than core is beneficial. A QSAR study of the inhibition of the usual five isoforms found that topological features such as aromatic count and fraction of sp^2 atoms had a significant effect on inhibition potency in a dataset of over 17,000 molecules [50]. The initial analysis of the same dataset found a number of substructural elements that are associated with reduced or enhanced inhibition (Fig. 8) [51]. Beneficial elements include hydroxyls, amines, carboxylic acids, imides, ureas, and oxolanes. By contrast, thiophenes, pyrimidines, indoles, benzodioxoles, and naphthalenes are associated with enhanced inhibition.

The impact of changing the shape of a molecule can be deduced in a very clean fashion by investigating the change in potency observed when a chiral compound is inverted. All the physical properties of a molecule remain identical, only the shape changes. Leach et al. reported that measured pIC_{50} differences between enantiomers have 95% confidence interval range of 0.68 log units for 1A2, 0.62 log units for 2C9, 0.72 log units for 2C19, 0.98 log units for 2D6, and 0.66 log units for 3A4. Even when experimental variation is factored out of these differences, 95% confidence interval ranges of 0.56, 0.44, 0.64, 0.78, and 0.23 log units respectively are found. These differences are larger than most of those predicted to be caused by substantial changes in molecular weight or lipophilicity [12]. When trying to avoid inhibition of CYPs, the shape of a molecule is a dominant factor. How to bring about a change in shape should be considered with higher priority than reducing molecular weight and lipophilicity, at least in so far as avoiding CYP

Advantageous in five usual isoforms

Advantageous in four (not 2D6) usual isoforms

Detrimental in five usual isoforms
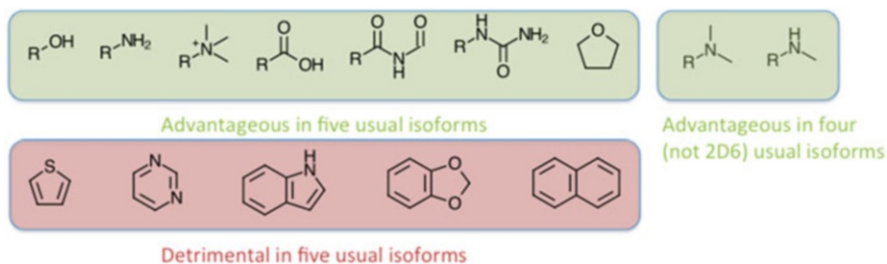
Fig. 8 Examples of substructures that are found to reduce (in *green*) or increase (in *red*) inhibition of CYPs according to Veith et al. [51]

inhibition is concerned; the influence of those properties on the rest of a molecule's profile is beyond doubt [10–12, 52]. A fuller understanding of the set of interactions in three dimensions that cause binding between inhibitors and the CYPs should aid in selecting how to alter the molecules shape to best effect. Crystal structures of the protein–ligand complex should aid this process.

4.3.1 Strategic Approach

Attempting to reduce lipophilicity and molecular weight is generally beneficial to a molecule's pharmaceutical properties. The influence of lipophilicity and molecular weight on CYP inhibition is weak. Changes to the shape of a molecule are more likely to have a substantial effect. In particular, the removal of flat, aromatic regions should be considered, and the introduction of chirality or twisting could be attempted. Obtaining a crystal structure of the complex of a representative of the series with the relevant CYP could be considered in particularly serious cases.

5 Metal Chelation

The active site of the CYPs contains an iron atom encircled by a porphyrin-based heme and linked to the protein through a deprotonated cysteine atom. As shown in Fig. 3, the iron atom samples three different oxidation states during the catalytic cycle although only two of these are in coordination states that have one apex of the octahedral arrangement either vacant or with an exchangeable (water) ligand present. The Fe^{II} and Fe^{III} oxidation states each have three spin states available (singlet, triplet, and quintet for Fe^{II} and doublet, quartet, and sextet for Fe^{III}) [34, 39, 41, 53]. Each of these spin states will have a different propensity to form interactions with particular ligand types. This variability may be one further aspect of the CYPs that allow such a broad range of chemotypes to cause inhibition.

The spectroscopic characterization of iron–porphyrin complexes is long established and, indeed, it is one of the characteristic spectroscopic features of
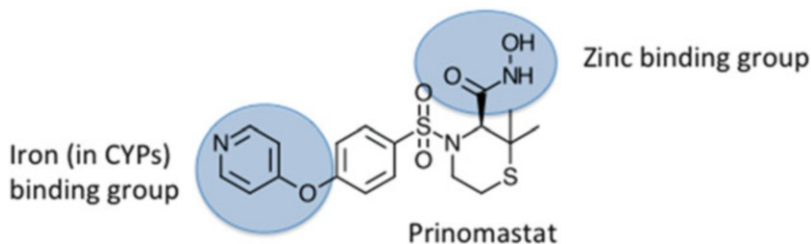
Fig. 9 The two metal binding groups in Prinomastat are highlighted

such systems that leads to the P450 label of these enzymes [54–56]. Changes in the nature of the coordination of the iron atom perturb the populations of the various spin states and lead to distinctive spectroscopic changes. Hence, inhibitors that bind directly to iron can be distinguished spectroscopically from those that do not. Inhibitors that have their effect without binding to iron are styled type I and those that do bind type II [57]. It is to be expected that a direct bond to iron should lead to enhanced inhibition although, as with all interaction types, this will depend upon whether the metal–ligand bond can offset the desolvation penalty for both the metal and the ligand [58].

As previously mentioned, the steric environment around the iron atom is particularly crowded and rigid and, hence, it is difficult to envisage anything other than a mono-dentate binding mode. Thus, some inhibitors that employ groups deliberately chosen to bind to metals by exploiting the entropic advantage that can arise through bidentate (or higher oligodentate) interactions may not bind to the CYPs using the metal binding group. For instance, the matrix metalloproteinase inhibitor Prinomastat (Fig. 9) includes a hydroxamate, which is a well-known zinc binding group. The crystal structure of this compound in complex with CYP 2D6 reveals that rather than using this bidentate metal binding group to interact with the iron, the alternative pyridyl nitrogen is employed [28].

It is interactions with aromatic nitrogen atoms like that in Prinomastat that are the predominant feature of inhibitors that exploit binding to iron in CYP inhibition. It is well understood now that reducing the molecular weight and lipophilicity of compounds can improve many of their pharmaceutical properties and their likelihood of success in the clinic [10–12, 52, 59, 60]. One of the means by which one might think to reduce lipophilicity is to introduce polar groups, which inevitably increases molecular weight. Alternatively, the removal of lipophilic groups might be considered, albeit with a risk of reduced potency. One highly attractive option is to transform aromatic carbon atoms to aromatic nitrogen, which lowers lipophilicity with only a tiny increase in molecular weight. It is a common procedure during medicinal chemistry programs to attempt the introduction of one or more nitrogen atoms into a lead series with the aim of finding a location in which this feature does not substantially diminish the potency of the series [61]. One of the few negative consequences of introducing such a nitrogen is that it may enhance inhibition of CYPs by binding to the iron.
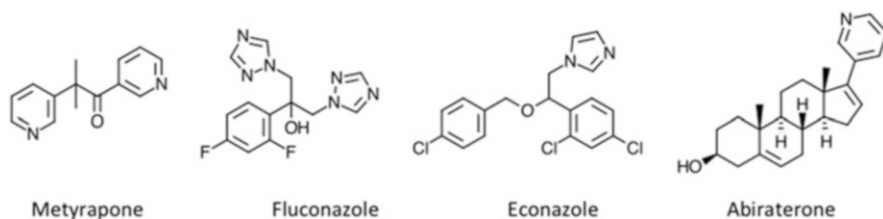
Fig. 10 Examples of compounds designed to include metal binding elements

Binding to iron by inhibitors containing aromatic nitrogen atoms has been long known [62], with this feature being deliberately present in a number of marketed compounds which have CYPs as their target, e.g., metyrapone, fluconazole, econazole, and abiraterone (Fig. 10) [63–66].

The extra inhibition caused by compounds containing pyridines, imidazoles, and triazoles compared to other compounds with identical lipophilicities was quantified by Riley et al. [67]. By comparing a dataset containing 11 compounds that each had one or other of those substructures and a dataset containing 19 compounds that did not, potency of inhibition of 3A4 is enhanced by 1.3 log units (~20 fold) at the same measured $logD_{7.4}$ for compounds containing one of the metal binding groups. Such an approach was extended further by Leach and Kidley, who studied all five of the usual isoforms in the AstraZeneca database and Ritchie et al. who examined the GSK database [34, 46].

Leach and Kidley studied 14 different heteroaromatic rings and found that even when weighted for the differences in their lipophilicity, compounds containing certain heteroaromatics were significantly more likely to be potent inhibitors of CYPs than compounds that did not contain them. In particular, imidazole, oxazole, isoquinoline, pyridine, and pyrimidine were found to be more potent than expected across all isoforms. These are shown in Fig. 11 in which the atom likely to form a bond to iron is indicated in red. When examining a similar set of heteroaromatics, Ritchie et al. focused on inhibition of 2C9, 2C19, and 3A4 which they had previously found to be more sensitive to the presence of heteroaromatic and heteroaliphatic rings [48]. Leach and Kidley observed the same enhanced sensitivity for 2C9, 2C19, and 3A4 compared to 1A2 and 2D6 and attribute this to differences in the degree of occlusion of the iron atom in the various observed protein crystal structures. Ritchie et al. did not seek out ring types that can bind to iron but rather those that showed enhanced inhibition. They found that thiophene, tetrazole, 1,2,3-triazole, isoxazole, pyridine, imidazole, and thiazole show a degree of enhanced inhibition of P450. As indicated in Fig. 11, most of these also have a nitrogen atom capable of forming an interaction with iron. The two studies contrast in whether they constrained the atoms adjacent to the indicated atom to be unsubstituted (Leach and Kidley) or not (Ritchie et al.). This difference provides some of the rationale for the differences observed. The effect of fluorination or methylation adjacent to a pyridyl nitrogen was studied by Leach and Kidley who found that both groups effectively eliminate the ability to
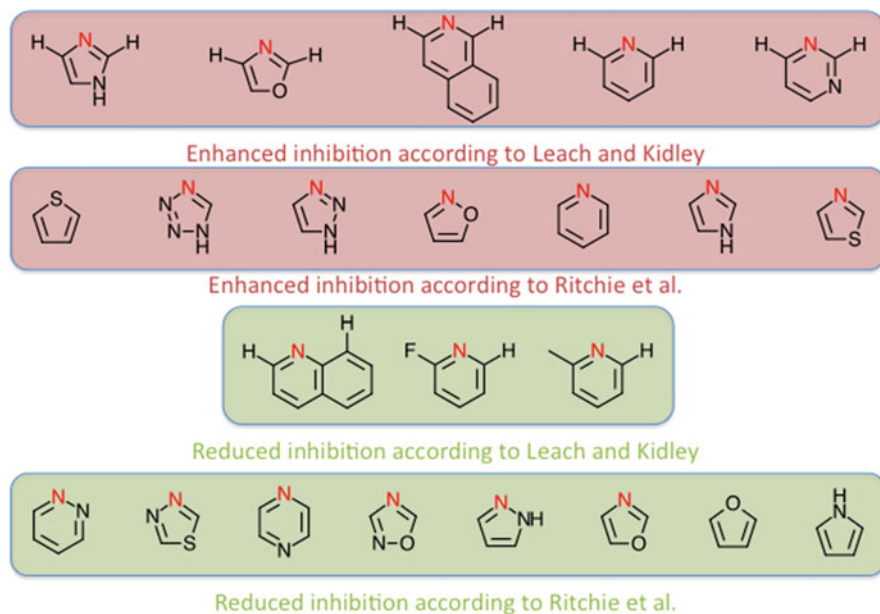
Fig. 11 Sets of heteroaromatic rings found to increase (in *red*) or decrease (in *green*) inhibition of CYPs by Leach and Kidley or Ritchie et al. [34, 46]

complex the heme iron [34]. This steric effect does not diminish the nitrogen's ability to form hydrogen bonds with water to the same extent such that pyridyls with ortho-binding groups are less likely to be inhibitors than general compounds of the same lipophilicity. Both the AstraZeneca and GSK studies also found heteroaromatic rings that contain a potential iron binding group that are actually less likely to be inhibitors than other compounds of the same lipophilicity, and these are also indicated in Fig. 11. The studies agree on most heterocycles but diverge on the conclusions concerning oxazole and tetrazole. Such differences can be a useful prompt for discussion [68].

Leach and Kidley also performed quantum mechanical calculations in order to be able to make predictions about heteroaromatic rings that are not currently well represented in the dataset and to provide further understanding of the binding event between iron and nitrogen. The geometries of the complexes of imidazoles and pyridines were compared to those observed in crystal structures of the complexes of inhibitors and CYPs. These showed that the computed geometry for complexes to iron as Fe^{III} in its doublet state was in best agreement and that rotation about the Fe–N axis has essentially no barrier making this a very flexible interaction type. Computed gas phase complexation energies alone were not able to reproduce trends in binding energies. However, differences between the complexation energy to iron and hydrogen bonding to water were able to correct this. It should be noted that the case that this provided the largest correction to was oxazole which the results of Ritchie et al. suggest may not be as potent an inhibitor as found by Leach and

Kidley. Having a quantum mechanical procedure enabled predictions to be made for heteroaromatic rings that have not yet been reported but were suggested by Pitt et al. [69].

Quantum mechanical studies had also been performed by Balding et al. who examined potential binding modes of imidazole, 1,2,4-triazole, and pyridine containing compounds [53]. Rather than simply studying isolated molecules or complexes, they studied the process that happens when the aza nitrogen approaches the complex of Fe^{III} porphyrin with water bound. They found that during the initial approach, a hydrogen bond is formed between the nitrogen atom and the iron bound water. A structure in which this interaction is present had been reported for fluconazole with a CYP from *Mycobacterium tuberculosis* [70]. The process of displacement requires the water to leave, forming a pentacoordinate iron atom, before the nitrogen atom is able to approach and form a bond to iron. They found that during this dissociation, it is possible that the spin states approach one another closely in energy and a spin flip is possible (from the ground state doublet of the hexacoordinate complex) unless the nitrogen ligand is already proximal as it retains the doublet ground state.

5.1 Strategic Approach

A number of approaches can be taken to reduce CYP inhibition for compounds in which iron binding is thought likely. If the heteroaromatic is one of those mentioned as detrimental, the introduction of substituents (even just F or Me) adjacent to the likely iron binding nitrogen atom or exchanging the heteroaromatic for one of the alternatives identified as beneficial could be considered. Alternatively, if a number of transformed examples are possible but their iron binding liability is not known, they could be subject to quantum mechanical calculations.

6 Reactive Metabolites and Mechanism-Based Inhibition

Thus far, the focus has been on compounds that inhibit CYPs in a reversible way. It is in the nature of the CYPs that they transform some compounds (usually by oxidation) as described above such that the resulting compound is reactive. If this metabolite reacts with the CYP that generated it, then this is a mechanism-based inhibitor (MBI). MBIs were discussed in great detail by Orr et al., but earlier surveys also highlight similar points [71–74]. CYP MBIs carry even more safety risks than reversible inhibitors. By irreversibly preventing their own metabolism, they can reach elevated, dangerous levels upon repeat dosing. The effect of their inhibition does not necessarily reduce once treatment is eliminated, and so prolonged effects can ensue; the elimination of the effect depends upon the turnover of the CYP. Covalently modified proteins can be immunogenic, and antibodies can be raised to the complex of MBIs with CYPs leading to dangerous immune

Fig. 12 Substructures reported to occur in MBIs of CYPs and which generate metabolites that react with the heme [71]

responses. Detection of MBIs of CYPs requires assays to be run at different time points to monitor the decrease in activity with time and requires the presence of NADPH and the requisite machinery for turnover of the enzyme.

There are particular substructures that have been associated with MBI, and these are shown in Figs. 12, 13, and 14 along with the presumed metabolite that causes the irreversible inhibition [71]. It is often the case that the substrate that becomes the reactive species is itself the product of metabolism, notably phenolic hydroxyls which originate from oxidation of the aromatic ring and primary amines resulting from dealkylation of secondary and tertiary amines. There are three parts of the CYP that can form irreversible interactions: the iron, the heme, and the protein. Each of these is implicated as the site of reaction of some of the metabolites indicated in Figs. 12, 13, and 14. These figures are derived from the review of Orr et al. and show conceptual ideas of what the liability of each substructure might be although the details will vary from molecule to molecule and between CYP isoforms [71]. Examples of each based on a range of drugs and medicinal chemistry programs are also provided by Orr et al. [71].

The heme including its nitrogens is not particularly reactive, but the generation of highly reactive species in close proximity to it prompts this otherwise inert group to engage in reaction. Figure 12 shows some substructures that have been found to

Fig. 13 Substructures reported to occur in MBIs of CYPs and which generate metabolites that react with the protein [71]

generate species that react with the heme group. These tend to involve highly reactive electrophiles or radicals. Compounds that act in this way are detectable because of the changes caused to the UV/vis spectra of the system. By contrast, the reactions that modify the protein need not significantly alter these spectra. In Fig. 13, a number of the substructures that generate species that react with the protein are shown. These tend to be less reactive electrophiles. The final class of MBIs usually identified are those that lead to very tight binding to the iron atom shown in Fig. 14. In particular, nitroso compounds and some carbenes are able to bind very tightly to iron in its Fe^{II} state. These tend to bind so tightly that they cannot be displaced by CO or O_2 and hence halt the catalytic cycle in its early stages. These complexes provide a very distinctive UV/vis spectrum because the iron is in its Fe^{II} state [56].

Fig. 14 Substructures reported to occur in MBIs of CYPs and which generate metabolites that form effectively irreversible interactions with the iron atom [71]
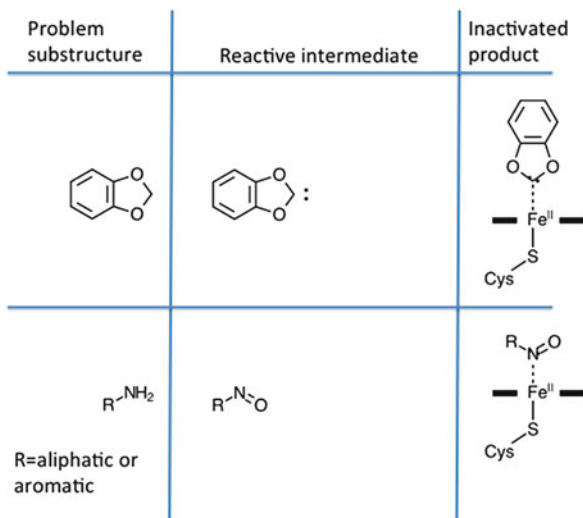
A number of attempts to eliminate or decrease MBI have been discussed in the literature. Two broad strategies are employed. The first is to reduce the metabolism that generates the reactive intermediate. This could be done by lowering the lipophilicity thereby reducing the rate of the transformation to the reactive intermediate. Whereas the link between lipophilicity and CYP inhibition is not very strong, the link to the degree of metabolism is tighter [12, 52]. In this approach, the problematic substructure in the molecule may be retained, but the degree to which it generates any reactive species is attenuated. An example of this is shown in Fig. 15 taken from the STAT6 inhibitor program at Astellas [75]. The morpholine **10** shows clear MBI of 3A4, but when the morpholine is transformed to the N-hydroxyethyl group in **11**, the lipophilicity is reduced by over a log unit and TDI could not be detected. Alternatively, the metabolism could be redirected elsewhere in the molecule. For instance, the PI3-kinase δ inhibitor **12** (Fig. 16) is hypothesized to undergo activation of the indole moiety through epoxidation and formation of a p-quinoneimine and consequent reaction with the protein of 3A4 [76]. No metabolism is detected at the piperazine, but by transforming this to the fused morpholine–piperazine in **13**, metabolism is directed to this moiety and TDI significantly reduced. This strategy presumably benefits from the reduced lipophilicity of the metabolites of **13** which are therefore less likely to undergo metabolism to form the reactive intermediate. A final approach in this category is to block the metabolism that leads to the reactive group. This is illustrated by potassium channel KCNQ2 opener **14** (Fig. 17) which is activated via initial oxidation to **15** then quinoneiminium formation [73]. This can be blocked by introduction of a fluorine atom at the site that is oxidized, as in **16** [77].

The second strategy is to remove the substructure that is responsible for generating the reactive intermediate or to modify it in such a way that it can no
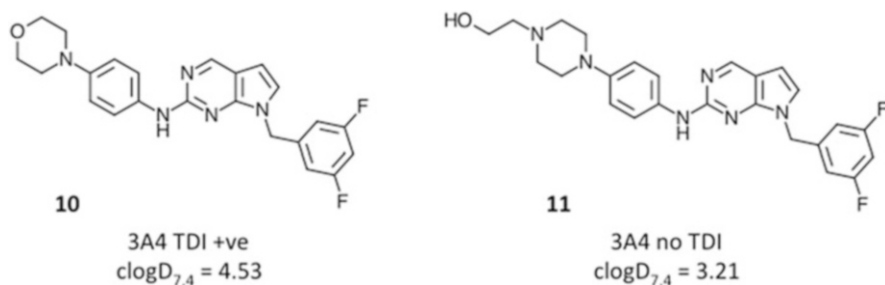
Fig. 15 An example in which mechanism-based inhibition is reduced by decreasing the rate of metabolism by modulating the lipophilicity of the molecule
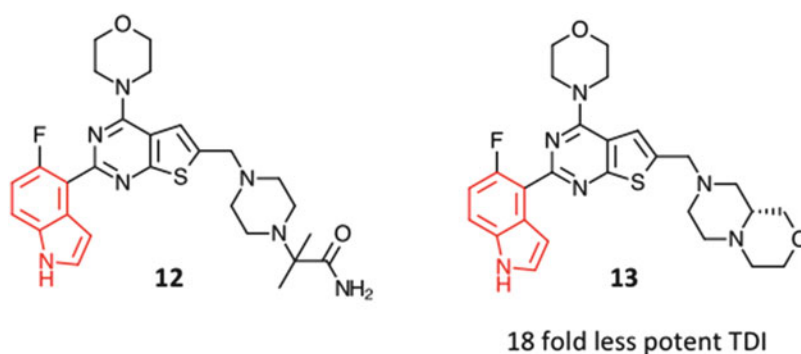
Fig. 16 An example in which metabolism is redirected away from the substructure that causes mechanism-based inhibition (the indole highlighted in *red*)
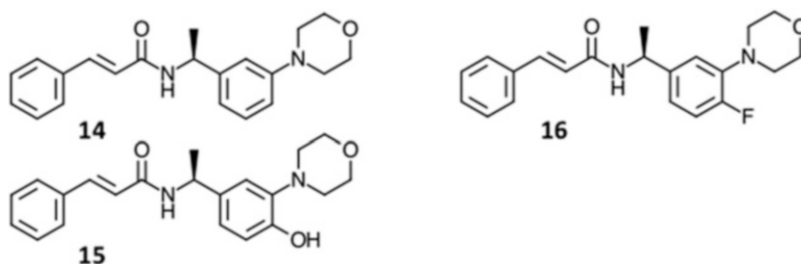
Fig. 17 An example in which metabolism that causes mechanism-based inhibition is blocked

longer undergo the requisite activation. An example of this is taken from the optimization of some PYK2 inhibitors [78]. Compound **17** containing the *p*-diaminophenyl group highlighted in red can generate p-diimine species analogous to a *p*-quinone. This in turn can react with the 3A4 protein. By simply inverting the amide, the substructure that causes the liability is removed, and the PYK2 activity is retained (Fig. 18).
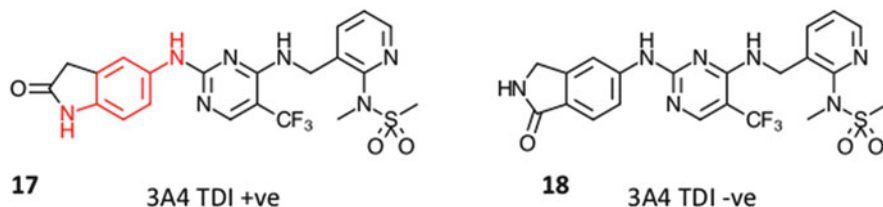
Fig. 18 An example in which the substructure responsible for mechanism-based inhibition is removed
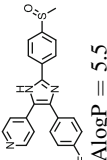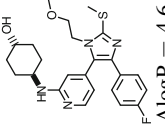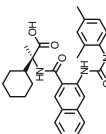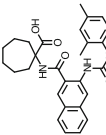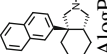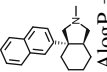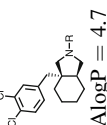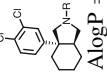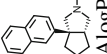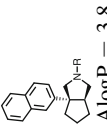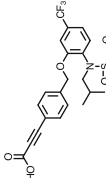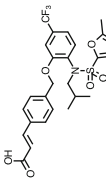
6.1 Strategic Approach

Reduce the rate of the metabolism that generates the reactive intermediate by reducing lipophilicity, introducing blocking groups, or redirecting the metabolism elsewhere. If possible, eliminate the substructure that is transformed into the reactive intermediate.

7 Examples of Successful Reduction of P450 Liability
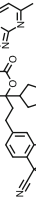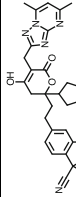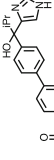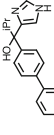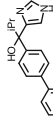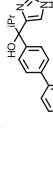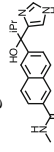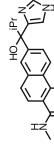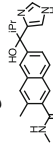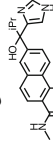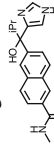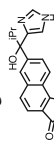
There are a number of literature surveys of structural changes that have effected beneficial changes to molecules. Among these, the survey of bioisosteres of Meanwell is a good recent example and includes a number of structural changes that reduce CYP inhibition [79]. For examples of how to modulate MBI, the review of Orr is a very comprehensive survey [71]. This section complements the earlier publications by providing a selection of interesting examples abstracted from the literature. These show significant changes in CYP inhibition against one or several isoforms and are grouped by the present author according to the factor that is likely to be most significant in causing the change. The first tables [1–3] illustrate improvements in CYP inhibition caused by changes in shape, those caused by changes in lipophilicity and those caused by a combination of these two factors. Table 4 includes improvements caused by changing or removing a metal binding group, some of these examples also exploit a change in shape. Table 5 contains a small selection of examples where changing the charge type of a molecule has a large influence on CYP inhibition. These include a small number of repeats from earlier tables. Table 6 contains a range of changes in reversible CYP inhibition that could not readily be rationalized. Table 7 contains some examples of variations in TDI caused by either reducing metabolism or by eliminating the metabolic group. Wherever possible a computed lipophilicity (AlogP) is provided to enable the reader to judge the degree to which lipophilicity is responsible for any change. The isoform being varied is indicated in each case. Where potency is assessed by full IC_{50} measurement, these values are quoted (usually in μM units). Where only a single concentration has been evaluated, a percentage inhibition value is quoted along with the concentration at which it is measured.
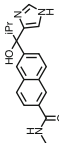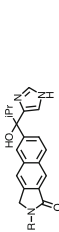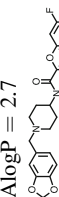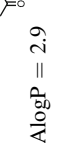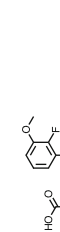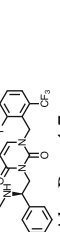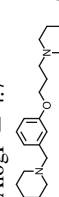
Table 1 Modulation of reversible CYP inhibition dominated by shape changes
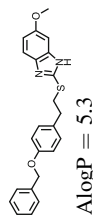
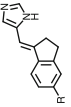| Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 5.5	1A2: 38% 2C9: 87% 2C19: 92% 2D6: 73% 3A4: 77% At 10 µM	AlogP = 4.6	1A2: 23% 2C9: 40% 2C19: 58% 2D6: 24% 3A4: 80% At 10 µM	Large substituent Laufer [16]
2	AlogP = 6.1	2C9: 9.2 µM	AlogP = 5.9	2C9: >33 µM	Change in topology Cyclohexyl/cyclopentyl/ cyclopropyl/gem-di-n-propyl also >33 Sparks [80]
3	AlogP = 4.2	2D6: 0.6 µM	AlogP = 4.2	2D6: 13.6 µM	Enantiomers Shao [81]
4	AlogP = 4.7	R = H 2D6: <0.05 µM R = Me 2D6: <0.05 µM	AlogP = 4.4	R = H 2D6: 7.9 µM R = Me 2D6: 1.1 µM	Change in topology Shao [81]
5	AlogP = 3.8	R = H 1A2: 4.1 µM R = H 2C9: >25 µM R = H 2C19: >25 µM R = H 2D6: 9.4 µM R = H 3A4: 2.7 µM R = Me 2D6: 2.6 µM	AlogP = 3.8	R = H 1A2: 0.7 µM R = H 2C9: 14.4 µM R = H 2C19: 7.2 µM R = H 2D6: 10.8 µM R = H 3A4: 10.7 µM R = Me 2D6: 12.8 µM	Enantiomers Shao [81]
6	AlogP = 6.4	1A2: 12% 2C9: 47% 2C19: 19% 2D6: 6% 3A4: 42% At 3 µM	AlogP = 6.4	1A2: 2% 2C9: 9% 2C19: −4% 2D6: −3% 3A4: 11% At 3 µM	Alkyne–alkene Naganawa [82]

(continued)
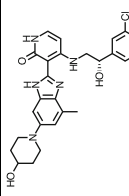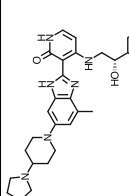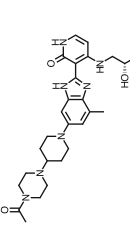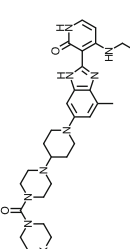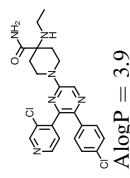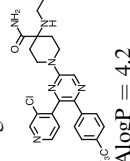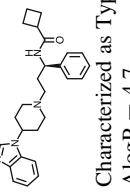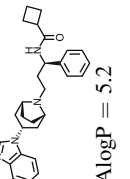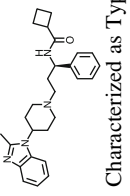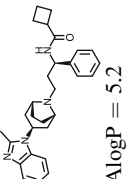
Table 1 (continued)

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
7	AlogP = 5.1	2D6: 0.3 μM	AlogP = 5.9	2D6: >6 μM	Large substituent Li [83]
8	AlogP = 3.6	3A4: 1.0 μM	AlogP = 3.6	3A4: 7.0 μM	Para–meta Kaku [84]
9	AlogP = 2.7	R = H 3A4: 1.2 μM R = OMe 3A4: 2.0 μM R = F 3A4: 2.1 μM R = Cl 3A4: 2.4 μM	AlogP = 3.4 (Cl)	R = H 3A4: 3.7 μM R = OMe 3A4: 4.9 μM R = F 3A4: 5.3 μM R = Cl 3A4: 7.6 μM	Meta–para Kaku [84]
10	AlogP = 3.0	3A4: 3.6 μM	AlogP = 3.0	R = Cl 3A4: >10.0 μM R = Me 3A4: >10.0 μM	Twist induced by ortho-substituent Kaku [85]
11	AlogP = 3.0	3A4: 3.6 μM	AlogP = 3.0	3A4: >10.0 μM	Regioisomeric: methyl moved around ring Kaku [85]
12	AlogP = 2.7	3A4: 3.6 μM	AlogP = 2.8 (Me)	R = H 3A4: 6.7 μM R = Me 3A4: >10.0 μM	Cyclization Kaku [85]

No.	Structure (left)	3A4/2D6 (left)	Structure (right)	3A4/2D6 (right)	Note
13	AlogP = 2.7	3A4: 3.6 μM	AlogP = 2.8 (Me)	R = H 3A4: >10.0 μM; R = Me 3A4: 5.4 μM	Cyclization Kaku [85]
14	AlogP = 2.9	3A4: 23.5 μM	AlogP = 3.3 (O); AlogP = 3.3 (NMe)	X = O 3A4: 62.1 μM; X = NMe 3A4: 60.0 μM	Iyengar [86]
15	AlogP = 4.7	3A4: 36.0 μM	AlogP = 5.5	3A4: 56.0 μM	Large substituent Chen [87]
16	AlogP = 4.0	2D6: 3.0 μM	AlogP = 4.0	2D6: 14.0 μM	Meta–para isomers Berlin [88]
17	AlogP = 4.8	2D6: 0.3 μM	AlogP = 4.8	2D6: 2.0 μM	Para–meta isomers Berlin [88]
18	AlogP = 4.0	3A4: 0.5 μM	AlogP = 4.6	3A4: 3.1 μM	Ring breaking Zimmermann [89]

(continued)

Table 1 (continued)

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
19	AlogP = 3.5	3A4: 0.4 μM	AlogP = 4.6	3A4: 4.5 μM	Morpholine in place of pyrrolidine equally effective Zimmermann [89]
20	AlogP = 3.6	3A4: 0.8 μM	AlogP = 3.5	X = O 3A4: 19.0 μM X = NMe 3A4: 40.0 μM	Change in topology Zimmermann [89]
21	AlogP = 3.9	3A4: 54% At 5 μM	AlogP = 4.2	3A4: <5% At 5 μM	Wustrow [90]
22	Characterized as Type I AlogP = 4.7	2D6: Active but unspecified	AlogP = 5.2	2D6: Inactive	Change in topology Rationalized as steric blocking of pharmacophore recognition element (basic nitrogen) Wood [91]
23	Characterized as Type I AlogP = 4.7	2D6: Active but unspecified	AlogP = 5.2	2D6: Inactive	Change in topology Rationalized as steric blocking of pharmacophore recognition element (basic nitrogen) Wood [91]
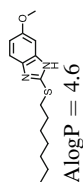
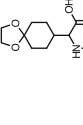| #	Structure 1	Data 1	Structure 2	Data 2	Comment
24	AlogP = 4.6	1A2: 0.4 µM 2B6: 15.4 µM 2C8: 2.4 µM 2C9: 2.9 µM 2C19: 6.3 µM 2D6: 2.6 µM 2E1: 50.0 µM	AlogP = 5.3	1A2: 50.0 µM 2B6: 50.0 µM 2C8: 0.7 µM 2C9: 0.9 µM 2C19: 50.0 µM 2D6: 0.1 µM 2E1: 50.0 µM	Large substituent Foti [92]
25	AlogP = 5.6	1A2: 0.7 µM 2B6: 1.1 µM 2C8: 0.3 µM 2C9: 0.8 µM 2C19: 2.2 µM 2D6: 0.02 µM 2E1: 50.0 µM	AlogP = 5.7	1A2: 0.6 µM 2B6: 2.0 µM 2C8: 50.0 µM 2C9: 1.0 µM 2C19: 0.07 µM 2D6: 1.5 µM 2E1: 50.0 µM	Change in topology Foti [92]
26	AlogP = 5.5 (H)	R = H 11B1: 3.3 nM R = H 11B2: 9.6 nM R = H CYP17: 13% R = H CYP19: 190 nM R = CN CYP17: 3% R = CN CYP19: 810 nM		R = H 11B1: 31.4 nM R = H 11B2: 24.8 nM R = H CYP17: 13% R = H CYP19: 226 nM R = CN CYP17: 2% R = CN CYP19: 970 nM	Double bond isomers Ulmschneider [93]
27	AlogP = 5.5	1A2: 13% 2B6: 8% 2C9: 96% 2C19: 74% 2D6: <5% 3A4: 60% At 1 µM	AlogP = 5.5	1A2: 23% 2B6: <5% 2C9: 51% 2C19: <5% 2D6: <5% 3A4: 30% At 1 µM	Large substituent Lucas [94]
28	AlogP = 5.5 (H)	R = H CYP19: 0.071 µM R = NO$_2$ CYP19: 0.045 µM R = Br CYP19: 0.44 µM R = CN CYP19: 0.069 µM	AlogP = 5.5 (H)	R = H CYP19: 0.55 µM R = NO$_2$ CYP19: 0.47 µM R = Br CYP19: 4.1 µM R = CN CYP19: 1.8 µM	Large substituent Gobbi [95]
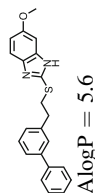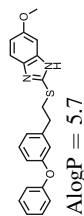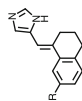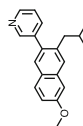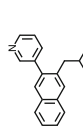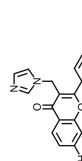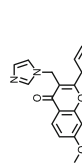
(continued)

Table 1 (continued)

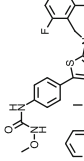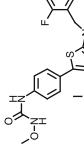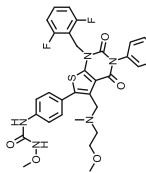| Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
29		R = H 11B1: 6.1 nM R = H 11B2: 11.0 nM R = H CYP17: 1% R = H CYP19: 130 nM R = CN 11B1: 12.3 nM R = CN 11B2: 35.7 nM R = CN CYP17: 4% R = CN CYP19: 15 nM R = F 11B1: 11.2 nM R = F 11B2: 13.9 nM R = F CYP17: 11% R = F CYP19: 20 nM R = Cl 11B1: 19.5 nM R = Cl 11B2: 3.7 nM R = Cl CYP17: 26% R = Cl CYP19: 39 nM R = Br 11B1: 23.5 nM R = Br 11B2: 10.3 nM R = Br CYP17: 15% R = Br CYP19: 100 nM At 2.5 μM		R = H 11B1: 25.9 nM R = H 11B2: 41.0 nM R = H CYP17: 11% R = H CYP19: 955 nM R = CN 11B1: 15.0 nM R = CN 11B2: 35.9 nM R = CN CYP17: 21% R = CN CYP19: 119 nM R = F 11B1: 20.6 nM R = F 11B2: 16.7 nM R = F CYP17: 16% R = F CYP19: 218 nM R = Cl 11B1: 28.7 nM R = Cl 11B2: 88.8 nM R = Cl CYP17: 37% R = Cl CYP19: 330 nM R = Br 11B1: 26.2 nM R = Br 11B2: 92.8 nM R = Br CYP17: 13% R = Br CYP19: 27 nM At 2.5 μM	Double bond isomers Modeling to rationalize effect of change in shape Metal chelation ability changed by change in shape Ulmschneider [93]

Table 2 Modulation of reversible CYP inhibition dominated by lipophilicity reduction
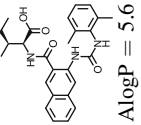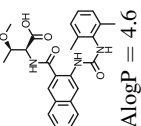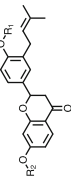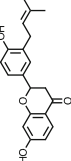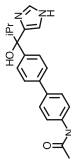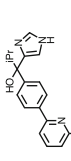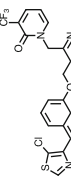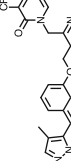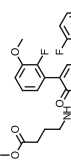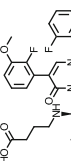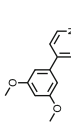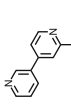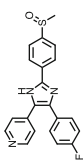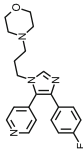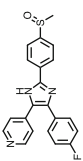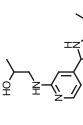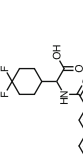
Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 5.6	2C9: 1.9 μM	AlogP = 4.6	2C9: 32.0 μM	Introduce polarity Sparks [96]
2	AlogP = 3.6	R1 = H, R2 = Me CYP19: 4.8 μM AlogP = 3.9 R1 = Me, R2 = H CYP19: 4.1 μM AlogP = 3.9 R1 = Me, R2 = Me CYP19: 3.7 μM AlogP = 4.2	AlogP = 3.6	CYP19: 41 μM	Remove lipophilicity Min [97]
3	AlogP = 3.6	3A4: 1.0 μM	AlogP = 3.0	3A4: >10.0 μM	Introduce polarity Kaku [84]
4	AlogP = 6.6	2C19: 0.3 μM	AlogP = 5.5	2C19: less potent	Remove lipophilicity Gibson [98]
5	AlogP = 6.6	3A4: 0.44 μM	AlogP = 5.5	3A4: 56.0 μM	Remove lipophilicity *Neutral to acid Chen [87]
6	AlogP = 3.8	3A4: 0.38 μM	AlogP = 3.2	3A4: >20.0 μM	Introduce polarity Remove lipophilicity Zhang [99]

Table 3 Modulation of reversible CYP inhibition by a combination of shape changes and lipophilicity reduction

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 5.5	1A2: 61% 2C9: 75% 2C19: 85% 2D6: 67% 3A4: 61% At 10 µM	AlogP = 3.5	1A2: <50% 2C9: <50% 2C19: <50% 2D6: 86% 3A4: <50% At 10 µM	Move substituent Introduce polarity Adams [15]
2	AlogP = 5.5	1A2: 38% 2C9: 87% 2C19: 92% 2D6: 73% 3A4: 77% At 10 µM	AlogP = 3.6	1A2: 29% 2C9: 39% 2C19: 25% 2D6: 5% 3A4: 33% At 10 µM	Large substituent Introduce polarity Laufer [16]
3	AlogP = 6.7	2C9: 0.7 µM	AlogP = 5.8	2C9: 6.7 µM	Cyclization Introduce polarity Sparks [80]
4	AlogP = 6.5	3A4: 66% At 10 µM	AlogP = 5.7	3A4: 14% At 10 µM	Change topology Introduce polarity Miwa [100]
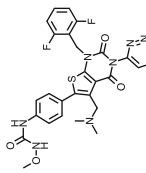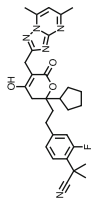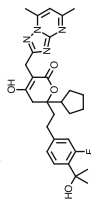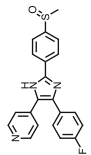
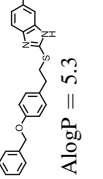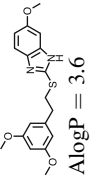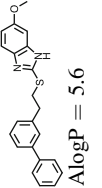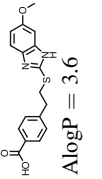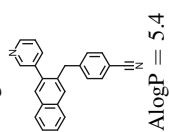| #	Structure	Data	Structure	Data	Tactics / Reference
5	AlogP = 5.7	3A4: 14% At 10 μM	AlogP = 4.5	3A4: 0% At 10 μM	Large substituent Introduce polarity Miwa [100]
6	AlogP = 5.1	2D6: 0.3 μM	AlogP = 4.4	2D6: >6 μM	Change topology Introduce polarity Li [83]
7	AlogP = 5.5	1A2: 61% 2C9: 75% 2C19: 85% 2D6: 67% 3A4: 61% At 10 μM	AlogP = 5.1	1A2: 0% 2C9: 19% 2C19: 19% 2D6: 0% 3A4: 38% At 10 μM	Large substituent Introduce polarity Laufer [101]
8	AlogP = 4.4	3A4: 0.014 μM	AlogP = 3.4	3A4: 15.0 μM	Change topology Introduce polarity *Neutral to acid Holsworth [102]
9	AlogP = 5.8	3A4: 0.084 μM	AlogP = 3.8	3A4: 4.0 μM	Move substituent Introduce polarity Degnan [103]
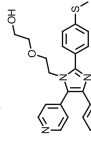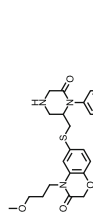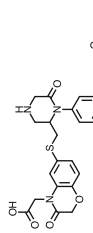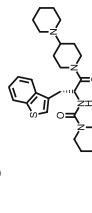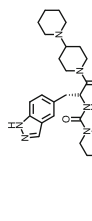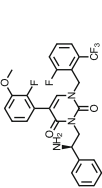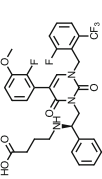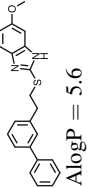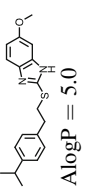
(continued)

Table 3 (continued)

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
10	AlogP = 5.0	3A4: 0.23 μM	AlogP = 4.9	3A4: 56.0 μM	Large substituent Introduce polarity Chen [87]
11	AlogP = 5.6	1A2: 0.7 μM 2B6: 1.1 μM 2C8: 0.3 μM 2C9: 0.8 μM 2C19: 2.2 μM 2D6: 0.02 μM 2E1: 50.0 μM	AlogP = 5.3	1A2: 50.0 μM 2B6: 50.0 μM 2C8: 0.7 μM 2C9: 0.9 μM 2C19: 50.0 μM 2D6: 0.1 μM 2E1: 50.0 μM	Change topology Introduce polarity Foti [92]
12	AlogP = 5.0	1A2: 0.7 μM 2B6: 1.1 μM 2C8: 0.3 μM 2C9: 0.8 μM 2C19: 2.2 μM 2D6: 0.02 μM 2E1: 50.0 μM	AlogP = 3.6	1A2: 10.7 μM 2B6: 20.0 μM 2C8: 3.4 μM 2C9: 2.3 μM 2C19: 50.0 μM 2D6: 9.1 μM 2E1: 50.0 μM	Move substituent Remove lipophilicity Foti [92]
13	AlogP = 5.6	1A2: 0.7 μM 2B6: 1.1 μM 2C8: 0.3 μM 2C9: 0.8 μM 2C19: 2.2 μM 2D6: 0.02 μM 2E1: 50.0 μM	AlogP = 3.6	1A2: 50.0 μM 2B6: 50.0 μM 2C8: 50.0 μM 2C9: 50.0 μM 2C19: 50.0 μM 2D6: 50.0 μM 2E1: 50.0 μM	Move substituent Introduce polarity Remove lipophilicity *Neutral to acid Foti [92]
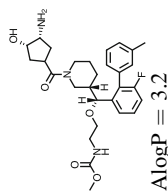
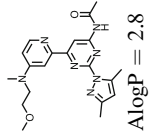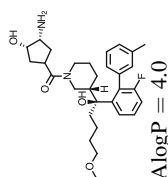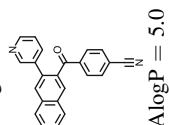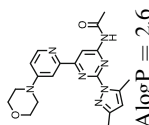|			
14	AlogP = 3.2 3A4: 0.5 μM	AlogP = 4.0 3A4: 25.0 μM	Change topology Introduce polarity Yuan [104]
15	AlogP = 5.4 1A2: 41% 2B6: 24% 2C9: 78% 2C19: 62% 2D6: 23% 3A4: 9% At 1 μM 3A4: 2.5 μM	AlogP = 5.0 1A2: 22% 2B6: 14% 2C9: 35% 2C19: <5% 2D6: 23% 3A4: 23% At 1 μM 3A4: >20.0 μM	Change topology Introduce polarity Lucas [94]
16	AlogP = 2.8	AlogP = 2.6	Cyclization Remove lipophilicity Zhang [99]
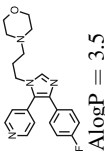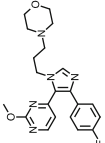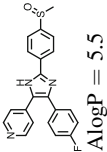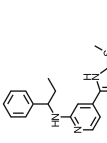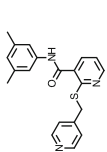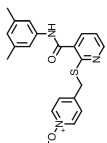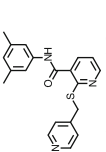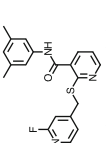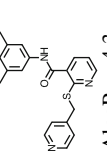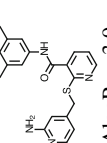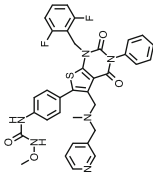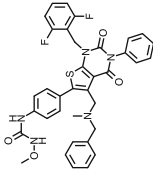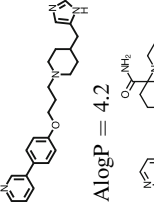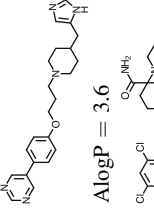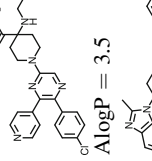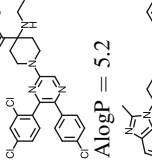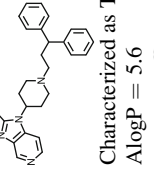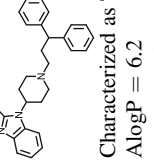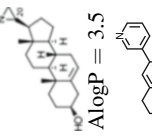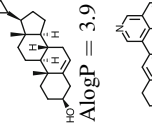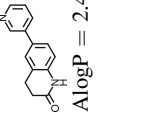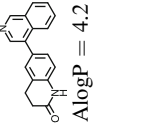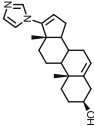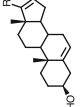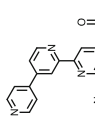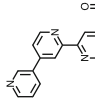
Table 4 Modulation of reversible CYP inhibition dominated by changes to metal binding ability

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 3.5	2D6: 86% At 10 μM	AlogP = 2.9	2D6: 7% At 10 μM	Steric block Modulation Adams [15]
2	AlogP = 5.5	1A2: 38% 2C9: 87% 2C19: 92% 2D6: 73% 3A4: 77% At 10 μM	AlogP = 6.0	1A2: 55% 2C9: 80% 2C19: 88% 2D6: 73% 3A4: 81% At 10 μM	Attempted steric block to prevent metal binding did not work Laufer [16]
3	AlogP = 4.3	3A4: 96% At 10 μM	AlogP = 3.5	3A4: 16% At 10 μM	Removal Tajima [105]
4	AlogP = 4.3	3A4: 96% At 10 μM	AlogP = 4.4	3A4: 40% At 10 μM	Steric block Cl, CN, SMe less effective than F Tajima [105]
5	AlogP = 4.3	3A4: 96% At 10 μM	AlogP = 3.9	3A4: 21% At 10 μM	Steric block NHAc equally effective Tajima [105]

#	Structure A	Structure B	Strategy / Reference
6	AlogP = 6.5; 3A4: 66% At 10 μM	AlogP = 7.1; 3A4: 36% At 10 μM	Removal Miwa [100]
7	AlogP = 4.2; 2D6: 0.3 μM	AlogP = 3.6; 2D6: 4.0 μM	Modulation Berlin [88]
8	AlogP = 3.5; 3A4: 71% At 5 μM	AlogP = 5.2; 3A4: <5% At 5 μM	Removal Wustrow [90]
9	Characterized as Type II; AlogP = 5.6; 2D6: 0.04 μM	Characterized as Type I; AlogP = 6.2; 2D6: 0.71 μM	Removal Wood [91]
10	AlogP = 3.5; 20R CYP17: 1.52 μM; 20S CYP17: 0.54 μM	AlogP = 3.9; 20R CYP17: >2.5 μM; 20S CYP17: 1.52 μM	Removal UV/vis detection of metal binding Reversible binding Hartmann [106]
11	AlogP = 2.4; 1A2: 2.0 μM; 2B6: <50.0 μM; 2C9: 58.9 μM; 2C19: <200 μM; 2D6: 171 μM; 3A4: 127 μM	AlogP = 4.2; 1A2: >150 μM; 2B6: <100 μM; 2C9: 2.9 μM; 2C19: 9.2 μM; 2D6: <50 μM; 3A4: <50 μM	Modulation Lucas [107]

(continued)
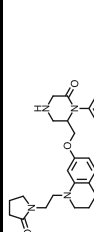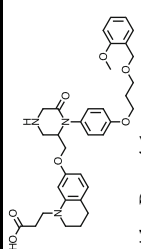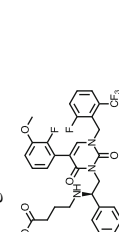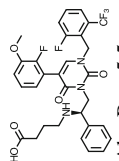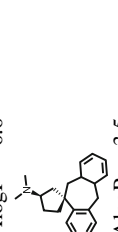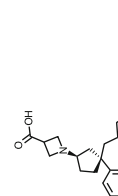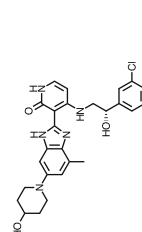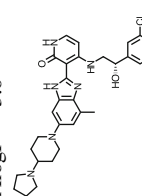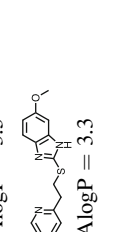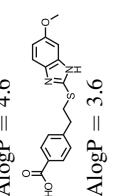
Table 4 (continued)

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
12		CYP17: 0.05 µM		3.8 µM, 0.8 µM; 3.7 µM, 0.8 µM; 2.6 µM, 0.5 µM; 1.4 µM, 0.4 µM; 1.3 µM, 0.3 µM	Modulation Clement [108]
13	AlogP = 3.2	3A4: 8.5 µM	AlogP = 3.2	3A4: >20.0 µM	Steric block (change in shape) Zhang [99]

Table 5 Modulation of reversible CYP inhibition dominated by changes in charge type

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 4.3	3A4: 0.042 μM	AlogP = 4.1	3A4: 10.0 μM	Neutral to acid Holsworth [102]
2	AlogP = 6.6	3A4: 0.44 μM	AlogP = 5.5	3A4: 56.0 μM	Neutral to acid Chen [87]
3	AlogP = 3.5	2D6: <0.1 μM	AlogP = 3.6	2D6: >10.0 μM	Base to acid Range of other acid containing amine derivatives also inactive Gianotti [109]
4	AlogP = 3.5	3A4: 0.4 μM	AlogP = 4.6	3A4: 4.5 μM	Neutral to base Zimmermann [89]
5	AlogP = 3.3	1A2: 0.7 μM 2B6: 1.1 μM 2C8: 0.3 μM 2C9: 0.8 μM 2C19: 2.2 μM	AlogP = 3.6	1A2: 50.0 μM 2B6: 50.0 μM 2C8: 50.0 μM 2C9: 50.0 μM 2C19: 50.0 μM	Neutral (weak base) to acid Foti [92]

(continued)

Table 5 (continued)
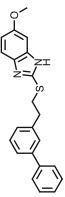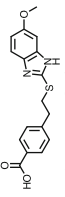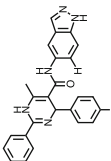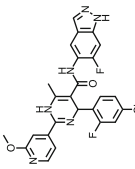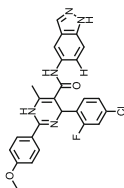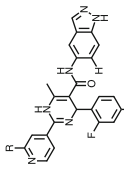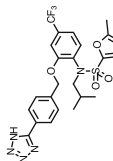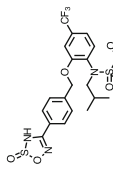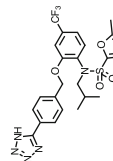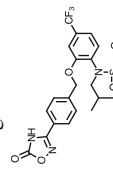
Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
6	AlogP = 5.6	2D6: 0.02 μM 2E1: 50.0 μM 1A2: 0.7 μM 2B6: 1.1 μM 2C8: 0.3 μM 2C9: 0.8 μM 2C19: 2.2 μM 2D6: 0.02 μM 2E1: 50.0 μM	AlogP = 3.6	2D6: 50.0 μM 2E1: 50.0 μM 1A2: 50.0 μM 2B6: 50.0 μM 2C8: 50.0 μM 2C9: 50.0 μM 2C19: 50.0 μM 2D6: 50.0 μM 2E1: 50.0 μM	Neutral to acid Foti [92]

Table 6 Modulation of reversible CYP inhibition by means not readily classified

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 3.8	2C9: 0.64 μM 2D6: 0.09 μM 3A4: 0.35 μM	AlogP = 4.0	2C9: 8.4 μM 2D6: 2.1 μM 3A4: 5.3 μM	Sehon [110]
2	AlogP = 4.4	2D6: 0.2 μM 3A4: 2.7 μM	R = Cl AlogP = 4.5 R = OMe AlogP = 3.8	R = Cl 2D6: 5.7 μM R = Cl 3A4: 5.5 μM R = OMe 2D6: 3.7 μM R = OMe 3A4: 4.3 μM	Sehon [110]
3	AlogP = 6.0	1A2: 12% 2C9: 98% 2C19: 98% 2D6: 3% 3A4: 94% At 3 μM	AlogP = 6.6	1A2: 22% 2C9: 9% 2C19: 33% 2D6: 2% 3A4: 68% At 3 μM	Naganawa [82]
4	AlogP = 6.0	1A2: 12% 2C9: 98% 2C19: 98% 2D6: 3% 3A4: 94% At 3 μM	AlogP = 6.2	1A2: 10% 2C9: 33% 2C19: 29% 2D6: 4% 3A4: 32% At 3 μM	Naganawa [82]

(continued)

Table 6 (continued)
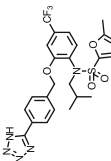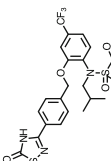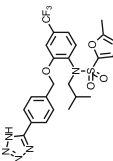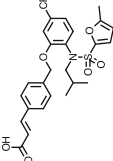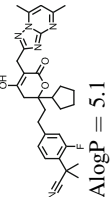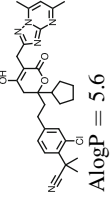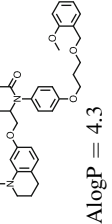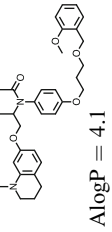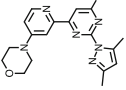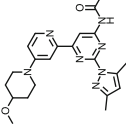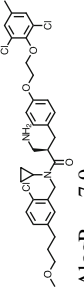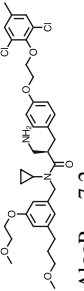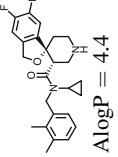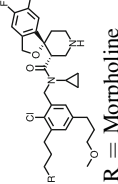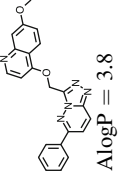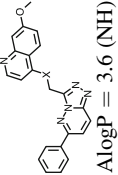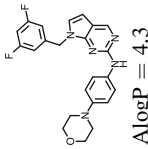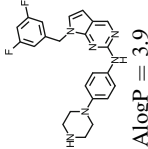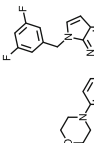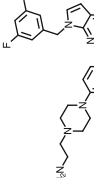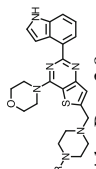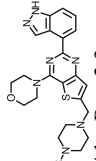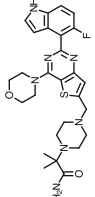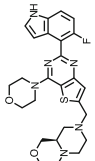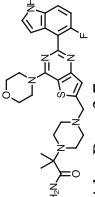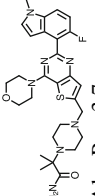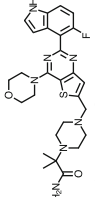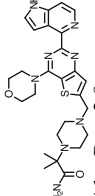
Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
5	AlogP = 6.0	1A2: 12% 2C9: 98% 2C19: 98% 2D6: 3% 3A4: 94% At 3 μM	AlogP = 6.7	1A2: 6% 2C9: -9% 2C19: 33% 2D6: 1% 3A4: 15% At 3 μM	Naganawa [82]
6	AlogP = 6.0	1A2: 12% 2C9: 98% 2C19: 98% 2D6: 3% 3A4: 94% At 3 μM	AlogP = 6.4	1A2: 2% 2C9: 9% 2C19: -4% 2D6: -3% 3A4: 11% At 3 μM	Naganawa [82]
7	AlogP = 5.1	2D6: 0.3 μM	AlogP = 5.6	2D6: 3.9 μM	Li [83]
8	AlogP = 4.3	3A4: 0.042 μM	AlogP = 4.1	3A4: 3.8 μM	Holsworth [102]
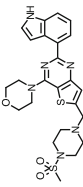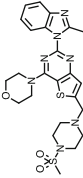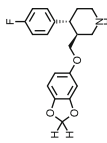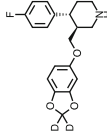
Table 7 Modulation of mechanism-based CYP inhibition

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
1	AlogP = 2.6	3A4: TDI active	AlogP = 2.7	3A4: TDI inactive	Reduce metabolism—block Zhang [99]
2	AlogP = 7.9	3A4: TDI active	AlogP = 7.2	3A4: TDI inactive	Reduce metabolism—reduce lipophilicity Chen [111]
3	AlogP = 4.4	3A4: 69% loss of activity after 30 min	R = Morpholine AlogP = 4.6 R = NEt2 AlogP = 5.7 R = NHSO2Me AlogP = 4.9	R = Morpholine 18% loss R = NEt2 AlogP = 17% loss R = NHSO2Me AlogP = 22% loss	Reduce metabolism—reduce lipophilicity (local not global) Chen [112]
4	AlogP = 3.8	3A4: 7.8 fold increase in IC50 after pre-incubation	AlogP = 3.6 (NH) AlogP = 4.5 (S)	X = NH or S 3A4: No decrease in IC50 after pre-incubation	Reduce metabolism—block Boezio [113]
5	AlogP = 4.3	3A4: TDI active	AlogP = 3.9	3A4: TDI inactive	Reduce metabolism—reduce lipophilicity and block Nagashima [75]

(continued)

Table 7 (continued)

Example number	Problem molecule	Potency of problem molecule	Improved molecule	Potency of improved molecule	Comments and reference
6	AlogP = 4.3	3A4: TDI active	AlogP = 3.6	3A4: TDI inactive	Reduce metabolism—reduce lipophilicity and block Nagashima [75]
7	AlogP = 2.9 (CMe_2CONH_2)	R = CMe_2CONH_2 3A4: TDI active; R = SO_2Me 3A4: TDI active	AlogP = 2.3 (CMe_2CONH_2)	R = CMe_2CONH_2 3A4: TDI inactive; R = SO_2Me 3A4: TDI inactive	Reduce metabolism—reduce lipophilicity. Remove—vulnerable double bond (epoxide formation a liability). Indole metabolism leads to MBI. Safina [76]
8		3A4: TDI active		3A4: TDI inactive	Redirect metabolism. Indole metabolism leads to MBI. Safina [76]
9	AlogP = 3.7	3A4: TDI active	AlogP = 3.7	3A4: TDI inactive	Remove—vulnerable NH (N-oxidation a liability). Indole metabolism leads to MBI
10	AlogP = 3.7	3A4: TDI active	AlogP = 2.9	3A4: TDI inactive	Reduce metabolism—reduce lipophilicity and block (oxidation to phenol a precursor to quinoneimine formation)

11	3A4: TDI active	3A4: TDI inactive	Remove—vulnerable double bond that yields phenol Indole metabolism leads to MBI Safina [76]
	AlogP = 3.7	AlogP = 3.4	Remove—indole origin of MBI
12	2D6: TDI active	2D6: TDI inactive	Reduce metabolism— block using kinetic isotope effect Yarnell [114]
	AlogP = 3.3	AlogP = 3.3	

Tactics highlighted in Tables 1 and 3 include the movement of substituents around rings, inversion of chirality, double bond isomerism, the introduction of large substituents, changing topology, cyclization, and ring breaking. Table 2 includes examples of the introduction of polarity, the removal of lipophilicity a subset of which also include changes in charge type. Table 4 includes examples in which the metal binder is removed, sterically blocked or modulated. Table 5 is dominated by changes in which an acid is introduced to a molecule but also includes one example where a base was introduced. Table 6 includes an interesting set in which tetrazole is exchanged for a range of alternative common acid isosteres, for which many possible causes, including lipophilicity and pKa, could contribute to the observed changes. Finally, a small selection of structural changes that modulate MBIs is shown, a much more thorough set of such changes are provided by Orr et al. [71].

References

1. Li AP (2008) Drug–drug interactions in pharmaceutical development. Wiley New York; Chichester
2. Guengerich FP (2005) Human cytochrome P450 enzymes. In: Ortiz de Montellano PR (ed) Cytochrome P450: structure, mechanism, and biochemistry, 3rd edn. New York: Kluwer Academic/Plenum, pp 377–530
3. Obach RS (2008) In inhibition of drug-metabolizing enzymes and drug-drug interactions in drug discovery and development. Pharmacology 75–93
4. Bibi Z (2008) Role of cytochrome P450 in drug interactions. Nutr Metab 5:27
5. Po AL, Zhang WY (1998 Jun 20) What lessons can be learnt from withdrawal of mibefradil from the market? Lancet 351(9119):1829–1830
6. Tung RD, Chandorkar G, Perni RB, inventors (2005) Anonymous pharmaceutical compositions comprising HCV NS3/4A protease inhibitor and cytochrome P 450 monooxygenase inhibitor for HCV treatment. WO2005042020A2. 2005 05/12
7. Norbeck DW, Kempf DJ, Leonard JM, Bertz RJ, inventors (1997) Anonymous use of ritonavir (ABT-538) for improving the pharmacokinetics of drugs metabolized by cytochrome P450 in a method of treating aids. WO9701349A1. 1997 01/16
8. Lumpkin MM, Alpert S (2012) Risk of drug interactions with St. John's Wort and Indinavir and other Drugs. http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformation-forPatientsandProviders/DrugSafetyInformationforHeathcareProfessionals/PublicHealthAd-visories/ucm052238.htm. Accessed 10 September 2012
9. Morgan ET (2001 March 01) Regulation of cytochrome P450 by inflammatory mediators: why and how? Drug Metab Dispos 29(3):207–212
10. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Delivery Rev 23(1–3):3–25
11. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discovery 6(11):881–890

12. Gleeson MP (2008) Generation of a set of simple, interpretable ADMET rules of thumb. J Med Chem 51(4):817–834
13. Hann MM (2011) Molecular obesity, potency and other addictions in drug discovery. Med Chem Commun 2(5):349–355
14. Testa B, Jenner P (1981) Inhibitors of cytochrome P-450s and their mechanism of action. Drug Metab Rev 12(1):1–117
15. Adams JL, Boehm JC, Kassis S, Gorycki PD, Webb EF, Hall R et al (1998) Pyrimidinyli-midazole inhibitors of CSBP/p38 kinase demonstrating decreased inhibition of hepatic cytochrome P450 enzymes. Bioorg Med Chem Lett 8(22):3111–3116
16. Laufer SA, Hauser DRJ, Domeyer DM, Kinkel K, Liedtke AJ (2008) Design, synthesis, and biological evaluation of novel tri- and tetrasubstituted imidazoles as highly potent and specific ATP-mimetic inhibitors of p38 MAP kinase: focus on optimized interactions with the enzyme's surface-exposed front region. J Med Chem 51(14):4122–4149
17. Ekroos M, Sjoegren T (2006) Structural basis for ligand promiscuity in cytochrome P 450 3A4. Proc Natl Acad Sci USA 103(37):13682–13687
18. Sevirioukova IF, Poulos TL (2010) Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. Proc Natl Acad Sci USA 1–6:6
19. Yano JK, Wester MR, Schoch GA, Griffin KJ, Stout CD, Johnson EF (2004) The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-. ANG. resolution. J Biol Chem 279(37):38091–38094
20. Bren U, Oostenbrink C (2012) Cytochrome P450 3A4 inhibition by ketoconazole: tackling the problem of ligand cooperativity using molecular dynamics simulations and free-energy calculations. J Chem Inf Model 52(6):1573–1582
21. Williams PA, Cosme J, Ward A, Angove HC, Matak VD, Jhoti H (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature 424(6947):464–468
22. Wester MR, Yano JK, Schoch GA, Yang C, Griffin KJ, Stout CD et al (2004) The structure of human cytochrome P 450 2C9 complexed with flurbiprofen at 2.0-.ANG. resolution. J Biol Chem 279(34):35630–35637
23. Sansen S, Yano JK, Reynald RL, Schoch GA, Griffin KJ, Stout CD et al (2007) Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. J Biol Chem 282(19):14348–14355
24. Porubsky PR, Meneely KM, Scott EE (2008 November 28) Structures of human cytochrome P-450 2E1. J Biol Chem 283(48):33698–33707
25. Porubsky PR, Battaile KP, Scott EE (2010) Human cytochrome P450 2E1 structures with fatty acid analogs reveal a previously unobserved binding mode. J Biol Chem 285 (29):22282–22290
26. DeVore NM, Meneely KM, Bart AG, Stephens ES, Battaile KP, Scott EE (2012) Structural comparison of cytochromes P450 2A6, 2A13, and 2E1 with pilocarpine. FEBS J 279 (9):1621–1631
27. Rowland P, Blaney FE, Smyth MG, Jones JJ, Leydon VR, Oxbrow AK et al (2006) Crystal structure of human cytochrome P 450 2D6. J Biol Chem 281(11):7614–7622
28. Wang A, Savas U, Hsu M, Stout CD, Johnson EF (2012 March 30) Crystal structure of human cytochrome P450 2D6 with prinomastat bound. J Biol Chem 287(14):10834–10843
29. Schoch GA, Yano JK, Wester MR, Griffin KJ, Stout CD, Johnson EF (2004 March 05) Structure of human microsomal cytochrome P450 2C8. J Biol Chem 279(10):9497–9503
30. Schoch GA, Yano JK, Sansen S, Dansette PM, Stout CD, Johnson EF (2008 June 20) Determinants of cytochrome P450 2C8 substrate binding. J Biol Chem 283(25):17227–17237
31. DeVore NM, Smith BD, Urban MJ, Scott EE (2008) Key residues controlling phenacetin metabolism by human cytochrome P450 2A Enzymes. Drug Metab Dispos 36 (12):2582–2590
32. Sansen S, Hsu M, Stout CD, Johnson EF (2007) Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants. Arch Biochem Biophys 464(2):197–206

33. Yano JK, Hsu MH, Griffin KJ, Stout CD, Johnson EF (2005 Sep) Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. Nat Struct Mol Biol 12(9):822–823

34. Leach AG, Kidley NJ (2011) Quantitatively interpreted enhanced inhibition of cytochrome P450s by heteroaromatic rings containing nitrogen. J Chem Inf Model 51:1048–1063

35. Guengerich FP (2001) Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. Chem Res Toxicol 14(6):611–650

36. Oxidative, reductive, and hydrolytic metabolism of drugs. Section Title: Pharmacology; 2008.

37. Rittle J, Green MT (2010 November 12) Cytochrome P450 compound I: capture, characterization, and C–H bond activation kinetics. Science 330(6006):933–937

38. Sligar SG (2010) Glimpsing the critical intermediate in cytochrome P450 oxidations. Science 330(6006):924–925

39. Meunier B, de Visser SP, Shaik S (2004) Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. Chem Rev (Washington, DC) 104(9):3947–3980

40. Newcomb M, Toy PH (2000 Jul) Hypersensitive radical probes and the mechanisms of cytochrome P450-catalyzed hydroxylation reactions. Acc Chem Res 33(7):449–455

41. Shaik S, Kumar D, de Visser SP, Altun A, Thiel W. Theoretical perspective on the structure and mechanism of cytochrome P 450 enzymes. Chem Rev (Washington, DC) 105 (6):2279–2328

42. Shaik S, Cohen S, Wang Y, Chen H, Kumar D, Thiel W (2010) P 450 Enzymes: their structure, reactivity, and selectivity – modeled by QM/MM calculations. Chem Rev (Washington, DC) 110(2):949–1017

43. Afzelius L, Arnby CH, Broo A, Carlsson L, Isaksson C, Jurva U et al (2007) State-of-the-art tools for computational site of metabolism predictions: comparative analysis, mechanistical insights, and future applications. Drug Metab Rev 39(1):61–86

44. Refsgaard HHF, Jensen BF, Christensen IT, Hagen N, Brockhoff PB (2006) In silico prediction of cytochrome P450 inhibitors. Drug Dev Res 67(5):417–429

45. Viswanadhan VN, Ghose AK, Revankar GR, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J Chem Inf Comput Sci 29(3):163–172

46. Ritchie TJ, Macdonald SJF, Peace S, Pickett SD, Luscombe CN (2012) The developability of heteroaromatic and heteroaliphatic rings – do some have a better pedigree as potential drug molecules than others? Med Chem Commun 3(9):1062–1069

47. Ritchie TJ, Macdonald SJ (2009 Nov) The impact of aromatic ring count on compound developability–are too many aromatic rings a liability in drug design? Drug Discov Today 14 (21–22):1011–1020

48. Ritchie TJ, Macdonald SJ, Young RJ, Pickett SD (2011 Feb) The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. Drug Discov Today 16(3–4):164–171

49. Lovering F, Bikker J, Humblet C (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52(21):6752–6756

50. Sun H, Veith H, Xia M, Austin CP, Huang R (2011) Predictive models for cytochrome P450 isozymes based on quantitative high throughput screening data. J Chem Inf Model 51 (10):2474–2481

51. Veith H, Southall N, Huang R, James T, Fayne D, Artemenko N et al (2009 Nov) Comprehensive characterization of cytochrome P450 isozyme selectivity across chemical libraries. Nat Biotechnol 27(11):1050–1055

52. Waring MJ (2010) Lipophilicity in drug discovery. Expert Opin Drug Discovery 5 (3):235–248

53. Balding PR, Porro CS, McLean KJ, Sutcliffe MJ, Marechal J, Munro AW et al (2008) How do azoles inhibit cytochrome P450 enzymes? a density functional study. J Phys Chem A 112 (50):12911–12918
54. Klingenberg M (1958 Jun) Pigments of rat liver microsomes. Arch Biochem Biophys 75 (2):376–386
55. OMURA T, SATO R (1962 Apr) A new cytochrome in liver microsomes. J Biol Chem 237:1375–1376
56. The Porphyrins vol 3 : Physical Chemistry Part A. New York: Academic; 1978.
57. Jefcoate CR (1978) Measurement of substrate and inhibitor binding to microsomal cytochrome P-450 by optical-difference spectroscopy. Biomemb C Methods Enzymol 52:258–279
58. Hunter CA (2004) Quantifying intermolecular interactions: guidelines for the molecular recognition toolbox. Angew Chem Int Ed 43(40):5310–5324
59. Hughes JD, Blagg J, Price DA, Bailey S, DeCrescenzo GA, Devraj RV et al (2008) Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett 18(17):4872–4875
60. Waring MJ (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower log D limits based on permeability. Bioorg Med Chem Lett 19(10):2844–2851
61. Scott JS, Birch AM, Brocklehurst KJ, Broo A, Brown HS, Butlin RJ et al (2012 Jun 14) Use of small-molecule crystal structures to address solubility in a novel series of G protein coupled receptor 119 agonists: optimization of a lead and in vivo evaluation. J Med Chem 55 (11):5361–5379
62. Schenkman JB, Remmer H, Estabrook RW (1967 Mar) Spectral studies of drug interaction with hepatic microsomal cytochrome. Mol Pharmacol 3(2):113–123
63. Attard G, Belldegrun AS, De Bono JS (2005) Selective blockade of androgenic steroid synthesis by novel lyase inhibitors as a therapeutic strategy for treating metastatic prostate cancer. BJU Int 96(9):1241–1246
64. Temple TE, Liddle GW (1970) Inhibitors of adrenal steroid biosynthesis. Annu Rev Pharmacol 10:199–218
65. Ahmad Z, Sharma S, Khuller GK (2006 Aug) Azole antifungals as novel chemotherapeutic agents against murine tuberculosis. FEMS Microbiol Lett 261(2):181–186
66. Zhang W, Ramamoorthy Y, Kilicarslan T, Nolte H, Tyndale RF, Sellers EM (2002 Mar) Inhibition of cytochromes P450 by antifungal imidazole derivatives. Drug Metab Dispos 30 (3):314–318
67. Riley RJ, Parker AJ, Trigg S, Manners CN (2001 May) Development of a generalized, quantitative physicochemical model of CYP3A4 inhibition for use in early drug discovery. Pharm Res 18(5):652–655
68. Lowe D (2012) In the pipeline. http://pipeline.corante.com/archives/2012/07/18/the_best_rings_to_put_in_your_molecules.php. Accessed 19 September 2012
69. Pitt WR, Parry DM, Perry BG, Groom CR (2009) Heteroaromatic rings of the future. J Med Chem 52(9):2952–2963
70. Seward HE, Roujeinikova A, McLean KJ, Munro AW, Leys D (2006 Dec 22) Crystal structure of the Mycobacterium tuberculosis P450 CYP121-fluconazole complex reveals new azole drug-P450 binding mode. J Biol Chem 281(51):39437–39443
71. Orr ST, Ripp SL, Ballard TE, Henderson JL, Scott DO, Obach RS et al (2012 Jun 14) Mechanism-based inactivation (MBI) of cytochrome P450 enzymes: structure-activity relationships and discovery strategies to mitigate drug-drug interaction risks. J Med Chem 55(11):4896–4933
72. Guengerich FP (2008) Cytochrome P450 and chemical toxicology. Chem Res Toxicol 21 (1):70–83
73. Fontana E, Dansette PM, Poli SM (2005 Oct) Cytochrome p450 enzymes mechanism based inhibitors: common sub-structures and reactivity. Curr Drug Metab 6(5):413–454

74. Zhou S, Yung Chan S, Cher Goh B, Chan E, Duan W, Huang M et al (2005) Mechanism-based inhibition of cytochrome P450 3A4 by therapeutic drugs. Clin Pharmacokinet 44 (3):279–304

75. Nagashima S, Hondo T, Nagata H, Ogiyama T, Maeda J, Hoshii H et al (2009) Novel 7H-pyrrolo[2,3-d]pyrimidine derivatives as potent and orally active STAT6 inhibitors. Bioorg Med Chem 17(19):6926–6936

76. Safina BS, Baker S, Baumgardner M, Blaney PM, Chan BK, Chen YH et al (2012 Jun 28) Discovery of novel PI3-kinase delta specific inhibitors for the treatment of rheumatoid arthritis: taming CYP3A4 time-dependent inhibition. J Med Chem 55(12):5887–5900

77. Wu YJ, Davis CD, Dworetzky S, Fitzpatrick WC, Harden D, He H et al (2003 Aug 28) Fluorine substitution can block CYP3A4 metabolism-dependent inhibition: identification of (S)-N-[1-(4-fluoro-3- morpholin-4-ylphenyl)ethyl]-3- (4-fluorophenyl)acrylamide as an orally bioavailable KCNQ2 opener devoid of CYP3A4 metabolism-dependent inhibition. J Med Chem 46(18):3778–3781

78. Walker DP, Bi FC, Kalgutkar AS, Bauman JN, Zhao SX, Soglia JR et al (2008 Dec 1) Trifluoromethylpyrimidine-based inhibitors of proline-rich tyrosine kinase 2 (PYK2): structure-activity relationships and strategies for the elimination of reactive metabolite formation. Bioorg Med Chem Lett 18(23):6071–6077

79. Meanwell NA (2011) Synopsis of some recent tactical application of bioisosteres in drug design. J Med Chem 54(8):2529–2591

80. Sparks SM, Banker P, Bickett DM, Carter HL, Clancy DC, Dickerson SH et al (2009) Anthranilamide-based glycogen phosphorylase inhibitors for the treatment of type 2 diabetes: 1. Identification of 1-amino-1-cycloalkyl carboxylic acid headgroups. Bioorg Med Chem Lett 19(3):976–980

81. Shao L, Hewitt MC, Malcolm SC, Wang F, Ma J, Campbell UC et al (2011) Synthesis and pharmacological characterization of bicyclic 3-aryl octahydrocyclopenta[c]pyrrole analogues as triple reuptake inhibitors. J Med Chem 54(15):5283–5295

82. Naganawa A, Matsui T, Ima M, Saito T, Murota M, Aratani Y et al (2006) Further optimization of sulfonamide analogs as EP1 receptor antagonists: synthesis and evaluation of bioisosteres for the carboxylic acid group. Bioorg Med Chem 14(21):7121–7137

83. Li H, Tatlock J, Linton A, Gonzalez J, Jewell T, Patel L et al (2009) Discovery of (R)-6-cyclopentyl-6-(2-(2,6-diethylpyridin-4-yl)ethyl)-3-((5,7-dimethyl-[1,2,4]triazolo[1,5-a] pyrimidin-2-yl)methyl)-4-hydroxy-5,6-dihydropyran-2-one (PF-00868554) as a potent and orally available hepatitis C virus polymerase inhibitor. J Med Chem 52(5):1255–1258

84. Kaku T, Tsujimoto S, Matsunaga N, Tanaka T, Hara T, Yamaoka M et al (2011) 17,20-Lyase inhibitors. Part 3: Design, synthesis, and structure-activity relationships of biphenylylmethylimidazole derivatives as novel 17,20-lyase inhibitors. Bioorg Med Chem 19(7):2428–2442

85. Kaku T, Matsunaga N, Ojida A, Tanaka T, Hara T, Yamaoka M et al (2011) 17,20-Lyase inhibitors. Part 4: design, synthesis and structure-activity relationships of naphthylmethylimidazole derivatives as novel 17,20-lyase inhibitors. Bioorg Med Chem 19(5):1751–1770

86. Iyengar RR, Lynch JK, Mulhern MM, Judd AS, Freeman JC, Gao J et al (2007) An evaluation of 3,4-methylenedioxy phenyl replacements in the aminopiperidine chromone class of MCHr1 antagonists. Bioorg Med Chem Lett 17(4):874–878

87. Chen C, Wu D, Guo Z, Xie Q, Reinhart GJ, Madan A et al (2008) Discovery of Sodium R-(+)-4-{2-[5-(2-Fluoro-3-methoxyphenyl)-3-(2-fluoro-6-[trifluoromethyl]benzyl)-4-methyl-2,6-dioxo-3,6-dihydro-2H-pyrimidin-1-yl]-1-phenylethylamino}butyrate (Elagolix), a potent and orally available nonpeptide antagonist of the human gonadotropin-releasing hormone receptor. J Med Chem 51(23):7478–7485

88. Berlin M, Ting PC, Vaccaro WD, Aslanian R, McCormick KD, Lee JF et al (2006) Reduction of CYP450 inhibition in the 4-[(1H-imidazol-4-yl)methyl]piperidine series of histamine H3 receptor antagonists. Bioorg Med Chem Lett 16(4):989–994

89. Zimmermann K, Wittman MD, Saulnier MG, Velaparthi U, Langley DR, Sang X et al (2008) Balancing oral exposure with Cyp3A4 inhibition in benzimidazole-based IGF-IR inhibitors. Bioorg Med Chem Lett 18(14):4075–4080

90. Wustrow DJ, Maynard GD, Yuan J, Zhao H, Mao J, Guo Q et al (2008) Aminopyrazine CB1 receptor inverse agonists. Bioorg Med Chem Lett 18(11):3376–3381

91. Wood A, Armour D (2005) The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. Prog Med Chem 43:239–271

92. Foti RS, Rock DA, Han X, Flowers RA, Wienkers LC, Wahlstrom JL (2012) Ligand-based design of a potent and selective inhibitor of cytochrome P450 2C19. J Med Chem 55 (3):1205–1214

93. Ulmschneider S, Mueller-Vieira U, Mitrenga M, Hartmann RW, Oberwinkler-Marchais S, Klein CD et al (2005) Synthesis and evaluation of imidazolylmethylenetetrahydro-naphthalenes and imidazolylmethyleneindanes: potent inhibitors of aldosterone synthase. J Med Chem 48(6):1796–1805

94. Lucas S, Heim R, Negri M, Antes I, Ries C, Schewe KE et al (2008) Novel aldosterone synthase inhibitors with extended carbocyclic skeleton by a combined ligand-based and structure-based drug design approach. J Med Chem 51(19):6138–6149

95. Gobbi S, Cavalli A, Rampa A, Belluti F, Piazzi L, Paluszcak A et al (2006) Lead optimization providing a series of flavone derivatives as potent nonsteroidal inhibitors of the cytochrome P450 aromatase enzyme. J Med Chem 49(15):4777–4780

96. Sparks SM, Banker P, Bickett DM, Clancy DC, Dickerson SH, Garrido DM et al (2009) Anthranilamide-based glycogen phosphorylase inhibitors for the treatment of Type 2 diabetes: 2 Optimization of serine and threonine ether amino acid residues. Bioorg Med Chem Lett 19(3):981–985

97. Min KH, Xia Y, Kim EK, Jin Y, Kaur N, Kim ES et al (2009) A novel class of highly potent multidrug resistance reversal agents: disubstituted adamantyl derivatives. Bioorg Med Chem Lett 19(18):5376–5379

98. Gibson C, Schnatbaum K, Pfeifer JR, Locardi E, Paschke M, Reimer U et al (2009) Novel small molecule bradykinin B2 receptor antagonists. J Med Chem 52(14):4370–4379

99. Zhang X, Tellew JE, Luo Z, Moorjani M, Lin E, Lanier MC et al (2008) Lead optimization of 4-Acetylamino-2-(3,5-dimethylpyrazol-1-yl)-6-pyridylpyrimidines as A2A adenosine receptor antagonists for the treatment of Parkinson's disease. J Med Chem 51(22):7099–7110

100. Miwa K, Hitaka T, Imada T, Sasaki S, Yoshimatsu M, Kusaka M et al (2011) Discovery of 1-{4-[1-(2,6-difluorobenzyl)-5-[(dimethylamino)methyl]-3-(6-methoxypyridazin-3-yl)-2,4-dioxo-1,2,3,4-tetrahydrothieno[2,3-d]pyrimidin-6-yl]phenyl}-3-methoxyurea (TAK-385) as a potent, orally active, non-peptide antagonist of the human gonadotropin-releasing hormone receptor. J Med Chem 54(14):4998–5012

101. Laufer SA, Zimmermann W, Ruff KJ (2004) Tetrasubstituted imidazole inhibitors of cytokine release: probing substituents in the N-1 position. J Med Chem 47(25):6311–6325

102. Holsworth DD, Cai C, Cheng X, Cody WL, Downing DM, Erasga N et al (2006) ketopiperazine-based renin inhibitors: optimization of the "C" ring. Bioorg Med Chem Lett 16(9):2500–2504

103. Degnan AP, Chaturvedula PV, Conway CM, Cook DA, Davis CD, Denton R et al (2008) Discovery of (R)-4-(8-Fluoro-2-oxo-1,2-dihydroquinazolin-3(4H)-yl)-N-(3-(7-methyl-1H-indazol-5-yl)-1-oxo-1-(4-(piperidin-1-yl)piperidin-1-yl)propan-2-yl)piperidine-1-carboxamide (BMS-694153): a potent antagonist of the human calcitonin gene-related peptide receptor for migraine with rapid and efficient intranasal exposure. J Med Chem 51 (16):4858–4861

104. Yuan J, Simpson RD, Zhao W, Tice CM, Xu Z, Cacatian S et al (2011) Biphenyl/diphenyl ether renin inhibitors: filling the S1 pocket of renin via the S3 pocket. Bioorg Med Chem Lett 21(16):4836–4843

105. Tajima H, Honda T, Kawashima K, Sasabuchi Y, Yamamoto M, Ban M et al (2011) Pyridylmethylthio derivatives as VEGF inhibitors: Part 2. Bioorg Med Chem Lett 21 (4):1232–1235
106. Hartmann RW, Hector M, Wachall BG, Palusczak A, Palzer M, Huch V et al (2000) Synthesis and evaluation of 17-aliphatic heterocycle-substituted steroidal inhibitors of 17Î ± -Hydroxylase/C17-20-Lyase (P450 17). J Med Chem 43(23):4437–4445
107. Lucas S, Heim R, Ries C, Schewe KE, Birk B, Hartmann RW (2008) In vivo active aldosterone synthase inhibitors with improved selectivity: lead optimization providing a series of pyridine substituted 3,4-dihydro-1H-quinolin-2-one derivatives. J Med Chem 51 (24):8077–8087
108. Clement OO, Freeman CM, Hartmann RW, Handratta VD, Vasaitis TS, Brodie AMH et al (2003) Three dimensional pharmacophore modeling of human CYP17 inhibitors. Potential agents for prostate cancer therapy. J Med Chem 46(12):2345–2351
109. Gianotti M, Botta M, Brough S, Carletti R, Castiglioni E, Corti C et al (2010) Novel spirotetracyclic zwitterionic dual H(1)/5-HT(2A) receptor antagonists for the treatment of sleep disorders. J Med Chem 53(21):7778–7795
110. Sehon CA, Wang GZ, Viet AQ, Goodman KB, Dowdell SE, Elkins PA et al (2008) Potent, selective and orally bioavailable dihydropyrimidine inhibitors of Rho kinase (ROCK1) as potential therapeutic agents for cardiovascular diseases. J Med Chem 51(21):6631–6634
111. Chen A, Bayly C, Bezencon O, Richard-Bildstein S, Dube D, Dube L et al (2010) Design and optimization of a substituted amino propanamide series of renin inhibitors for the treatment of hypertension. Bioorg Med Chem Lett 20(7):2204–2209
112. Chen A, Aspiotis R, Campeau L, Cauchon E, Chefson A, Ducharme Y et al (2011) Renin inhibitors for the treatment of hypertension: design and optimization of a novel series of spirocyclic piperidines. Bioorg Med Chem Lett 21(24):7399–7404
113. Boezio AA, Berry L, Albrecht BK, Bauer D, Bellon SF, Bode C et al (2009) Discovery and optimization of potent and selective triazolopyridazine series of c-Met inhibitors. Bioorg Med Chem Lett 19(22):6307–6312
114. Yarnell A (2009) Heavy-hydrogen drugs turn heads, again. Chem Eng News Arch 87 (25):36–39

Avoiding PXR and CAR Activation and CYP3A4 Enzyme Induction

Michael W. Sinz

Abstract Avoiding drug–drug interactions is an important aspect of today's drug discovery and development process. The most significant of these interactions occur through changes in the enzyme level of CYP3A4 which is involved in the metabolism of many drugs. Increases in the expression of CYP3A4 mRNA and enzyme activity can occur through several mechanisms, the most predominant of which is the activation of transcription factors, such as the nuclear hormone receptors pregnane X receptor (PXR) and constitutive androstane receptor (CAR). Through an understanding of interactions between drugs (ligands) and the ligand binding pockets of these receptors, several laboratories have attenuated the binding interactions and significantly reduced the potential for CYP3A4 induction and ultimately drug–drug interactions.

Keywords Constitutive androstane receptor, CYP3A4, Drug discovery, Drug-drug interactions, Pregnane X receptor, Structure-activity-relationship

Contents

M.W. Sinz (✉)
Bristol Myers-Squibb Research and Development, 5 Research Parkway, Wallingford, CT 06492, USA
e-mail: michael.sinz@bms.com

1 Introduction

In today's healthcare system, patients routinely take multiple medications on a daily basis and, as such, the risk of adverse events due to drug–drug interactions (DDIs) has increased significantly. DDIs are a concern for clinicians and pharmaceutical companies as they can have a significant impact on patient safety as well as a negative impact on the financial return on investment if the marketability of a new drug is restricted due to DDIs. Serious DDIs have contributed to drugs being withdrawn from the market, changes to labeling, dose adjustments, and in some situations non-approval notification from regulatory agencies [1, 2]. Consequently, there is a compelling need in the pharmaceutical industry to avoid, as much as possible, the development of new drugs that have a potential to cause DDIs.

Drug–drug interactions occur when the efficacy or toxicity of one drug is unduly affected by coadministration of a second drug. These interactions occur as a result of changing concentrations of drug-metabolizing enzymes or drug transporters, both of which play a significant role in the absorption, distribution, metabolism, and elimination (ADME) of drugs. Inhibition of enzyme activity leads to reduced metabolism of the parent drug and a decrease in metabolite formation, while enzyme induction leads to an increase in parent drug metabolism and increased formation of metabolites. In the case of enzyme inhibition or reduced enzyme activity, the perpetrator drug (drug that causes the enzyme inhibition) causes a significant increase in the exposure of the victim drug resulting in exaggerated pharmacology or increased off-target toxicity. For example, when the CYP3A4 substrate sildenafil (Viagra®) is coadministered with the antiretroviral protease inhibitor combination of saquinavir/ritonavir (both inhibitors of CYP3A4), there is an 11-fold increase in the AUC (area under the curve) of sildenafil. This significant increase in sildenafil systemic exposure necessitates limiting the dose of sildenafil when given in combination with potent inhibitors of CYP3A4, such as saquinavir/ritonavir [3]. In one of the most significant interactions, the AUC of ramelteon (Rozerem®) increases 190-fold when coadministered with fluvoxamine. The significant magnitude of this interaction stems from two properties of ramelteon: (1) the elimination of ramelteon is solely by metabolism and (2) ramelteon is predominately metabolized by CYP1A2 and to a lesser extent by CYP2C19 and CYP3A4. Fluvoxamine is a potent inhibitor of CYP1A2 and can also inhibit CYP2C19 and CYP3A4; therefore coadministration of fluvoxamine abolishes all pathways of ramelteon elimination and significantly increases the systemic exposure of ramelteon [4].

In the case of enzyme induction or increased enzyme activity, the perpetrator drug (drug that causes the enzyme induction) causes a significant decrease in the exposure of the victim drug resulting in a loss of efficacy. Perpetrators of enzyme induction can be drugs, natural products, or endogenous substrates, such as hormones. A number of drug-related clinical DDIs are illustrated in the literature with midazolam as the victim since the disposition of this drug is very sensitive to changes in CYP3A4 enzyme activity. For example, rifampicin, carbamazepine,
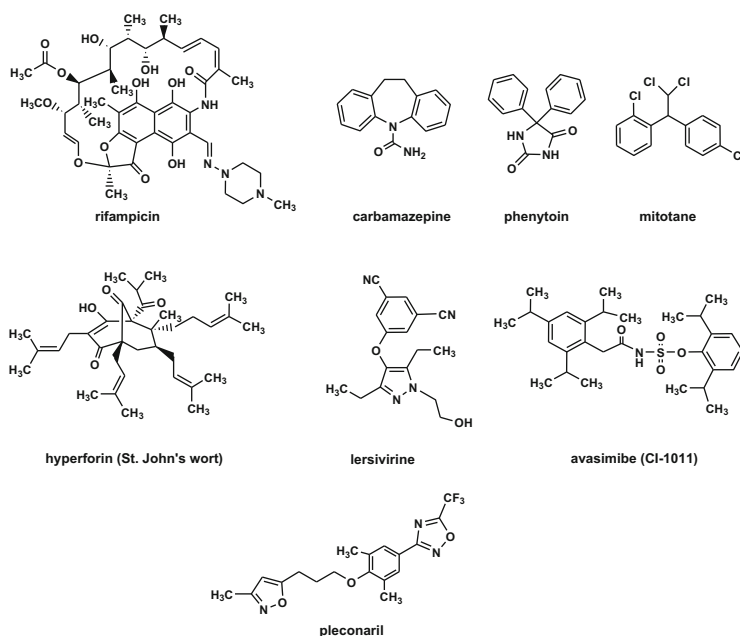
Fig. 1 Structures of PXR and CAR activators that cause CYP3A4 enzyme induction and DDIs

phenytoin, and pleconaril give rise to 96%, 94%, 94%, and 35% reductions, respectively, in the midazolam AUC when coadministered (Fig. 1) [5–8]. In addition to drugs, many natural products also cause DDIs, including the herbal antidepressant St. John's Wort which contains the CYP3A4 inducer hyperforin [9]. Natural products or herbal remedies typically result in highly variable DDI responses due to the varying amount of inducer or inhibitor in each source or lot of material. For example, St. John's Wort has been shown to cause decreases in the AUC of midazolam that range from 21% to 80% and is dependent on the amount of hyperforin in each preparation (Fig. 1) [8]. Progesterone is an example of an endogenous compound that induces drug-metabolizing enzymes [10]. Progesterone levels increase significantly during pregnancy and lead to induction of the phase II conjugating enzyme UGT1A1 (uridine diphosphate glucuronosyltransferase 1A1). Coadministration of the antihypertensive agent labetalol to pregnant woman did not lead to the appropriate efficacy. In this situation, it was found that labetalol plasma exposure in pregnant woman dramatically decreased due to increased metabolism of the drug which is metabolized by UGT1A1 [11].

In addition to the aforementioned scenarios where the perpetrator causes the induction and a second victim drug is affected, there are situations where a single drug acts as both perpetrator and victim, referred to as autoinduction. The anticonvulsant carbamazepine causes CYP3A4 enzyme induction and this induction increases the elimination of carbamazepine itself because carbamazepine is predominately metabolized by CYP3A4 [12]. The antimalarial drug artemisinin also exhibits this

Table 1 List of clinically significant CYP3A4 DDIs (victim drugs) due to enzyme induction after oral administration of rifampicin[a]

Victim drugs	Percent decrease in victim drug AUC[b]
Budesonide, voriconazole, (R)-verapamil, midazolam, (S)-verapamil, nilvadipine, casopitant, mirodenafil, ruboxistaurin, alfentanil, triazolam, levomethadyl, buspirone, telaprevir, nifedipine, brotizolam, simvastatin, ivacaftor, alprazolam, itraconazole, tolvaptan, toremifene, tamoxifen, oxycodone, ticagrelor, lersivirine, praziquantel, gefitinib, propafenone, quinine, quinidine, lurasidone, dienogest, nilotinib, dasatinib, zopiclone, amprenavir, ketoconazole, ramelteon, codeine, dronedarone, roflumilast, atorvastatin, gemigliptin	>80–100
atazanavir, rilpivirine, repaglinide, axitinib, macitentan, methadone, saxagliptin, diazepam, ebastine, (S)-warfarin, imatinib, dextromethorphan, cyclosporine, zolpidem, gliclazide, bupropion, saquinavir, (R)-warfarin, ruxolitinib, mycophenolic acid, propranolol, tacrolimus, zibotentan, mefloquine, lopinavir, dabigatran, rosiglitazone, ethinyl estradiol, ondansetron, etoricoxib, glyburide, celecoxib, tolbutamide, maraviroc, carvedilol, fentanyl, everolimus, ritonavir, caffeine, disopyramide, norethindrone, antipyrine, doxycycline, bosentan, tertatolol, ezetimibe, warfarin, efavirenz, aliskiren, celiprolol, tizanidine, pioglitazone, ranitidine, (S)-bupropion, prednisolone, zidovudine	>50–80
rivaroxaban, vandetanib, trimethoprim, (R)-bupropion, nevirapine, risperidone, deferasirox, novobiocin, lamotrigine, pefloxacin, raltegravir, mexiletine, acetaminophen, losartan, talinolol, (S)-talinolol, glimepiride, bisoprolol, metoprolol, (R)-talinolol, linezolid, repaglinide, moxifloxacin, pravastatin, digoxin, temsirolimus, theophylline, tocainide, morphine, flurbiprofen, cortisol, fluconazole, dicloxacillin, sulfamethoxazole, glipizide, nateglinide, ornidazole	>20–50

[a]Data abstracted from the University of Washington Drug Interaction Database [8]
[b]All compounds are rank ordered by the maximum change reported in the literature

effect by inducing CYP2B6, which is significantly involved in artemisinin elimination [13]. Autoinduction is a phenomenon that can occur anytime a drug induces an enzyme which is also predominately involved in its own metabolic clearance. The enzyme being induced generally needs to be a major elimination pathway for the inducer to have an impact on its own metabolism and efficacy. Therefore, when understanding the overall liability of an enzyme inducer, a complete understanding of the major elimination pathways and enzymes involved in elimination of the perpetrator drug is important to assess potential autoinduction and loss of efficacy.

Although enzyme induction can occur with nearly all drug-metabolizing enzymes and many drug transporters, the elimination pathway with the greatest impact in precipitating a DDI is induction of the phase I oxidative enzyme, CYP3A4. CYP3A4 is involved in the metabolism or elimination of approximately 50% of marketed drugs [14]. Therefore changes in the activity of this enzyme will have an effect on a significant number of drugs currently prescribed. Table 1 lists

Table 2 List of clinically significant CYP3A4 enzyme inducers (perpetrator drugs) changing the AUC of orally administered midazolam[a]

Precipitant (inducer)	Therapeutic class of inducer	Percent change in midazolam AUC[b]
Rifampicin	Antibiotic	−98 S[c]
Mitotane	Antineoplastic	−95 S
Carbamazepine and/or phenytoin	Anticonvulsants	−94 S
Avasimibe	Antilipemic	−94 S
St. John's Wort	Herbal medication	−80 S
Lersivirine	Nonnucleoside reverse-transcriptase inhibitor (NNRTI)	−51 M
Pleconaril	Antiviral	−35 W
Ginseng	Herbal medication	−34 W
Ginkgo	Herbal medication	−34 W
Armodafinil (R-modafinil)	Psychostimulant	−32 W
Desvenlafaxine	Serotonin-norepinephrine reuptake inhibitor (SNRI)	−31 W
Danshen	Herbal medication	−28 W
Clobazam	Benzodiazepine	−28 W
Echinacea	Herbal medication	−27 W
Quercetin	Food product	−24 W
Glycyrrhizin	Herbal medication	−23 W

[a]Data abstracted from the University of Washington Drug Interaction Database [8]
[b]Maximum change reported in the literature
[c]S, strong inducer; M, moderate inducer; W, weak inducer

many of the victim drugs (CYP3A4 substrates) that have exposure decreases when coadministered with the potent CYP3A4 enzyme inducer rifampicin. There are many drugs, such as simvastatin, warfarin, ethinyl estradiol, and cyclosporine, which have significant pharmacological consequences (i.e., loss of efficacy) when co-dosed with rifampicin or other potent CYP3A4 enzyme inducers. Table 2 lists many of the known, clinically relevant CYP3A4 enzyme inducers along with the measured AUC decrease when each inducer is co-dosed with midazolam (a sensitive probe substrate which measures CYP3A4 enzyme activity in vitro and in vivo). The magnitude of each interaction is categorized according to the guideline recommendations of the FDA and EMA which classify strong, moderate, or weak inducers as those that decrease AUC of a victim drug (e.g., midazolam) by >80%, >50–80%, and >20–50%, respectively [15, 16].

All of the aforementioned examples illustrate the end result of enzyme induction, that being a CYP3A4-mediated clinically relevant DDI which has significant consequences for patient safety and efficacy. The following sections describe the mechanisms of CYP450 induction with an emphasis on mechanisms related to induction of CYP3A4 and how to screen and predict CYP3A4 DDIs. Moreover, examples of SAR that provide insight into the nature of CYP3A4 inducers as well as
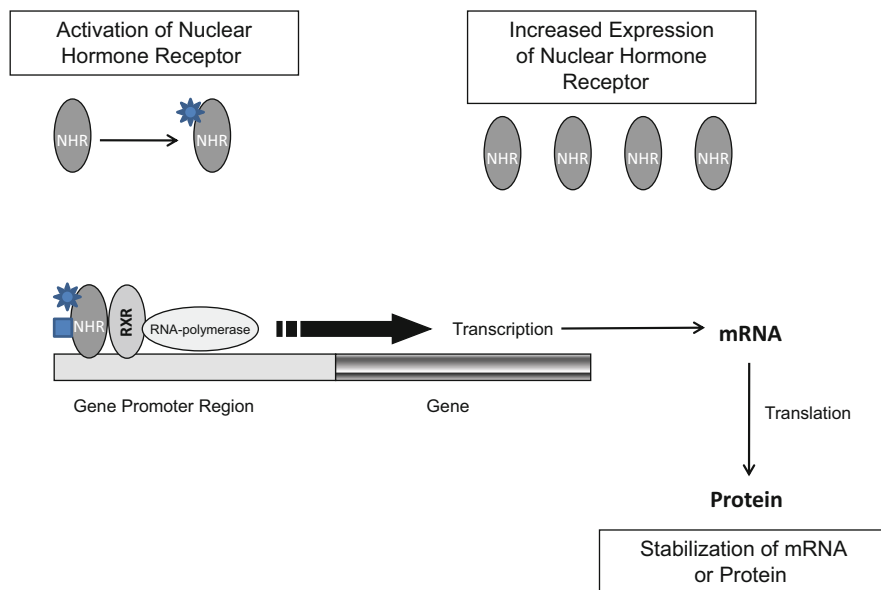
Fig. 2 Illustration of the four mechanisms of enzyme induction: (1) receptor activation, (2) increased expression of NHR, (3) stabilization of mRNA, and (4) stabilization of enzyme

practices that have eliminated or attenuated a compound's propensity for CYP3A4 enzyme induction are described. Decades of research evaluating drugs that inhibit CYP450 enzymes have led to a wealth of useful information that can provide guidance to the medicinal chemist in avoiding this liability and similar information is arising with enzyme induction, in particular for CYP3A4 [17].

2 Mechanisms of CYP450 Enzyme Induction

There are several known mechanisms of enzyme induction: (1) activation of transcription factors, (2) increased expression of transcription factors, and (3) stabilization of mRNA or protein, summarized in Fig. 2. By far the predominant mechanism responsible for inducing the expression of drug-metabolizing enzymes is activation of transcription factors, such as PXR, CAR, and AhR (pregnane X receptor, constitutive androstane receptor, and aryl hydrocarbon receptor, respectively). Table 3 lists the major enzymes/transporters induced by the transcription factors along with their tissue distribution. PXR and CAR are predominately responsible for the increased expression and activity of CYP3A4.

An example of PXR-mediated CYP3A4 induction is illustrated by the drug rifampicin (rifampin) used to treat tuberculosis. As can be seen in Table 2, rifampicin significantly reduces the exposure of CYP3A4 substrates like midazolam (98%). Rifampicin binds to PXR as an agonist and "activates" the receptor. In brief, this activation process involves displacing corepressors bound to PXR and the

Table 3 Transcription factors responsible for induction of drug-metabolizing enzymes and drug transporters and their tissue distribution

Transcription factor	Induced enzymes/transporters	Tissue distribution[a]
PXR	CYP2A6, 2B6, 2C8, 2C9, 2C19, **3A4**, 3A7, 4F12, 7A1; CES2; SULT2A, UGT1A1, 1A3, 1A4, 1A6, GST1A; MDR1, MRP2, MRP3, OATP2	Liver > small intestine > kidney
CAR	**CYP2B6**, 2C8, 2C9, 2C19, 3A4; FMO5; UGT1A1, SULT1A1; MDR1, MRP2, MRP3, MRP4	Liver > kidney > small intestine
AhR	**CYP1A1**, **1A2**, 1B1; UGT1A1, 1A3, 1A4, 1A6, 1A9; NQO1, GSTA1, GSTA2; BCRP	Lung > liver = kidney > small intestine

[a]mRNA expression or transcription factor [25, 78–81]

Bold indicates major enzyme induced

recruitment of coactivators, as well as inducing binding to a heterodimer partner, RXR (retinoid X receptor). This complex then initiates transcription of PXR target genes, as indicated in Table 3 (more details of this mechanism will be described in subsequent sections describing PXR and CAR).

Increased expression of a transcription factor is another mechanism by which CYP3A4 can be induced; however, this is a minor mechanism for CYP3A4 and not a prevalent mechanism for other enzymes/transporters, as summarized in Fig. 2. Pascussi et al. demonstrated that dexamethasone induces CYP3A4 in human hepatocytes via two mechanisms: (1) activation of the PXR receptor at high concentrations of drug, >10 μM and (2) activation of the glucocorticoid receptor (GR) which in turn induces transcription of PXR and ultimately increases the expression of CYP3A4, at low concentrations of dexamethasone (~0.1 μM) [18]. The induction of CYP3A4 by dexamethasone at low concentrations is considered modest with about a four-fold increase in CYP3A4 mRNA expression, whereas the induction response by dexamethasone at higher concentrations results in much greater mRNA expression (~16-fold). In addition to PXR, Li et al. have suggested that increased expression of CAR may be responsible (in part) for the CYP3A4 induction observed for a series of IGF-1R antagonists in human hepatocytes, HepG2, and Huh7 cells [19]. Cells treated with BMS-665351 demonstrated 3–6-fold increases in CAR mRNA expression along with >20-fold increases in CYP3A4 mRNA expression.

The final pair of mechanisms, stabilization of mRNA or protein, is also minor mechanisms for induction of drug-metabolizing enzymes and appears to be most relevant for CYP2E1 (Fig. 2). In the case of CYP2E1, protein and activity are stabilized by organic solvents such as acetone, ethanol, and pyrazole which are all substrates of CYP2E1. As a result of substrate binding to the enzyme, the enzyme is partially protected from endogenous mechanisms of degradation. When the rate of enzyme synthesis remains the same and the rate of degradation is reduced, the enzyme concentration increases, leading to induction. Moreover, when rats are chemically made diabetic, CYP2E1 protein and mRNA are increased. The increased mRNA is not due to increased transcription but due to stabilization of

the RNA itself. These mechanisms of CYP2E1 induction are nicely reviewed in an article by Gonzalez [20].

3 How to Avoid CYP3A4 Enzyme Induction and Drug–Drug Interactions

For centuries, leaders of armies have known that to defeat your enemy you must know your enemy. The better you know your enemy the more you are prepared to defeat them. The same reasoning can be used to avoid CYP3A4 enzyme induction; by understanding the transcription factors involved, how they interact with drugs, and their mechanisms of activation, the better you can develop compounds that do not elicit this liability.

Over the past decade, important advances have been made in our understanding of the mechanisms that regulate the expression of drug-metabolizing enzymes and transporters, in particular CYP3A4. These mechanisms have at their root the transcription factors PXR and CAR which are also known as nuclear hormone receptors. The following sections will describe the characteristics of each receptor, their mechanisms and binding partners, and the xenobiotics (agonists) that bind to each receptor, along with various examples where structure activity information or specific chemical modifications have disrupted the interactions between agonist and receptor thereby eliminating or reducing the liability of CYP3A4 induction.

3.1 Pregnane X Receptor (PXR)

The pregnane X receptor (PXR, NR1I2) is a member of the nuclear receptor superfamily which is ligand-regulated DNA-binding transcription factors. Activation of PXR is the primary mechanism by which CYP3A4 is induced. The activation of PXR and subsequent induction of CYP3A4 and other proteins is considered a defense mechanism by which xenobiotics, such as herbal extracts or drugs taken to treat disease, are recognized as foreign and should be quickly eliminated from the body [9]. Given the role of PXR as a protective mechanism against foreign substances, it is not surprising that the largest concentrations of PXR are found in the intestine, the initial barrier to entering the body, and the liver, the main organ for elimination (Table 3).

The major structural features of the PXR receptor are illustrated in Fig. 3a. At the N-terminus is the activation function 1 (AF-1) region that regulates the receptor in a ligand-independent manner by making functional interactions with transcriptional machinery. The DNA-binding domain (DBD) directs the receptor to its target genes by binding to specific DNA response elements found within the promoter region of target genes (Fig. 3). Near the C-terminus is the ligand binding domain (LBD) which contains the binding pocket for agonists/antagonists and the ligand-dependent activation factor, AF-2, region. In between the DBD and the LBD is a flexible hinge region.
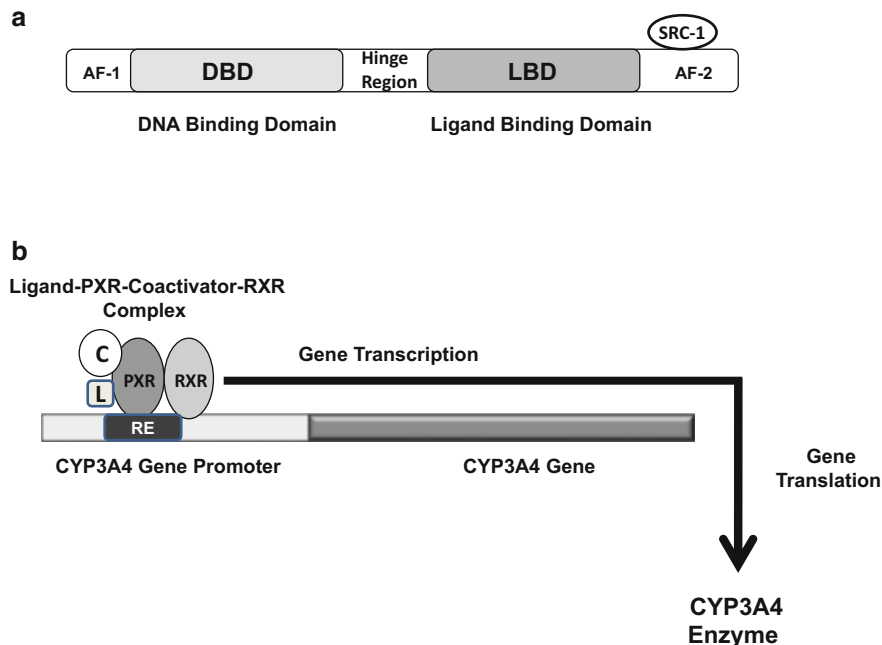
Fig. 3 (**a**) Representation of the pregnane X receptor: AF1 region, DNA-binding domain (DBD), ligand binding domain (LBD), AF-2 region, and SRC-1 coactivator binding site. (**b**) Illustration of complex formation for CYP3A4 transcriptional activation with heterodimer partners PXR and RXR along with ligand (*box*), SRC-1 (*circle*), and RNA polymerase

The key to understanding PXR binding and activation is the structure of the LBD. Like many other nuclear receptors, the PXR LBD is generally hydrophobic; however, in contrast to other receptors, the binding pocket is much larger, more flexible, and can accommodate molecules of varying sizes and shapes. Also uncharacteristic of other receptors are additional amino acid residues lining the bottom of the LBD which have been attributed to the receptor's flexibility and ability to change size. It has been reported that the binding pocket of PXR can change its internal volume from 1,200 to 1,600 A^3 depending on the size of the ligand [21]. These characteristics of the LBD help explain the promiscuous nature of PXR and why many drugs have a propensity to bind to this receptor. Several groups have screened large libraries of drugs or drug-like molecules and found that between 11% and 16% of molecules are reasonably good ligands for PXR. These percentages would suggest that the odds of activating PXR are fairly high and this is a liability that needs to be evaluated during drug discovery campaigns [22–24].

The LBD is elliptical in shape and lined with 28 amino acids of which 20 are hydrophobic, four are polar, and four are potentially charged [25]. The polar residues (Ser208, Ser247, Cys284, Gln285), the charged residues (Glu321, His327, His407, Arg410), and the hydrophobic residues are all fairly well distributed throughout the binding pocket, as summarized in Fig. 4. Although

Fig. 4 Representation of the LBD of PXR and identified interaction points. *Boxes* represent H-bonding interactions and shaded half-circles represent hydrophobic regions
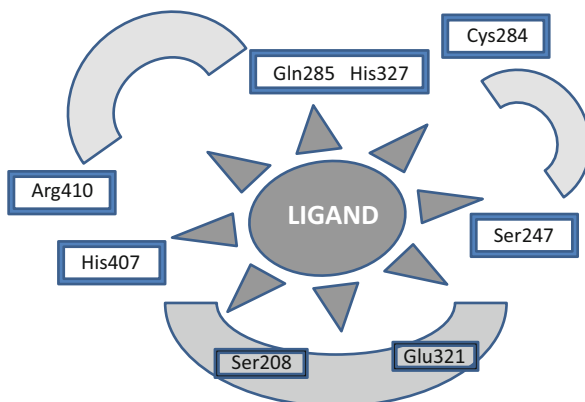
PXR ligands have been shown to interact with many of these LBD amino acids, binding studies with SR12813, hyperforin, and rifampicin have demonstrated that six key amino acids are involved in the binding of all three drugs. These amino acids include hydrogen bonding with Ser207, Gln285, and His407 as well as hydrophobic interactions with Met243, Trp299, and Phe420 [26]. The later sections will illustrate how interactions with these significant amino acid residues alter the binding and activation of PXR.

Another unique aspect of the PXR LBD is the significant differences in the amino acids lining the pocket, which leads to significant differences in ligand-activated responses between animals and humans [27–29]. The amino acid similarity between human PXR and rat, mouse, rabbit, dog, pig, and monkey PXR is 76%, 77%, 82%, 83%, 87%, and 95%, respectively. These differences within the LBD of PXR lead to significant differential binding (or no binding) of ligands between species. In contrast to the LBD, the DBD of PXR is fairly well conserved across multiple animal species with >92% sequence homology [27]. Given the significant differences in binding between species, human-based models, such as human PXR and human hepatocytes, must be employed to accurately assess PXR agonists and human CYP3A4 enzyme induction potential. However, the most pronounced homologies in LBD sequence to human PXR are the rhesus and cynomolgus monkeys which are 95% similar to the human form of PXR. It has been shown that monkey in vitro and in vivo models respond in a similar fashion to human in vitro and in vivo models with known human PXR agonists, indicating that these animal models are useful in vivo tools to assess human CYP3A4 drug–drug interaction potential [28, 30, 31].

Binding of a ligand to the binding pocket is merely the first step in a series of events, including conformational changes of the receptor that ultimately leads to induction of enzymes and transporters. The transcriptional activity of PXR is mediated by nuclear receptor coactivators, such as SRC-1 of the steroid receptor coactivator family which binds to PXR after ligand activation at the AF-2 region [32]. These coactivators contain chromatin-modifying enzymes used to activate transcription by stabilizing the ligand-receptor complex and initiating transcription. In addition to recruitment of coactivators, ligand binding also leads to conformational changes that displace

corepressors, such as nuclear receptor corepressor (NCoR) and silencing mediator of retinoid and thyroid (SMRT) which are responsible for repressing the activity of PXR. Once activated and translocated to the nucleus, the transcriptional activation is further facilitated by binding of the heterodimer partner RXR which aids in DNA binding and transcription. Targeting of PXR to specific genes is mediated by the DBD of PXR and specific nucleotide sequences known as response elements found within the promoter region of specific genes. Once PXR is fully activated and bound to the promoter region of its target gene(s), transcription is initiated leading to an increase in mRNA expression and eventually translation of larger quantities of enzyme, as summarized in Fig. 3b.

In addition to agonists or activators of PXR, there are also inhibitors and antagonists of PXR that exist naturally or are part of the drug pharmacopeia. Many of the inhibitors of PXR mediate their action by interfering with recruitment of coactivators, such as SRC1 and HNF4α (hepatocyte nuclear factor 4-alpha); these include sesamin – a major constituent of sesame seeds, ketoconazole, sulforaphane, and camptothecin [33–36]. Interestingly, camptothecin inhibits PXR activity while its close analog irinotecan (CPT-11) is an agonist of PXR [34]. Coumestrol, which is a potent agonist of the estrogen receptor, has been found to be a modest antagonist of PXR, $IC_{50} = 12$ μM [33]. Other compounds also inhibit PXR such as ecteinascidin-743 (ET-743), a marine-derived antineo-plastic agent, and the HIV protease inhibitor A-792611 but the underlying mechanisms are not yet fully understood [33].

3.1.1   PXR Assays

There are multiple assays that can be employed to assess PXR activation and CYP3A4 enzyme induction that range from simple to complex [37–39]. The simple high-throughput assays generally involve direct measurements of ligand binding or activation of PXR. Common assays that measure the binding of drugs to the PXR LBD include scintillation proximity (SPA) and time-resolved fluorescence resonance energy transfer (TR-FRET), both of which measure the degree of dis-placement of a radiolabeled ligand (^{3}H-SR12813) or acceptor fluorophore (fluorescein), respectively, upon binding of a second agonist [38]. These assays can be extremely high throughput; for example, SPAs are typically run in a 384-well format, while Shukla et al. were able to screen 8,280 compounds at multiple concentrations using a TR-FRET assay in a 1,536-well format [22, 40]. One significant disadvantage of these nuclear receptor binding assays is that they are unable to distinguish between agonists and antagonists. For this reason binding assays are typically used to initially identify possible PXR agonists. Once the agonist properties are confirmed, these binding assays can be used for structure activity relationships (SAR) or rank ordering of agonist affinity to the receptor.

Another simple assay is the PXR transactivation assay. Although there are many variations of the assay, they are all cell based and consist of an expression vector containing response elements from the CYP3A4 promoter region coupled to a

reporter gene (typically luciferase) and a human PXR expression vector. In these assays, ligands enter the cell and bind to the LBD of PXR. This bound complex then recruits coactivators, dislodges corepressors, and binds with its heterodimer partner, RXR, preparing the activated receptor to bind to a target gene promoter. The activated complex initiates transcription of the luciferase gene creating greater quantities of luciferase mRNA and enzyme. The degree of PXR activation is proportional to the amount of luciferase enzyme produced, which is measured by the metabolism of luciferin to oxyluciferin and light.

The general assay is run in 96- or 384-well formats and compounds are tested over a large concentration range in duplicate or triplicate wells. Typically 8–10 concentrations are tested to produce an informative concentration-response plot [38, 41]. The PXR ligand rifampicin (10 μM) is generally included as a positive control for assessment of assay performance and as a benchmark compound for comparator purposes [24, 40]. Appropriate measures should also be implemented as part of the assay to assess possible cytotoxicity and compound solubility. Cytotoxicity or test compound insolubility leads to attenuated or false-negative responses, which will confound discovery optimization by providing misleading SAR and erroneously altering the rank order of compounds. Conveniently, several versions of the PXR transactivation assay have become commercially available (Puracyp and Indigo Biosciences).

Agonist properties are measured by determining the potency (EC_{50}) and efficacy (E_{max}) of a compound. For higher-throughput screening, data can be collected on larger numbers of compounds tested at a single concentration (measuring fold activation compared to a DMSO control or percent activation of a positive control, such as rifampicin). Several methods have been employed to predict CYP3A4 induction potential from PXR transactivation data. In general, these methods involve the use of EC_{50}, E_{max}, and information about the anticipated efficacious drug concentration at steady state (C_{ss}) or maximum systemic concentration (C_{max}) [24, 39, 42, 43].

Overall the PXR binding and transactivation assays are useful tools in identifying compounds that interact with PXR. Further, the PXR transactivation assay performs well when identifying PXR agonists and in the binning or rank ordering of compounds in their ability to increase the transcription of CYP3A4, as well as an informative tool in drug discovery for SAR. However, these assays may not sufficiently recapitulate the in vivo situation in order to accurately predict a DDI that results from PXR-mediated CYP3A4 enzyme induction. Nonetheless, these assays have established themselves as useful liability screens and mechanistic tools for evaluating large numbers of compounds in discovery lead optimization by eliminating or attenuating the potential for CYP3A4 DDIs.

There are several additional models that can be employed to assess PXR agonists and predict more complex PXR-mediated DDIs. Unlike the previously mentioned PXR assays that are simple in nature and directly measure the interactions of agonists with PXR, these more complex models require functional PXR and measure responses that are much farther downstream from the initial PXR binding-activation event. In addition, these models tend to be more physiologically

relevant by employing endogenous cofactors at nominal concentrations, albeit they are generally considered to be low throughput.

Primary human hepatocytes are the most useful and accepted in vitro model to predict PXR-mediated CYP3A4 DDIs. Hepatocytes express many transcription factors such as PXR and CAR that regulate the expression of CYP3A4. The hepatocyte model has been shown to effectively replicate the human in vivo induction response and is recognized as an effective model to predict CYP3A4 induction potential [44–46]. Moreover, only human hepatocytes can be used in this model due to the species differences in the LBD of PXR, as described previously. Although this assay is considered to be low throughput, it can be used to some extent in drug discovery to gain a preliminary assessment of PXR-mediated CYP3A4 induction. Immortalized human hepatocytes (e.g., HepG2, BC2, HepaRG, or Fa2N-4) have also been used for screening purposes as they contain various levels of human PXR expression and offer the advantage of convenience compared to primary human hepatocytes with respect to availability and reproducibility [37, 44, 47].

Thus far, all of the models and assays described have been in vitro in nature, making them useful tools for SAR, rank ordering, and predicting PXR-mediated CYP3A4 DDIs. However, in vitro experiments have limitations, the most significant of which is the static nature of the assays in which drug concentrations do not change over time. This is in contrast to the situation in vivo where drug concentrations rise and fall with time and, as such, drug exposure to PXR changes with time. Several laboratories have evaluated the use of in vivo models (mice and monkeys) to predict DDIs more accurately. Due to species differences mice cannot be used to assess human PXR interactions or induction; however, humanized mice (mice in which the mouse PXR has been eliminated and human PXR inserted) have been shown to be reasonable in vivo models of human PXR response. Several laboratories have been able to demonstrate that in vivo dosing of rifampicin to humanized PXR mice leads to pharmacokinetic changes similar to those found in patients taking rifampicin and CYP3A4 substrates [48, 49]. In addition, monkeys have been proposed as an in vivo model to predict human PXR-mediated CYP3A4 induction. Both the cynomolgus and rhesus monkey PXRs are ~95% homologous to human PXR and demonstrate a high degree of correlation in transactivation properties, as well as induction profiles in primary hepatocytes and in vivo pharmacokinetic changes due to CYP3A induction. Kim et al. demonstrated with rifampicin and hyperforin that both monkey and human PXR transactivation and primary hepatocytes responded in a similar fashion (efficacy and potency) [28]. Also, in vivo induction of CYP3A4 in the cynomolgus monkey by rifampicin and hyperforin demonstrated comparable pharmacokinetic changes to those found in humans [28].

3.1.2   Practical Examples of Attenuating PXR Activation and CYP3A4 Induction

There are several examples in the literature of laboratories using PXR transactivation assay data as a means to identify and attenuate or eliminate PXR activation from a series of drug discovery candidates. The LBD or pocket of human PXR has been described as large, hydrophobic, and flexible. Using this information along with crystal structure data with bound ligands, it has been determined that PXR ligands tend to have a flexible structure, one hydrogen-bond acceptor and several hydrophobic regions that interact with amino acids in the LBDs [50, 51]. Many of the known ligands of PXR express these structural features, which are summarized in Fig. 1, but, unfortunately, so do many drugs. Several strategies to diminish a compound's interaction with the PXR LBD through destabilizing the ligand-receptor interaction have been described by Gao et al. The strategies to reduce PXR interactions include introduction of polar groups to disrupt hydrophobic interactions, increasing steric hindrance (decreasing flexibility) to disrupt specific binding interactions within the LBD, and elimination of key hydrogen-bond acceptors. The following examples from the literature illustrate how these changes, as well as others, can change ligand binding properties with human PXR.

Triazole Derivatives

Using the approach of increasing polarity, Gao was able to eliminate the PXR activation from the series of triazole derivatives depicted in Fig. 5 [50]. PXR docking results indicated that the triazole heterocycle provided a H-bond acceptor that interacts with H407, the *para*-chloro group makes a hydrophobic interaction with Phe429, and the biphenyl moiety interacts with another hydrophobic region that includes Trp299. The transactivation results could have also been explained by increasing basicity with the first three compounds; however, polarity appears to be the driving force for binding disruption as shown by changes to the sulfone derivatives. Increasing the polarity by replacing the terminal phenyl ring of the biphenyl group with a pyridine or imidazole heterocycle or substituting the phenyl ring with an ethylsulfonyl moiety decreased the degree of PXR transactivation relative to the positive control rifampicin. In addition, in the case of the ethylsulfonyl-substituted compounds, steric hindrance may have also played a role in disrupting the drug-PXR interaction. Curiously, deploying the ethylsulfone group in the *para*- or *meta*-position led to no PXR transactivation, while an ethylsulfone at the *ortho*-position resulted in increased PXR transactivation. Additional docking results indicate that although the ethylsulfone moiety increases polarity, when deployed at the *ortho*-position, it introduces a new H-bond interaction with Gln285, thereby restabilizing the compound in the ligand binding pocket. All of these PXR transactivation results were subsequently replicated in
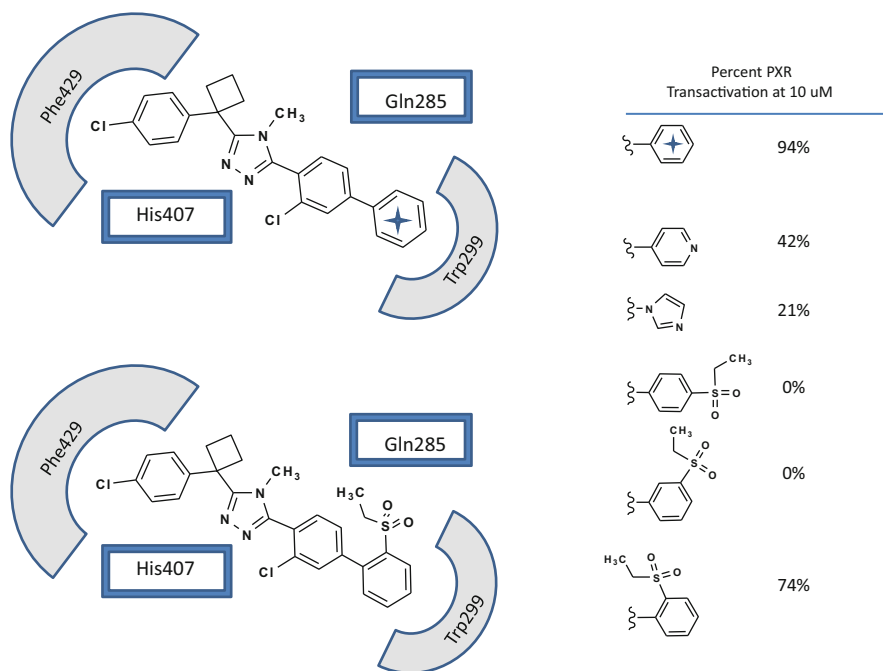
Fig. 5 PXR transactivation results for a series of triazole derivatives and critical interaction points within the PXR LBD

cultures of human hepatocytes with the same trend demonstrated by measuring CYP3A4 mRNA expression and enzyme activity.

T-Type Calcium Channel Antagonists

A series of 4,4-disubstituted quinazolinones as T-type calcium channel antagonists with varying degrees of PXR transactivation (as well as several other metabolic issues) was reported by Schlegal et al. [52]. Compound **1**, Fig. 6, was the starting point for the program and this compound activated PXR to 63% of the rifampicin response at 10 μM. Various modifications, such as halogen substitutions and changes to the ethyl substituent (propyl, nitrile, and halogenation), had little effect on the extent of transactivation; however, compound **2**, which incorporates multiple fluorine substituents, exhibited significantly reduced PXR activation to 35% of control. Further reductions in PXR transactivation were made by the addition of the 4-pyridylmethyl N-oxide moiety as exemplified by compounds **3** and **4**. Likely the added bulk (steric effect) and/or zwitterionic nature of the N-oxide had an effect on the binding affinity of these compounds to PXR. However, combination of the structural modifications that led to reduced PXR transactivation in compounds **2–4** did not lead to further reductions in PXR transactivation when hybridized in **5** but led to an increase in transactivation (56%). In this set of examples, it would appear
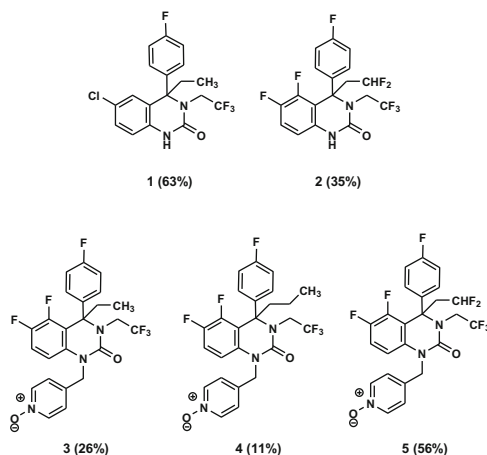
Fig. 6 PXR transactivation results for a series of T-type calcium channel antagonists (% PXR transactivation)
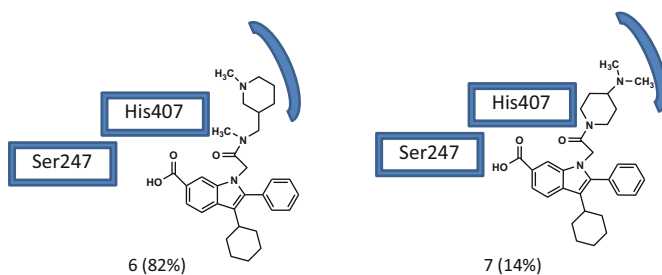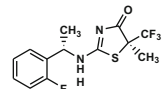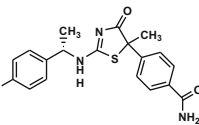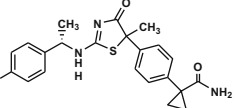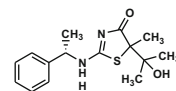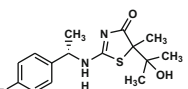
Fig. 7 PXR transactivation results for a series of hepatitis C virus NS5B polymerase inhibitors and key interaction points within the PXR LBD (% PXR transactivation)

that steric and charge differences altered the binding within the binding pocket of PXR. CYP3A4 enzyme induction in primary human hepatocytes was not reported in the article so it is unclear how these PXR results translate to actual DDI potential.

Inhibitors of Hepatitis C Virus NS5B Polymerase

Harper et al. rationalized the PXR transactivation properties of a series of indole-*N*-acetamides being developed as potent allosteric inhibitors of HCV NS5B polymerase [53]. Compound **6** (Fig. 7) is reprehensive of a series of compounds (phenyl shown; *meta*-fluorophenyl and *para*-chlorophenyl not shown) that activated PXR to 82%, 72%, and 92% of the rifampicin response at 10 μM, respectively. Within the series, it was determined that the PXR agonist properties were associated with the acetamide side chain attached at N1 and that by removing the sp^3 carbon link to the piperidine ring, PXR transactivation could be reduced. Compound **7** represents a series of

Fig. 8 PXR transactivation results for a series of thiazole-based 11β-hydroxysteroid dehydrogenase type 1 inhibitors
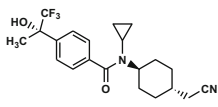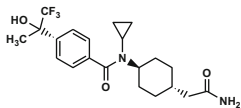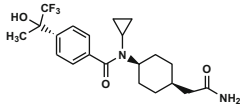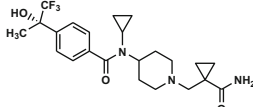
compounds (phenyl as shown; *meta*-fluorophenyl and *para*-chlorophenyl, not shown) that have reduced PXR activation down to 14%, 30%, and 15% of the rifampicin response at 10 μM, respectively. Subsequent molecular modeling by Gao et al. indicated that the carboxylic acid and amide groups were responsible for H-bond interactions with amino acid residues Ser247 and His407 within the LBD, respectively [50]. They also suggested that removal of the sp^3 carbon linker increased the rigidity of the molecule so that the dimethylamine moiety buttresses against the interior of the LBD, moving the molecule out of a favorable binding position to engage Ser247 and His407.

11-Beta-hydroxysteroid Dehydrogenase Type 1 (11β-HSD1) Inhibitors

A series of 2-amino-1,3-thiazol-4(5*H*)-ones were profiled as potent activators of human PXR with **8** (45% PXR transactivation at 20 μM compared to rifampicin) the initial compound used as a starting point for attenuation of the PXR activity (Fig. 8) [54]. By altering several structural features around the molecules, Fotsch et al. found that replacing the trifluoromethyl group with a more bulky and polar phenyl amide or alkyl alcohol group in addition to the introduction of a *para*-fluoro substituent to the phenyl ring had a significant impact on reducing the PXR activity, as exemplified by compound **9** which exhibits only 9% PXR transactivation. Further increases in steric bulk with the phenyl amide however had detrimental effects and significantly increased the degree of PXR transactivation (compound **10** with 97% PXR transactivation). Likely the additional length introduces an interaction with a
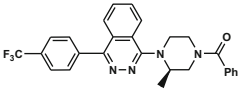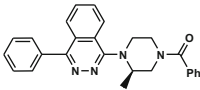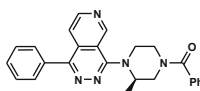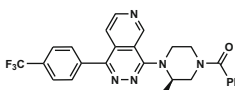
Fig. 9 PXR transactivation results for a series of benzamide-based 11β-hydroxysteroid dehydrogenase type 1 inhibitors (% transactivation)

H-bond donor in PXR. Other changes also affected the PXR activity, such as a halogen substitution to the phenyl ring on the left side of the molecule. Interestingly the unsubstituted phenyl derivative **11** exhibited 68% PXR transactivation. Ultimately **12**, which combines a modest sized tether, a tertiary alcohol, and *para*-fluoro substitution, was found to have negligible PXR activation at just 6% of control. Further studies demonstrated that the reduced PXR activity led to a significant decrease in CYP3A4 mRNA levels and enzyme activity when incubated with primary human hepatocytes but did not entirely eliminate the effects. Nevertheless, compound **12** did have the appropriate balance of PXR transactivation and other biochemical properties for further evaluation.

Rew et al. evaluated the PXR transactivation activity of a series of cyclohexyl and piperidyl benzamide derivatives of 11β-HSD1 inhibitors that are summarized in Fig. 9 [55]. It was determined that increasing the polarity of the cyclohexyl substituent from nitrile (**13**) to an amide (**14**) led to a significant decrease in PXR transactivation from 75% to 5%. In addition, the configuration around the cyclohexyl ring demonstrated dramatic differences in PXR activation. Compounds **14** and **15** with *trans*- and *cis*-configurations demonstrated 5% and 100% PXR transactivation, respectively. Further modification by replacing the lipophilic cyclohexane ring with a more polar piperidine heterocycle and extending the chain linking polar amide group at the 4-position of the ring removed the stereochemical center to afford **16** which was found to have low PXR transactivation (4%) and an appropriate balance of other properties, such as ADME, potency, and cytotoxicity.

Fig. 10 PXR transactivation results for a series of hedgehog signaling pathway antagonists

Antagonists of the Hedgehog Signaling Pathway

Kaizerman et al. evaluated the PXR activity of a series of antagonists of the hedgehog signaling pathway captured in Fig. 10 that started with **17** which exhibited 81% PXR transactivation compared to rifampicin, [56]. The researchers found that increasing the polarity of the left-hand side phenyl ring significantly reduced the PXR transactivation, effects that correlated reasonably well with the logD (pH 7.4). In fact, the simple change of removing the trifluoromethyl group, thereby creating the phenyl analog **18**, diminished PXR transactivation to 30%, while further increases in polarity were able to reduce the PXR transactivation to 3–8%. Unfortunately, changes at this position resulted in overall unfavorable biochemical properties and these molecules could not be progressed. Synthesis of a series of pyrido[3,4-*d*]pyridazines also lead to decreases in PXR transactivation, presumably by increasing polarity in the same region of the PXR LBD. Compounds **19** and **20** (Fig. 10) were found to transactivate PXR 5 and 9%, respectively. Although both compounds had favorable biochemical and PXR responses, compound **20** was found to possess significantly better biochemical properties than **19** and was selected for further evaluation.

Insulin-Like Growth Factor 1 Receptor (IGF-1R) Inhibitors

Zimmermann et al. evaluated the PXR activity of a series of benzimidazole-based IGF-1R inhibitors that are summarized in Fig. 11 [57]. Compound **21** was promising from a potency standpoint until it was determined that its PXR transactivation ranged from 90% to 123% of the rifampicin response. Moreover, CYP3A4 mRNA expression in human hepatocytes was similar to that of rifampicin, indicating that the compound would be expected to cause significant DDIs with coadministered drugs metabolized by CYP3A4. The researchers determined in a
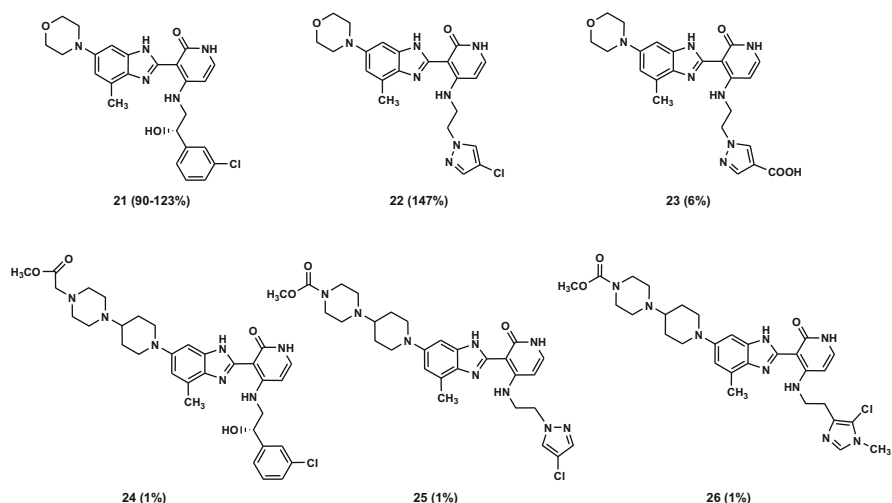
Fig. 11 PXR transactivation results for a series of insulin-like growth factor1 receptor inhibitors (% PXR transactivation)

series of analogs that polarity was a significant player in attenuating the PXR interaction. Starting with compound **22**, which also has high PXR transactivation (147%), and modifying only the pyrazole C4 substituent from chlorine to a series of more polar groups, the PXR transactivation was eventually reduced to 6% in **23**. The reduction in PXR transactivation of this series of compounds correlated well with the logD (pH 7.4) of the C4 substituents. Drawing attention to the other end of the molecule, the morpholine moiety would be expected to interact with a hydrophobic pocket in the LBD. Therefore, using the ideas of creating steric bulk, reducing flexibility and increasing polarity would be expected to lead to attenuation of the PXR interaction. These changes proved to be beneficial in reducing the PXR activity as shown, with compound **24** exhibiting only 1% PXR transactivation (Fig. 11). Unfortunately, **24** did not have the appropriate biochemical properties to advance. Subsequently, further modifications on both ends of the molecules led to analogs **25** and **26** which incorporated the features of polarity and bulk. These compounds showed low PXR transactivation (1%) along with favorable potency toward IGF-1R and improved ADME properties. Further studies with **25** and **26** in primary human hepatocytes led to the unusual discovery that both compounds despite being devoid of PXR activity still demonstrated induction of CYP3A4 mRNA and enzyme activity similar to that of the positive control rifampicin. Although the interaction with PXR was eliminated, the compounds would still be expected to cause significant DDIs with CYP3A4 substrates. This example illustrates why it is important to always confirm that reductions in PXR transactivation lead to significant reductions in CYP3A4 induction in human hepatocytes. This phenomenon is discussed further in the section describing CAR activation.
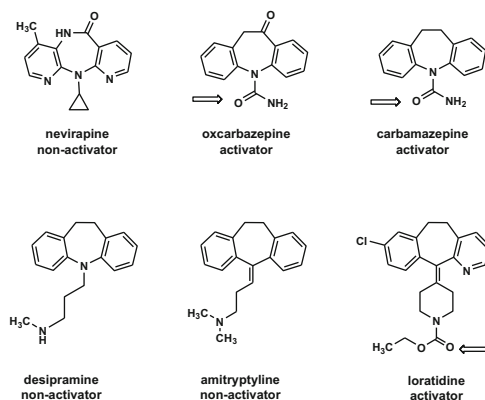
Fig. 12 Structural changes in analogous molecules that modulate PXR activation

Additional Examples of Differential PXR Interactions

Figure 12 illustrates two examples of PXR non-activators and PXR activators in which small structural changes lead to markedly different PXR binding properties [58]. Nevirapine is considered a PXR non-activator, while both carbamazepine and oxcarbazepine are considered PXR activators. The urea carbonyl groups on both carbamazepine and oxcarbazepine were shown by molecular modeling to interact with and form a H-bond with Gln285 in the PXR binding pocket. It was also suggested that the second carbonyl on oxcarbazepine may form a H-bond with His407 but that the analogous carbonyl of nevirapine was unable to properly interact with this residue. In a similar fashion, the carbonyl moiety of the carbamate of loratadine, a PXR activator, is predicted to interact with Gln285 of PXR while the PXR non-activators desipramine and amitriptyline are unable to establish this H-bond interaction. Other subtle structural differences were noted between the PXR activators paclitaxel and hydrocortisol (Fig. 13) after molecular modeling conducted by Schuster and Langer [51]. They noted that paclitaxel and hydrocortisone activated PXR but the structural analogs docetaxel and cortisone were much weaker. They hypothesized that the phenyl group of paclitaxel was able to form π–π stacking interactions with Tyr306 but docetaxel with its *O*-tert-butyl group could not. Similarly, while the hydroxyl moiety of hydrocortisone was able to H-bond within the model, cortisone (even with the carbonyl group having the ability to H-bond) was not in the proper orientation to do so [51].

The species differences in PXR LBDs were discussed previously and the following example illustrates both species and isomeric differences in PXR activation. Wipf et al., while screening a series of amino acid derivatives, found that one of the analogs not only demonstrated a species difference between mouse and human PXR activation but that the two enantiomers of the compound also demonstrated different profiles (Fig. 14). Compound **27** (+ isomer) was shown to significantly activate human PXR while compound **28** (− isomer) was a weak
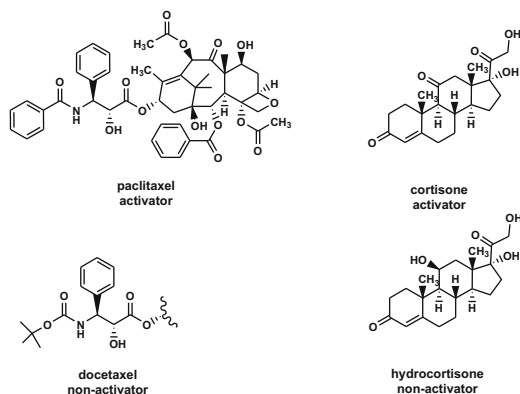
Fig. 13 Structural changes that modulate PXR effects within analogous series. Paclitaxel and cortisone are PXR activators but structural modifications afford docetaxel and hydrocortisone which do not activate PXR
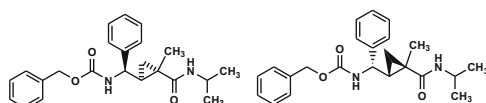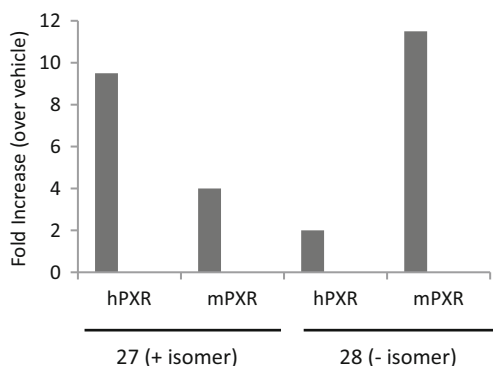
Fig. 14 An example of stereochemical modification which causes both PXR activation changes as well as marked species differences in response

activator of PXR. Conversely, compound **28** significantly activated mouse PXR, while compound **27** was a weak mouse PXR activator [59]. This example is helpful and illustrates that PXR activation may be modulated by different isomers and, assuming that each isomer has similar pharmacological properties, the isomer with lower PXR activation can be selected for further development.

Fig. 15 An illustration of the major features of the CAR receptor (*upper*) and key interaction points within the CAR LBD for the binding of 5β-pregnanedione and CITCO (*lower*)
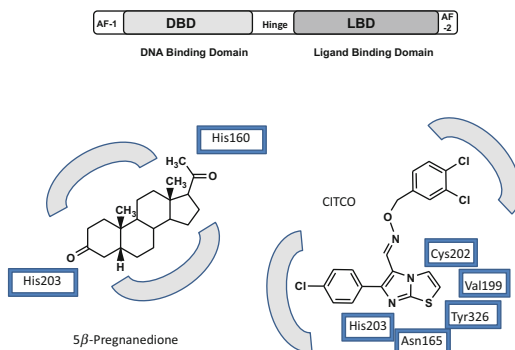
3.2 Constitutive Androstane Receptor (CAR)

The constitutive androstane receptor (CAR, NR1I3) is a member of the nuclear receptor superfamily primarily responsible for the induction of CYP2B6, but also CYP3A4. The term CAR was originally introduced to represent the constitutive activated receptor because of its partnering with the hetereodimer RXR and activation of target gene transcription in the absence of a ligand. Upon further screening for CAR ligands, the two androstane metabolites andostranol and androstenol were identified as endogenous ligands (albeit antagonists) and the name was changed to the constitutive androstane receptor [60]. The largest concentration of CAR is found in the liver with significantly lower amounts in the intestine. A library of 2,000 FDA-approved drugs were initially screened by docking and ligand-based structure activity modeling followed by evaluation in several in vitro models of CAR activation, including the expression of CYP2B6 in human primary hepatocytes. Out of the 2,000 original compounds, 19 had optimal properties based on modeling to interact with CAR and five of those compounds were shown to have in vitro properties consistent with induction of CYP2B6 [61]. The low percent of CAR activators found in this study (0.25%) suggest that the odds of activating CAR are much lower than for PXR; nevertheless, there are a number of known CAR activators that cause significant DDIs such that it is important to evaluate this liability during a drug discovery campaign when induction of CYP2B6 and/or CYP3A4 are observed.

The major features of the CAR receptor are illustrated in Fig. 15. There are great similarities of CAR receptor structure to that of PXR, such as the N-terminus AF-1 region that regulates the receptor in a ligand-independent manner by making functional interactions with transcriptional machinery. The DNA-binding domain (DBD) directs the receptor to its target genes by binding to specific DNA response elements found within the promoter region of target genes. In between the DBD and the ligand binding domain (LBD) is a flexible hinge region. Near the C-terminus is the LBD which contains the binding pocket and the ligand-dependent activation factor, AF-2, region. One significant difference between CAR and PXR is the extremely short AF-2 region of CAR, with only eight amino acids. The short

AF-2 region is implicated in the constitutive activity of the receptor by forming a charge clamp which is closely associated with the core of the CAR LBD.

The activation and LBD of CAR have unique characteristics from many of the other nuclear receptors. Similar to the PXR LBD, the CAR LBD is generally hydrophobic; however, in contrast, the ligand binding pocket is much smaller (675 A^3) and has limited flexibility, and ligand binding leads to little change in the ligand binding domain [62]. The LBD of CAR is round in shape and lined with 31 amino acids. The LBD contains a number of polar and charged residues, such as Cys202, Asn165, Tyr224, Tyr326, His160, and His203, and a number of hydrophobic residues all fairly well distributed throughout the binding pocket. Not much is known about the key CAR residues involved in ligand binding; however, there is some information related to the binding of 5β-pregnanedione and CITCO (6-(4-chlorophenyl)-imidazo[2,1-*b*]thiazole-5-carbaldehyde *O*-(3,4-dichlorobenzyl) oxime) based on crystal structure data. The structure and binding interactions of 5β-pregnanedione are shown in Fig. 15. 5β-Pregnanedione appears to bind most significantly with the two carbonyls engaging in H-bond interactions with His160 and His203, while the remaining hydrophobic interactions border the sides of the molecule [62]. CITCO binding with the CAR LBD is not as well defined; however, both the *para*-chlorophenyl and dichlorophenyl oxime linker regions appear to reside in two separate hydrophobic pockets. The imidazo[2,1-*b*]thiazole ring system appears to be in a more hydrophilic region of the pocket surrounded by several polar residues, such as Asn165*, Val199, Cys202, His203*, and Tyr326* (*possible electrostatic interactions) [63].

As with PXR, there are significant differences in ligand-activated responses of CAR between animals and humans. The amino acid similarity to human CAR when compared to rat, mouse, dog, pig, and monkey CAR is 77%, 73%, 84%, 82%, and 94%, respectively [25]. When comparing the mouse CAR LBD, which is thought to be square in shape, a third of the mouse amino acids are different from the human CAR LBD amino acids [62]. These differences lead to significant differential binding of ligands between species. For example, CITCO is a good ligand for human CAR while it does not activate mouse CAR and TCPOBOP (1,4-*bis* [2-(3,5-dichlorpyridyloxy)]benzene) is a strong mouse CAR activator, but not a human CAR activator. As with PXR, given the significant differences in binding between species, human-based models must be employed to accurately assess CAR activation and ultimately human CYP2B6 and CYP3A4 enzyme induction potential.

As previously discussed, PXR activation is dependent only upon ligand binding, while CAR can be activated by one of two mechanisms: ligand-dependent binding or a ligand-independent mechanism. The activation of CAR can occur by direct ligand binding and translocation, as in the case of CITCO, similar to the examples shown with PXR. The second mechanism does not require ligand binding, but does lead to CAR translocation to the nucleus. For example, phenobarbital does not bind to CAR but does cause significant CAR translocation and induction of CYP2B6 and CYP3A4. This ligand-independent mechanism is thought to involve a phosphorylation-dependent signaling pathway [60]. In both activation mechanisms, CAR is sequestered in the cytoplasm by protein complexes (heat shock protein
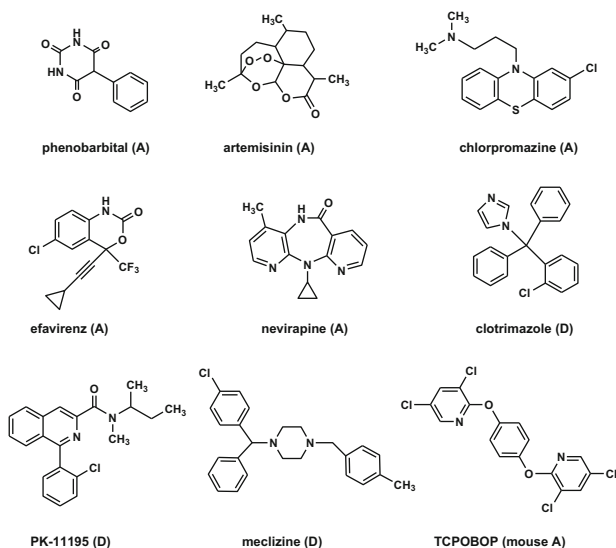
Fig. 16 Structures of compounds that are activators (A) and deactivators or inverse agonists (D) of CAR

90 (Hsp90) and cytoplasmic CAR retaining protein (CCRP)) which prohibit CAR translocation to the nucleus. In the presence of a CAR activator (ligand-dependent or ligand-independent), CAR is dephosphorylated, dissociates from the protein complex retaining it in the cytoplasm, displaces corepressors, recruits coactivators, and heterodimerizes with RXR [64]. Once the CAR complex is properly assembled and translocated to the nucleus, gene transcription can initiate at the appropriate target genes (e.g., CYP2B6 and CYP3A4).

There are a number of CAR activators including CITCO, phenobarbital, artemisinin, phenytoin, carbamazepine, efavirenz, nevirapine, and chlorpromazine, depicted in Figs. 1 and 16. However, due to the constitutive activity of CAR in the absence of a ligand, ligand binding to CAR can result in agonism or inverse agonism (an inverse agonist binds to the receptor and reduces the constitutive activity thereby producing an effect opposite of an agonist). Figure 16 illustrates the structures of several known human inverse agonists (also referred to as CAR deactivators), including clotrimazole, meclizine, and PK11195 [65]. Many of the CAR deactivators, such as SO7662, have been shown to disrupt or eliminate interactions with other CAR components necessary for full activity and result in little to no CAR transcriptional activity [66].

3.2.1 CAR Assays

Whereas CAR has two potential mechanisms leading to translocation (ligand-dependent and ligand-independent), the models used to evaluate CAR are somewhat different than PXR. Simple LBD assays, such as SPA assays, have been

shown to identify ligand-dependent activators of CAR; however, these assays would not be able to detect compounds such as phenobarbital that act through a ligand-independent mechanism [38]. Moreover, the expression of CAR is lost in most immortalized cell lines and transfected wild-type CAR (known as CAR1) spontaneously accumulates in the nucleus due to its strong constitutive activity which prohibits the use of CAR1 in simple, cell-based reporter assays. Whereas these are the primary assay for identification and rank ordering of PXR agonists, these types of assays are not appropriate for wild-type CAR. Nevertheless, several mutant or mutated versions of CAR (CAR3, CAR1+A, CAR(+3aa)) have been shown to have reduced constitutive activity and can be used in cell-based reporter assays with either expressed promoters for CYP2B6 or CYP3A4 [61, 67, 68]. Conveniently, a version of the mutant CAR3 assay has become commercially available (Indigo Biosciences). More recently, however, Kublbeck et al. were able to demonstrate that wild-type CAR1 transfected into a C3A hepatoma cell line was able to effectively respond to CAR activators [69]. The authors surmise that CAR1 was able to retain enough cytoplasmic retention due to endogenous components (or lack thereof) unique to the C3A cell line.

Primary human hepatocytes can also be used to evaluate CAR activators; however, due to the significant overlap of target genes between CAR and PXR, it is sometimes difficult to understand through which mechanism CYP3A4 is being induced [70]. In most situations, CAR predominately activates expression of CYP2B6 with moderate overlap to CYP3A4 while the opposite occurs with PXR agonists. Therefore, the differential expression of these two gene targets can often help identify whether PXR or CAR is the primary driver of an induction effect. A novel tool for screening CAR activators in primary human hepatocytes is the use of an adenovirus expressing enhanced yellow fluorescent protein-tagged CAR1 (Ad/EYFP-hCAR) which infects the hepatocytes with high efficiency [71]. The majority of the Ad/EYFP-hCAR is expressed in the cytoplasm of non-induced hepatocytes and is translocated to the nucleus in response to activators of CAR. Most importantly, this model responds to both direct ligand binding and ligand-independent mechanisms of CAR activation; however, this assay has the disadvantage of seemingly lower throughput. Nevertheless, this assay has become one of the more standard methods to identify human CAR activators. An illustration of the nuclear translocation observed by confocal microscopy is shown in Fig. 17 with DMSO vehicle- and phenobarbital-treated cultured primary human hepatocytes. In addition to the in vitro assay described above, this Ad/EYFP-hCAR has also been used in vivo. For example, $CAR^{-/-}$ mice (CAR-deficient mice) can be injected in the tail vein with the Ad/EYFP-hCAR and subsequently treated with vehicle or various CAR activators by IP injection. After several hours, the livers are removed, frozen, and sliced, and hepatocytes visualized by confocal laser scanning microscopy [67]. Similar to the in vitro assay, this in vivo method demonstrated similar translocation of CAR by activators, but not the vehicle-treated animals or animals treated with the PXR agonist rifampicin.
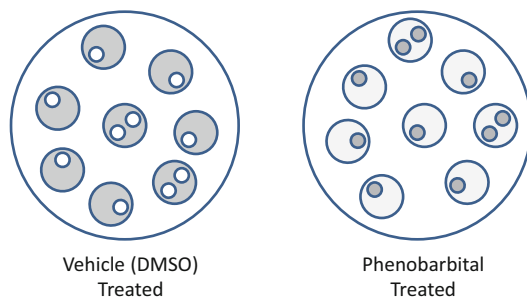
Fig. 17 An illustration of the translocation of CAR from the cytoplasm to the nucleus of primary human hepatocytes as observed by confocal microscopy. The *grey* areas represent CAR localization in cytoplasm of DMSO-treated cells and in the nucleus of phenobarbital-treated cells

3.2.2 Identifying and/or Attenuating CAR Activation and CYP3A4 Induction

In contrast to the situation with PXR ligands, there are few examples of SAR associated with attenuation of CYP3A4 induction by CAR activation, mostly due to a lack of robust assays and the complexity of CAR activation (direct and indirect activation). In addition, fewer crystal structures of human CAR are available for analysis and comparison (two CAR structures and seven PXR structures have been solved) [72]. Finally, most SAR examples involve interactions with rat and mouse CAR that are unlikely to translate to human CAR [73, 74]. Much of the research dedicated to CAR has been in the area of identifying CAR activators with different structural motifs designed to broaden our basic understanding of structure and CAR activation [61, 75, 76]. In a limited example using an in silico approach relying on compound docking, Jyrkkarinne et al. hypothesized that the binding of CAR activators may not be as flexible as PXR agonists. Their research indicates that CAR agonists need to fit into a deep pocket within the LBD in order to be effective activators and a lack of binding to this pocket leads to no or little CAR activation [77]. In addition, as was seen previously in the crystal structures of 5β-pregnanedione and CITCO, interactions with some key amino acid residues are important and disruption of these interactions would be expected to attenuate or eliminate activation of CAR (but only those compounds that bind directly to CAR, not ligand-independent activators).

CYP3A4 Induction by Insulin-Like Growth Factor 1 Receptor (IGF-1R) Inhibitors

Studies with compounds **25** and **26** (Fig. 11) in primary human hepatocytes led to the discovery that both compounds, despite being devoid of PXR activity, still demonstrated induction of CYP3A4 mRNA and enzyme activity similar to that of the positive control rifampicin. This phenomenon is not unusual as there are examples of compounds that either do not bind to PXR or bind weakly, such as

phenobarbital, but cause CYP3A4 induction through other nuclear receptors (e.g., CAR) [24]. Subsequent studies revealed that compound **26** did not cause CYP3A4 induction in hepatocytes through the known mechanisms of CAR (direct ligand binding-activation or ligand-independent translocation). Using both a CAR3 reporter and an Ad/EYFP-hCAR-infected human hepatocyte assay, it was shown that the compound did not directly or indirectly activate or translocate CAR [19]. Moreover, CYP3A4 mRNA expression was always greater than CYP2B6 mRNA expression, which is in contrast to conventional CAR mechanisms. Eventually it was demonstrated in immortalized liver cell lines and primary human hepatocytes that **26** was able to increase the expression of CAR mRNA, as well as CYP3A4 mRNA. The increased expression and transcriptional activity of CAR manifested by **26** in cell-based models could be the cause of the CYP3A4 enzyme induction (similar to the situation of increased expression of PXR and CYP3A4 induction); however, further investigation is necessary to determine if this is the primary mechanism or merely one of several mechanisms by which CYP3A4 is induced by these compounds.

4   Summary

Over the years many drugs and natural products have been found that activate PXR due to the promiscuous nature of its LBD. Fortunately many PXR activators, although they bind to and activate PXR, do so at higher than normal or supra-therapeutic concentrations. Therefore, many compounds have the ability to activate PXR but few demonstrate DDIs at therapeutic drug concentrations [24]. For those drugs that are identified as PXR activators, several strategies to diminish interactions with PXR through destabilizing the ligand-PXR interaction have been described by Gao and others [50]. It is worth repeating those strategies here: (1) introduction of polar groups to disrupt hydrophobic interactions, (2) increasing steric hindrance (decreasing flexibility) to disrupt specific binding interactions within the LBD, (3) elimination of key hydrogen-bond acceptors, and (4) possible stereochemical changes.

References

1. Huang SM, Lesko LJ (2004) Drug–drug, drug–dietary supplement, and drug–citrus fruit and other food interactions: what have we learned? J Clin Pharmacol 44:559–569
2. Huang SM, Strong JM, Zhang L et al (2008) New era in drug interaction evaluation: US Food and Drug Administration update on CYP enzymes, transporters, and the guidance process. J Clin Pharmacol 48:662–670
3. Muirhead GJ, Wulff MB, Fielding A et al (2000) Pharmacokinetic interactions between sildenafil and saquinavir/ritonavir. Br J Clin Pharmacol 50:99–107
4. Obach RS, Ryder TF (2010) Metabolism of ramelteon in human liver microsomes and correlation with the effect of fluvoxamine on ramelteon pharmacokinetics. Drug Metab Dispos 38:1381–1391

5. Backman JT, Olkkola KT, Neuvonen PJ (1996) Rifampin drastically reduces plasma concentrations and effects of oral midazolam. Clin Pharmacol Ther 59:7–13

6. Backman JT, Olkkola KT, Ojala M et al (1996) Concentrations and effects of oral midazolam are greatly reduced in patients treated with carbamazepine or phenytoin. Epilepsia 37:253–257

7. Ma JD, Nafziger AN, Rhodes G et al (2006) Duration of pleconaril effect on cytochrome P450 3A activity in healthy adults using the oral biomarker midazolam. Drug Metab Dispos 34:783–785

8. Levy R (2012) Metabolism and transport drug–drug interaction database, 2005–2012 edn. University of Washington, Washington

9. Chang TK (2009) Activation of pregnane X receptor (PXR) and constitutive androstane receptor (CAR) by herbal medicines. AAPS J 11:590–601

10. Choi SY, Koh KH, Jeong H (2013) Isoform-specific regulation of cytochromes p450 expression by estradiol and progesterone. Drug Metab Dispos 41:263–269

11. Jeong H, Choi S, Song JW et al (2008) Regulation of UDP-glucuronosyltransferase (UGT) 1A1 by progesterone and its impact on labetalol elimination. Xenobiotica 38:62–75

12. Bertilsson L, Hojer B, Tybring G et al (1980) Autoinduction of carbamazepine metabolism in children examined by a stable isotope technique. Clin Pharmacol Ther 27:83–88

13. Simonsson US, Jansson B, Hai TN et al (2003) Artemisinin autoinduction is caused by involvement of cytochrome P450 2B6 but not 2C9. Clin Pharmacol Ther 74:32–43

14. Wilkinson GR (2001) Pharmacokinetics: the dynamics of drug absorption, distribution, and elimination. In: Hardman J, Limbird L, Goodman Gilman A (eds) The pharmacological basis of therapeutics, 10th edn. McGraw-Hill, New York

15. Ema (2012) Guideline on the investigation of drug interactions. Guideline on the investigation of drug interactions [Online]. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf Accessed September 2012

16. FDA (2012) Guidance for industry: drug interaction studies-study design, data analysis, implications for dosing, and labeling recommendations. Guidance for industry: drug interaction studies-study design, data analysis, implications for dosing, and labeling recommendations [Online]. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292362.pdf

17. Kumar S, Sharma R, Roychowdhury A (2012) Modulation of cytochrome-P450 inhibition (CYP) in drug discovery: a medicinal chemistry perspective. Curr Med Chem 19:3605–3621

18. Pascussi JM, Drocourt L, Gerbal-Chaloin S et al (2001) Dual effect of dexamethasone on CYP3A4 gene expression in human hepatocytes. Sequential role of glucocorticoid receptor and pregnane X receptor. Eur J Biochem 268:6346–6358

19. Li L, Sinz MW, Zimmermann K et al (2012) An insulin-like growth factor 1 receptor inhibitor induces CYP3A4 expression through a pregnane X receptor-independent, noncanonical constitutive androstane receptor-related mechanism. J Pharmacol Exp Ther 340:688–697

20. Gonzalez FJ (2007) The 2006 Bernard B. Brodie Award Lecture. Cyp2e1. Drug Metab Dispos 35:1–8

21. Timsit YE, Negishi M (2007) CAR and PXR: the xenobiotic-sensing receptors. Steroids 72:231–246

22. Shukla SJ, Nguyen DT, Macarthur R et al (2009) Identification of pregnane X receptor ligands using time-resolved fluorescence resonance energy transfer and quantitative high-throughput screening. Assay Drug Dev Technol 7:143–169

23. Shukla SJ, Sakamuru S, Huang R et al (2011) Identification of clinically used drugs that activate pregnane X receptors. Drug Metab Dispos 39:151–159

24. Sinz M, Kim S, Zhu Z et al (2006) Evaluation of 170 xenobiotics as transactivators of human pregnane X receptor (hPXR) and correlation to known CYP3A4 drug interactions. Curr Drug Metab 7:375–388

25. Di Masi A, De Marinis E, Ascenzi P et al (2009) Nuclear receptors CAR and PXR: molecular, functional, and biomedical aspects. Mol Aspects Med 30:297–343

26. Poso A, Honkakoski P (2006) Ligand recognition by drug-activated nuclear receptors PXR and CAR: structural, site-directed mutagenesis and molecular modeling studies. Mini Rev Med Chem 6:937–947
27. Jones SA, Moore LB, Shenk JL et al (2000) The pregnane X receptor: a promiscuous xenobiotic receptor that has diverged during evolution. Mol Endocrinol 14:27–39
28. Kim S, Dinchuk JE, Anthony MN et al (2010) Evaluation of cynomolgus monkey pregnane X receptor, primary hepatocyte, and in vivo pharmacokinetic changes in predicting human CYP3A4 induction. Drug Metab Dispos 38:16–24
29. Kliewer SA, Goodwin B, Willson TM (2002) The nuclear pregnane X receptor: a key regulator of xenobiotic metabolism. Endocr Rev 23:687–702
30. Nishimura M, Koeda A, Suganuma Y et al (2007) Comparison of inducibility of CYP1A and CYP3A mRNAs by prototypical inducers in primary cultures of human, cynomolgus monkey, and rat hepatocytes. Drug Metab Pharmacokinet 22:178–186
31. Prueksaritanont T, Kuo Y, Tang C et al (2006) In vitro and in vivo CYP3A64 induction and inhibition studies in rhesus monkeys: a preclinical approach for CYP3A-mediated drug interaction studies. Drug Metab Dispos 34:1546–1555
32. Carnahan VE, Redinbo MR (2005) Structure and function of the human nuclear xenobiotic receptor PXR. Curr Drug Metab 6:357–367
33. Chen T (2008) Nuclear receptor drug discovery. Curr Opin Chem Biol 12:418–426
34. Chen Y, Tang Y, Robbins GT et al (2010) Camptothecin attenuates cytochrome P450 3A4 induction by blocking the activation of human pregnane X receptor. J Pharmacol Exp Ther 334:999–1008
35. Huang H, Wang H, Sinz M et al (2007) Inhibition of drug metabolism by blocking the activation of nuclear receptors by ketoconazole. Oncogene 26:258–268
36. Lim YP, Ma CY, Liu CL et al (2012) Sesamin: a naturally occurring lignan inhibits CYP3A4 by antagonizing the pregnane X receptor activation. Evid Based Complement Alternat Med 2012:242810
37. Mcginnity DF, Zhang G, Kenny JR et al (2009) Evaluation of multiple in vitro systems for assessment of CYP3A4 induction in drug discovery: human hepatocytes, pregnane X receptor reporter gene, and Fa2N-4 and HepaRG cells. Drug Metab Dispos 37:1259–1268
38. Raucy JL, Lasker JM (2010) Current in vitro high throughput screening approaches to assess nuclear receptor activation. Curr Drug Metab 11:806–814
39. Sinz M, Wallace G, Sahi J (2008) Current industrial practices in assessing CYP450 enzyme induction: preclinical and clinical. AAPS J 10:391–400
40. Zhu Z, Kim S, Chen T et al (2004) Correlation of high-throughput pregnane X receptor (PXR) transactivation and binding assays. J Biomol Screen 9:533–540
41. Zhu Z, Puglisi J, Connors D et al (2007) Use of cryopreserved transiently transfected cells in high-throughput pregnane X receptor transactivation assay. J Biomol Screen 12:248–254
42. Cui X, Thomas A, Gerlach V et al (2008) Application and interpretation of hPXR screening data: Validation of reporter signal requirements for prediction of clinically relevant CYP3A4 inducers. Biochem Pharmacol 76:680–689
43. El-Sankary W, Gibson GG, Ayrton A et al (2001) Use of a reporter gene assay to predict and rank the potency and efficacy of CYP3A4 inducers. Drug Metab Dispos 29:1499–1504
44. Chu V, Einolf HJ, Evers R et al (2009) In vitro and in vivo induction of cytochrome p450: a survey of the current practices and recommendations: a pharmaceutical research and manufacturers of America perspective. Drug Metab Dispos 37:1339–1354
45. Fahmi OA, Ripp SL (2010) Evaluation of models for predicting drug–drug interactions due to induction. Expert Opin Drug Metab Toxicol 6:1399–1416
46. Hewitt NJ, De Kanter R, Lecluyse E (2007) Induction of drug metabolizing enzymes: a survey of in vitro methodologies and interpretations used in the pharmaceutical industry–do they comply with FDA recommendations? Chem Biol Interact 168:51–65
47. Sinz M, Kim S (2006) Stem cells, immortalized cells, and primary cells in ADMET assays. Drug Discov Today Technol 3:79–85

48. Hasegawa M, Kapelyukh Y, Tahara H et al (2011) Quantitative prediction of human pregnane X receptor and cytochrome P450 3A4 mediated drug-drug interaction in a novel multiple humanized mouse line. Mol Pharmacol 80:518–528
49. Kim S, Pray D, Zheng M et al (2008) Quantitative relationship between rifampicin exposure and induction of Cyp3a11 in SXR humanized mice: extrapolation to human CYP3A4 induction potential. Drug Metab Lett 2:169–175
50. Gao YD, Olson SH, Balkovec JM et al (2007) Attenuating pregnane X receptor (PXR) activation: a molecular modelling approach. Xenobiotica 37:124–138
51. Schuster D, Langer T (2005) The identification of ligand features essential for PXR activation by pharmacophore modeling. J Chem Inf Model 45:431–439
52. Schlegel KA, Yang ZQ, Reger TS et al (2010) Discovery and expanded SAR of 4,4-disubstituted quinazolin-2-ones as potent T-type calcium channel antagonists. Bioorg Med Chem Lett 20:5147–5152
53. Harper S, Avolio S, Pacini B et al (2005) Potent inhibitors of subgenomic hepatitis C virus RNA replication through optimization of indole-N-acetamide allosteric inhibitors of the viral NS5B polymerase. J Med Chem 48:4547–4557
54. Fotsch C, Bartberger MD, Bercot EA et al (2008) Further studies with the 2-amino-1,3-thiazol-4(5H)-one class of 11beta-hydroxysteroid dehydrogenase type 1 inhibitors: reducing pregnane X receptor activity and exploring activity in a monkey pharmacodynamic model. J Med Chem 51:7953–7967
55. Rew Y, Mcminn DL, Wang Z et al (2009) Discovery and optimization of piperidyl benzamide derivatives as a novel class of 11beta-HSD1 inhibitors. Bioorg Med Chem Lett 19:1797–1801
56. Kaizerman JA, Aaron W, An S et al (2010) Addressing PXR liabilities of phthalazine-based hedgehog/smoothened antagonists using novel pyridopyridazines. Bioorg Med Chem Lett 20:4607–4610
57. Zimmermann K, Wittman MD, Saulnier MG et al (2010) SAR of PXR transactivation in benzimidazole-based IGF-1R kinase inhibitors. Bioorg Med Chem Lett 20:1744–1748
58. Khandelwal A, Krasowski MD, Reschly EJ et al (2008) Machine learning methods and docking for predicting human pregnane X receptor activation. Chem Res Toxicol 21:1457–1467
59. Wipf P, Gong H, Janjic JM et al (2007) New opportunities for pregnane X receptor (PXR) targeting in drug development. lessons from Enantio- and species-specific PXR ligands identified from a discovery library of amino acid analogues. Mini Rev Med Chem 7:617–625
60. Wang H, Lecluyse EL (2003) Role of orphan nuclear receptors in the regulation of drug-metabolising enzymes. Clin Pharmacokinet 42:1331–1357
61. Lynch C, Pan Y, Li L et al (2012) Identification of novel activators of constitutive androstane receptor from FDA-approved drugs by integrated computational and biological approaches. Pharm Res 30:489–501
62. Zhou X, Xu H (2009) Structure and function of PXR and CAR. In: Xie W (ed) Nuclear receptors in drug metabolism. Wiley, Hoboken
63. Xu RX, Lambert MH, Wisely BB et al (2004) A structural basis for constitutive activity in the human CAR/RXRalpha heterodimer. Mol Cell 16:919–928
64. Kachaylo EM, Pustylnyak VO, Lyakhovich VV et al (2011) Constitutive androstane receptor (CAR) is a xenosensor and target for therapy. Biochemistry (Mosc) 76:1087–1097
65. Anderson LE, Dring AM, Hamel LD et al (2011) Modulation of constitutive androstane receptor (CAR) and pregnane X receptor (PXR) by 6-arylpyrrolo[2,1-d][1,5]benzothiazepine derivatives, ligands of peripheral benzodiazepine receptor (PBR). Toxicol Lett 202:148–154
66. Kublbeck J, Jyrkkarinne J, Molnar F et al (2011) New in vitro tools to study human constitutive androstane receptor (CAR) biology: discovery and comparison of human CAR inverse agonists. Mol Pharm 8:2424–2433
67. Faucette SR, Zhang TC, Moore R et al (2007) Relative activation of human pregnane X receptor versus constitutive androstane receptor defines distinct classes of CYP2B6 and CYP3A4 inducers. J Pharmacol Exp Ther 320:72–80

68. Kanno Y, Inouye Y (2010) A consecutive three alanine residue insertion mutant of human CAR: a novel CAR ligand screening system in HepG2 cells. J Toxicol Sci 35:515–525
69. Kublbeck J, Laitinen T, Jyrkkarinne J et al (2011) Use of comprehensive screening methods to detect selective human CAR activators. Biochem Pharmacol 82:1994–2007
70. Pascussi JM, Gerbal-Chaloin S, Duret C et al (2008) The tangle of nuclear receptors that controls xenobiotic metabolism and transport: crosstalk and consequences. Annu Rev Pharmacol Toxicol 48:1–32
71. Li H, Chen T, Cottrell J et al (2009) Nuclear translocation of adenoviral-enhanced yellow fluorescent protein-tagged-human constitutive androstane receptor (hCAR): a novel tool for screening hCAR activators in human primary hepatocytes. Drug Metab Dispos 37:1098–1106
72. Wu B, Li S, Dong D (2013) 3D structure and ligand specificity of nuclear xenobiotic receptors CAR, PXR and VDR. Drug Discov Today 18:574–581
73. Pustylnyak V, Yarushkin A, Kachaylo E et al (2011) Effect of several analogs of 2,4,6-triphenyldioxane-1,3 on constitutive androstane receptor activation. Chem Biol Interact 192: 177–183
74. Repo S, Jyrkkarinne J, Pulkkinen JT et al (2008) Ligand specificity of constitutive androstane receptor as probed by induced-fit docking and mutagenesis. J Med Chem 51:7119–7131
75. Dring AM, Anderson LE, Qamar S et al (2010) Rational quantitative structure-activity relationship (RQSAR) screen for PXR and CAR isoform-specific nuclear receptor ligands. Chem Biol Interact 188:512–525
76. Kublbeck J, Jyrkkarinne J, Poso A et al (2008) Discovery of substituted sulfonamides and thiazolidin-4-one derivatives as agonists of human constitutive androstane receptor. Biochem Pharmacol 76:1288–1297
77. Jyrkkarinne J, Windshugel B, Ronkko T et al (2008) Insights into ligand-elicited activation of human constitutive androstane receptor based on novel agonists and three-dimensional quantitative structure-activity relationship. J Med Chem 51:7181–7192
78. Kohle C, Bock KW (2009) Coordinate regulation of human drug-metabolizing enzymes, and conjugate transporters by the Ah receptor, pregnane X receptor and constitutive androstane receptor. Biochem Pharmacol 77:689–699
79. Li H, Wang H (2010) Activation of xenobiotic receptors: driving into the nucleus. Expert Opin Drug Metab Toxicol 6:409–426
80. Pavek P, Dvorak Z (2008) Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. Curr Drug Metab 9:129–143
81. Tolson AH, Wang H (2010) Regulation of drug-metabolizing enzymes by xenobiotic receptors: PXR and CAR. Adv Drug Deliv Rev 62:1238–1249

Strategies for Minimisation of the Cholestatic Liver Injury Liability Posed by Drug-Induced Bile Salt Export Pump (BSEP) Inhibition

J. Gerry Kenna, Simone H. Stahl, and Tobias Noeske

Abstract Hepatobiliary uptake and efflux transporter proteins play key roles in the formation of bile, which is a vital function of the liver. The ATP-dependent bile salt export pump (BSEP) excretes bile salts from hepatocytes into bile. Inherited BSEP mutations in humans cause intrahepatic accumulation of bile salts, which results in cholestatic liver injury. Furthermore, inhibition of BSEP activity is considered one of a number of key initiating mechanisms by which drugs may cause liver injury (drug-induced liver injury, DILI) in the human population. DILI is an important cause of serious drug-induced illness and is a leading cause of drug attrition during development and of drug withdrawal and restrictive labelling post-marketing. In this chapter we summarise the evidence that BSEP inhibition is a drug-related DILI risk factor, we describe experimental approaches (in silico, in vitro and in vivo) which may be used to predict and quantify this process during drug discovery and development and we discuss data interpretation. We also outline an approach by which assessment of BSEP inhibition in drug discovery can be used to reduce the likelihood that DILI may arise during development. In addition, we consider the current state of computational predictive modelling of BSEP inhibition and discuss the influence of physicochemical parameters.

Keywords Bile acids, Bile salt export pump, Cholestasis, Computational modelling, Drug-induced liver injury

---

J.G. Kenna (✉)
Safety Science Consultant, Macclesfield, UK
e-mail: jgerrykenna@gmail.com

S.H. Stahl
DMPK, Drug Safety and Metabolism, AstraZeneca R&D Alderley Park, Macclesfield, Cheshire, UK

T. Noeske
Discovery Safety, Drug Safety and Metabolism, AstraZeneca R&D Mölndal, Mölndal, Sweden

Contents

Abbreviations

ABC	ATP-binding cassette
ALT	Alanine aminotransferase
BCRP	Breast cancer resistance protein
BRIC2	Benign recurrent intrahepatic cholestasis type 2
BSEP/Bsep	Bile salt export pump
CDF	5(6)-Carboxy-2′,7′-dichlorofluorescein
CLF	Cholyllysylfluorescein
DILI	Drug-induced liver injury
IADR	Idiosyncratic adverse drug reaction
ICP	Intrahepatic cholestasis of pregnancy
MDCK	Madin–Darby canine kidney
MDR	Multidrug resistance
MRP	Multidrug resistance-associated protein 2
NTCP	Sodium taurocholate co-transporting polypeptide
OATP	Organic anion transporting polypeptide
OPLS-DA	Orthogonal partial least-squares projection to latent structures discriminant analysis
P-gp	P-glycoprotein, ABCB1
PFIC2	Progressive familial intrahepatic cholestasis type 2
QSAR	Quantitative structure–activity relationship
SCH	Sandwich-cultured hepatocytes
SLC	Solute carrier
TAP	Transporter associated with antigen processing

1 Introduction

1.1 *Role of BSEP in Bile Formation and Bile Flow*

The bile salt export pump (BSEP in humans, ABCB11; Bsep in animal species, Abcb11) is a member of the MDR/TAP subfamily of the ATP-binding cassette (ABC) transporter superfamily. It is a liver-specific membrane protein which is expressed on the apical domain of the hepatocyte plasma membrane, plays a key role in bile formation [1] and is highly conserved throughout evolution [2]. Bile is a digestive fluid that is produced within the liver and is released into the duodenum (following its storage and concentration in the gallbladder, in humans and many other animal species), where it is responsible for emulsification and absorption of lipids and fat soluble nutrients. Its principal components are bile acids, cholesterol, phospholipids, conjugated bile pigments, inorganic electrolytes and water [3]. Bile acids comprise a large family of molecules, which are produced via cytochrome P450-mediated oxidation of cholesterol within the liver to primary bile acids that may be conjugated in the liver to form glycine or taurine conjugates (termed "bile salts") and/or may be dehydroxylated to secondary bile acids by intestinal bacteria [3]. BSEP mediates active efflux of bile acids from hepatocytes into bile and is the primary transporter responsible for bile acid-dependent bile flow. BSEP acts in concert with other ABC transporters present on the apical plasma membrane domain of hepatocytes, which mediate efflux into bile of cholesterol, phospholipids and bile pigments. These include multidrug resistance-associated protein 2 (MRP2; ABCC2), which transports many different conjugated endobiotics and drug conjugates [4]; P-glycoprotein (P-gp/MDR1; ABCB1), which has broad substrate specificity and transports some endobiotics and many different drugs [5]; and the phospholipid translocator (MDR3 in humans, previously termed MDR2; ABCB4) [6]. Whereas MRP2 and P-gp mediate the efflux of numerous drugs or drug conjugates from hepatocytes into bile, BSEP is not considered to play a major role in drug disposition.

In humans, the primary bile acids are cholic acid and chenodeoxycholic acid, although numerous other bile acid species are also produced and excreted in bile. Markedly different bile acid compositions have been observed in other animal species and some species differences in BSEP/Bsep substrate kinetics have been described [1, 7]. Following their excretion from hepatocytes, bile acids and other biliary components are reabsorbed from the intestinal lumen and then transported back to the liver via the portal vein, where they re-enter hepatocytes via facilitated carriers located on the basolateral domain of the plasma membrane. This process is termed enterohepatic circulation and involves plasma membrane transporters expressed on hepatocytes, biliary epithelial cells and enterocytes [7, 8]. Transport of bile acids into hepatocytes across the basolateral plasma membrane domain is mediated by the sodium taurocholate co-transporting polypeptide (NTCP; SLC10A1) and organic anion transporting polypeptides (OATPs; SLCOs), which are members of the solute carrier (SLC) family [8]. This is illustrated schematically

Fig. 1 Major transporters which mediate bile formation and vectorial bile flow in human hepatocytes
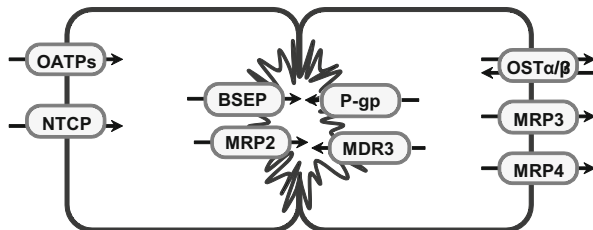
in Fig. 1, which also highlights that the ABC transporters MRP3 (ABCC3) and MRP4 (ABCC4) are expressed on the basolateral plasma membrane domain of hepatocytes.

MRP3 and MRP4 are structurally related to MRP2 and are expressed at low levels in hepatocytes under normal physiological conditions. However, when bile flow is impaired, their expression is markedly increased, as is the expression of BSEP and the enzymes involved in bile acid formation, while expression of NTCP and OATPs is decreased [7–9]. Regulation of bile formation and biliary transporter expression is mediated via both transcriptional and post-transcriptional mechanisms and is believed to play a protective role in enabling hepatocytes to maintain sub-toxic concentrations of bile acids, which are cytotoxic at high concentration, and other biliary components [7–9]. Transcriptional regulation of bile acid transporter expression is mediated by ligand interactions between bile acids and nuclear hormone receptors, most notably via the farnesoid X receptor, and also involves crosstalk with other nuclear hormone receptors [9].

1.2 Impaired BSEP Activity and Human Cholestatic Liver Disease

Impairment of bile flow is termed cholestasis and results in accumulation of bile acids and other biliary constituents within the liver. In humans, an especially important mechanism of cholestasis is due to genetically inherited defects in BSEP. The most severe functional consequence is progressive familial intrahepatic cholestasis type 2 (PFIC2). This is characterised by early onset cholestasis soon after birth and subsequent progressive degenerative liver injury, which is fatal unless treated by liver transplantation [10]. PFIC2 is a consequence of mutations or single nucleotide polymorphisms (SNPs) in the BSEP gene which result in either reduced levels of mRNA transcription or translation or reduced protein stability or activity [10–12]. The resulting liver injury has been attributed to intracellular accumulation of cytotoxic bile constituents [13]. Less functionally severe *ABCB11* gene mutations result in benign recurrent intrahepatic cholestasis type 2 (BRIC2) and in intrahepatic cholestasis of pregnancy (ICP), which are characterised by non-progressive cholestasis [14, 15]. Studies undertaken in cell lines transfected with rat Bsep variants which incorporated human PFIC2, BRIC2 and ICP mutations

revealed a correlation between the observed clinical phenotype and the amount of mature protein expressed at the cell surface, which was attributed to rapid intracellular degradation of the variant proteins[16, 17].

1.3 Bsep Knockout Mouse

When fed a normal diet, $Bsep^{-/-}$ knockout mice were found to exhibit mild cholestasis and markedly altered bile composition (reduced bile acid levels, high levels of tetra-hydroxylated bile acids which were not present in bile from wild-type mice, plus increased cholesterol and phospholipid content) [18]. In addition, the $Bsep^{-/-}$ mice exhibited marked adaptive up-regulation of P-gp and less marked up-regulation of Mdr2 and Mrp2, which was presumed to enable them to cope with Bsep deficiency. This was because severe cholestatic liver injury was observed when the animals were fed a bile acid-enriched diet [19], or when knockout mice which were deficient in both Bsep and P-glycoprotein were fed a normal diet [20]. Furthermore, when the original strain of $Bsep^{-/-}$ mice was backcrossed with C57BL/6J mice for ten generations to produce C57BL/6 $Bsep^{-/-}$ mice, the resulting animals developed progressive cholestatic liver injury which resembled PFIC2 in humans when fed a normal diet [21]. The onset of cholestasis in C57BL/6 $Bsep^{-/-}$ mice was preceded by impaired hepatic fatty acid β-oxidation and elevated levels of expression of various anti-oxidative response genes. This suggested that oxidative stress, caused by elevated fatty acid oxidation which results in the generation of reactive oxygen species, could have contributed to the mechanism of liver injury in C57BL/6 $Bsep^{-/-}$ mice [21].

1.4 Drug-Induced Liver Injury (DILI)

Toxicity to the liver is a relatively frequent finding during preclinical safety testing of potential candidate drugs in animals. Typically, liver injury observed in animals exposed to test compounds is a dose-dependent and reproducible finding. If the safety margin vs. the expected human efficacious dose is low and the type of liver injury is considered likely to raise clinical concern were it to occur in humans (e.g., acute hepatocellular necrosis), compound termination prior to clinical evaluation may occur. In addition, many drugs do not cause evidence of liver damage in preclinical species yet may cause liver injury in humans. For the vast majority of drugs, this type of liver injury occurs infrequently and only in certain susceptible drug-treated patients, does not exhibit overt dose dependence and is termed idiosyncratic DILI [22, 23]. The pattern of liver injury may be cholestatic, hepatocellular or mixed hepatocellular/cholestatic [23].

The most concerning clinical outcome of idiosyncratic DILI is acute life-threatening liver failure, which arises when large proportions of hepatocytes are
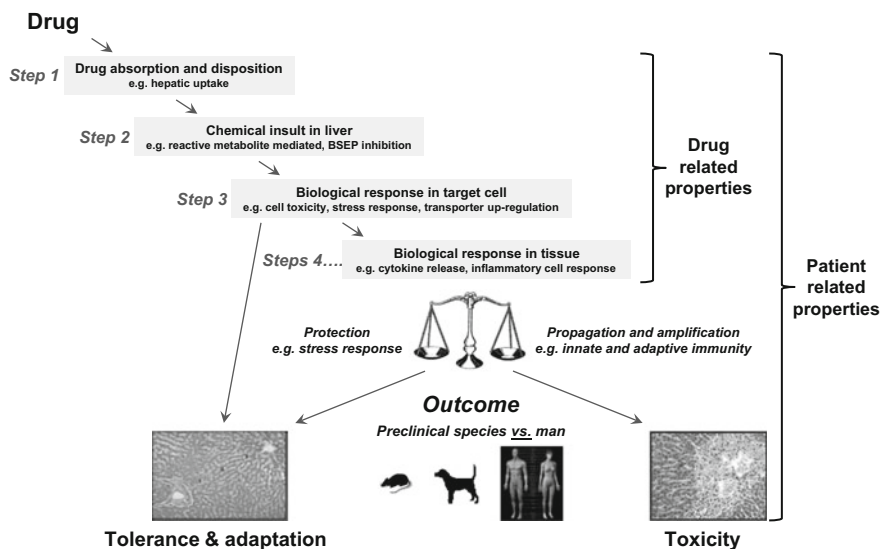
Drug

Fig. 2 DILI is complex and involves multiple contributory mechanisms

damaged. However, numerous drugs cause acute or chronic damage to cells of the liver which results in symptomatic liver injury, but not liver failure, and some drugs cause asymptomatic and mild liver damage which results in abnormal serum clinical chemistry (most commonly, elevated concentrations of the enzyme alanine aminotransferase (ALT) which is released from damaged hepatocytes). Idiosyncratic DILI is a major cause of serious illness in man and, when detected during drug development, is a leading cause of terminated development or failed registration of otherwise promising new therapies [24]. Furthermore, idiosyncratic DILI caused by licensed drugs is an important cause of drug withdrawal or cautionary and restrictive drug labelling [23, 25].

Extensive investigations of the mechanisms which underlie DILI in animals and humans, and the basis of susceptibility, have been undertaken and have revealed that this is a complex multi-step process which involves both drug-related and patient-related factors. This is illustrated in Fig. 2.

Drug-related processes determine whether or not individual drugs have the propensity to cause DILI in vivo, in preclinical test species or in humans. These arise via interaction of drugs and/or their metabolites with biochemical processes within liver cells, which initiate and propagate tissue injury or trigger protective responses. Notable examples include drug metabolism within hepatocytes to chemically reactive intermediates, impairment of mitochondrial function, initiation of oxidative stress and activation of inflammatory processes and cell stress responses [26–29]. Patient-related processes play a major role in influencing whether or not drug-related processes result in DILI in vivo. These include genotype, underlying

disease state, co-medications which cause drug–drug interactions, adaptive immune responsiveness, gender, age, ethnicity and various other demographic factors [30, 31].

1.5 BSEP Inhibition by Drugs

Numerous methods have been used to investigate modulation by drugs of BSEP/ Bsep activity in vitro and to assess functional consequences arising from modulation of the activity of the transporter in preclinical species and humans in vivo. These are summarised and discussed in Sect. 2. Data obtained using these approaches have provided evidence that BSEP inhibition is an important drug-related adverse property which contributes to cholestatic idiosyncratic DILI caused by many drugs in humans. This is discussed in Sect. 3, as are approaches which can be used to assess and minimise the potential liability posed by BSEP inhibition during drug discovery and drug development. The current status of BSEP structure–activity modelling is reviewed in Sect. 4. Key gaps in our current understanding and future needs are highlighted in Sect. 5.

2 Tools for Evaluation of BSEP Function and Its Inhibition by Drugs

2.1 In Vitro Tools

A wide variety of different experimental systems may be used to assess inhibition of BSEP/Bsep activity by drugs in vitro. These include membrane vesicles prepared from ex vivo liver samples or from cell lines transfected with BSEP/Bsep from various species [32–43], primary hepatocytes cultured in sandwich configuration [44, 45] and polarised epithelial Madin–Darby canine kidney (MDCK) cell lines co-transfected with Ntcp and Bsep and cultured as monolayers [46].

2.1.1 Membrane Vesicles

Two assay formats may be used to assess effects of test compounds on BSEP/Bsep activity in membrane vesicles. These are the ATPase assay and the vesicular transport assay. The most frequently used approach is the vesicular transport assay. Typically, this utilises inside-out membrane vesicles prepared from cells transfected with a plasmid vector which includes the appropriate cDNA sequence and quantifies the ATP-dependent accumulation of a probe substrate into inside-out

Fig. 3 Assessment of
BSEP inhibition using
membrane vesicles
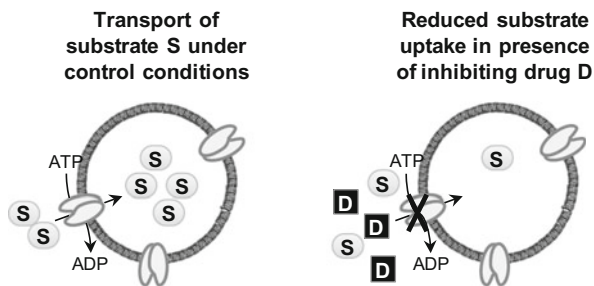
vesicles in the absence and presence of test compound. The assay principle is
illustrated in Fig. 3.

Inhibition of BSEP transport activity in the presence of the test compound results
in reduced uptake of the probe substrate into the vesicles, while use of vesicles from
mock-transfected cells and studies undertaken using AMP in place of ATP enable
correction for possible artefacts which are unrelated to inhibition of BSEP activity.
Membrane vesicle assays enable direct determination of concentration–response
relationships, determination of apparent IC_{50} values and estimation of key kinetic
parameters (i.e. K_m, K_i, V_{max}). In addition, the approach can be adapted to 96-well
plate format, enabling convenient and rapid generation of data with substantial
numbers of compounds. For BSEP, the most commonly used probe substrate is
taurocholate, which is one of the important endogenous substrates of this trans-
porter and is available commercially in its tritiated form as $[^3H]$-taurocholate.
A detailed description of the vesicle methodology is provided in recent reviews
and detailed protocols are available [47–49]. This approach has been used to
investigate a large number of drugs and endogenous molecules in order to
identify compounds which interfere with BSEP function, to study the relationship
between BSEP inhibition and DILI and to explore structure–activity relationships
[32, 38, 42, 43, 50]. Typical concentration–response curves obtained with
inhibitory and non-inhibitory drugs are shown in Fig. 4.

Typically, insect cells such as *Spodoptera frugiperda* (e.g., *Sf* 9 or *Sf* 21) or High
Five are used as BSEP expression systems as they are easy to manipulate, the
required technologies are well established and large quantities of membranes can be
generated. Vesicles prepared from transfected mammalian cells (e.g., HEK) can
also be used but typically give a lower yield of vesicles, although approaches for
larger-scale vesicle preparation have been described [51]. It has been found that
cholesterol content can influence the activity of BSEP and other plasma membrane
transporters and use of vesicles prepared from mammalian cells provides a more
relevant membrane lipid composition than is provided by insect cell membranes
[41, 52]. However, a study of BSEP inhibition by a small set of compounds in
insect-derived vesicles with and without cholesterol supplementation demonstrated
no significant differences in transport inhibition potencies between the two types of
vesicles [52]. Other investigators also observed no significant differences when a
variety of test compounds were evaluated for transport and transporter interactions
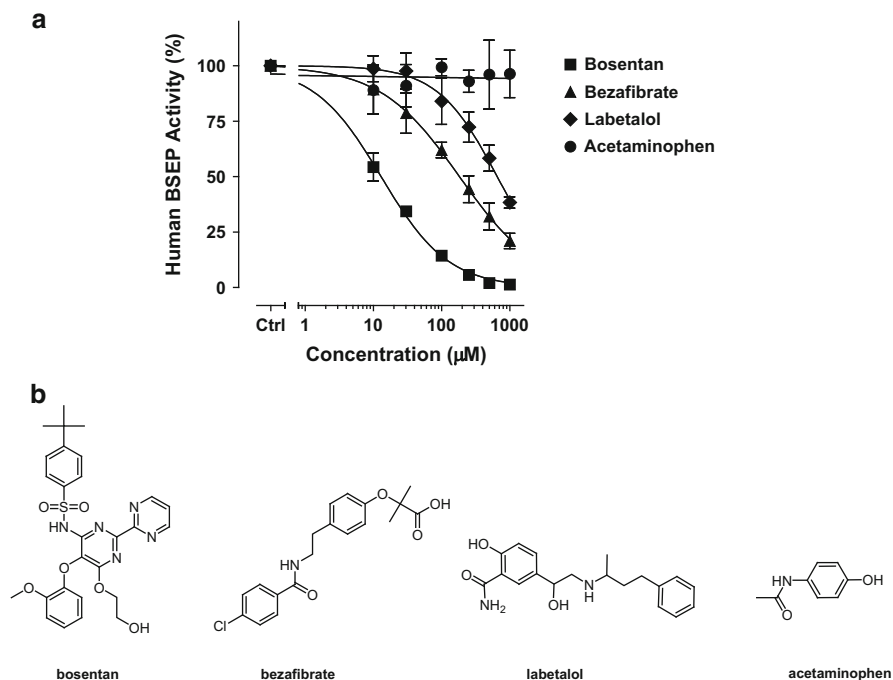
Fig. 4 (**a**) Illustrative concentration–response data from a BSEP membrane vesicle inhibition assay with four drugs (from [50], reprinted with permission from American Society for Pharmacology and Experimental Therapeutics). (**b**) Chemical structures of the drugs

using insect and mammalian cell-derived vesicles which expressed MRP2 or BCRP (the breast cancer resistance protein; BCRP/ABCG2) [53].

Vesicular transporter studies may also be undertaken using apical plasma membrane vesicles prepared from liver tissue [54]. Such membranes express multiple hepatic transporters at physiologically relevant levels in a native membrane environment, although their preparation is technically challenging. The potential value that can be obtained from their use is exemplified by studies undertaken with estradiol-17β-glucuronide, which plays an important role in human intrahepatic cholestasis of pregnancy [55]. This compound has been shown to inhibit Bsep activity in rat liver canalicular membrane vesicles and in membrane vesicles from insect *Sf* 9 cells which had been co-transfected with both Bsep and Mrp2. However, inhibition of Bsep activity by estradiol-17β-glucuronide was not observed when studies were undertaken using membranes from insect cells transfected with Bsep alone, or liver canalicular membrane vesicles from Mrp2 deficient rats [32]. It was inferred that estradiol-17β-glucuronide trans-inhibits Bsep only after its secretion into bile canaliculi, which is mediated by Mrp2 [32].

Since BSEP transporter activity is coupled to ATP hydrolysis, BSEP interactions may also be assessed indirectly by quantification of inhibition of ATPase activity. To distinguish between enzymatically mediated and spontaneous ATP hydrolysis,

inhibition of vanadate-sensitive ATPase activity is assessed. While apparent IC_{50} values can differ between ATPase and vesicle systems, in a study undertaken by Kis et al. [41], the rank order of inhibition potencies was similar between the two systems for murine Bsep. The ATPase assay is suitable for high-throughput applications and does not require the use of radioactivity. However, relatively few investigators have used this approach to explore potential BSEP inhibition by drugs since it provides only indirect data indicative of compound-induced modulation of BSEP activity.

In addition to their use for evaluation of BSEP inhibition, membrane vesicles are a versatile tool which enable investigation of transport of compounds by BSEP. This has been demonstrated for a range of endogenous substrates [35, 36] and for some drugs, e.g., pravastatin and vinblastine [56, 57], although overall BSEP is not considered to play an important role in drug disposition. Another area where vesicle assays are especially useful is in investigations of the impact of BSEP single nucleotide polymorphisms [58, 59] on transport activity and its modulation by test compounds.

However, it is important to be aware that compounds which are highly permeable may return false-negative results in vesicle substrate assays, since they may diffuse out of the vesicles passively. In addition, partitioning into membrane lipids may pose problems for highly lipophilic compounds.

2.1.2   Polarised Epithelial Cells Transfected with BSEP

Polarised cells which express BSEP on the apical domain of the plasma membrane, in the correct orientation, provide a uniquely valuable tool for investigation of the role of BSEP in cellular efflux of test compounds. In order to avoid artefacts which may arise due to other transporters, it is important to select cells which have low endogenous expression of relevant transporters (MRP2, P-gp) and to undertake "negative control" studies with mock-transfected cells. However, when using cell lines transfected with BSEP to investigate the role of the transporter in vectorial transport of conjugated bile salts and other endogenous substrates, it is important to ensure that the cells also express relevant basolateral solute carriers which mediate their cellular uptake (which in hepatocytes is mediated by NTCP and OATP, as described previously) since these substrates have low passive permeability across cell membranes. This limitation can be avoided by co-expression of BSEP and sinusoidal uptake transporters in polarised cell lines. Cells transfected with multiple genes have been used to investigate the role of sinusoidal uptake and apical efflux transporters other than BSEP in hepatic uptake and clearance of a variety of drugs and also to explore the interdependency between transporters and drug-metabolising enzymes. Notable examples include MDCK cell lines which co-expressed both the uptake transporter OATP1B3 (SLC21A8, formerly termed OATP8) and the efflux transporter MRP2 [60] and, more recently, triply or quadruply transfected MDCK cell lines which co-expressed OATP1B1, UDP-glucuronosyltransferase 1A1, MRP2 and cytochrome P450 3A4 [61, 62].

A general note of caution when interpreting data obtained using such models is that the levels of expression of proteins derived from the transfected genes may differ markedly from their expression levels in the liver in vivo.

Doubly transfected cells which co-express human NTCP and BSEP, or their rat orthologues, in MDCK cells [46] and porcine kidney-derived LLC-PK1 cells [63, 64], have been used to investigate the bile acid substrate specificity of these transporters and to assess transport of fluorescently tagged bile acids. Due to the polarised expression of NTCP and BSEP in these cells, transwell devices may be used to study basal-to-apical transport and apical efflux of probe substrates. For the probe substrate taurocholate, both parameters were shown to be reduced when cells were exposed to drugs which cause cholestasis (e.g., rifampicin, cyclosporine A) [64]. This approach has the potential to provide information on the effects of drugs on transporter function under conditions which are more physiologically relevant than those utilised in isolated membrane vesicle assays. However, this cell model system does not permit precise quantification of the potency of BSEP inhibition unless the intracellular concentration of the test compound is determined.

2.1.3 Cultured Hepatocytes

Although hepatocytes can be isolated from the livers of numerous animal species and humans at high yield and with high viability, this process results in loss of polarised expression of basolateral and apical plasma membrane transporters. Therefore, freshly isolated hepatocytes cannot be used to explore vectorial transport processes. This limitation can be circumvented by culturing isolated hepatocytes in 3-dimensional sandwich configuration. The term "sandwich" refers to the culture configuration achieved when hepatocytes are plated on collagen-coated plates and then overlaid with either gelled collagen or Matrigel™. Under such conditions, hepatocytes re-polarise and express functionally active transporter proteins at the corresponding plasma membrane domains. Apical domains are localised to plasma membrane domains between adjacent hepatocytes, which are surrounded by impermeable tight junctions that result in the formation of sealed "canalicular pockets," while basolateral domains are expressed at the interface between hepatocytes and the cell overlay. Methods for preparation of sandwich-cultured hepatocytes (SCH) derived from primary hepatocytes (cryopreserved or freshly isolated) of a range of species, including human and rat, have been established. In addition to their expression of functionally active basolateral and apical transporter proteins, SCH express some phase 1 and 2 drug-metabolising enzymes [65]. Therefore they comprise a physiologically relevant model which allows the assessment of various important hepatic drug elimination pathways (uptake, metabolism, sinusoidal efflux, biliary efflux) in one in vitro system [65].

The technique used most frequently to quantify apical efflux transporter activity in SCH requires evaluation of compound efflux from the cells in buffer which contains Ca^{2+}/Mg^{2+} (where tight junctions are sealed) and in Ca^{2+}/Mg^{2+}-free buffer (which disrupts protein–protein interactions between connexins and releases
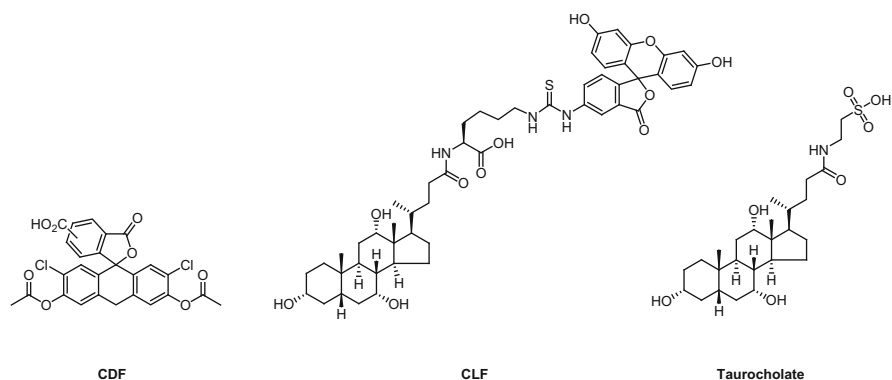
Fig. 5 Chemical structures of taurocholate and the probe substrates 5(6)-carboxy-2′,7′-dichlorofluorescein (CDF) and cholyllysylfluorescein (CLF)

material otherwise enclosed within sealed canalicular pockets). Subtraction of the extent of compound efflux into culture medium containing Ca^{2+}/Mg^{2+} (i.e. transport mediated by basolateral transporters) from the extent of efflux in Ca^{2+}/Mg^{2+}-free medium (i.e. transport mediated by both basolateral and apical transporters) enables indirect quantification of apical transporter-mediated cell efflux. A close correlation has been observed between the extents of apical excretion of a variety of test compounds in this model in vitro and biliary excretion of the compounds in vivo, in bile duct cannulated rats [66]. Furthermore, numerous drugs which were shown to inhibit BSEP activity in inverted membrane vesicle assays have been shown to inhibit apical excretion of the BSEP probe substrate [^{3}H]-taurocholate from rat and human SCH in vitro. Examples include cyclosporine A, CI-1034, glibenclamide, the macrolide antibiotics troleandomycin and erythromycin estolate and nefazodone [67, 68].

Direct assessment of apical transporter activity in SCH may be undertaken by live cell imaging of apical efflux of fluorescent probe substrates. Such studies have been undertaken using the probe substrates 5(6)-carboxy-2′,7′-dichlorofluorescein (CDF), which is added to cells in its membrane permeable, non-fluorescent diacetate form and is cleaved intracellularly to release fluorescent CDF [69, 70], and cholyllysylfluorescein (CLF), which is a fluorescent analogue of the natural bile salt cholylglycine that exhibits in vivo clearance in humans similar to that of endogenous bile acids [71–73]. The chemical structures of CDF and CLF, and of the endogenous BSEP substrate taurocholate, are shown in Fig. 5.

Both CDF and CLF have been shown to accumulate in sealed canalicular pockets between hepatocytes following their administration to SCH in medium containing Ca^{2+}/Mg^{2+}, but not in Ca^{2+}/Mg^{2+}-free medium. In addition, canalicular excretion of CLF from human or rat SCH was found to be inhibited in the presence of a variety of compounds which inhibit BSEP/Bsep activity (cyclosporine A, troglitazone, tiapride or troleandomycin) [44, 73]. In the case of CLF, this inhibition could be quantified by use of a high-content cell imaging algorithm [73].

The available data indicate that biliary efflux of both CDF and CLF from hepatocytes is mediated primarily by Mrp2 [69, 70, 74], even though earlier data had suggested that biliary CLF efflux was mediated by BSEP [71, 72]. Alternative fluorescent probes have been described which exhibit high affinity for BSEP (e.g., [75]), although these have yet to be used in SCH imaging studies.

SCH provide higher throughput and have much lower compound requirements than in vivo biliary transport studies. They also enable cross-species investigations of biliary transport and transporter interactions using a convenient system, which is more physiologically relevant than inverted membrane vesicle assays or transfected cell expression models. In addition, since SCH express drug-metabolising enzymes, they may be suitable for the investigation of transporter interactions caused by metabolites in addition to parent compounds. However, studies with SCH require a substantial resource and expertise and are dependent upon access to plateable primary hepatocytes of high quality.

2.2   Ex Vivo and In Vivo Tools

Several approaches can be used to investigate effects of test compounds on bile formation and BSEP/Bsep activity ex vivo or in vivo. These methods are used to explore and understand functional consequences arising from transporter interactions.

2.2.1   Isolated Perfused Liver

The isolated perfused liver [76] is an ex vivo model which maintains the hepatic architecture and zonal nature of the liver, as well as the natural composition of parenchymal and non-parenchymal cells and cell–cell interactions, and is not influenced by extra-hepatic factors. This approach enables assessment of hepatic uptake, metabolism and the biliary and basolateral efflux of test compounds and has been utilised to investigate drug clearance in liver from mouse, rat and other preclinical species. The technique requires experience and surgical skills and a number of key parameters need to be controlled carefully to ensure that reproducible results are obtained (e.g., perfusion buffer, flow rates, temperature, oxygenation). Typically, blood vessels (either venous alone or venous and arterial) are cannulated and connected to a perfusion apparatus and the bile duct is also cannulated, enabling measurement of bile flow and biliary excretion. The perfusion medium can be administered in a recirculating or a single pass mode and contains probe substrate (e.g., $[^3\text{H}]$-taurocholate, when exploring bile acid clearance) plus test compound. Effects on specific transporters can be investigated to some degree, although since the preparation retains the complement of transporters expressed in the liver, the data that can be obtained are determined by the transporter specificity of the probe substrate used. Examples of BSEP inhibitors which have

been investigated in the isolated rat perfused liver model are troglitazone, which was found to cause a marked decrease in bile flow due to its cholestatic action [77], and rifamycin SV, which inhibited the biliary excretion of taurocholic acid [78].

2.2.2 Bile Duct Cannulation

Cannulation of the common bile duct (or gall bladder) permits in vivo quantification of biliary clearance of drugs, their metabolites or endogenous molecules and therefore enables direct assessment of effects of test compounds on bile flow and bile composition. Bile duct cannulation is used most frequently in mice and rats [79, 80], but can also be performed in a range of other species such as dog [81], monkey [82], rabbit [83], mini pig [84] and humans [85]. In the preclinical setting, the procedure requires surgical preparation under anaesthesia. Experiments in anaesthetised animals can be performed for a few hours, whereas studies in conscious animals can be undertaken for durations that vary from several days up to months. The conscious model requires externalisation of catheters, to allow sample collection and free movement of animals, and administration of bile acids is required to avoid bile acid depletion and to maintain normal digestive function in studies of longer duration. Typically, bile flow is measured gravimetrically, by determining the amount of bile collected within short defined time intervals over the duration of the experiment, while analysis of bile composition can provide detailed and valuable mechanistic insights into effects arising from administration of test compounds (e.g., [86–88]). However, the technique is technically challenging and labour intensive and its value can be limited by marked inter-subject variability.

2.2.3 Plasma Biomarkers of In Vivo BSEP Inhibition

Elevated plasma or serum bile acid concentrations have been observed following administration of compounds which inhibit BSEP/Bsep in vivo. The total pool of 3-hydroxy bile acids may be measured in these studies using an enzymatic clinical chemistry assay. Troglitazone and cyclosporine A are examples of cholestatic drugs which cause a rapid and dose-dependent increase in plasma bile acid levels after intravenous administration to rats [34]. Oral administration to rats of nefazodone (a potent BSEP inhibitor which has a black box warning for liver injury) caused a transient increase in plasma bile acid levels one hour post-dose, which was not observed when rats were dosed with the structurally related drug buspirone (a less potent BSEP inhibitor than nefazodone that has not been reported to cause liver injury but has been associated with infrequent serum transaminitis, which suggests that mild liver injury may occur) [68]. Test compounds which have been reported to inhibit BSEP/Bsep and to cause in vivo bile acid elevations are summarised in Table 1 and their chemical structures are illustrated in Fig. 6.

Table 1 Examples of in vivo plasma bile acid elevations caused by test compounds which inhibit BSEP

Compound	BSEP inhibition in vitro (data are IC_{50} values in μM from vesicles, unless indicated otherwise)	Species exhibiting bile acid increase in vivo	Label (where other outcome not available)	References
Cyclosporine A	0.5 (h) 0.6 (r)	Rat	Warnings: hepatotoxicity	[34, 43]
Troglitazone	2.7 (h) 10.6 (r)	Rat	Withdrawn due to liver injury	[34, 43, 89]
Bosentan	38.1 (h) 30.6 (r)	Human, rat	Black box warning for liver injury	[33]
CI-1034	90% inhibition at 50 μM (h, SCH)	Rat	Development stopped, transaminitis in phase 2 studies	[44]
Glibenclamide	5.3 (h) 2.8 (r)	Rat	Adverse reaction: cholestatic jaundice and hepatitis	[34, 44, 89]
Nefazodone	4.2 (h) 17.4 (r) 100% inhibition at 100 μM (h, SCH)	Rat	Black box warning for liver injury	[43, 68]
Buspirone	104.5 (h) 369.8 (r) 10% inhibition at 100 μM (h, SCH)	Rat	Adverse reaction: infrequent transaminitis	[43, 68]
CP-724,714	16 (h) 70–80% inhibition at 50 μM (h, SCH)	Dog	Discontinued from clinical development due to liver injury	[90]
Chemokine receptor antagonist	129.7 (r)	Rat	Discontinued from clinical development due to rapid in vivo clearance	[91]

h human, *r* rat, *SCH* sandwich-cultured hepatocytes

These data raise the possibility that evaluation of serum or plasma bile acids might provide a useful noninvasive biomarker of in vivo BSEP/Bsep inhibition, which can aid in establishing the rank order of cholestatic potential of test compounds and might, therefore, be valuable if used to aid compound selection during drug discovery [92]. However, since transporters other than BSEP also play important roles in bile acid clearance (e.g., the hepatic uptake transporters NTCP and OATP1B1), inhibition of such transporters also has the potential to cause elevated
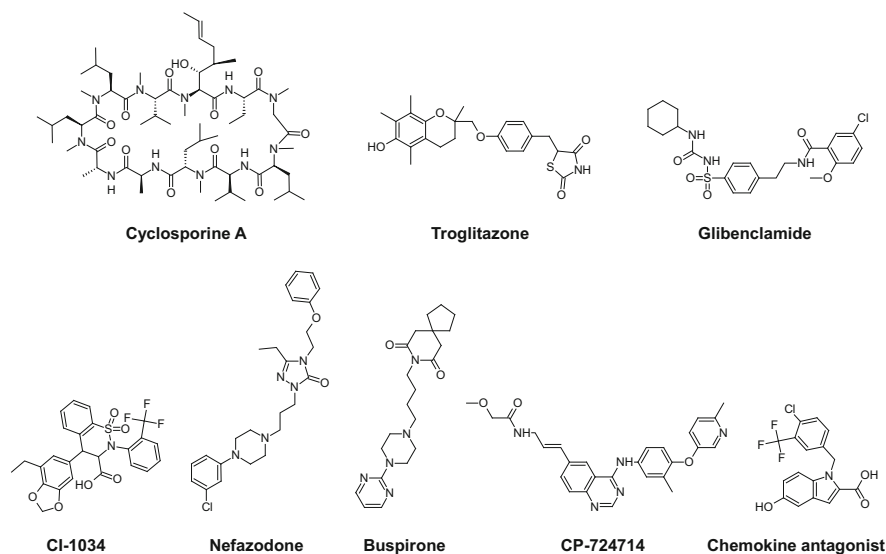
Fig. 6 Chemical structures of the compounds which are listed in Table 1. Note that the structure of bosentan is shown in Fig. 4b

serum or plasma bile acid levels and therefore needs to be taken into account when interpreting such data. Recently it has been reported that bile acid levels were increased in plasma and urine from rats several days after in vivo administration of a variety of hepatotoxic compounds, some of which do not inhibit BSEP (e.g., acetaminophen, carbamazepine) [93]. These data indicate that elevated total plasma bile acid levels should not be considered a specific biomarker of BSEP inhibition, but instead may provide a sensitive index of impaired liver function which arises via multiple mechanisms. They also highlight the importance of evaluating and interpreting plasma bile acids alongside other markers of liver dysfunction and not in isolation. Nonetheless, it is conceivable that further work may identify particular bile acid species, or combinations of individual bile acids, which provide a more specific biomarker of BSEP inhibition in vivo than total plasma bile acids.

3 Interpretation of BSEP Inhibition Data

3.1 BSEP Inhibition Is a Human DILI Risk Factor

Many pharmaceuticals which cause DILI have been shown to inhibit BSEP/Bsep activity in vitro and to cause functional effects in vivo in experimental animals which are indicative of BSEP inhibition (see Sect. 2). Systematic investigations of the relationship between BSEP inhibition and DILI in humans have been conducted

Fig. 7 Comparison between potency of in vitro BSEP inhibition, maximum unbound concentration in peripheral blood ($C_{max,u}$) and DILI caused by pharmaceuticals. Each point represents an individual drug (adapted from [43], with permission from American Society for Pharmacology and Experimental Therapeutics)
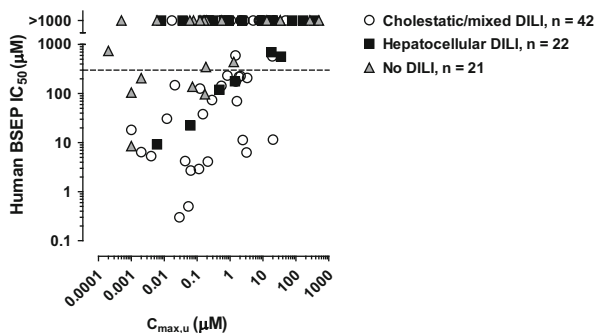
by two groups of investigators, both of whom quantified BSEP inhibition in vitro by drugs that caused clinically concerning DILI and by widely used drugs not reported to cause DILI [42, 43]. These analyses were undertaken using membrane vesicle assays based on the insect cell transient transfection system and evaluated >250 drugs. In both studies, the frequency and potency of BSEP inhibition by drugs which caused DILI was markedly greater than that observed with drugs which did not cause DILI. Furthermore, one of the studies reported a higher frequency and potency of BSEP inhibition by drugs which caused cholestatic or mixed DILI than by drugs which caused hepatocellular DILI [43]. This is consistent with the cholestatic pattern of DILI observed in humans who have defective BSEP expression due to genetic causes [14, 15]. In both studies, drugs were identified which exhibited potent BSEP inhibition but did not cause DILI in humans. However, it was also observed that the majority of these drugs are administered to humans at low oral doses either via non-oral routes which result in low plasma exposure (see Fig. 7) or only via very short courses of treatment [43].

The data obtained in one of the two studies and illustrated in Fig. 7 suggested that an in vitro apparent BSEP inhibition IC_{50} value of <300 μM provided useful discrimination between many drugs which caused cholestatic DILI and the majority of drugs which did not cause DILI [43]. This value is markedly higher than the unbound plasma concentrations of the majority of drugs and also the somewhat higher concentrations typically observed at the hepatic inlet, when compared with peripheral blood [94]. In comparison, most of the drugs tested in the other study which were associated with DILI in humans and inhibited BSEP in vitro had apparent BSEP IC_{50} values <25 μM [42]. When calculating apparent IC_{50} values in both studies, it was assumed that all of the drugs added to the in vitro incubations were free in solution and that no binding to protein or lipid had occurred. Many of the tested drugs exhibit high plasma protein binding; therefore, the true in vitro IC_{50} values are likely to be much lower than the apparent values that were calculated. Furthermore, many drugs accumulate within hepatocytes at concentrations that are much higher than extracellular concentrations [95] and for some drugs it has been found that their metabolites (which are not produced in the membrane vesicle assay model) are more potent BSEP inhibitors than the parent

compounds (e.g., troglitazone [89], sulindac [96]) and/or that BSEP inhibition requires co-expression of MRP2 (which is not present in insect cells transfected with BSEP) [32]. In addition, the fraction of transport inhibition required to cause functional alterations in bile flow in vivo that contribute to DILI progression following long-term dosing with drugs is unknown and could be much lower than the observed IC_{50} values (e.g., IC_{10} values). Therefore, the relationship between potency of in vitro BSEP inhibition by drugs, plasma drug exposure and clinically significant effects on BSEP function in vivo remains poorly defined and merits further investigation.

Consideration of physicochemical properties revealed that all of the 33 drugs which exhibited apparent BSEP IC_{50} <300 μM in one of the studies [43] had molecular weight >250 Da, ClogP >1.5 and non-polar surface area (NPSA) >180 Å. Subsequently, this association was confirmed in a more comprehensive assessment of the relationship between in vitro BSEP inhibition caused by >600 compounds and physicochemical properties [50], which is discussed in more detail in Sect. 5.

The drugs tested in the studies summarised above [42, 43] which inhibited BSEP and caused DILI have been reported to cause idiosyncratic DILI in humans and, with very few exceptions, have not been reported to cause liver injury in preclinical species. This implies that the level of BSEP inhibition achieved in vivo is insufficient to cause liver damage in animals or in non-susceptible humans. The reasons why liver injury occurs in susceptible patients remain to be determined. When considering the human safety implications of BSEP inhibition data, it is important to bear in mind that BSEP inhibition is one of the several drug-related mechanisms which have been implicated in initiation of DILI (see Fig. 2). Other important drug-related mechanisms include formation of chemically reactive metabolites, mitochondrial injury and activation of both innate and adaptive immune responses [26–29].

The value for human risk assessment that can be gained from evaluation of these additional drug-related processes alongside BSEP inhibition has been highlighted in a recent study of 36 drugs, 27 of which caused severe idiosyncratic adverse drug reactions (IADRs) in humans (most frequently DILI) and 9 of which did not [97]. Each drug was tested in vitro for: cytotoxicity to the human liver-derived THLE cell line, or to a THLE cell line which expressed human cytochrome P450 3A4 (to enable drug metabolism mediated by this enzyme); mitochondrial toxicity to HepG2 cells cultured in media which contained galactose in place of glucose to potentiate mitochondrially mediated toxicity; inhibition of BSEP activity; inhibition of Mrp2 activity; and covalent binding of radiolabelled drug to human hepatocyte proteins. Data from the toxicity assays were assigned a binary score of 1 or 0, depending upon whether or not a threshold potency value indicating in vitro activity was observed, and the aggregated in vitro panel score for each drug was calculated. The human hepatocyte covalent binding data were adjusted for fractional turnover of the drugs under the in vitro assay conditions and also for the daily human dose, which enabled calculation of the daily covalent binding burden
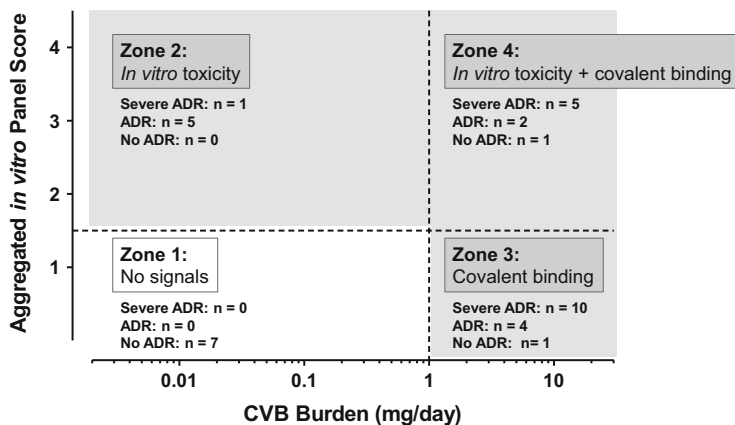
Fig. 8 Hazard matrix which summarises the relationship between multiple in vitro toxicity assay signals, estimated daily covalent binding burden to hepatocyte proteins and ADR risk (modified from [97])

(CVB burden). For each drug, the aggregated in vitro panel score was then plotted against CVB burden. This resulted in a two-dimensional hazard matrix, which is illustrated schematically in Fig. 8. Seven of the 9 drugs which did not cause concerning IADRs (78%) were located in Zone 1 (i.e. exhibited few in vitro toxicity assay signals and low CVB burden). In contrast, the 27 drugs which caused IADRs were all (100%) located in Zones 2, 3 or 4 (i.e. exhibited multiple in vitro toxicity assay signals and/or high CVB burden). These results support the complex, multi-step model of DILI initiation and progression that is outlined in Fig. 2 [97].

3.2 Recommendations from the International Transporter Consortium and European Medicines Agency

The International Transporter Consortium includes leading scientists from academia, industry and regulatory authorities who provide advice that may be incorporated into regulatory guidance. In their latest series of White Papers, this Consortium has acknowledged the emerging data linking BSEP inhibition with cholestatic DILI and has highlighted that currently it is impossible to define a value for BSEP inhibition that can be considered predictive of BSEP-mediated DILI [98]. Furthermore, it was noted that a strategy has not yet been defined which can accurately assess the clinical relevance of BSEP inhibition. In view of these limitations, the Consortium has not recommended proactive testing of BSEP inhibition during drug development. However, it has recommended that evaluation of the potential contribution of BSEP inhibition should be considered if evidence of cholestatic DILI is apparent from clinical studies or preclinical safety studies. In addition, the Consortium recommended that monitoring of serum bile acid levels
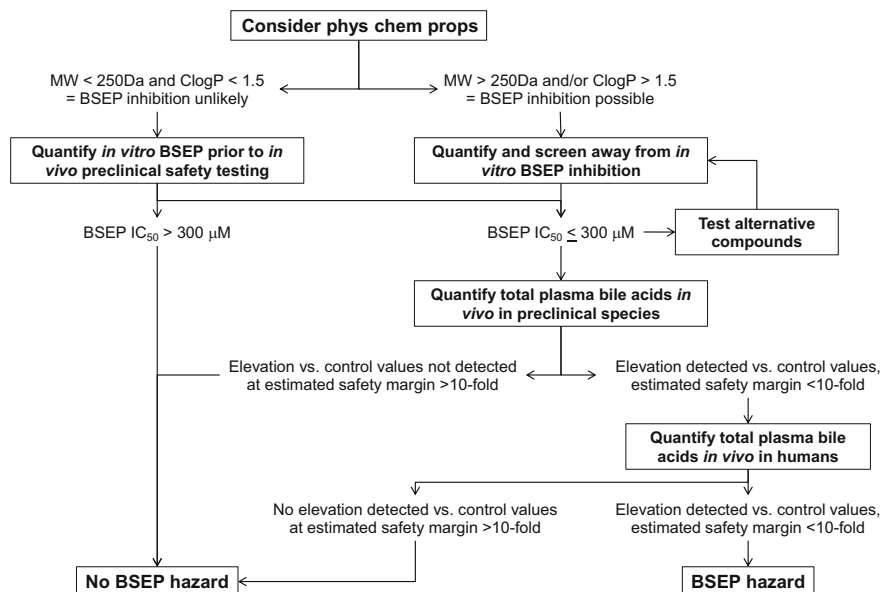
Fig. 9 Proposed decision tree for in vitro BSEP assessment and data interpretation

should be undertaken in vivo for compounds which have been identified as Bsep inhibitors in rats, as should evaluation of alkaline phosphatase and transaminase levels [98]. If BSEP inhibition is observed under such circumstances, biochemical evaluation of serum bile acids has been recommended to aid ongoing clinical safety assessment [98].

The current European Medicines Agency guideline states that if in vitro studies indicate BSEP inhibition, adequate biochemical monitoring including serum bile salts is recommended during drug development [99].

3.3 Assessment of BSEP Inhibition During Drug Discovery

While the recommendations from the International Transporter Consortium are appropriate to support clinical development of drugs which have exhibited evidence of cholestatic liver injury in vivo and have been found to inhibit BSEP activity, they do not address the needs of drug discovery. At this stage, chemical choice is available; therefore, it is important to design and select candidate drugs which exhibit an optimal balance between desirable pharmacology and pharmacokinetic properties and least possible propensity to cause clinically and commercially undesirable drug–drug interactions and toxicity. If compounds which have reduced (or ideally no) propensity to cause potent BSEP inhibition can be designed and selected, the need to take account of and manage possible DILI caused by this mechanism will be avoided or minimised in drug development. With this in mind, a

BSEP assessment decision tree which may be suitable for generic use in drug discovery is illustrated in Fig. 9.

The primary objective is to enable early identification of drug series which have the propensity to cause potentially concerning levels of in vitro BSEP inhibition (e.g., $IC_{50} < 300$ μM), so that in silico tools and in vitro screening can be used to support the design and selection of compounds which exhibit the least possible BSEP inhibition potency. Progression of candidate drugs which inhibit BSEP activity in vitro may be appropriate when chemical choice has been exhausted, the dose and clinical exposure is expected to be low or the anticipated benefit–risk is considered likely to justify clinical evaluation of a compound that may cause DILI. Since the risk of BSEP inhibition posed by compounds with molecular weight <250 Da and ClogP <1.5 appears to be low [43, 50], the value of testing such molecules for BSEP inhibition is questionable and it is more logical to focus resources on testing molecules with molecular weight >250 Da and/or ClogP >1.5. The potency cut-off value of 300 μM assumes that the *Sf* 21 insect expression membrane vesicle assay [42, 43], which is suited to high-volume data generation and utilises the physiologically relevant substrate [³H]-taurocholate, is used to determine BSEP inhibition. Other cut-off values may be more appropriate if other in vitro assays or BSEP probe substrates are utilised. For example, Thompson et al. [97] utilised the fluorescent BSEP probe substrate (7β-[(4-nitro-2,1,3-benzoxadiazol-7-yl)amino] taurocholate), which necessitated the use of a higher substrate concentration to yield acceptable signal to noise ratio than was used in the [³H]-taurocholate assay and resulted in higher apparent BSEP IC_{50} values than were observed when using the radiolabelled substrate [43].

In view of the uncertain relationship between BSEP inhibition in vitro and functional effects in vivo, quantification of plasma or serum bile acid concentrations is required to provide insight into whether BSEP inhibition observed in vitro results in adverse functional consequences in vivo and to determine whether in vivo no effect levels and safety margins can be identified. These should be evaluated alongside conventional plasma biomarkers of liver injury (i.e. increased levels of plasma alanine aminotransferase, alkaline phosphatase and other enzymes released from damaged hepatocytes, plus elevated bilirubin arising due to its impaired hepatic clearance into bile) and altered liver histopathology, as proposed by the ITC [98]. A tenfold in vivo exposure margin is proposed for identification of compounds which inhibit BSEP in vitro, but are administered in vivo under conditions which result in low plasma exposure that can be considered unlikely to raise safety concerns.

When assessing the in vitro DILI risk posed by test compounds, it is recommended that the hazard matrix approach described by Thompson et al. [97] and illustrated in Fig. 6 is utilised. This approach requires evaluation of several additional parameters, i.e. metabolism-independent cell toxicity, CYP3A4-mediated cell toxicity, mitochondrially mediated cell injury, inhibition of Mrp2 activity and covalent binding of radiolabelled drug to human hepatocyte proteins, and is, therefore, relatively resource intensive. It has the potential to provide valuable insight into the propensity of drugs to cause DILI and other clinically

concerning IADRs, which otherwise may not be detected until late in development or even post-licensing [97].

4 Computational Approaches and Structure–Activity Relationship (SAR)

4.1 Introduction to Computational Modelling Approaches

Computational methods are widely used in predictive safety assessment, with the ultimate goal of filtering out compounds with safety liabilities at an early stage in drug discovery [100]. Such methods include but are not limited to statistical- and structure-based models, expert systems and quantitative SAR (QSAR) models. Early QSAR models were of the multiple linear regression type and assumed a linear relationship between physicochemical and structural descriptors which encode chemical information and biological responses [101]. In practice, correlations between these parameters often are non-linear. Therefore, the need for more advanced non-linear regression models such as support vector machine, random forest or artificial neural networks has been recognised and such models are now being developed and applied widely in drug discovery [102, 103].

Computational approaches have been used to gain improved understanding of how chemical compounds interfere with ABC transporters and to aid in the design of compounds which are devoid of such interactions. P-gp is the most extensively investigated biliary transporter and has been the focus of the majority of computational efforts in this area. Recently, QSAR models based on around 200 tested compounds were published in two different studies for P-gp. In one study, different machine learning methods were compared to achieve the highest accuracy in distinguishing between P-gp substrates and non-substrates [104]. In the second study, Bikadi et al. built a QSAR model based on a compound set of similar size and augmented their analysis by establishing a homology model for P-gp to dock ligands into the binding pocket [105]. Although computational modelling of MRP2 interactions is currently less advanced, it is notable that Pedersen et al. employed the concept of a multivariate orthogonal partial least-squares projection to latent structures discriminant analysis (OPLS-DA) model to compare MRP2 inhibitors and non-inhibitors and identified molecular descriptors as major determinants of the inhibitory effect [106].

4.2 BSEP Inhibition SAR Models

Hirano et al. reported the first application of computational approaches to BSEP inhibitor profiling in 2006 [38, 107]. In this study, BSEP inhibition by a limited yet
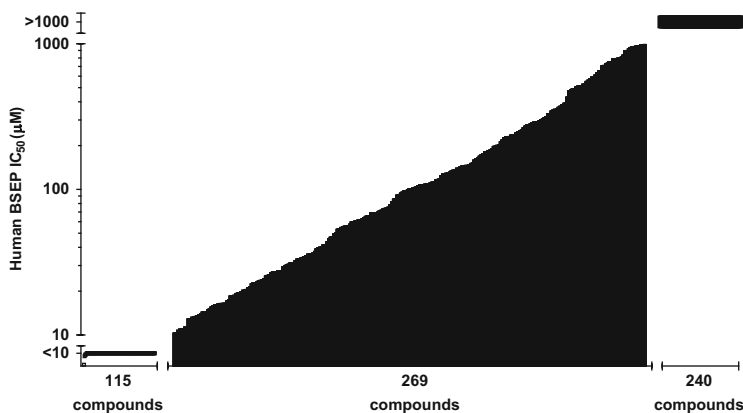
Fig. 10 Distribution of BSEP IC$_{50}$ values obtained with 624 test compounds. Data are geometric means from at least three independent experiments and the compounds have been ranked by potency (from [50], reprinted with permission from American Society for Pharmacology and Experimental Therapeutics)

diverse set of 37 drugs from seven different therapeutic classes, which ranged from neurotransmitters and ion channel blockers to anti-cancer drugs, was analysed using membrane vesicles from transfected *Sf*9 insect cells. Chemical fragmentation descriptors were extracted for each of the drugs and multiple linear regression analysis was used to explore possible associations between BSEP inhibition and the descriptors. A model which incorporated six fragmentation codes was developed, which provided a high correlation with the experimentally observed in vitro BSEP inhibition data ($r^2 = 0.95$). Five of the descriptors correlated positively with inhibition (which included ring-linking groups containing one carbon atom, thioester groups bound to heterocyclic carbon and aromatic rings), while hydroxyl groups bonded to aliphatic carbon correlated negatively. Although this analysis is very encouraging, its practical value is unclear due to the small number of compounds in the training set which were used to build the model and by the absence of an independent test set of compounds.

Due to these limitations, we have recently described the use of several different approaches to undertake a more rigorous evaluation of the relationship between physicochemical features and BSEP inhibition by test compounds [50]. Our analysis was based on a reasonably large dataset totalling 624 compounds, each of which was tested for BSEP inhibition using a membrane vesicle assay that utilised an insect cell expression system and [^{3}H]-taurocholate as the probe substrate. The tested compounds had a broad range of apparent BSEP IC$_{50}$ potency values (see Fig. 10).

In order to build and test QSAR models, the dataset was split randomly into a training set and a test set and an IC$_{50}$ cut-off value of 300 μM was selected to discriminate between inhibitors and non-inhibitors. This IC$_{50}$ cut-off value was selected because data obtained by Dawson et al. [43] indicated that it provided a

Table 2 Distribution of BSEP inhibitors across different ion classes (data taken from [50])

Ion class	BSEP non-inhibitors, IC$_{50}$ > 300 μM ($N = 299$) (%)	BSEP inhibitors, IC$_{50}$ < 300 μM ($N = 325$) (%)
Acid	22.7	20.0
Base	34.8	38.5
Cationic	3.3	0.6
Neutral	32.1	37.8
Zwitterion	7.0	3.1

useful discrimination between many drugs which cause cholestatic DILI and the majority of the tested non-hepatotoxic drugs. The approaches employed to build classification models included simple recursive partitioning and more elaborate, non-linear machine learning methods (support vector machine and random forest). The recursive partitioning approach was employed to identify key properties influencing BSEP inhibition and the machine learning methods were used to build computational models capable of predicting BSEP inhibition for a given compound. In addition, molecular pair analysis was used to derive SAR for BSEP inhibitors and to explore structural modifications which were associated with reduced BSEP inhibition potency. BSEP inhibition was observed for many different pharmacologically and structurally distinct drugs, which were distributed across all ion classes (see Table 2).

A key finding from our QSAR analysis [50] was that lipophilicity and molecular weight were significantly correlated with BSEP inhibition, which confirmed and extended a previous analysis of 87 compounds [43]. For compounds where molecular weight was greater than 309, the most important parameter influencing BSEP inhibitory potency was lipophilicity, as calculated by ClogP. Molecular weight and ClogP were used to construct a simple recursive partitioning scheme, which resulted in an improved classification of BSEP inhibitors ($r^2 = 0.5$ in the training set, $r^2 = 0.36$ in the test set) when compared with any single property and provided a benchmark against which more elaborate modelling strategies could be compared. Evaluation of a collection of molecular descriptor sets and modelling algorithms yielded a further improvement in classification of a collection of 187 compounds. The source of this improvement appeared to originate from the ability of the QSAR model to identify low-molecular-weight BSEP inhibitors, where the simple recursive partitioning scheme failed. Furthermore, for basic compounds a statistically significant relationship was observed between the acid dissociation constants and BSEP inhibition, indicating that more basic compounds and those with more hydrogen-bond donors tend to be less potent inhibitors.

Key similarities and differences between the study carried out by Hirano et al. [38, 107] and our analyses [50] can be summarised as follows. Hirano et al. used a relatively small dataset of compounds to derive information on structural features influencing BSEP inhibition. We used a set almost 20 times as large and employed not one but three different computational approaches to improve the understanding of chemical properties which distinguish between

BSEP inhibitors and non-inhibitors. A direct comparison of both study results can be done by comparing the six fragmentation codes (i.e. chemical features) of the Hirano group with the results of our matched molecular pair analysis. No similar or identical substructural features being responsible for BSEP inhibition could be found between the two studies. This lack of correlation is likely to be attributable to both the small dataset investigated by Hirano et al. and the limited number of matched pairs that could be identified in our study.

4.3 Limitations of Current BSEP SAR Models and Opportunities for Improvement

The ultimate goal of assessing BSEP inhibition is to predict cholestatic drug-induced liver injury. At present, it is unclear whether an in vitro BSEP inhibition IC_{50} threshold value can be identified which accurately distinguishes between drugs which have the propensity to cause cholestatic DILI and drugs which do not. In our QSAR analysis [50], we used a threshold value of 300 μM, as proposed by Dawson et al. [43]. This provided a relatively balanced dataset of positive and negative compounds, which aided our analyses. However, when using this threshold value, we observed that many of the tested compounds were classified as BSEP inhibitors (see Fig. 10). This highlights the urgent need for an improved understanding of the relationship between in vitro potency of BSEP inhibition and in vivo adverse functional consequences. Another limitation of the categorical QSAR modelling we have undertaken is the so-called "black box" behaviour of the machine learning method that was applied. For the current model, it is almost impossible to assess why a compound is predicted to be a BSEP inhibitor. Nevertheless, the information provided by the model is helpful if used by medicinal chemists when attempting to design and synthesise test compounds which are devoid of BSEP inhibition.

One obvious next step when developing more predictive QSAR models for BSEP inhibition would be the analysis of data from larger sets of compounds. A set of chemically diverse compounds which comprises thousands of entries with IC_{50} values that range over at least three (and ideally five) orders of magnitude would enable generation of a reasonably good regression model with the potential of high predictive power. In itself, this will not solve the problem of the limited interpretability of some of the machine learning methods that were applied. Use of predicted significant substructures or properties may provide one useful way to tackle this issue. The Computational Toxicology group at AstraZeneca is working on this field [108] and recently published an article that introduces the principle of heuristic localised inverse QSAR to enable assessment of the relative ability of the descriptors to influence the biological response in an area localised around predicted compounds [109]. This concept can be used to guide structural modifications that affect the biological response in the desired direction. Another more general problem is the limited ability of current transporter models to distinguish

between BSEP inhibitors and transported substrates [110]. It can be expected that analysis of larger datasets of diverse compounds, including substrates and inhibitors, is likely to help address this issue.

A further limitation of the current approaches is that they do not take account of the possibility that multiple mechanisms may contribute to BSEP inhibition. This is likely in view of the complex processes involved in cellular BSEP expression and activity, defective BSEP expression in cholestatic liver injury and BSEP inhibition by compounds [111–113]. This situation is undesirable and will need to be addressed in due course. It will require access to data from more physiologically relevant test systems than the insect cell expression membrane vesicle assays used currently in many studies.

5 Conclusions

Transporter proteins perform vital functions throughout the body and are also key determinants of drug disposition, excretion and toxicity. Inhibition of the hepatic efflux transporter BSEP has been identified in a number of in vitro studies as a drug-related risk factor for DILI. It has to be recognised that DILI is a complex, multifactorial process and that BSEP inhibition is only one of the multiple mechanisms which contribute to the initiation of DILI. Assessment of BSEP inhibition during drug discovery and development provides the opportunity to reduce the potential for this important safety liability to occur. An understanding of the contribution of physicochemical and molecular properties to the BSEP inhibition potential of new chemical entities is in its infancy. Nevertheless, the SAR models available currently allow computational evaluation of BSEP inhibition potential and are able to inform and guide an in vitro testing strategy. Simple in vitro assays which use membrane vesicles from transfected insect cell lines can be utilised at sufficient throughput to influence compound selection in early drug discovery, while more complex in vitro and in vivo models enable a more detailed assessment at later stages of the drug discovery process and in drug development.

However, important challenges remain to be addressed, in particular around the interpretation of in vitro data and their translatability to in vivo adverse functional consequences in nonclinical animal studies and, ultimately, in humans. Future research needs to focus on improving human risk assessment of compounds which exhibit BSEP inhibition in vitro. Key questions here include : What level and potency of BSEP inhibition in vitro translates into functional inhibition in vivo? How can an understanding of a compound's pharmacokinetic properties be used to enhance interpretation of the significance of in vitro data? What measures of in vivo drug exposure are the most appropriate to take into account when assessing safety margins? What is the relationship between BSEP inhibitory potency in vivo in healthy, non-susceptible humans and liver injury in patients who develop DILI?

Mechanistic mathematical modelling and systems modelling approaches are now being utilised to address the risk of DILI arising from other mechanisms,

most notably via reactive metabolite formation (e.g., [114, 115]), and may offer an opportunity to address some of these questions and improve the prediction of DILI potential caused by BSEP inhibition.

The ITC has highlighted BSEP as one of the emerging transporters that needs to be considered when evaluating drug safety. It can be expected that this will stimulate a discussion amongst transporter scientists on how best to tackle and manage safety issues arising from BSEP inhibition during the process of drug discovery and development.

References

1. Lam P, Soroka CJ, Boyer JL (2010) The bile salt export pump: clinical and experimental aspects of genetic and acquired cholestatic liver disease. Semin Liver Dis 30:125–133
2. Cai SY, Wang L, Ballatori N, Boyer JL (2001) Bile salt export pump is highly conserved during vertebrate evolution and its expression is inhibited by PFIC type II mutations. Am J Physiol Gastrointest Liver Physiol 281:G316–G322
3. Hofmann AF (2004) Bile composition. In: Johnson L (ed) Encyclopedia of gastroenterology. Elsevier, New York, pp 176–184, http://dx.doi.org/10.1016/B0-12-386860-2/00063-0
4. Gerk PM, Vore M (2002) Regulation of expression of the multidrug resistance-associated protein 2 (MRP2) and its role in drug disposition. J Pharmacol Exp Ther 302:407–415
5. Zhou SF (2008) Structure, function and regulation of P-glycoprotein and its clinical relevance in drug disposition. Xenobiotica 38:802–832
6. Oude Elferink RPJ, Paulusma CC (2007) Function and pathophysiological importance of ABCB4 (MDR3 P-glycoprotein). Pflugers Arch 453:601–610
7. Hofmann AF, Hagey LR (2008) Bile acids: chemistry, pathochemistry, biology, pathobiology, and therapeutics. Cell Mol Life Sci 65:2461–2483
8. Kullak-Ublick GA, Stieger B, Meier PJ (2004) Enterohepatic bile salt transporters in normal physiology and liver disease. Gastroenterology 126:322–342
9. Gonzalez FJ (2012) Nuclear receptor control of enterohepatic circulation. Compr Physiol 2:2811–2828
10. Strautnieks SS, Bull LN, Knisely AS, Kocoshis SA, Dahl N, Arnell H, Sokal E, Dahan K, Childs S, Ling V, Tanner MS, Kagalwalla AF, Németh A, Pawlowska J, Baker A, Mieli-Vergani G, Freimer NB, Gardiner RM, Thompson RJ (1998) A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. Nat Genet 20:233–238
11. Byrne JA, Strautnieks SS, Ihrke G, Pagani F, Knisely AS, Linton KJ, Mieli-Vergani G, Thompson RJ (2009) Missense mutations and single nucleotide polymorphisms in ABCB11 impair bile salt export pump processing and function or disrupt pre-messenger RNA splicing. Hepatology 49:553–567
12. Ho RH, Leake BF, Kilkenny DM, Meyer Zu Schwabedissen HE, Glaeser H, Kroetz DL, Kim RB (2010) Polymorphic variants in the human bile salt export pump (BSEP; ABCB11): functional characterization and interindividual variability. Pharmacogenet Genomics 20:45–57
13. Perez MJ, Briz O (2009) Bile-acid-induced cell injury and protection. World J Gastroenterol 15:1677–1689
14. Noe J, Kullak-Ublick GA, Jochum W, Stieger B, Kerb R, Haberl M, Müllhaupt B, Meier PJ, Pauli-Magnus C (2005) Impaired expression and function of the bile salt export pump due to three novel ABCB11 mutations in intrahepatic cholestasis. J Hepatol 43:536–543
15. Stieger B, Meier Y, Meier PJ (2007) The bile salt export pump. Pflugers Arch 453:611–620

16. Lam P, Pearson CL, Soroka CJ, Xu S, Mennone A, Boyer JL (2007) Levels of plasma membrane expression in progressive and benign mutations of the bile salt export pump (Bsep/Abcb11) correlate with severity of cholestatic diseases. Am J Physiol Cell Physiol 293:C1709–C1716

17. Kagawa T, Watanabe N, Mochizuki K, Numari A, Ikeno Y, Itoh J, Tanaka H, Arias IM, Mine T (2008) Phenotypic differences in PFIC2 and BRIC2 correlate with protein stability of mutant Bsep and impaired taurocholate secretion in MDCK II cells. Am J Physiol Gastrointest Liver Physiol 294:G58–G67

18. Wang R, Salem M, Yousef IM, Tuchweber B, Lam P, Childs SJ, Helgason CD, Ackerley C, Phillips MJ, Ling V (2001) Targeted inactivation of sister of P-glycoprotein gene (spgp) in mice results in nonprogressive but persistent intrahepatic cholestasis. Proc Natl Acad Sci U S A 98:2011–2016

19. Wang R, Lam P, Liu L, Forrest D, Yousef IM, Mignault D, Phillips MJ, Ling V (2003) Severe cholestasis induced by cholic acid feeding in knockout mice of sister of P-glycoprotein. Hepatology 38:1489–1499

20. Wang R, Chen HL, Liu L, Sheps JA, Phillips MJ, Ling V (2009) Compensatory role of P-glycoproteins in knockout mice lacking the bile salt export pump. Hepatology 50:948–956

21. Zhang Y, Li F, Patterson AD, Wang Y, Krausz KW, Neale G, Thomas S, Nachagari D, Vogel P, Vore M, Gonzalez FJ, Schuetz JD (2012) Abcb11 deficiency induces cholestasis coupled to impaired β-fatty acid oxidation in mice. J Biol Chem 287:24784–24794

22. Greaves P, Williams A, Eve M (2004) First dose of potential new medicines to humans: how animals help. Nat Rev Drug Discov 3:226–236

23. Kaplowitz N, DeLeve L (2013) Drug-induced liver injury, 3rd edn. Intra Healthcare, New York

24. Abboud G, Kaplowitz N (2007) Drug-induced liver injury. Drug Saf 30:277–294

25. Kola I, Landis J (2004) Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 3:711–715

26. Dykens JA, Will Y (2007) The significance of mitochondrial toxicity testing in drug development. Drug Discov Today 12:777–785

27. Greer ML, Barber J, Eakins J, Kenna JG (2010) Cell-based approaches for evaluation of drug-induced liver injury. Toxicology 268:125–131

28. Roth RA, Ganey PE (2010) Intrinsic versus idiosyncratic drug-induced hepatotoxicity – two villains or one? J Pharmacol Exp Ther 332:692–697

29. Thompson RA, Isin EM, Li Y, Weaver R, Weidolf L, Wilson I, Claesson A, Page K, Dolgos H, Kenna JG (2011) Risk assessment and mitigation strategies for reactive metabolites in drug discovery and development. Chem Biol Interact 192:65–71

30. Pachkoria K, Lucena MI, Molokhia M, Cueto R, Carballo AS, Carvajal A, Andrade RJ (2007) Genetic and molecular factors in drug-induced liver injury: a review. Curr Drug Saf 2:97–112

31. Daly AK, Day CP (2012) Genetic association studies in drug-induced liver injury. Drug Metab Rev 44:116–126

32. Stieger B, Fattinger K, Madon J, Kullak-Ublick GA, Meier PJ (2000) Drug- and estrogen-induced cholestasis through inhibition of the hepatocellular bile salt export pump (Bsep) of rat liver. Gastroenterology 118:422–430

33. Fattinger K, Funk C, Pantze M, Weber C, Reichen J, Stieger B, Meier PJ (2001) The endothelin antagonist bosentan inhibits the canalicular bile salt export pump: a potential mechanism for hepatic adverse reactions. Clin Pharmacol Ther 69:223–231

34. Funk C, Ponelle C, Scheuermann G, Pantze M (2001) Cholestatic potential of troglitazone as a possible factor contributing to troglitazone-induced hepatotoxicity: in vivo and in vitro interaction at the canalicular bile salt export pump (Bsep) in the rat. Mol Pharmacol 59:627–635

35. Noé J, Stieger B, Meier PJ (2002) Functional expression of the canalicular bile salt export pump of human liver. Gastroenterology 123:1659–1666

36. Byrne JA, Strautnieks SS, Mieli-Vergani G, Higgins CF, Linton KJ, Thompson RJ (2002) The human bile salt export pump: characterization of substrate specificity and identification of inhibitors. Gastroenterology 123:1649–1658

37. Horikawa M, Kato Y, Tyson CA, Sugiyama Y (2003) Potential cholestatic activity of various therapeutic agents assessed by bile canalicular membrane vesicles isolated from rats and humans. Drug Metab Pharmacokinet 18:16–22

38. Hirano H, Kurata A, Onishi Y, Sakurai A, Saito H, Nakagawa H, Nagakura M, Tarui S, Kanamori Y, Kitajima M, Ishikawa T (2006) High-speed screening and QSAR analysis of human ATP-binding cassette transporter ABCB11 (bile salt export pump) to predict drug-induced intrahepatic cholestasis. Mol Pharm 3:252–265

39. Mano Y, Usui T, Kamimura H (2007) Effects of bosentan, an endothelin receptor antagonist, on bile salt export pump and multidrug resistance-associated protein 2. Biopharm Drug Dispos 28:13–18

40. Yabuuchi H, Tanaka K, Maeda M, Takemura M, Oka M, Ohashi R, Tamai I (2008) Cloning of the dog bile salt export pump (BSEP; ABCB11) and functional comparison with the human and rat proteins. Biopharm Drug Dispos 29:441–448

41. Kis E, Rajnai Z, Ioja E, Herédi Szabó K, Nagy T, Méhn D, Krajcsi P (2009) Mouse Bsep ATPase assay: a nonradioactive tool for assessment of the cholestatic potential of drugs. J Biomol Screen 14:10–15

42. Morgan RE, Trauner M, van Staden CJ, Lee PH, Ramachandran B, Eschenberg M, Afshari CA, Qualls CW Jr, Lightfoot-Dunn R, Hamadeh HK (2010) Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. Toxicol Sci 118:485–500

43. Dawson S, Stahl S, Paul N, Barber J, Kenna JG (2012) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. Drug Metab Dispos 40:130–138

44. Kostrubsky VE, Strom SC, Hanson J, Urda E, Rose K, Burliegh J, Zocharski P, Cai H, Sinclair JF, Sahi J (2003) Evaluation of hepatotoxic potential of drugs by inhibition of bile-acid transport in cultured primary human hepatocytes and intact rats. Toxicol Sci 76:220–228

45. Swift B, Pfeifer ND, Brouwer KL (2010) Sandwich-cultured hepatocytes: an in vitro model to evaluate hepatobiliary transporter-based drug interactions and hepatotoxicity. Drug Metab Rev 42:446–471

46. Mita S, Suzuki H, Akita H, Stieger B, Meier PJ, Hofmann AF, Sugiyama Y (2005) Vectorial transport of bile salts across MDCK cells expressing both rat Na + −taurocholate cotransporting polypeptide and rat bile salt export pump. Am J Physiol Gastrointest Liver Physiol 288:G159–G167

47. Herédi-Szabó K, Kis E, Krajcsi P (2012) The vesicular transport assay: validated in vitro methods to study drug-mediated inhibition of canalicular efflux transporters ABCB11/BSEP and ABCC2/MRP2. Curr Protoc Toxicol Chapter 23:Unit 23.4

48. Saito H, Osumi M, Hirano H, Shin W, Nakamura R, Ishikawa T (2009) Technical pitfalls and improvements for high-speed screening and QSAR analysis to predict inhibitors of the human bile salt export pump (ABCB11/BSEP). AAPS J 11:581–589

49. van Staden CJ, Morgan RE, Ramachandran B, Chen Y, Lee PH, Hamadeh HK (2012) Membrane vesicle ABC transporter assays for drug safety assessment. Curr Protoc Toxicol Chapter 23:Unit 23.5

50. Warner DJ, Chen H, Cantin LD, Kenna JG, Stahl S, Walker CL, Noeske T (2012) Mitigating the inhibition of human bile salt export pump by drugs: opportunities provided by physico-chemical property modulation, in silico modeling, and structural modification. Drug Metab Dispos 40:2332–2341

51. Karlsson JE, Heddle C, Rozkov A, Rotticci-Mulder J, Tuvesson O, Hilgendorf C, Andersson TB (2010) High-activity P-glycoprotein, multidrug resistance protein 2, and breast cancer resistance protein membrane vesicles prepared from transiently transfected human embryonic kidney 293-Epstein-Barr virus nuclear antigen cells. Drug Metab Dispos 38:705–714

52. Kis E, Ioja E, Nagy T, Szente L, Herédi-Szabó K, Krajcsi P (2009) Effect of membrane cholesterol on BSEP/Bsep activity: species specificity studies for substrates and inhibitors. Drug Metab Dispos 37:1878–1886
53. Elsby R, Smith V, Fox L, Stresser D, Butters C, Sharma P, Surry DD (2011) Validation of membrane vesicle-based breast cancer resistance protein and multidrug resistance protein 2 assays to assess drug transport and the potential for drug-drug interaction to support regulatory submissions. Xenobiotica 41:764–783
54. Meier PJ, Boyer JL (1990) Preparation of basolateral (sinusoidal) and canalicular plasma membrane vesicles for the study of hepatic transport processes. Methods Enzymol 192:534–545
55. Pauli-Magnus C, Meier PJ, Stieger B (2010) Genetic determinants of drug-induced cholestasis and intrahepatic cholestasis of pregnancy. Semin Liver Dis 30:147–159
56. Hirano M, Maeda K, Hayashi H, Kusuhara H, Sugiyama Y (2005) Bile salt export pump (BSEP/ABCB11) can transport a nonbile acid substrate, pravastatin. J Pharmacol Exp Ther 314:876–882
57. Lecureur V, Sun D, Hargrove P, Schuetz EG, Kim RB, Lan LB, Schuetz JD (2000) Cloning and expression of murine sister of P-glycoprotein reveals a more discriminating transporter than MDR1/P-glycoprotein. Mol Pharmacol 57:24–35
58. Wang L, Soroka CJ, Boyer JL (2002) The role of bile salt export pump mutations in progressive familial intrahepatic cholestasis type II. J Clin Invest 110:965–972
59. Hayashi H, Takada T, Suzuki H, Akita H, Sugiyama Y (2005) Two common PFIC2 mutations are associated with the impaired membrane trafficking of BSEP/ABCB11. Hepatology 41:916–924
60. Cui Y, König J, Keppler D (2001) Vectorial transport by double-transfected cells expressing the human uptake transporter SLC21A8 and the apical export pump ABCC2. Mol Pharmacol 60:934–943
61. Fahrmayr C, König J, Auge D, Mieth M, Fromm MF (2012) Identification of drugs and drug metabolites as substrates of multidrug resistance protein 2 (MRP2) using triple-transfected MDCK-OATP1B1-UGT1A1-MRP2 cells. Br J Pharmacol 165:1836–1847
62. Fahrmayr C, König J, Auge D, Mieth M, Münch K, Segrestaa J, Pfeifer T, Treiber A, Fromm M (2013) Phase I and II metabolism and MRP2-mediated export of bosentan in a MDCKII-OATP1B1-CYP3A4-UGT1A1-MRP2 quadruple-transfected cell line. Br J Pharmacol 169:21–33
63. Mita S, Suzuki H, Akita H, Hayashi H, Onuki R, Hofmann AF, Sugiyama Y (2006) Vectorial transport of unconjugated and conjugated bile salts by monolayers of LLC-PK1 cells doubly transfected with human NTCP and BSEP or with rat Ntcp and Bsep. Am J Physiol Gastrointest Liver Physiol 290:G550–G556
64. Mita S, Suzuki H, Akita H, Hayashi H, Onuki R, Hofmann AF, Sugiyama Y (2006) Inhibition of bile acid transport across Na+/taurocholate cotransporting polypeptide (SLC10A1) and bile salt export pump (ABCB 11)-coexpressing LLC-PK1 cells by cholestasis-inducing drugs. Drug Metab Dispos 34:1575–1581
65. De Bruyn T, Chatterjee S, Fattah S, Keemink J, Nicolaï J, Augustijns P, Annaert P (2013) Sandwich-cultured hepatocytes: utility for in vitro exploration of hepatobiliary drug disposition and drug-induced hepatotoxicity. Expert Opin Drug Metab Toxicol 9:589–616
66. Liu X, Chism JP, LeCluyse EL, Brouwer KR, Brouwer KL (1999) Correlation of biliary excretion in sandwich-cultured rat hepatocytes and in vivo in rats. Drug Metab Dispos 27:637–644
67. Chu X, Korzekwa K, Elsby R, Fenner K, Galetin A, Lai Y, Matsson P, Moss A, Nagar S, Rosania GR, Bai JP, Polli JW, Sugiyama Y, Brouwer KL (2013) Intracellular drug concentrations and transporters: measurement, modeling, and implications for the liver. Clin Pharmacol Ther 94:126–141

68. Kostrubsky SE, Strom SC, Kalgutkar AS, Kulkarni S, Atherton J, Mireles R, Feng B, Kubik R, Hanson J, Urda E, Mutlib AE (2006) Inhibition of hepatobiliary transport as a predictive method for clinical hepatotoxicity of nefazodone. Toxicol Sci 90:451–459
69. Zamek-Gliszczynski MJ, Xiong H, Patel NJ, Turncliff RZ, Pollack GM, Brouwer KL (2003) Pharmacokinetics of 5 (and 6)-carboxy-2,7-dichlorofluorescein and its diacetate promoiety in the liver. J Pharmacol Exp Ther 304:801–809
70. Nakanishi T, Shibue Y, Fukuyama Y, Yoshida K, Fukuda H, Shirasaka Y, Ikumi T (2011) Quantitative time-lapse imaging-based analysis of drug-drug interaction mediated by hepatobiliary transporter, multidrug resistance-associated protein 2, in sandwich-cultured rat hepatocytes. Drug Metab Dispos 39:984–991
71. Milkiewicz P, Baiocchi L, Mills CO, Ahmed M, Khalaf H, Keogh A, Baker J, Elias E (1997) Plasma clearance of cholyl-lysyl-fluorescein: a pilot study in humans. J Hepatol 27:1106–1109
72. Mills CO, Milkiewicz P, Müller M, Roma MG, Havinga R, Coleman R, Kuipers F, Jansen PL, Elias E (1999) Different pathways of canalicular secretion of sulfated and non-sulfated fluorescent bile acids: a study in isolated hepatocyte couplets and TR⁻ rats. J Hepatol 31:678–684
73. Hopwood J, Summers C, Pognan F, Barrett G, Jones K, Laine R, Kenna G (2006) A novel method for quantification of canalicular transporter inhibition in primary rat hepatocyte sandwich cultures. Toxicology 226:66–67
74. de Waart DR, Häusler S, Vlaming MLH, Kunne C, Hänggi E, Gruss H-J, Oude Elferink RPJ, Stieger B (2010) Hepatic transport mechanisms of cholyl-L-lysyl-fluorescein. J Pharmacol Exp Ther 334:78–86
75. Yamaguchi K, Murai T, Yabuuchi H, Hui SP, Kurosawa T (2010) Measurement of bile salt export pump transport activities using a fluorescent bile acid derivative. Drug Metab Pharmacokinet 25:214–219
76. Ferrigno A, Richelmi P, Vairetti M (2013) Troubleshooting and improving the mouse and rat isolated perfused liver preparation. J Pharmacol Toxicol Methods 67:107–114
77. Preininger K, Stingl H, Englisch R, Furnsinn C, Graf J, Waldhausl W, Roden M (1999) Acute troglitazone action in isolated perfused rat liver. Br J Pharmacol 126:372–378
78. Kroker R, Anwer MS, Hegner D (1978) The interaction of rifamycin SV with hepatic transport of taurocholic acid in the isolated perfused rat liver. Naunyn Schmiedebergs Arch Pharmacol 302:323–327
79. Wang YM, Reuning RH (1994) A comparison of two surgical techniques for preparation of rats with chronic bile duct cannulae for the investigation of enterohepatic circulation. Lab Anim Sci 44:479–485
80. van Wijk H, Donachie P, Mann DL, McMahon H, Robb D (2001) A novel bile duct cannulation method with tail cuff exteriorization allowing continuous intravenous infusion and enterohepatic recirculation in the unrestrained rat. Lab Anim 35:325–333
81. Marshall RW, Moreno OM, Brodie DA (1964) Chronic bile duct cannulation in the dog. J Appl Physiol 19:1191–1192
82. Meszaros J, Nimmerfall F, Rosenthaler J, Weber H (1975) Permanent bile duct cannulation in the monkey. A model for studying intestinal absorption. Eur J Pharmacol 32:233–242
83. West WL, Cheatham LR, Gaillard ET, Wright M (2002) A chronic bile duct and intravenous cannulation model in conscious rabbits for pharmacokinetic studies. J Invest Surg 15:81–89
84. Merle-Melet M, Bresler L, Didelot JP, Jehl F, Gerard A, Boissel P (1994) A surgical model for studying biliary excretion of drugs in the awake micropig yucatan. J Exp Anim Sci 36:201–208
85. Ghibellini G, Leslie EM, Brouwer KL (2006) Methods to evaluate biliary excretion of drugs in humans: an updated review. Mol Pharm 3:198–211
86. Böhme M, Müller M, Leier I, Jedlitschky G, Keppler D (1994) Cholestasis caused by inhibition of the adenosine triphosphate-dependent bile salt transport in rat liver. Gastroenterology 107:255–265

87. Fouassier L, Kinnman N, Lefèvre G, Lasnier E, Rey C, Poupon R, Elferink RP, Housset C (2002) Contribution of mrp2 in alterations of canalicular bile formation by the endothelin antagonist bosentan. J Hepatol 37:184–191

88. Yoshikado T, Takada T, Yamamoto H, Tan JK, Ito K, Santa T, Suzuki H (2013) Ticlopidine, a cholestatic liver injury-inducible drug, causes dysfunction of bile formation via diminished biliary secretion of phospholipids: involvement of biliary-excreted glutathione-conjugated ticlopidine metabolites. Mol Pharmacol 83:552–562

89. Funk C, Pantze M, Jehle L, Ponelle C, Scheuermann G, Lazendic M, Gasser R (2001) Troglitazone-induced intrahepatic cholestasis by an interference with the hepatobiliary export of bile acids in male and female rats. Correlation with the gender difference in troglitazone sulfate formation and the inhibition of the canalicular bile salt export pump (Bsep) by troglitazone and troglitazone sulfate. Toxicology 167:83–98

90. Feng B, Xu JJ, Bi YA, Mireles R, Davidson R, Duignan DB, Campbell S, Kostrubsky VE, Dunn MC, Smith AR, Wang HF (2009) Role of hepatic transporters in the disposition and hepatotoxicity of a HER2 tyrosine kinase inhibitor CP-724,714. Toxicol Sci 108:492–500

91. Ulloa JL, Stahl S, Yates J, Woodhouse N, Kenna JG, Jones HB, Waterton JC, Hockings PD (2013) Assessment of gadoxetate DCE-MRI as a biomarker of hepatobiliary transporter inhibition. NMR Biomed 26(10):1258–1270

92. Pähler A, Funk C (2008) Drug-induced hepatotoxicity: learning from recent cases of drug attrition. In: Fishbein JC (ed) Advances in Molecular Toxicology, vol 2, 1st edn. Elsevier, Oxford, pp 25–56

93. Yamazaki M, Miyake M, Sato H, Masutomi N, Tsutsui N, Adam KP, Alexander DC, Lawton KA, Milburn MV, Ryals JA, Wulff JE, Guo L (2013) Perturbation of bile acid homeostasis is an early pathogenesis event of drug induced liver injury in rats. Toxicol Appl Pharmacol 268:79–89

94. Ito K, Chiba K, Horikawa M, Ishigami M, Mizuno N, Aoki J, Gotoh Y, Iwatsubo T, Kanamitsu S, Kato M, Kawahara I, Niinuma K, Nishino A, Sato N, Tsukamoto Y, Ueda K, Itoh T, Sugiyama Y (2002) Which concentration of the inhibitor should be used to predict in vivo drug interactions from in vitro data? AAPS PharmSci 4:53–60

95. Grime K, Webborn PJ, Riley RJ (2008) Functional consequences of active hepatic uptake on cytochrome P450 inhibition in rat and human hepatocytes. Drug Metab Dispos 36:1670–1678

96. Lee JK, Paine MF, Brouwer KL (2010) Sulindac and its metabolites inhibit multiple transport proteins in rat and human hepatocytes. J Pharmacol Exp Ther 334:410–418

97. Thompson RA, Isin EM, Li Y, Weidolf L, Page K, Wilson I, Swallow S, Middleton B, Stahl S, Foster AJ, Dolgos H, Weaver R, Kenna JG (2012) In vitro approach to assess the potential for risk of idiosyncratic adverse reactions caused by candidate drugs. Chem Res Toxicol 25:1616–1632

98. Hillgren KM, Keppler D, Zur AA, Giacomini KM, Stieger B, Cass CE, Zhang L (2013) Emerging transporters of clinical importance: an update from the International Transporter Consortium. Clin Pharmacol Ther 94:52–63

99. European Medicines Agency Committee for Human Medicinal Products (CHMP) Guideline on the Investigation of Drug Interactions CPMP/EWP/560/95/Rev. 1 Corr.* http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500129606.pdf

100. Nigsch F, Lounkine E, McCarren P, Cornett B, Glick M, Azzaoui K, Urban L, Marc P, Müller A, Hahne F, Heard DJ, Jenkins JL (2011) Computational methods for early predictive safety assessment from biological and chemical data. Expert Opin Drug Metab Toxicol 7:1497–1511

101. Fujita T, Hansch C (1967) Analysis of the structure-activity relationship of the sulfonamide drugs using substituent constants. J Med Chem 10:991–1000

102. Liu P, Long W (2009) Current mathematical methods used in QSAR/QSPR studies. Int J Mol Sci 10:1978–1998

103. Michielan L, Moro S (2010) Pharmaceutical perspectives of nonlinear QSAR strategies. J Chem Inf Model 50:961–978

104. Broccatelli F (2012) QSAR models for P-glycoprotein transport based on a highly consistent data set. J Chem Inf Model 52:2462–2470
105. Bikadi Z, Hazai I, Malik D, Jemnitz K, Veres Z, Hari P, Ni Z, Loo TW, Clarke DM, Hazai E, Mao Q (2011) Predicting P-glycoprotein-mediated drug transport based on support vector machine and three-dimensional crystal structure of P-glycoprotein. PLoS One 6:e25815, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3186768/
106. Pedersen JM, Matsson P, Bergström CA, Norinder U, Hoogstraate J, Artursson P (2008) Prediction and identification of drug interactions with the human ATP-binding cassette transporter multidrug-resistance associated protein 2 (MRP2; ABCC2). J Med Chem 51:3275–3287
107. Sakurai A, Kurata A, Onishi Y, Hirano H, Ishikawa T (2007) Prediction of drug-induced intrahepatic cholestasis: in vitro screening and QSAR analysis of drugs inhibiting the human bile salt export pump. Expert Opin Drug Saf 6:71–86
108. Carlsson L, Helgee EA, Boyer S (2009) Interpretation of nonlinear QSAR models applied to Ames mutagenicity data. J Chem Inf Model 49:2551–2558
109. Stålring J, Almeida PR, Carlsson L, Helgee Ahlberg E, Hasselgren C, Boyer S (2013) Localized heuristic inverse quantitative structure activity relationship with bulk descriptors using numerical gradients. J Chem Inf Model 53(8):2001–2017
110. Matsson P, Artursson P (2013) Computational prospecting for drug-transporter interactions. Clin Pharmacol Ther 94:30–32
111. Wang L, Dong H, Soroka CJ, Wei N, Boyer JL, Hochstrasser M (2008) Degradation of the bile salt export pump at endoplasmic reticulum in progressive familial intrahepatic cholestasis type II (PFIC II). Hepatology 48:1558–1569
112. Crocenzi FA, Sánchez Pozzi EJ, Ruiz ML, Zucchetti AE, Roma MG, Mottino AD, Vore M (2010) Ca^{2+}-dependent protein kinase C isoforms are critical to estradiol 17β-D-glucuronide-induced cholestasis in the rat. Hepatology 48:1885–1895
113. Lam P, Xu S, Soroka CJ, Boyer JL (2012) A C-terminal tyrosine-based motif in the bile salt export pump directs clathrin-dependent endocytosis. Hepatology 55:1901–1911
114. Geenen S, Taylor PN, Snoep JL, Wilson ID, Kenna JG, Westerhoff HV (2012) Systems biology tools for toxicology. Arch Toxicol 86:1251–1271
115. Bhattacharya S, Shoda LK, Zhang Q, Woods CG, Howell BA, Siler SQ, Woodhead JL, Yang Y, McMullen P, Watkins PB, Andersen ME (2012) Modeling drug- and chemical-induced hepatotoxicity with systems biology approaches. Front Physiol 3:462. doi:10.3389/fphys.2012.00462

Drug Discovery vs hERG

Derek J Leishman and Zoran Rankovic

Abstract A number of marketed drugs have been withdrawn and numerous clinical and preclinical compounds discontinued due to prolongation of the time interval between Q and T waves of the electrocardiogram (long QT syndrome). Drug-induced QT interval prolongation, which may lead to life-threatening cardiac arrhythmia torsades de pointes (TdP), has been linked to blockade of a cardiac potassium channel, a product of the human ether-à-go-go-related gene (hERG). Consequently, drug discovery efforts across the pharmaceutical industry have been utilizing hERG in vitro assays to predict and minimize the risk of QT prolongation. This chapter discusses in silico, in vitro, ex vivo, and in vivo methods employed to measure blockade of hERG and QT prolongation, as well as regulatory recommendations, and medicinal chemistry strategies utilized to circumvent interactions with hERG.

Contents

D.J. Leishman and Z. Rankovic (✉)
Eli Lilly, 893 South Delaware Street, Indianapolis, IN 46285, USA
e-mail: rankovic_zoran@lilly.com

Abbreviations

ADME   Absorption, distribution, metabolism, and elimination
CNS   Central nervous system
FDA   Food and Drug Administration
ICH   International conference on harmonization
IP   Inflection point

1 Introduction

The assessment of a compound's affinity for and ability to block the human ether-a-go-go-related gene (hERG) potassium channel (Kv11.1) has become a standard and required assay in drug development. This off-target pharmacology is remarkably common. Estimates made when hERG block rose to prominence, through examining either published accounts or a random screening exercise, suggest more than 65% of tested compounds blocked the channel or bound to the channel with a potency of 10 μM or less. The ability to inhibit the hERG channel is therefore one of the most common off-target activities a medicinal chemist is likely to face.

In the late 1980s and early 1990s several pharmaceutical companies were exploring class III anti-arrhythmic drugs as a therapy for atrial fibrillation. These were drugs designed to prolong the refractory period of the cardiac muscle. The atrial refractory period was prolonged by such drugs. The QT interval of the electrocardiogram (Fig. 1), a measure of ventricular depolarization and repolarization, was also prolonged. Thus there was only very modest atrial selectivity for these agents. Electrophysiological evaluation in isolated cardiac myocytes identified an ionic current in the heart with the properties of a delayed rectifying potassium current. Studies also demonstrated that the current could be separated into two components biophysically and pharmacologically [1]. One of these components was rapidly activating and became known as IKr. This current was sensitive to the class III anti-arrhythmic compounds which were being developed. In the 1990s it also became apparent that some noncardiovascular compounds which prolonged the QT interval and had been associated with the cardiac polymorphic ventricular tachyarrhythmia torsade des pointes (TdP) also blocked IKr [2]. Given the difficulty of isolating IKr current for study, the assay of choice to evaluate delayed repolarization through blockade of IKr was examination of action potential duration in dog Purkinje fibers. Purkinje fibers respond to IKr blockade with a prolongation of 100% or more whilst other cardiac tissues such as guinea pig

Fig. 1 The electrocardiogram illustrating the components; P, Q, R, S, T. The key interval from the onset of Q until the end of T is also shown. This interval represents the depolarization and repolarization time for the ventricles of the heart
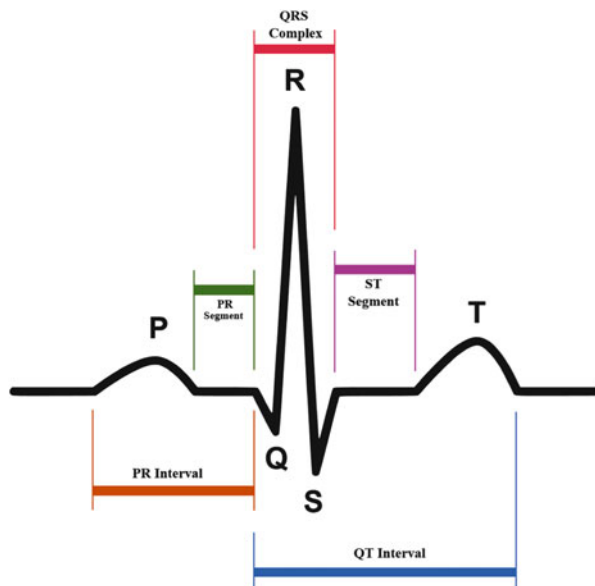
papillary muscles prolonged by around 30%. The sensitivity of the Purkinje fiber resulted in this assay being suggested alongside in vivo QTc prolongation (QTc is the QT interval corrected for the effect of heart rate) as nonclinical assessments for the ability to prolong cardiac repolarization for all new medicines in a CMP Points to Consider document first released for comment in 1996 and finalized in 1997 [3]. Meanwhile in 1995 the hERG potassium channel was first isolated and quickly identified as the specific molecular target of the class III anti-arrhythmic compounds and the vast majority of those noncardiovascular drugs associated with QTc prolongation and TdP [4]. In the intervening years a hERG potassium channel assessment has replaced the dog Purkinje fiber action potential duration study as the method of choice and required assay to examine the potential for delayed cardiac repolarization in current globally recognized testing guidelines [5].

Prior to the existence of the current international nonclinical and clinical regulatory guidelines, a number of drugs were withdrawn from the market owing to the association with TdP. One might question how such a pharmacologically promiscuous activity may have been missed in prior nonclinical and clinical testing paradigms. There are a number of reasons for this related to species differences in sensitivity to IKr blockade and the relative infrequency of a TdP event even on a background of prolonged QT interval. Firstly neither rats nor mice use the current IKr to any significant extent in cardiac repolarization. They are, therefore, insensitive to drug exposures which might be expected to completely block the IKr current for the entire duration of chronic toxicology studies. Fatal arrhythmias would not occur in such studies and an effect on morbidity would have been the primary potentially observable event. Secondly the dog has a relatively short QT interval compared to its heart rate (around 60% of the human value at comparable heart

rates). Thus even marked prolongation only results in QT intervals which are still far shorter than the interval between successive beats and does not result in arrhythmia. Again there would be no impact on the morbidity of dogs owing to IKr inhibition. Thirdly in monkey, where the relative QT interval compared to the interval between beats is more like man, TdP can occur. However the TdP arrhythmia in its definition is self-reverting and would be expected to return to normal sinus rhythm after a short period. The arrhythmia becomes fatal when it degenerates into ventricular fibrillation rather than self-reverting. In a model of TdP in the monkey, where sensitivity to QT prolongation and arrhythmia is increased by AV node ablation, TdP can be induced but is usually not fatal [6]. Thus whilst TdP may induce fainting, owing to ineffective circulation, it would rarely be fatal in monkey. Overall in these animal species the selective and complete block of IKr is unlikely to lead to death. In contrast, it is unlikely that the same could be said of other cardiac ionic currents such as the sodium and calcium currents or the potassium current IK1 which controls resting membrane potential; these are likely to be fatal in commonly tested animal species. Thus in nonclinical studies the effects would primarily be detectable in in vitro assays and sensitive QTc assessments (which had not been routine). In man despite the fact that many compounds do block IKr there remain relatively few very potent blockers and few which can give 50% or more block within tenfold of the plasma exposures achieved routinely in man [7, 8]. Even deliberate targeting of IKr with class III anti-arrhythmic compounds which prolong the QT interval in man by an average of 40 ms at a therapeutic dose [9] is associated with TdP in only a small proportion of patients [10]. A literature review suggests that 95% of cases of TdP with cardiovascular and noncardiovascular drugs occur when the QT or QTc interval is 500 ms or more [11]. A QTc interval of 500 ms is around 100 ms longer than the normal duration for most people.

Overall this important off-target pharmacology has led to fatalities amongst patients and the withdrawal of drugs from the market. Around 1% of all drugs approved since 1950 have been withdrawn owing to an association with TdP, around 3% are probably associated with TdP to some extent, and up to 10% are associated with TdP under some circumstances (for a list of drugs associated with TdP, see www.AZCERT.org). No drugs approved since the introduction of the current regulatory guidelines ICH S7B and ICH E14 have been withdrawn owing to an association with QTc prolongation or TdP. The current concern is that this success may have come at the cost of many hERG blocking or QTc prolonging agents which may not have caused TdP and which may have been successful therapies for unmet medical needs.

2   hERG Physiology and Pharmacology

Cardiac electrophysiology and electrophysiology in general are characterized by nomenclature dominated by description. The term IKr was derived from the observations that current (I) was being measured, it was potassium (K) current,
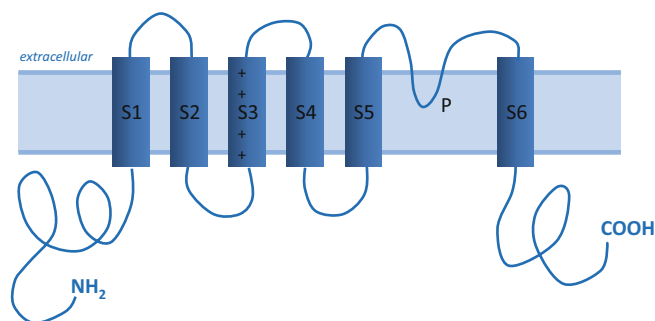
hERG channel

Fig. 2 A diagram depicting a single hERG subunit containing six alpha-helical transmembrane domains, S1–S6. Each hERG molecule has 1,159 amino acids and contains six membrane-spanning domains. A homotetramer of four hERG molecules form the pore α-subunit of I_{kr} potassium channel [12]. A reentrant loop forms the K^+ selective pore within the channel

and it was the rapidly activating component (r) of the delayed rectifying current. In the field of ion channel genetics many ion channels were recognized by the phenotype observed in drosophila that had a mutation in a channel or where a phenotype was triggered pharmacologically and that effect was associated with an ion channel. In one such example fruit flies exposed to ether were observed to shake and move in a way which reminded the observer of the dancing in the then popular Hollywood "Whiskey a Go Go" dance club. The isolated protein was named ether-à-go-go, it was a potassium channel component abbreviated to *eag*. A human analog was isolated and named human ether-à-go-go-related gene (hERG Fig. 2.) [12, 13]; subsequently other species including mouse ether-à-go-go (mERG [14]) and dog ether-à-go-go (cERG [15]) have been evaluated. In considering the physiology and pharmacology of hERG this raises a very important distinction and potential knowledge gap. The ionic current IKr is the current which flows through the native cardiac channel. The protein encoded by hERG will form a potassium channel when expressed in heterologous expression systems. A total of four hERG proteins come together to form the channel. The current which flows through this hERG potassium channel is similar to but not identical to IKr [16]. The native cardiac channel has not been fully described and in general will contain four alpha subunits (hERG), four beta subunits, and other ancillary proteins. Our understanding of the constituents of the native channel has come from studies of hereditary long QT syndromes (LQTs) which have identified subunits associated with the current. The hERG protein also exists in different subtypes [17] which raises the possibility of heteromeric forms of channel. The extent to which the potential differences between composition of the native channel and a hERG-only expressed channel impact conclusions on pharmacology remains unclear. However no compound has been identified to block hERG current and not block IKr and vice versa.

It is likely that the differences would be observed in more subtle variations between potency estimates in assays examining hERG current and IKr.

In addition to its obvious role in cardiac repolarization in most species, the hERG channel has other physiological roles. Developmentally two interesting observations can be made. Firstly even in rat and mouse where in adults IKr plays little role in cardiac repolarization the hERG channel is involved in fetal cardiac electrophysiology and hERG blockers have been demonstrated to be teratogenic in animal studies [18]. Secondly all cancer cells tested to date appear to express hERG endogenously [19]. Taken together these observations might suggest a role for hERG in membrane potential control across many tissues during development.

The hERG channel subunits are also expressed in neuronal tissues in the brain and periphery. In these tissues roles such as adaptation of neuronal firing rates have been hypothesized [20].

The existence of hERG subtypes expressed differentially in different tissues and the potential diversity in the full composition of native hERG-based potassium channels suggest some functional selectivity of IKr blockers may be observed. The class III anti-arrhythmic compounds, which were potent selective IKr blockers, might have offered some insight here into effects on other tissues although practically the effect on the heart was very dominant. In PK/PD modeling it appears hERG blockade equivalent to only 5% may be associated with measurable QTc interval prolongation in man and that 50% of the maximal effect in man occurs at concentrations which would only block around 20% of the hERG current measured in vitro [21]. Thus in man the cardiac role of hERG is very dominant. There are no examples of noncardiac selective hERG blockers.

The IKr and hERG currents share other unusual and novel features. Potassium channels are outward and repolarizing owing to the concentration gradient between high potassium inside the cell and low potassium outside the cell. When the potassium outside the cell drops this would increase the gradient and would be expected to increase the potassium current – this does occur for most potassium channels. In the case of IKr and hERG current there is a role of potassium ions in channel activation and lower external potassium reduces the current [22]. Thus hypokalemia prolongs the QTc interval in man [23]. Acidosis also has an impact on hERG current since external protons also appear to influence channel properties – a modest fall in pH can dramatically impact IKr [24] and hERG current [25].

The hERG channel is remarkably pharmacologically promiscuous. When there is a paucity of potent selective blockers of other potassium channels it seems unusual that medicinal chemists can make potent and selective hERG blockers accidentally. As studies have explored the mechanism of block and the structural features of hERG channels an explanation of the pharmacological promiscuity has emerged.

2.1   Mechanisms of Channel Block

Voltage-dependent ion channels tend to have large portions of conserved structure. Common functions such as formation of a transmembrane pore and voltage sensing lead to conserved structure. Differences between the voltage-gated channels which

Fig. 3 Schematic depiction of the putative interactions between MK-499 (anti-arrhythmic) and the hERG channel. The alpha-subunit inner helices (S6) and loops extending from the pore helices (PH) to the selectivity filter form the channel inner cavity and the ligand-binding site. For clarity, S6 and PH domains of only two subunits of the tetrameric channel are shown (longitudinal view with respect to the cell membrane)
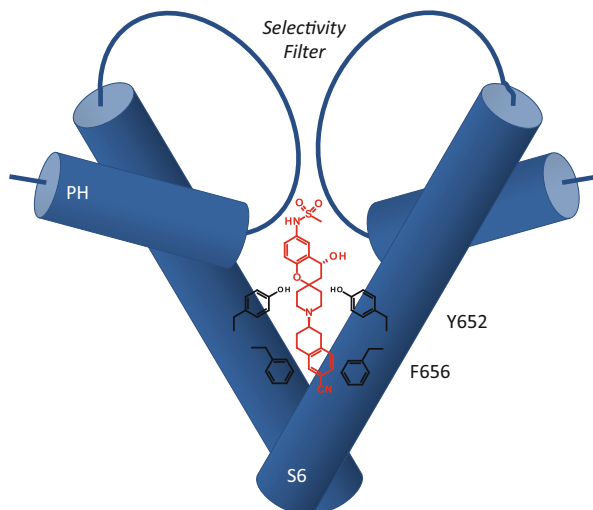
are sodium, calcium, potassium, or cation selective channels obviously exist in the selectivity filter of the pores. The mechanism by which a pore can select between potassium and the smaller sodium ion was part of a series of studies which earned Roderick Mackinnon a Nobel Prize. This pioneering work also helped uncover an intriguing difference between the hERG channel and other potassium channels. Whilst most voltage-dependent potassium channels have a proline–other amino acid–proline (PXP) motif of amino acids in the alpha helix forming the pore, the hERG channel lacks this motif [26]. The PXP motif kinks the alpha helix and limits the size of the so-called vestibule between the activation gate and the selectivity filter/pore of the potassium channel. As hERG lacks this PXP motif it lacks the kinked helix and has a much larger vestibule. This vestibule can accommodate drugs and these can even be trapped in the vestibule upon closure of the activation gate. The work of Sanguinetti and colleagues [27] using scanning alanine mutagenesis to alter successively all the amino acids which were thought to line the vestibule then demonstrated three key binding spots for compounds. One of these was deep in the vestibule near the pore and selectivity filter. Another was the tyrosine at position 652 (Y652) and the third was the phenylalanine at position 656 (F656). The potency of block for the class III anti-arrhythmic compound MK499 was reduced by mutation of any of the three locations. The potency of block for the noncardiovascular agents, cisapride and terfenadine, was not affected by change to the pore region but was affected by mutation of either Y652 or F656 (Fig. 3). Thus overall the mechanistic studies have shown that hERG has a large vestibule with key amino acids to interact with compounds. Combined with electrophysiological studies demonstrating that hERG block is use dependent a general mechanism of block has been described. The blocker crosses the cell membrane and accesses the vestibule when the activation gate opens. When it occupies the vestibule it impedes the passage of potassium ions and thus blocks current. The

compound may or may not be trapped in the channel when the activation gate closes. Subsequent evaluations of a number of different hERG blockers have shown that they are impacted by mutations at either or both Y652 or F656 [28, 29]. These include an evaluation of erythromycin which despite its size has hERG blocking potency also dependent on Y652 and F656 [30]. These studies have confirmed that there is a relatively limited mechanism for hERG block common across most compounds.

In common with other ion channels there are toxins which can block IKr/hERG current. The scorpion toxin, CnErg1, blocks hERG current [31]. A large protein molecule such as the toxin would not be anticipated to cross the cell membrane to block the hERG channel. Indeed the nature of the block with the toxin is clearly different. The block lacks the use dependence and is relieved by prolonged or large voltage [32]. These are features common to closed channel blockers. The toxin has been shown to bind to the channel outside of the cell membrane in a manner consistent with closed channel block. There are no reported accounts of small molecule closed channel hERG blockers.

There are other structural features of hERG which might suggest other mechanisms of modulation. There is a nucleotide-binding domain on the channel. Current through excised patches also tends to "run down" very quickly suggesting some feature of the internal milieu is important in maintaining hERG current. Intriguingly hERG channels expressed in human embryonic kidney cells are inhibited by both hypoglycemia and hyperglycemia via different second-messenger systems [33]. This and the observed effects of kinase inhibition [34] suggest other pharmacological mechanisms of modulation exist. There are accounts of hypo- and hyperglycemia in man being associated with QTc prolongation [35, 36]. These indirect modulatory influences raise a further opportunity. The possibility of increasing hERG current pharmacologically could exist. Indeed in the routine screening evaluations for hERG blockers a small number of potentiators of hERG current have emerged [37].

In the field of class I anti-arrhythmics the existence of distinct phenotypes has long existed. The first classification by Vaughn–Williams divided these into Ia, Ib, and Ic on a basis which might have included different ion channel effects over and above the common effect of blocking the cardiac sodium channel which they share. Subsequently Campbell [38] classified these on the basis of the kinetics of their sodium channel block – notably their off-rate or dissociation rate which is more robust as it is not concentration dependent. The compounds remain classified as Ia, Ib, and Ic but on a kinetic basis. The implications for effects in man and their clinical utility differ for the different subtypes. Electrophysiological evaluation of hERG blockers has suggested they differ with respect to effects on channel activation. They also differ kinetically. Cocaine has kinetics of interaction which are very quick [39]. The class III anti-arrhythmics dofetilide and E4031 have very slow kinetics. The kinetics of the majority of other compounds lies between these. Further consideration of the structural basis of hERG block with some compounds binding to all three key sites (pore, Y652 and F656), many binding to only 2, and some to only 1 further suggests there may be different phenotypes of hERG blocker.

No systematic attempt to classify hERG blockers exists and it is unclear if different phenotypes would have different clinical or safety implications.

A further consideration of hERG block mechanism concerns longer-term changes. It has become clear that in different pathophysiological conditions cardiac remodeling occurs and the proportions of expressed channels change which, in turn, changes the electrophysiological properties of the tissue. In the case of hERG it has been demonstrated that compounds can impact the expression of the channel primarily by disrupting the trafficking of the hERG protein from the endosome to the cell membrane. The compounds arsenic and pentamidine are notable examples of compounds which have this effect on hERG trafficking at concentrations more potent than for direct hERG blockade [40]. Other compounds have effects on hERG trafficking at concentrations similar to or less potent than for direct effects [41]. In the clinic pentamidine has been demonstrated to have a delayed QTc prolongation [40]. Whilst it has become clear that for a few compounds there is a very specific effect which disrupts hERG trafficking, the physicochemical properties of the majority of compounds impacting hERG are such that they are also implicated in lysosomal accumulation and effects on autophagy. Medicinal chemical strategies minimizing hERG block and effects such as phospholipidosis, binding to cytochrome P450 (CYP) and p-glycoprotein (PGP), are all likely to minimize the more nonspecific effects on trafficking. This would impact the prevalence of the more specific trafficking blockers. Specific effects on trafficking might then be deduced from their mechanism, detected in vitro or detected in an in vivo evaluation of sufficient duration.

2.2   In Vitro and In Vivo Assays

Since the CHMP Points to Consider document focused attention on routine evaluation of cardiac repolarization there has been considerable attention devoted to different assays both in vitro and in vivo. Technology has advanced over the intervening years and different assays are now commonplace compared to those initially used. The assays used comprise of those specifically designed to evaluate effects of compounds on the hERG channel and those which integrate the activity at hERG and other ion channels.

2.2.1   Direct Measurement of Effects on hERG

The technologies used to assess activity at other channels have been reviewed extensively (e.g., [42]). The hERG-specific assays can measure effects on the ionic current directly, be indirect measures or examine channel trafficking. The patch-clamp technique is the primary mechanism used to directly evaluate hERG channel current. The use of heterologous expression systems is preferred since these isolate hERG current in a background devoid of electrical activity and boost

the magnitude of the measured current. Whilst the magnitude of native IKr current in acutely isolated cardiac myocytes can measure around 50–100 pA, the hERG potassium current in HEK cells can measure 1–2 nA. The cardiac myocytes have additional sodium, calcium, and potassium channels all of which are voltage dependent and some means of isolating the hERG current is necessary. The myocyte isolation process also takes time on a daily basis making this a laborious process. The heterologous expression systems also facilitate the use of higher-throughput patch-clamp devices. Overall hERG channel patch-clamp studies are almost exclusively conducted in heterologous expression systems.

The specific parameter measured is also determined by the more unusual properties of the hERG current. Upon depolarization from a relatively hyperpolarized holding potential (e.g., −80 mV) the hERG channel opens allowing outward current to flow. However the channel inactivates very rapidly such that increasing depolarization does not increase the level of instantaneous current since it also inactivates strongly with increased depolarization. When the depolarization ends and the voltage is clamped back to more hyperpolarized levels, the inactivation is equally quickly released, current increases and then slowly diminishes as the channel slowly deactivates. This increase in current upon repolarization gives rise to what are known as tail currents. The pharmacological effect of compounds is generally measured as the extent of inhibition of the peak outward tail current. More sophisticated measures such as the half-maximal voltage for activation, inactivation, or deactivation can also be measured to examine the effects of compounds. The kinetics of effects on channel current and the voltage dependence of block are also parameters which have been measured. However the overwhelming parameter of choice is the extent of inhibition of the peak tail current. This yields percentage inhibition and where block exceeds 50% an IC_{50} can be determined along with a Hill slope or Hill coefficient. The Hill slope is usually close to 1; however, technical considerations can modify this. An example would be whether or not rundown occurs and whether this is corrected for. The rundown in current magnitude is an effect whereby the magnitude of the current diminishes with time, probably owing to the loss of an intracellular factor through dialysis with the solutions used to fill the patch-clamp pipette. The speed with which inhibition with drug reaches equilibrium is concentration dependent taking a long time with low concentrations. Thus the balance of drug effect and rundown decreases with increased concentration. Consequently, block at low concentrations can be overestimated and this flattens the Hill coefficient driving this to less than one. If a fixed duration of drug application is used (normally relatively short), this would tend to underestimate block at low concentrations making the Hill coefficient greater than one. These technical considerations can add variability to assessments of hERG potency although these are likely to be modest and account for less than threefold differences in potency estimates. Equally such variability raises the question of whether an apparent difference in potency between two compounds of three to fivefold is real.

The technical considerations can be relatively subtle since two laboratories given the same 12 drugs to examine and using identical cells and an identical

patch-clamp voltage protocol yielded different results. There was a systematic difference of around three- to fivefold between the laboratories and for one drug, pimozide, the difference was 30-fold [43].

Examination of the literature where a single compound has been tested in more than one laboratory suggests that some compounds are more affected by technical considerations than others. In a systematic study cisapride was shown to exhibit variable potency with protocol differences which was more dramatic than the drug dofetilide [44]. An explanation of these differences is lacking but differences in the kinetics of drug effect may be a factor.

The technology used for patch-clamp studies has changed in the last decade. Whilst manual patch-clamp recording from individual identified cells with a glass micropipette remains the gold standard technique, it has largely been replaced by higher-throughput planar patch technologies. In the latter, individual cells are not identified; rather, cells are seeded in a well in a specialized plate. A random cell will cover a small (1–2 μm) hole in the bottom of the well, and after disruption of the electrical integrity of the small patch of membrane obscuring the hole, the whole-cell current is measured in a manner analogous to the manual technique. Where this practice may depart from the manual technique is in the extent of the series resistance imposed by the electrical contiguity of the connection to the cells interior, the extent to which this can be compensated and then whether or not voltage control is permanently maintained over the cell. These factors can all have an impact on the measured potency of block and variability in this measure. This is an acknowledged compromise made in using higher-throughput technology in order to gain the increased capacity which allows more compounds to be tested earlier in the drug discovery process.

2.2.2 Indirect Measurement of Effects on hERG

Using more indirect measures of hERG block is also a compromise used primarily to gain access to higher capacity and lower cost testing. The most prevalent of these technologies currently is radioligand or fluorescence probe binding. In these assays a radiolabeled [45, 46] or fluorescently tagged [47] high-potency hERG blocker binds to the hERG channels in membranes made from cells heterologously expressing hERG. The potency with which test compound displaces the tagged binder is determined and expressed as an IC_{50} or Ki. Published studies with radiolabeled dofetilide [45] and astemizole [46] exist as well as studies with a fluorescently labeled dofetilide [47]. These technologies allow very high throughput especially if techniques such as scintillation proximity assays can be used to eliminate the filtration step otherwise required to separate bound from unbound radiolabel. Although nominally these techniques are indirect and it is possible that a compound could bind elsewhere to block hERG, the observation that all small molecules tested to date appear to bind to common amino acids in the vestibule where space is limited means that displacement assays are unlikely to miss hERG blockers. Theoretically it is possible that a binding displacer may be a potentiator of

current rather than a blocker. The relative paucity of potentiators and the likelihood that the compounds bind high in the vestibule impeding potassium ion flow suggests that the possibility of misclassifying a compound is low. Other indirect assays of hERG interaction do exist but few are less expensive or offer higher throughput than the binding displacement assays making the displacement assays the most commonly utilized assays in early hERG liability schemes.

Nowadays, early ion channel testing may also involve examination of additional ion channels such as cardiac calcium and sodium channels. Subsequent in vivo or in vitro assays are then used as a means of examining the integrated response to the compound in cardiac tissues.

2.2.3   Effects of hERG Block in In Vitro Tissue Assays

The effects of hERG blockers have been evaluated in a range of in vitro tissue assays. These can be considered to fall into two broad categories. Firstly there are integrative assays used to assess the likelihood of repolarization in vivo and in man – these are primarily native cardiac tissue assays examining repolarization duration. Secondly there are proarrhythmia assays designed to address the question of whether or not any given compound is likely to be proarrhythmic. These assays examine indices beyond solely examining a measure of cardiac repolarization – these have generally involved a higher degree of tissue organization such as in ventricular wedge preparations or whole-heart preparations. Focusing on the former the original CPMP Points to Consider document [4] took advantage of the high signal to noise ratio in the canine Purkinje fiber action potential duration to suggest this assay to detect IKr block. These integrative assays do allow an assessment of effects on duration with native ion channel composition under physiological conditions. Thus where different hERG evaluations may suggest different potencies for any given compound, the tissue action potential assay offers the possibility of determining the likely concentrations at which prolongations may occur. However, the broader integrative properties of these assays add a key complication. They also express the other ion channels key in determining the properties of the cardiac action potential. The canine Purkinje fiber, for instance, expresses sodium, calcium, and potassium channels and is a tissue in the conduction system of the heart. This means it expresses a large density of sodium channels and is sensitive to blockade of these channels. Blockade of sodium channels in canine Purkinje fibers leads to a slowing of the maximum upstroke velocity in the action potential and a slowing of conduction from the stimulating electrodes. It also causes a shortening of action potential duration. The influx of positively charged sodium ions down their concentration gradient contributes to the depolarization during a cardiac action potential both during the early more transient phase of sodium channel activation as well as during the plateau phase of the action potential owing to a small late-phase component of sodium current. Calcium ions entering through calcium channels are also important in supporting depolarization during the cardiac action potential. Thus a compound which had a propensity to block hERG and additional ion channels would have effects on the cardiac action potential

which are a summation of the activity at different ion channels according to the potency of block under the physiological conditions of the assay. In practice this has meant that some compounds which were able to prolong the QT interval in man and were associated with TdP did not cause any prolongation in canine Purkinje fiber action potentials [48] – so-called false-negative results. Compounds which did prolong action potential duration in Purkinje fiber robustly all also prolonged the QT interval in man. Thus this integrative assay is said to have strong positive predictive value but poor negative predictive value. The search continues for an integrative assay which minimizes the risk of false-negative results and optimizes the positive and negative predictive value. Isolated cardiac myocytes appear to offer an improved balance, prolongation being evident for compounds which failed to prolong Purkinje fiber action potential duration [49, 50]. The emergence of induced pluripotent stem cell-derived cardiomyocytes offers the potential for isolated human cardiac myocytes assays at a scale necessary for wide use and early evaluation.

2.2.4   The In Vivo Assessment of hERG Block

In vivo evaluation for hERG blockers generally focuses on the same measure used in man – QT interval prolongation. It focuses on species other than rats and mice since these do not express hERG in adult cardiac tissues to any significant level. The measurement of QT interval in animals suffers from the inter- and intraindividual variability observed in all species including man. The small number of animals in any given study compounds the issues of variability which can be offset to some extent by the density of data it is possible to capture with the telemetry systems routinely used in animal studies. It is possible to capture and use data from more than 90% of heartbeats in any period from hours to days. Following recommended industry best practices in design, execution, and analysis of the study [51] results in the ability to detect changes of 4 ms using only four animals in the study [52].

2.3   Building a Safety or Selectivity Margin

Clearly there are a number of ways to measure a compound's activity at the hERG channel. These are readily available. Given that choosing a compound with minimal liability to inhibit hERG channel current optimally takes place whilst there are compounds to choose from and potentially still being made, the testing paradigms can encompass very early data in the evaluation paradigm. After a compound is made and chosen for development then the assays can be used in a characterization of the compound to evaluate the potential effects in man. The overall testing paradigm can vary and needs to reflect components of how frequently potent hERG block occurs in the chemistry of interest, how prevalent other ion channel effects are in the same chemistry, and the risk:benefit assessment in a therapeutic

class. The following can be considered as a general framework which could be adapted to suit the specific needs of a drug discovery and development team.

The key difficulty in an assessment of hERG block liability is the uncertainty in the likely range of clinical exposures until the near final stages of clinical development. This makes an assessment of hERG block liability during the chemistry phase of drug discovery very challenging. This can be addressed by first considering this to be a question of intrinsic selectivity which transitions to a question of compound characterization and safety margin as human exposure data becomes available.

It has become clear that only a small amount of hERG block is necessary for a compound to be potentially associated with TdP and QT interval prolongation. Empirically it was noted that a margin between clinical exposure (unbound plasma concentrations) and hERG block IC_{50} of 30-fold was necessary to separate those compounds which were not associated with TdP from those which were [7, 53]. Later a margin of 45-fold was described as best predicting those compounds which would or would not prolong the QT interval in the rigorous clinical assessment described in regulatory requirements (ICH E14) [54]. Consideration of the PK/PD relationship for QT prolongation and pharmacological response characteristics in general makes sense of these empirical observations [20]. When the Hill slope for hERG block is around 1, as it generally is, a 30th of the IC_{50} relates to around 3% block of current and a 45th relates to around 2% block. Thus almost any appreciable block (>3%) of hERG current can be associated with QTc interval prolongation and has the potential to be associated with the arrhythmia TdP. Staying at concentrations below the concentration response curve for hERG is therefore recommended. In terms of intrinsic selectivity, this would suggest that the EC_{50} at a target of interest should be at least 30-fold lower than the hERG IC_{50} if a 50% effective level of compound is to be achieved before any appreciable and significant hERG block occurs. However, a 50% effect may not be sufficient for the desired robust therapeutic effect. If an effect level of 90% is desired, this equates to around tenfold higher than the EC_{50} (again assuming a Hill slope of 1) and would require a margin between target and hERG EC_{50}s of 300-fold. If an effect level of >97% is desired, then the margin between EC_{50}s would be around 1,000-fold. This 1,000-fold margin is a good rule of thumb (Fig. 4). There are other considerations which may modify this rule. If the target is in the CNS but penetration into that compartment is low, then an added margin may be required.

The variability in the estimates of hERG potency with different patch-clamp protocols impacts these assessments of intrinsic selectivity. The current standard patch-clamp protocols trend towards the more potent end of the potency assessments. Thus a margin of 1,000-fold may be more difficult than necessary to obtain but is likely to be conservative. A way in which to deal with the variability in potency estimates may be to use two protocols or two methods to determine the hERG block potency. In practical terms this often occurs by combining an early high-throughput assessment with a medium- or low-throughput assessment. It is possible to combine a binding displacement assay early with a later assessment of hERG patch clamp. The displacement assays trend towards lower potency estimates if they do not agree with the potency as estimated by patch clamp. Having access to the two assays means that
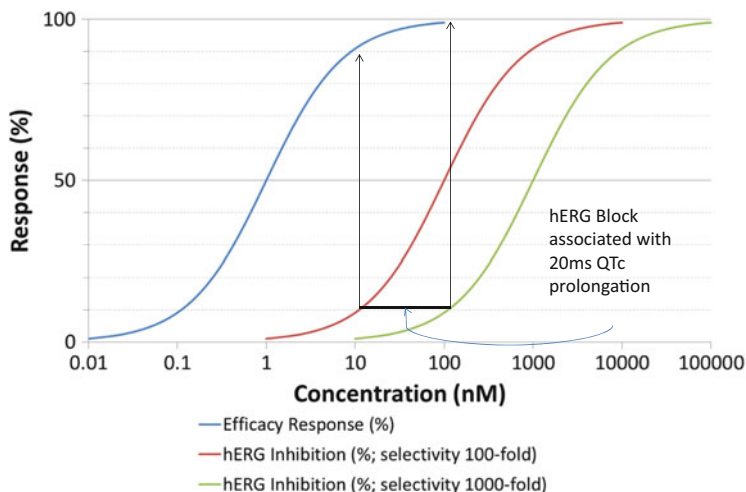
Fig. 4 Figure illustrating the extent of overlap in concentration–response curves with different degrees of separation. The extent of hERG block necessary to induce a 20 ms QTc change in man is illustrated along with the extent to which the same concentration would impact the desire pharmacological target. The figure illustrates the value of an intrinsic selectivity of 1,000-fold to separate the two pharmacologies and allow maximum engagement at target with minimum engagement of hERG

where assays agree (values are different by around threefold or less), one can be confident of likely concentrations associated with QT interval prolongation and one can have confidence in the 1,000-fold margin. Where the assays differ in potency estimate, some experience in taking such compounds to subsequent in vitro or in vivo evaluations is necessary to determine which assay to follow as more indicative of in vivo concentrations associated with QT interval prolongation. Having established a selectivity margin of 1,000-fold or more, a compound may progress with some confidence and a limited need for further follow-up until the regulatory required assays are conducted. If the selectivity margin is less than 1,000-fold or there are other drivers suggesting 1,000-fold may not be sufficient, assessments such as the in vivo assay can be conducted. In general these can be designed to reach plasma concentrations of maximum in vivo efficacy and/or between one fifth the hERG IC_{50} and the hERG IC_{50}. Using best practice principles these concentrations would be anticipated to give robust QT interval prolongation if the hERG potency assessments are predictive of in vivo effect.

3 Clinical and Regulatory Aspects of QT Prolongation

The first regulatory activity around QT prolongation and arrhythmia dates to the CPMP Points to Consider document in 1996. The assessment of the propensity to delay cardiac repolarization had not matured to sufficient consensus by 2000 when the original specific international safety pharmacology guidance, ICH S7A,

appeared [55]. Part B which focused on cardiac repolarization assessment was released in 2005 as ICH S7B [6]. This guidance document accompanied the clinical guidance ICH E14 [56]. The clinical guidance concerns the testing for the potential to prolong the QT interval in man and describes the Thorough QT (TQT) study. The TQT study is used prior to the larger-scale patient exposure in phase 3 clinical studies to determine the level of assessment of QT interval and arrhythmia which is required. ICH S7B describes the need for a hERG patch-clamp study and an in vivo evaluation of QTc prolongation. Other mechanistic studies may also be conducted as necessary. The accumulated data is to be included in an integrated assessment taking into account cardiac repolarization data along with relevant pharmacokinetic and related data. It disappointed many that in 2005 when ICH S7B and ICH E14 were released that the outcome of the nonclinical (S7B) studies were of no consequence regarding the need to conduct a TQT study. The implications of a positive TQT study were a more extensive QT testing requirement in phase 3 studies as well as restrictive labeling which could limit the patient population receiving drug and offer a differentiation opportunity in the therapeutic marketplace.

3.1 Emergence of a New Regulatory Paradigm

Some 8 years after the release of ICH E14, public discussions initiated by the Food and Drug Administration in the United States have suggested that a changed process is necessary and one in which the TQT study may not be routinely required [57]. These TQT studies are relatively expensive and require additional information from other studies in order to conduct and interpret effectively. It was thought that although almost all agents which were associated with TdP prolonged the QT interval and were hERG blockers, the reverse was not true. Not all hERG blockers will be associated with TdP and not all drugs prolonging the QTc interval will be associated with TdP. Thus the negative predictive value of assessment schemes based on hERG and QT interval prolongation may be strong but the positive predictive value is weak. Having accumulated experience from more 250 TQT studies and their associated nonclinical and early clinical QT assessment data, the hypothesis around the weak positive predictive power and strong negative predictive power has been borne out. No compound developed since ICH E14 appeared has had to be withdrawn owing to TdP; however, there is a concern that this has been at the expense of many false-positive signals and compounds with halted development owing to hERG block or QT interval prolongation. The suggested change to the process places much more emphasis on the nonclinical assessment. This combined with the earliest clinical assessments in phase 1 studies will determine the level of QT and arrhythmia assessment required in the latter stages of clinical development rather than relying on routine TQT studies. The nonclinical assessment would involve assessment at hERG and other key ion channels with those data being placed into context of cardiac repolarization through the use of computer simulation of cardiac myocyte action potentials or through myocyte testing. The exact measure which will be the primary index of proarrhythmia

Fig. 5 The hERG
pharmacophore which has
emerged over the last
decade. More subtle
structure–activity
relationships exist within
chemical series to tune
hERG affinity

C_0 4.5-6.5 Å C_1

5-9 Å   5.5-7 Å   4.5-7.5 Å

N

C_2
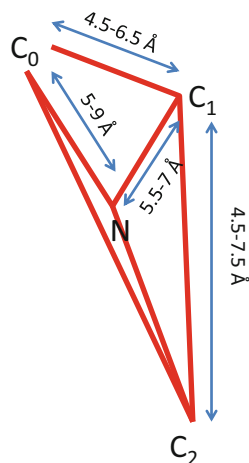
remains unclear. A simple index of prolongation may suffer the same weaknesses as QT interval prolongation itself.

4 In Silico Models

The first in silico hERG pharmacophore models appeared around 2000 [58, 59]. The basic pharmacophore proposed has been refined since and predicts a minimum structure which can potently block hERG [60]. Using this pharmacophoric model (Fig. 5) simple compounds were made to test this model and the importance of the components it describes [60]. The success of this study illustrates both the importance of the basic components shown in Fig. 1 and the impact additional features can have on driving hERG potency. One of the compounds in the small study described had a hERG IC_{50} value of 2.4 nM, one of the most potent hERG blockers published.

The decade or more of experience in routine hERG assessments using largely consistent techniques for long periods facilitates creation of data sets for in silico modeling. This has been further aided by using IC_{50} assessments rather than single-concentration testing. The largest pharmaceutical companies in particular have demonstrated the changes necessary to reduce hERG liability [61] and they possess very large databases on which to construct in silico models and use these to prescreen compounds [62]. The in silico models can be put to use very early, prior to any synthesis. The most simple may be in the context of a classification model based on the first practical studies which may be conducted, for instance, a classification model based on a binding displacement assay. If the target is a 1,000-fold margin and desired potency range for a target is 10 nM or less, then an appropriate cutoff may be 10 μM. Thus, compounds may be classified as likely to have a binding displacement $IC_{50} < 1$ μM, undeterminable, or >10 μM. Should the data fall in the middle grey category, some synthesis and testing would be warranted.

5 Medicinal Chemistry Strategies for Optimization of hERG Selectivity

The purpose of the following discussion is to summarize the approaches utilized to circumvent activity at hERG, identified through an extensive survey of the medicinal chemistry literature. Optimizations are recorded as pairs of compounds, which have been categorized in terms of the tactics employed to diminish hERG activity. In contrast to global computational hERG models where heterogeneous data sets are often used, this approach has the advantage of dealing with optimization pairs from the same chemical series, with data generated under identical assay conditions. Therefore, there exists a direct relationship between the start and end points of the optimization which has enabled determination of the most appropriate method to employ depending on the nature of the starting compound. The resulting analysis is aimed at producing a set of simple, empirical guidelines for attenuating hERG activity, which will be of utility to the practicing medicinal chemist.

The reported optimizations were grouped into the following classification categories based on common descriptors: control of LogP, attenuation of pK_a, incorporation of a negative charge (INC), and discrete structural modifications (DSM) [63]. Although this review attempts to focus on examples where there is only one clear site of modification, it should be noted that it is not possible to alter one parameter in isolation and this may confound interpretation of the controlling factor mitigating hERG activity. In addition, many of the research groups who encounter hERG issues employ several tactics to diminish this unwanted activity. Therefore, the categorization applied here is best regarded as a mnemonic rather than a rigorous classification.

5.1 Control of cLogP

Lipophilicity, as estimated by cLogP, is generally considered to be one of the most, if not the most important physicochemical property, the control of which is critical for ultimate success in drug discovery and development [64]. This property reflects the critical event of molecular desolvation in transfer from aqueous phases to cell membranes and to protein-binding sites, which are mostly hydrophobic in nature. Increase of ligand lipophilicity usually results in improved in vitro potency, which makes it a relatively straightforward and tempting medicinal chemistry optimization strategy. However, if lipophilicity is too high, there is an increased likelihood of not only poor solubility and high metabolic clearance but also binding to multiple targets and resultant pharmacologically based toxicology. Indeed, SAR data often show series-dependent correlation between cLogP and potency of hERG binding [63, 65]. This can be rationalized by the hERG homology models and mutagenesis data [66, 67] suggesting that the ligand-binding site is within a large hydrophobic cavity aligned with aromatic residues. Therefore, reduced lipophilicity could
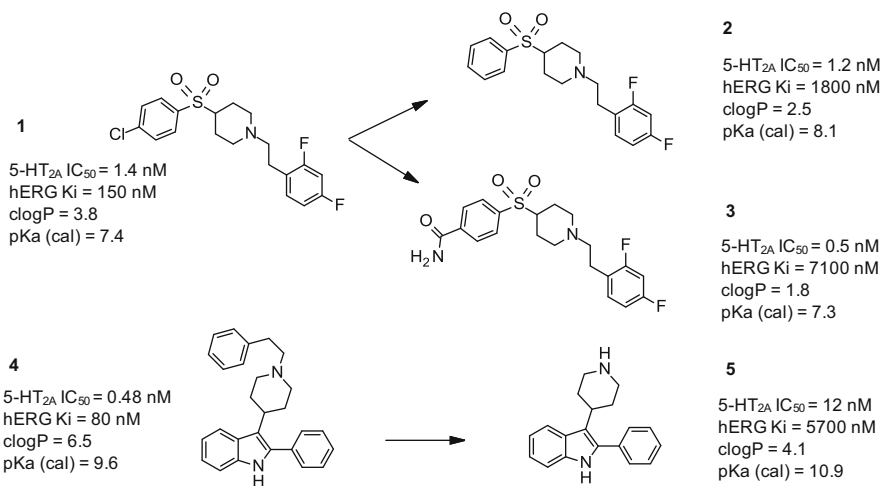
Fig. 6 Examples of the clogP control strategy

destabilize ligand interactions within the channel's hydrophobic cavity and consequently result in decreased hERG potency.

This strategy was very effectively employed by Fletcher and co-workers at Merck Sharp and Dohme seeking to improve selectivity over hERG in a series of 5-HT2$_A$ antagonists represented by lead compound **1** [68]. That was achieved by deletion of the Cl atom resulting in **2** with cLogP and hERG Ki values reduced by over a log unit and by its replacement with a polar carboxamide group to give **3** with even lower cLogP and hERG potency, 1.8 and 7.1 μM, respectively. Importantly, in both examples the primary 5-HT2$_A$ potency remained unaffected by these structural changes (Fig. 6).

In another series of 5-HT2$_A$ antagonists reported by a group at Merck Sharp and Dohme [69], the authors hypothesized that the phenethyl group in lead compound **4** conferred high activity at hERG (Ki = 80 nM). Indeed, deletion analogue **5** displayed significantly reduced hERG binding (Ki 5.7 μM) consistent with its lower cLogP (4.1), thus providing an attractive point for further optimization.

It is also important to remember that an aromatic ring is an important element of the hERG-binding pharmacophore [63, 66]. Reduction of the number of aromatic rings is frequently reported as a successful medicinal chemistry strategy not only in terms of improving selectivity against hERG but also CYPs, aqueous solubility, serum albumin binding, and a number of other parameters that impact a compound's overall developability. Indeed, the mean aromatic ring count was found to decline as compounds advance through clinical trials, suggesting that compounds with fewer aromatic rings are more likely to be successful in development [70]. The average number of aromatic rings in FDA-approved oral drugs is 1.6 [71]. This is consistent with findings of a complementary study that the fraction of sp^3-hybridized carbon atoms (Fsp^3 = number of sp^3-hybridized carbon atoms/total carbon atom count) increases with progression through the development process

6

MCHR1 = 7 nM
hERG = 160 nM
clogP = 4.7

7

MCHR1 = 12 nM
hERG = 1750 nM
clogP = 3.4

8

CCR5 2.6 nM
hERG 150 nM
clogP 6.0

9

CCR5 2.3 nM
hERG 570 nM
clogP 4.9

10

CCR5 0.9 nM
hERG >30000 nM
clogP 3.9

11

CCR5 39 nM
hERG 410 nM
clogP 3.5

12

Nav1.7 = 2900 nM
hERG = 160 nM
clogP = 4.4

13

Nav1.7 = 3100 nM
hERG = 320 nM
clogP = 3.7

14

Nav1.7 = 620 nM
hERG = 2900nM
clogP = 3.9
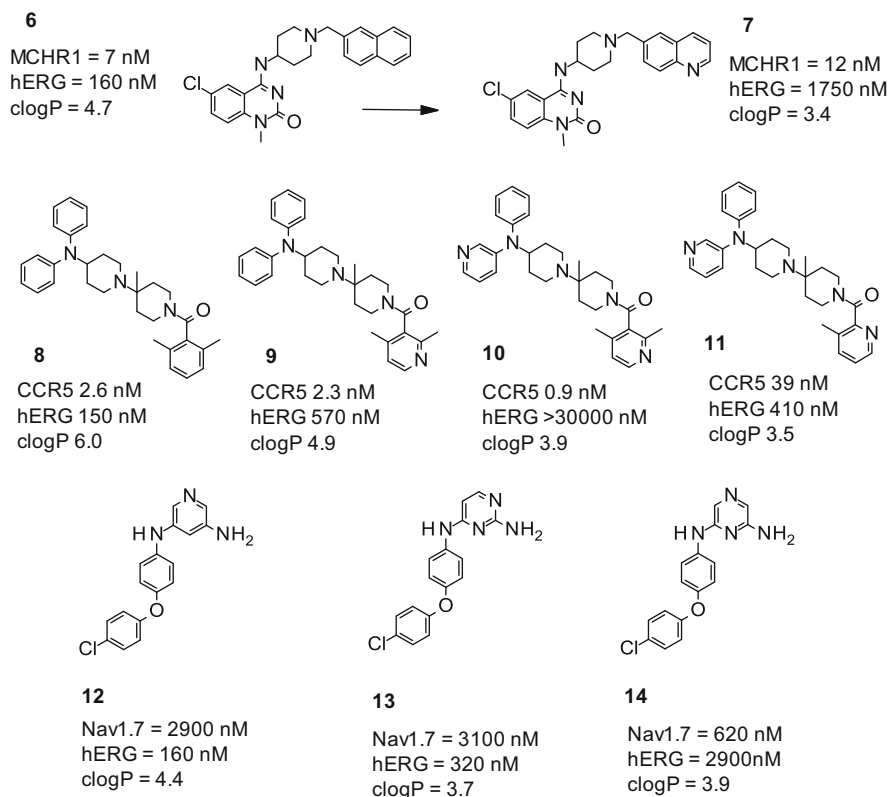
Fig. 7 Examples of the clogP control by replacing carboaromatics

[72]. Of course, it is not always possible to delete an aromatic group without significantly effecting activity for the primary target. In such cases replacing the carboaromatic (e.g., phenyl rings and benzo-fused ring systems) with a corresponding heteroaromatic groups may prove a more successful approach. It has been shown that the detrimental impact of increasing aromatic ring count is driven mainly by the carboaromatic ring component – for any given number of aromatic rings in a molecule, the higher the carboaromatic with respect to heteroaromatic content, the greater the risk of failure in development. It is therefore advisable not only to limit the overall aromatic rings in a molecule but also, where possible, to replace carboaromatics with heteroaromatic congeners [73].

This was precisely the approach adopted by Blackburn et al. in their efforts to improve hERG selectivity in a series of antagonists of the melanin-concentrating hormone receptor-1 (MCHR1), a particularly challenging target due to its close similarity to the hERG pharmacophore [74]. Replacement of the naphthyl group of **6** with a quinolyl group produced **7** with significantly reduced lipophilicity (1.3 log units) and improved selectivity over hERG from around 20 to almost 200-fold (Fig. 7).

For highly lipophilic molecules replacing only one of multiple carboaromatics may not be sufficient to significantly impact hERG binding. In an example reported by researchers at Novartis, exchange of the benzamide group in CCR5 antagonist **8** (cLogP 6.0; hERG 150 nM) with a corresponding pyridine moiety resulted in a modest fourfold reduction of hERG binding (**9**; LogP 4.9; hERG = 570 nM) [75]. To achieve a more significant effect on cLogP and hERG binding, incorporation of an additional pyridine group was required (**10**; cLogP 3.9; hERG >30 μM). Interestingly, a closely related positional analogue **11** is still a potent hERG binder (410 nM). A similarly dramatic difference in hERG binding displayed by heteroaromatic positional analogues has been observed in Nav1.7 antagonist SAR reported by a group at Amgen [76]. In this chemical series, pyridine **12** was found to be a potent hERG blocker (160 nM) with only modest Nav1.7 activity (2.9 μM). Replacement of the pyridine core of **12** by a pyrimidine moiety produced very little effect (**13**; hERG 320 nM; Nav1.7 3.1 μM), while a corresponding pyrazine replacement yielded derivative **14** with a dramatically reversed selectivity profile (hERG 2.9 μM; Nav1.7 620 nM). Since these are not unique examples of this kind in the literature, it is important when pursuing a carboaromatics replacement approach to explore alternative heteroatom positions.

5.2   Attenuation of pK_a

Although not a prerequisite to hERG blockade, many of the ligands that block hERG do contain a basic nitrogen that is likely to be protonated at physiological pH. Mutagenesis and homology modeling studies suggest that the contribution of a basic amine to hERG-binding affinity is related to π-cation interactions with aromatic residues within the channel's cavity [66, 77]. Consequently, lowering the pK_a of a basic nitrogen would reduce the proportion of molecules in the protonated form at physiological pH and would be expected to disrupt any putative π-cation interactions with the channel. In general, it has been observed that modification of pK_a can often have the effect of increasing the polarity of a compound (e.g., piperidine to piperazine) and thus can have the same net impact as LogP. In order to focus on pK_a-related effects in this section, we discuss examples with no significant change in cLogP (<1 log unit).

The pK_a lowering approach was adopted by Merck Sharp and Dohme scientists in their efforts to improve hERG selectivity within a new series of conformationally constrained hNK1 antagonists [78]. Lead compound N-2 methyl tetrazole analogue **15** (Fig. 8) had an attractive profile with high hNK1 affinity (IC_{50} = 1 nM) and long-lasting in vivo efficacy in the foot-tapping gerbil paradigm; however, it was compromised by a hERG liability (Ki = 0.1 μM). Introduction of a fluorine at the C-6 position of the azabicyclic ring to give **16** (calc pK_a = 5) maintained the hNK1 affinity observed in the parent des-fluoro analogue **15** (calc pK_a = 7.3) but greatly improved the selectivity over hERG (Ki > 10 μM).

Returning to the 5-HT_{2A} area, Fletcher et al. demonstrated how control of pK_a can be used to gain additional selectivity over hERG [68]. Compound **17** was found
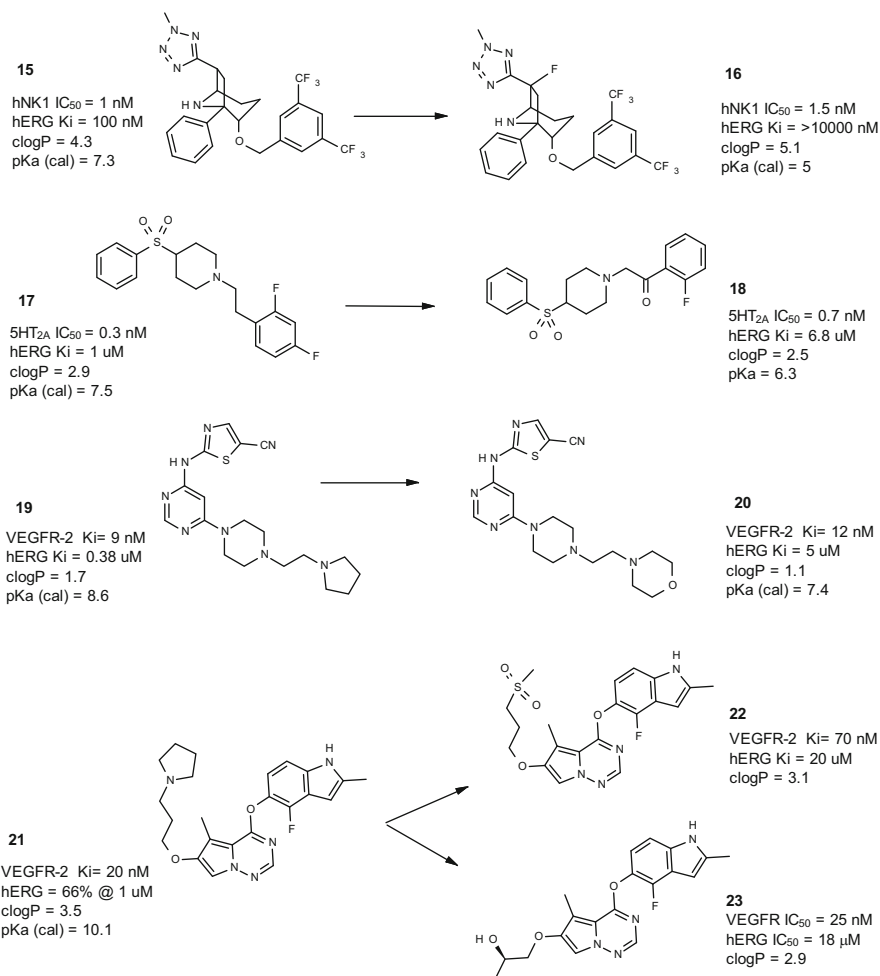
Fig. 8 Examples of the pK_a attenuation strategy

to prolong the QT interval in the anesthetized ferret by more than 10% at doses of three and 10 mg/kg/h. A range of modifications were made to **17** in order to further improve selectivity over hERG. Some success was achieved through the introduction of a ketone at the β-position to the amine (**18**, Fig. 8) where attenuation of the pK_a of the piperidine system yielded a commensurate reduction in activity at hERG. The optimized compound did not show QT prolongation in the anesthetized ferret model at doses up to 10 mg/kg/h and exhibited acceptable pharmacokinetic properties.

Similarly, in a series of VEGFR-2 (KDR) kinase inhibitors reported by Sisko et al., exchange of the terminal pyrrolidine system of **19** with a morpholine unit in **20** significantly reduced pK_a and provided a corresponding improvement in hERG
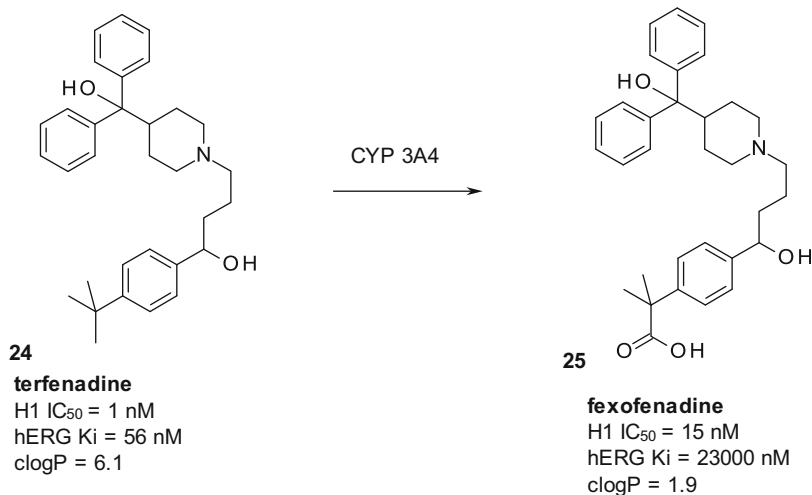
Fig. 9 Terfenadine 24 and its main metabolite, fexofenadine 25

activity [79]. Another series of VEGFR-2 kinase inhibitors which also displayed significant activity at the hERG channel were the pyrrolotriazine-based systems exemplified by 21 [80]. The authors hypothesized that the basic amine side chain was responsible for the anti-target activity, and they sought to replace it with alternative nonbasic solubilizing groups such as the sulfone in 22. Indeed, 22 exhibited considerably reduced activity at hERG as well as reduced inhibition of CYP3A4. The alkoxy region of the molecule proved to be highly tolerant of substitution which enabled the use of other solubilizing groups such as the secondary alcohol 23, which exhibited a similar profile (Fig. 3). Compound 23 was then derivatized as a prodrug which then became the development candidate [80].

5.3 Incorporation of a Negative Charge

As a potassium cation channel, hERG has evolved to stabilize positive charge within its central cavity, which may at least in part explain why many hERG blockers contain basic functionalities that can be protonated under normal physiological conditions. This may also explain the fact that the presence of functionality that is negatively charged at physiological pH, such as a carboxylate group, is almost universally detrimental to hERG binding [81].

This effect was observed for the first time with terfenadine 24 (Fig. 9), the progenitor in the second generation of antihistamines launched in 1982. Due to instances of cardiac arrhythmia, terfenadine was withdrawn from the market in 1997. However, it was subsequently discovered that its principal metabolite,
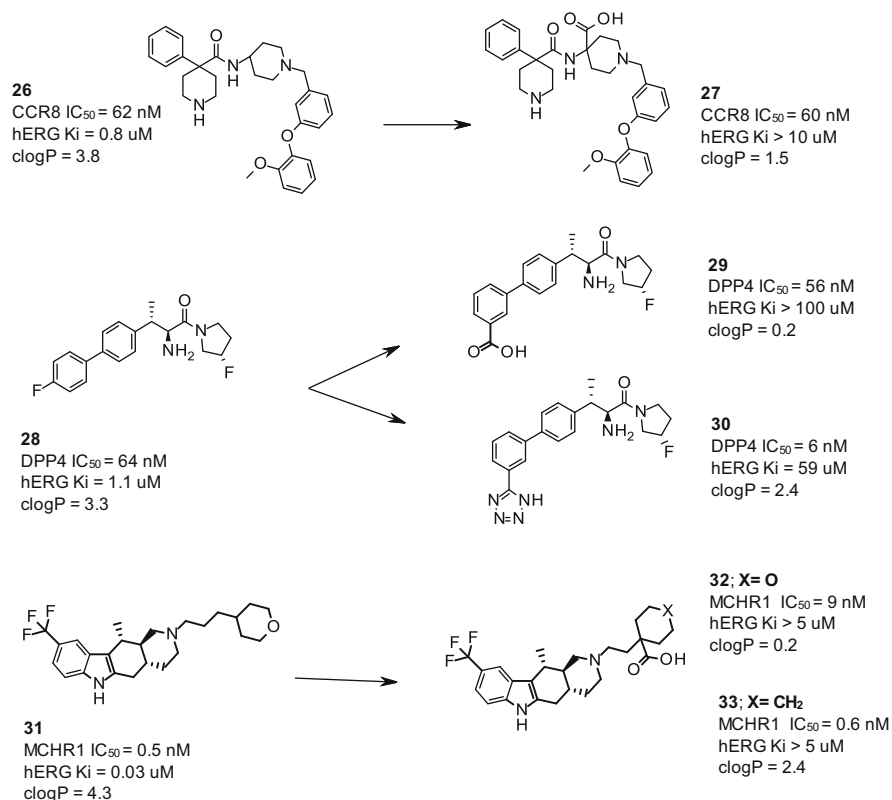
Fig. 10 Examples of the incorporation of negative charge (INC) approach

carboxylate **25,** not only accounts for all of the therapeutic effect of terfenadine, but it also displays significantly reduced hERG affinity and no effect on QT interval [82]. It is important to note that this profound effect on the hERG binding may also be, at least in part, attributed to the significant lipophilicity reduction that accompanies the incorporation of polar groups such as the carboxylate (over 4 log units for **25**). This metabolite was subsequently marketed as fexofenadine, the first of the third generation of antihistamines, characterized by a lack of central side effects intrinsic to the prior generations. The improved side-effect profile has been attributed to low brain exposure of the zwitterionic fexofenadine.

This serendipitous discovery has been subsequently used as one of the rational approaches to dial out hERG activity. In most cases the starting compound is an amine with hERG activity, into which a carboxylic group (or a suitable bioisostere) is incorporated to obtain a zwitterionic analogue with attenuated hERG potency.

For example, piperidine **26**, a potent lead compound in a CCR8 antagonist program reported by Ghosh et al. [83], was also a potent hERG blocker with a Ki value of 800 nM (Fig. 10). The authors found that incorporation of a carboxylic acid into the 4-position of the aminopiperidine of **26** to furnish a corresponding

zwitterionic analogue **27** practically eliminated hERG binding while retaining the CCR8 activity.

A group at Merck Sharp and Dohme pursued the same approach to mitigate the hERG liability in a series of DPP-4 inhibitors [84]. Again, the incorporation of a carboxylic acid into the basic amine-containing **28** gave zwitterionic **29** and abolished hERG binding without affecting the desired DPP-4 potency. Unfortunately carboxylate **29** displayed poor pharmacokinetic properties with oral bioavailability of less than 3%. Oral administration to portal vein-cannulated rats indicated that the low bioavailability is due to poor absorption. To improve the ADME profile, a range of carboxylic acid bioisosteres were synthesized; however, although they all retained high hERG selectivity, they also exhibited poor pharmacokinetic properties as exemplified by tetrazole analogue **30** (hERG 59 μM; $F = 3.5\%$), which underlines a potential challenge in this approach [85].

However, the INC-related ADME issues are not always insurmountable, as exemplified by a research group at Amgen developing MCHR1 antagonists for the treatment of obesity. As indicated earlier in the chapter, the high degree of similarity between MCHR1 and hERG pharmacophores presents a particular challenge in this field. Despite promising preclinical validation data and considerable drug discovery efforts across the pharmaceutical industry, only a very small number of MCHR1 antagonists progressed into the clinic, mostly due to hERG-related cardiovascular risks [86]. The hERG selectivity was also a major challenge in the Amgen MCHR1 antagonist series. The lead compound, **31**, displayed an excellent overall profile including high MCHR1 affinity ($IC_{50} = 0.5$ nM), good ADME properties, and high efficacy in reducing food consumption in an in vivo MCH-cannulated rat model [87]. However, it was also found to be a potent hERG blocker (Ki = 30 nM), which ultimately prevented its further development. To mitigate the hERG liability, the team employed INC strategy, focusing primarily on modifications around the tetrahydropyran (THP) moiety, the region of the molecule with the most tolerant SAR. Indeed, carboxylate analogue **32** with a shorter two-carbon linker showed significantly reduced hERG potency; however, MCHR1 potency was also decreased by 18-fold. Further optimization efforts led to the discovery of cyclohexyl carboxylate **33**, with improved MCHR1 affinity ($IC_{50} = 0.6$ nM) and maintained high selectivity over hERG ($IC_{50} > 5$ μM). This compound also displayed a broad selectivity and good pharmacokinetic profile with low clearance and oral bioavailability across preclinical species ranging from 29 to 79%. The compound showed no QTc change in dog after a single dose (300 mg/kg; C_{max} 33 μM, p.o.) or chronic treatment over 28 days (100 mg/kg; BID, p.o.). The MCHR1 is a CNS-expressed target, and the authors were concerned that due to its zwitterionic nature, compound **33** may lack sufficient brain exposure. As we have seen earlier in the fexofenadine example, zwitterions are generally associated with poor brain penetration for which reason their formation is often employed as a medicinal chemistry strategy when the objective is to design peripherally restricted molecules [88]. However, like several other examples reported in the literature [89, 90], compound **33** was found to be brain penetrant despite its zwitterionic nature, with a [CSF]/[plasma] ratio of 0.21 (10 mpk; p.o. rat). The compound was found to be efficacious in mouse high-fat diet-induced

obesity model at 3 and 10 mg/kg and was ultimately advanced for further clinical evaluation under the code name AMG 076.

5.4 Discrete Structural Modifications

Discrete, often peripheral modifications to a drug molecule can have a dramatic effect on hERG potency, a phenomenon potentially explained by disruption of the putative interactions with aromatic residues (F656 and Y652) lining the hERG channel [27]. Strategies adopted to mitigate hERG activity in this way are not restricted only to modifications to distal aryl rings but can also include introducing constraint and variation in stereochemistry. Criteria for including optimization pairs in this category are as follows: one clear site of modification, >1 log reduction in hERG activity, no significant reduction (<1 log unit) in key physicochemical parameters such as cLogP and pK_a, and maintained activity at the primary target.

A particularly interesting example in this category is a benzimidazole series of NMDA NR2B antagonists reported by researchers at Merck [91]. Lead compound **34** (Fig. 11), lacking a basic nitrogen that is often perceived as one of the key determinants of high hERG activity, displays high potency for the NR2B channel (Ki = 0.7 nM) but also shows high affinity for hERG in the MK-499-binding assay (IP = 120 nM). A simple change of the methyl sulfonamide of **34** to the ethyl sulfonamide in **35** resulted in a 13-fold decrease in hERG affinity. The reduction in hERG affinity may be attributed to specific structural changes to the periphery of the molecule potentially disrupting the interaction of the adjacent aromatic ring with the hERG channel binding site (Fig. 11).

As discussed earlier, attenuation of pK_a as a strategy initially pursued by Huscroft and his colleagues at Merck [78] to improve hERG selectivity in their new series of conformationally constrained hNK1 antagonists was successful. Introduction of a fluorine in the azabicyclic ring of **15** resulted in **16** with significantly improved selectivity over hERG (Fig. 8). However, this change was also accompanied by a reduction in duration of action in vivo. A closely related N-1 methyl analogue **36** displayed around 12-fold lower hERG binding (Ki = 1.25 μM), indicating the high sensitivity of hERG SAR to subtle structural changes in this series (Fig. 11). A more striking DSM example in this particular series is derived from introduction of an α-methyl substituent into the pendent benzyl ether side chain to give **37**, which resulted in attenuated hERG activity (Ki = 1.8 μM) whilst maintaining hNK1 affinity (0.5 nM) and in vivo efficacy [78].

In a CCR5 antagonist program, Shu and co-workers adopted a number of strategies in order to avoid hERG activity and found that peripheral structural effects were important in negating hERG activity [92]. Conversion of the *tert*-butyl group of **38** to the *sec*-butyl system of **39** is accompanied by a greater than tenfold reduction in the hERG activity with complete conservation of potency at the primary target.

Subtle structural modification was one of the strategies employed in an effort to design out hERG affinity within a series of macrocyclic farnesyltransferase
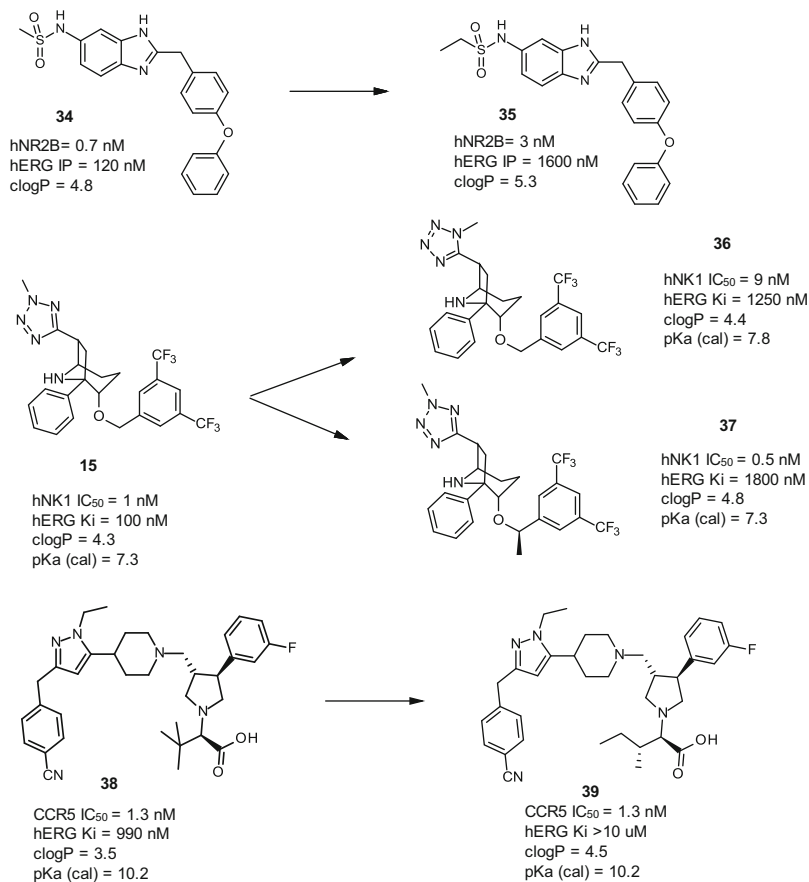
Fig. 11 Application of discrete structural modifications (DSM) to control hERG

inhibitors (FTIs) [93]. Compound **40** was one of the most potent FTIs in this series ($IC_{50} = 0.15$ nM); however, it was also one of the most potent hERG blockers, with an IP of 0.08 μM (Fig. 12). The compound was found to cause 10% prolongation of the QTc interval in the anesthetized dog at a plasma level of 2.4 μM which, although not as severe as hERG-binding data suggested, was deemed unacceptable. The hERG SAR within this macrocyclic series proved to be highly sensitive to minor structural and topological changes, as illustrated by diastereoisomer **41** showing significantly attenuated hERG binding (IP = 4.7 μM) in comparison to the parent compound, **40**. Expansion of the cyclopentyl to a cyclohexyl ring was also successful, delivering tetrahydronaphthalene derivative **42**, one of the most potent and selective FTIs in the series (hERG IP = 7 μM; FT $IC_{50} < 1$ nM).
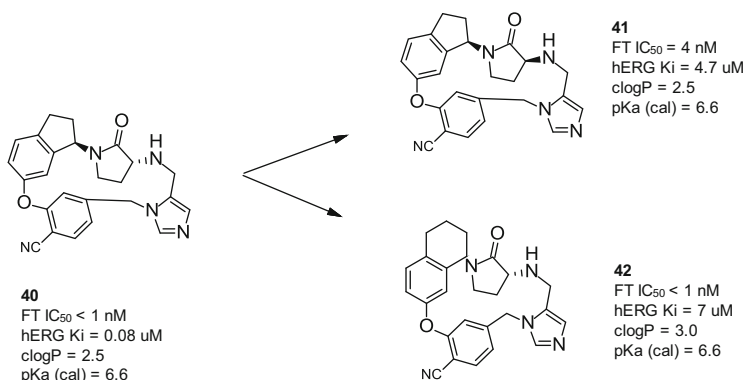
Fig. 12 Examples of the DSM approach: farnesyltransferase inhibitors [93]

5.5 Recommendations: Medicinal Chemistry vs hERG

On the basis of the extensive medicinal chemistry literature analysis discussed above, we propose the following recommendations for projects facing a hERG selectivity challenge [63]:

- Four strategies for the removal of hERG have been identified: control of LogP, attenuation of pK_a, incorporation of negative charge (INC), and discrete structural modifications (DSM). All four strategies were found to be equally efficacious in diminishing hERG activity.
- Consideration of cLogP of the starting compound enables selection of the most appropriate strategy for lowering hERG activity.
  - Where cLogP $\geq$ 3.0: seek to reduce this by, e.g., incorporation of heteroatoms, polar groups, or removal of lipophilic moieties. On average 1 log unit reduction in cLogP leads to 0.8 log unit reduction in hERG activity. If no correlation between hERG activity and LogP can be established in the series, then pursue the DSM strategy.
  - If cLogP < 3.0: DSM offer the highest probability of success.
- A corollary to the above is to remove or modify aryl moieties in the target molecule. This may have the effect of reducing cLogP and potentially disrupting π-stacking with the channel.
- Reducing pK_a of a basic nitrogen is a frequently employed strategy in attenuating hERG. This is often associated with reduction in LogP which may be the more relevant parameter.
- A basic nitrogen (although not a prerequisite) is often associated with hERG activity; therefore, alternative solubilizing groups should be considered.
- Literature evidence indicates that the INC approach resulting in zwitterion formation is often associated with issues related to membrane permeability, oral bioavailability, and brain exposure.

- hERG data should be interpreted in the context of measured solubility. Compounds with low aqueous solubility (<5 mg/L) may have significantly underestimated activity in the hERG assay.
- Only optically pure material should be tested in the hERG assay as the effects due to stereochemistry can be dramatic.
- To ensure the effects of DSM are identified and to establish reliable SAR, test a significant number of analogues within each chemical series against hERG.
- Where the amount of data available permits, consider development of a local (series specific) in silico model to guide medicinal chemistry efforts away from hERG activity.

  – Application of global in silico models is best reserved for prioritization of compounds for synthesis/acquisition from larger arrays.

- To ascertain effects of non-hERG-mediated QT prolongation, test key compounds in relevant ex vivo (e.g., dog Purkinje fibers) or in vivo models as early as possible.

6 Epilogue

The drug-induced blockade of the hERG-related component of the potassium current is thought to be a major reason for drug-induced arrhythmias in humans. Several previously approved drugs (e.g., terfenadine, cisapride, astemizole, and grepafloxacin) have been withdrawn from the market, and the use of others, such as thioridazine, haloperidol, sertindole, and pimozide, has been restricted. Importantly, no compound developed since ICH E14 was introduced has had to be withdrawn due to TdP. However, a concern has been raised that the current paradigm is costly and leads to false positives that may result in potentially successful drugs being discontinued from development. To address these concerns a new paradigm for cardiotoxicity risk assessment has been proposed based solely on nonclinical, in silico and in vitro solutions, with the ultimate goal of replacing the costly thorough QT study.

It should be noted, however, that the dramatic reduction of TdP occurrence in recent clinical development has been a result of not only the methods put in place to detect compounds with such risk but also a consequence of accumulated medicinal chemistry knowledge enabling the design of molecules with the reduced risks of hERG binding. There is now a significant body of information on the tactics for optimizing against the hERG channel. It is anticipated that the ever-burgeoning number of ion channel crystal structures, including mammalian channels, will have considerable impact on our understanding of drug–hERG interactions. Potential challenges in the area are likely to include the identification of additional ion channels implicated in drug-induced cardiac toxicity and subsequent optimization against these targets [94]. Similarly, increasing attention is being paid to other mechanisms which result in cardiotoxic end points such as inhibition of hERG channel trafficking, with compounds such as pentamidine producing QT prolongation via this mechanism rather than through inhibition of hERG [40].

References

1. Sanguinetti MC, Jurkewicz NK (1990) Two components of cardiac delayed rectifier K^+ current differential sensitivity to block by class III antiarrhythmic agents. J Gen Physiol 96:195–215
2. Crumb WJ, Wible B, Arnold DJ, Payne JP, Brown AM (1995) Blockade of multiple human cardiac potassium currents by the antihistamine terfenadine: possible mechanism for terfenadine-associated cardiotoxicity. Mol Pharmacol 47:181–190
3. Anonymous (1997) Committee for Proprietary Medicinal Products (CPMP). Points to consider: the assessment for the potential for QT interval prolongation by non-cardiovascular medicinal products (CPMP/986/96). European Agency for the Evaluation of Medicinal Products, London
4. Sanguinetti MC, Jiang C, Curran ME, Keating MT (1995) A mechanistic link between an inherited and an acquired cardiac arrhythmia: HERG encodes the IKr potassium channel. Cell 81:299–307
5. Anonymous (2005) ICH harmonised tripartite guideline the non-clinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals S7B
6. Sugiyama A (2008) Sensitive and reliable proarrhythmia in vivo animal models for predicting drug-induced torsades de pointes in patients with remodelled hearts. Br J Pharmacol 154:1528–1537
7. Redfern WS, Carlsson L, Davis AS, Lynch WG, MacKenzie I, Palethorpe S, Siegl PKS, Strang I, Sullivan AT, Wallis R, Camm AJ, Hammond TG (2003) Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. Cardiovasc Res 58:32–45
8. De Bruin ML, Pettersson M, Meyboom RHB, Hoes AW, Leufkens HGM (2005) Anti-HERG activity and the risk of drug-induced arrhythmias and sudden death. Eur Heart J 26:590–597
9. Allen MJ, Nichols DJ, Oliver SD (2000) The pharmacokinetics and pharmacodynamics of oral dofetilide after twice daily and three times daily dosing. Br J Clin Pharmacol 50:247–253
10. Singh S, Zoble RG, Yellen L, Brodsky MA, Feld GK, Berk M, Billing CB (2000) Efficacy and safety of oral dofetilide in converting to and maintaining sinus rhythm in patients with chronic atrial fibrillation or atrial flutter: the symptomatic atrial fibrillation investigative research on dofetilide (SAFIRE-D) study. Circulation 102:2385–2390
11. Bednar MM, Harrigan EP, Ruskin JN (2002) Torsades de pointes associated with nonantiarrhythmic drugs and observations on gender and QTc. Am J Cardiol 89:1316–1319
12. Trudeau MC, Warmke JW, Ganetzky B, Robertson GA (1995) HERG, a human inward rectifier in the voltage-gated potassium channel family. Science 269:92–95
13. Stansfeld PJ, Gedeck P, Gosling M, Cox B, Mitcheson JS, Sutcliffe MJ (2007) Drug block of the hERG potassium channel: insight from modeling. Proteins 68:568–580
14. London B, Trudeau MC, Newton KP, Beyer AK, Copeland NG, Gilbert DJ, Jenkins NA, Satler CA, Robertson GA (1997) Two isoforms of the mouse ether-à-go-go–related gene coassemble to form channels with properties similar to the rapidly activating component of the cardiac delayed rectifier K^+ current. Circ Res 81:870–878
15. Zehelein J, Zhang W, Koenen M, Graf M, Heinemann SH, Katus HA (2001) Molecular cloning and expression of cERG, the ether à go-go-related gene from canine myocardium. Pflugers Arch 442:188–191
16. Hancox JC, Levi AJ, Witchel HJ (1998) Time course and voltage dependence of expressed HERG current compared with native "rapid" delayed rectifier K current during the cardiac ventricular action potential. Pflugers Arch 436:843–853
17. Sanguinetti MC, Tristani-Firouzi M (2006) hERG potassium channels and cardiac arrhythmia. Nature 440:463–469
18. Azarbayjani F, Danielsson BR (2002) Embryonic arrhythmia by inhibition of herg channels: a common hypoxia-related teratogenic mechanism for antiepileptic drugs? Epilepsia 43:457–468

19. Bianchi L, Wible B, Arcangeli A, Tagliatatela M, Morra F, Castaido P, Crociani O, Rosati B, Faravelli L, Olivotto M, Wanke E (1998) hERG encodes a K^+ current highly conserved in tumors of different histogenesis: a selective advantage for cancer cells? Cancer Res 58:815–822

20. Chiesa N, Rosati B, Arcangeli A, Olivotto M, Wanke E (1997) A novel role for HERG K^+ channels: spike-frequency adaptation. J Physiol 501:313–318

21. Jonker DM, Kenna LA, Leishman D, Wallis R, Milligan PA, Jonsson EN (2005) A pharmacokinetic-pharmacodynamic model for the quantitative prediction of dofetilide clinical QT prolongation from human ether-à-go-go-related gene current inhibition data. Clin Pharmacol Ther 77:572–582

22. Lin C, Cvetanovic I, Ke X, Ranade V, Somberg J (2005) A mechanism for the potential proarrhythmic effect of acidosis, bradycardia, and hypokalemia on the blockade of human ether-a-go-go-related gene (HERG) channels. Am J Ther 12:328–336

23. Roden DM, Woosley RL, Primm RK (1986) Incidence and clinical features of the quinidine-associated long QT syndrome: implications for patient care. Am Heart J 111:1088–1093

24. Vereecke J, Carmeliet E (2000) The effect of external pH on the delayed rectifying K^+ current in cardiac ventricular myocytes. Pflugers Arch 439:739–751

25. Anumonwo JMB, Horta J, Delmar M, Taffet SM, Jalife J (1999) Proton and zinc effects on HERG currents. Biophys J 77:282–298

26. Shrivastava IH et al (2000) Structure and dynamics of K channel pore-lining helices: a comparative simulation study. Biophys J 78:79–92

27. Mitcheson JS, Chen J, Lin M, Culberson C, Sanguinetti MC (2000) A structural basis for drug-induced long QT syndrome. Proc Natl Acad Sci U S A 97:12329–12333

28. Milnes JT, Witchel HJ, Leaney JL, Leishman DJ, Hancox JC (2006) HERG K^+ channel blockade by the antipsychotic drug thioridazine: an obligatory role for the S6 helix residue F656. Biochem Biophys Res Commun 351:273–280

29. Duncan RS, McPate MJ, Ridley JM, Gao Z, James AF, Leishman DJ, Leaney JL, Witchel HJ, Hancox JC (2007) Inhibition of the HERG potassium channel by the tricyclic antidepressant doxepin. Biochem Pharmacol 74:425–437

30. Duncan RS, Ridley JM, Dempsey CE, Leishman DJ, Leaney JL, Hancox JC, Witchel HJ (2006) Erythromycin block of the HERG K^+ channel: accessibility to F656 and Y652. Biochem Biophys Res Commun 341:500–506

31. Hill AP, Sunde M, Campbell TJ, Vandenberg JI (2007) Mechanism of block of the hERG K^+ channel by the scorpion toxin CnErg1. Biophys J 92:3915–3929

32. Milnes JT, Dempsey CE, Ridley JM, Crociani O, Arcangeli A, Hancox JC, Witchel HJ (2003) Preferential closed channel blockade of HERG potassium currents by chemically synthesised BeKm-1 scorpion toxin. FEBS Lett 547:20–26

33. Zhang Y, Han H, Wang J, Wang H, Yang B, Wang Z (2003) Impairment of human ether-à-go-go-related gene (HERG) K^+ channel function by hypoglycemia and hyperglycemia: similar phenotypes but different mechanisms. J Biol Chem 278:10417–10426

34. Cockerill SL, Tobin AB, Torrecilla I, Willars GB, Standen NB, Mitcheson JS (2007) Modulation of hERG potassium currents in HEK-293 cells by protein kinase C. Evidence for direct phosphorylation of pore forming subunits. J Physiol 581:479–493

35. Marques JLB, George E, Peacey SR, Harris ND, Macdonald IA, Cochrane T, Heller SR (1997) Altered ventricular repolarization during hypoglycaemia in patients with diabetes. Diabet Med 14:648–654

36. Marfella R, Rossi F, Giugliano D (2001) Hyperglycemia and QT interval: time for re-evaluation. Diabetes Nutr Metab 14:63–65

37. Zeng H, Lozinskaya IM, Lin Z, Willette RN, Brooks DP, Xu X (2006) Mallotoxin is a novel human ether-a-go-go-related gene (hERG) potassium channel activator. J Pharmacol Exp Ther 319:957–962

38. Campbell TJ (1983) Importance of physico-chemical properties in determining the kinetics of the effects of class I antiarrhythmic drugs on maximum rate of depolarization in guinea-pig ventricle. Br J Pharmacol 80:33–40

39. Zhang S, Rajamani S, Chen Y, Gong Q, Rong Y, Zhou Z, Ruoho A, January CT (2001) Cocaine blocks HERG, but not KvLQT1 + minK, potassium channels. Mol Pharmacol 59:1069–1076

40. Kuryshev YA, Ficker E, Wang L, Hawryluk P, Dennis AT, Wible BA, Brown AM, Kang J, Chen X-L, Sawamura K, Reynolds W, Rampe R (2005) Pentamidine-induced long qt syndrome and block of hERG trafficking. J Pharmacol Exp Ther 312:316–323

41. Rajamani S, Eckhardt LL, Valdivia CR, Klemens CA, Gillman BM, Anderson CL, Holzem KM, Delisle BP, Anson BD, Makielski JC, January CT (2006) Drug-induced long QT syndrome: hERG K$^+$ channel block and disruption of protein trafficking by fluoxetine and norfluoxetine. Br J Pharmacol 149:481–489

42. Gintant GA, Su Z, Martin RL, Cox BF (2006) Utility of hERG assays as surrogate markers of delayed cardiac repolarization and QT safety. Toxicol Pathol 34:81–90

43. Hanson LA, Bass AS, Gintant G, Mittelstadt S, Rampe D, Thomas K (2006) ILSI-HESI cardiovascular safety subcommittee initiative: evaluation of three non-clinical models of QT prolongation. J Pharmacol Toxicol Methods 54:116–129

44. Milnes JT, Witchel HJ, Leaney JL, Leishman DJ, Hancox JC (2010) Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37°C: cisapride versus dofetilide. J Pharmacol Toxicol Methods 61:178–191

45. Chiu PJS, Marcoe KF, Bounds SE, Lin C-H, Feng J-J, Lin A, Cheng FC, Crumb WJ, Mitchell R (2004) Validation of a [^{3}H]-astemizole binding assay in HEK293 cells expressing HERG K$^+$ channels. J Pharmacol Sci 95:311–319

46. Diaz GJ, Daniell K, Leitza ST, Martin RL, Su Z, McDermott JS, Cox BF, Gintant GA (2004) The [^{3}H]-dofetilide binding assay is a predictive screening tool for hERG blockade and proarrhythmia: comparison of intact cell and membrane preparations and effects of altering [K$^+$]o. J Pharmacol Toxicol Methods 50:187–199

47. Deacon M, Singleton D, Szalkai N, Pasieczny R, Peacock C, Price D, Boyd J, Boyd H, Steidl-Nichols JV, Williams C (2007) Early evaluation of compound QT prolongation effects: a predictive 384-well fluorescence polarization binding assay for measuring hERG blockade. J Pharmacol Toxicol Methods 55:255–264

48. Leishman DJ, Helliwell R, Wakerell J, Wallis RM (2000) Effects of E-4031, cisapride, terfenadine and terodiline on cardiac repolarisation in canine Purkinje fibre and HERG channels expressed in HEK293 cells. Br J Pharmacol 133:130

49. Gintant GA, Limberis JT, McDermott JS, Wegner CD, Cox BF (2001) The canine Purkinje fiber: an in vitro model system for acquired long QT syndrome and drug-induced arrhythmogenesis. J Cardiovasc Pharmacol 37:607–618

50. Salata JJ, Jurkiewicz NK, Wallace AA, Stupienski RF, Guinosso PJ Jr, Lynch JJ Jr (1995) Cardiac electrophysiological actions of the histamine h1-receptor antagonists astemizole and terfenadine compared with chlorpheniramine and pyrilamine. Circ Res 76:110–119

51. Leishman DJ, Beck TW, Dybdal N, Gallacher DJ, Guth BD, Holbrook M, Roche B, Wallis RM (2010) Best practice in the conduct of key nonclinical cardiovascular assessments in drug development: current recommendations from the safety pharmacology society. J Pharmacol Toxicol Methods 65:93–101

52. Sivarajah A, Collins S, Sutton MR, Regan N, West H, Holbrook M, Edmunds N (2010) Cardiovascular safety assessments in the conscious telemetered dog: utilisation of super-intervals to enhance statistical power. J Pharmacol Toxicol Methods 62:12–19

53. Webster R, Leishman D, Walker D (2002) Towards a drug concentration effect relationship for QT prolongation and torsades de pointes. Curr Opin Drug Discov Devel 5:116–126

54. Gintant G (2011) An evaluation of hERG current assay performance: translating preclinical safety studies to clinical QT prolongation. Pharmacol Ther 129:109–119

55. Anonymous (2000) ICH harmonised tripartite guideline safety pharmacology studies for human pharmaceuticals S7A

56. Anonymous (2005) ICH harmonised tripartite guideline the clinical evaluation of QT/QTC interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs E14

57. Sager PT, Gintant G, Turner JR, Pettit S, Stockbridge N (2014) Rechanneling the cardiac proarrhythmia safety paradigm: a meeting report from the cardiac safety research consortium. Am Heart J 1–9
58. Ekins S, Crumb WJ, Sarazan RD, Wikel JH, Wrighton SA (2002) Three-dimensional quantitative structure-activity relationship for inhibition of human ether-à-go-go-related gene potassium channel. J Pharmacol Exp Ther 301:427–434
59. Cavalli A, Poluzzi E, De Ponti F, Recanatini M (2002) Toward a pharmacophore for drugs inducing the long QT syndrome: insights from a CoMFA study of HERG K(+) channel blockers. J Med Chem 45:3844–3853
60. Cavalli A, Buonfiglio R, Ianni C, Masetti M, Luisa Ceccarini L, Caves R, Chang MWY, Mitcheson JS, Roberti M, Recanatini M (2012) Computational design and discovery of "minimally structured" hERG blockers. J Med Chem 55:4010–4014
61. Price DA, Armour D, de Groot M, Leishman D, Napier C, Perros M, Stammen BL, Wood A (2006) Overcoming HERG affinity in the discovery of the CCR5 antagonist maraviroc. Bioorg Med Chem Lett 16:4633–4637
62. Pollard CE, Valentin J-P, Hammond TG (2008) Strategies to reduce the risk of drug-induced QT interval prolongation: a pharmaceutical company perspective. Br J Pharmacol 154:1538–1543
63. Jamieson C, Moir EM, Rankovic Z, Wishart G (2006) Medicinal chemistry of optimizations: highlights and hang-ups. J Med Chem 49:5029–5046
64. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 6:881–890
65. Waring MJ, Johnston C (2007) A quantitative assessment of hERG liability as a function of lipophilicity. Bioorg Med Chem Lett 17:1759–1764
66. Taboureau O, Jorgensen FS (2011) In silico predictions of hERG channel blockers in drug discovery: from ligand-based and target-based approaches to systems chemical biology. Comb Chem High Throughput Screen 14:375–387
67. Fernandez D, Ghanta A, Kauffman GW, Sanguinetti MC (2004) Physicochemical features of the hERG channel drug binding site. J Biol Chem 279:10120–10127
68. Fletcher SR, Burkamp F, Blurton P, Cheng SKF, Clarkson R, O'Connor D, Spinks D, Tudge M, van Niel MB, Patel S, Chapman K, Marwood R, Shepheard S, Bentley G, Cook GP, Bristow LJ, Castro JL, Hutson PH, MacLeod AM (2002) 4-(Phenylsulfonyl) piperidines: novel, selective, and bioavailable 5–HT2A receptor antagonists. J Med Chem 45:492–496
69. Rowley M, Hallet DJ, Goodacre S, Moyes C, Crawforth J, Sparey TJ, Patel S, Marwood R, Thomas S, Hitzel L, O'Connor D, Szeto N, Castro JL, Hutson PH, MacLeod AM (2001) 3-(4-Fluoropiperidin-3-yl)-2-phenylindoles as high affinity, selective, and orally bioavailable h5-HT$_{2A}$ receptor antagonists. J Med Chem 44:1603–1614
70. Ritchie TR, Macdonald SJF (2009) The impact of aromatic ring count on compound developability–are too many aromatic rings a liability in drug design. Drug Discov Today 14:1011–1020
71. Hann MM, Leach AR, 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
72. Lovering F, Bikker J, Humblet C (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52:6752–6756
73. Ritchie TR, Macdonald SJF, Peace S, Pickettc SD, Luscombec CN (2013) Increasing small molecule drug developability in sub-optimal chemical space. Med Chem Commun 4:673–680
74. Blackburn C, LaMarche MJ, Brown J, Lee Che J, Cullis CA, Lai S, Maguire M, Marsilje T, Geddes B, Govek E, Kadambi V, Doherty C, Dayton B, Brodjian S, Marsh KC, Collins CA, Kym PR (2006) Identification and characterisation of amino-piperidinequinolones and quinazolinones as MCHr1 antagonists. Bioorg Med Chem Lett 16:2621–2627
75. Thoma G, Beerli C, Bigaud M, Bruns C, Cooke NG, Streiff MB, Zerwes H-G (2008) Reduced cardiac side-effect potential by introduction of polar groups: discovery of NIBR-1282, an orally bioavailable CCR5 antagonist which is active in vivo. Bioorg Med Chem Lett 18:2000–2005

76. Bregman H, Nguyen HN, Feric E et al (2012) The discovery of aminopyrazines as novel, potent Nav1.7 antagonists: hit-to-lead identification and SAR. Bioorg Med Chem Lett 22:2033–2042

77. Perry M, Sanguinetti M, Mitcheson J (2010) Revealing the structural basis of action of hERG potassium channel activators and blockers. J Physiol 588:3157–3167

78. Huscroft IT, Carlson EJ, Chicchi GG, Kurtz MM, London C, Raubo P, Wheeldon A, Kulagowski JJ (2006) Phenyl-8-azabicyclo[3.2.1]octane ethers: a novel series of neurokinin (NK1) antagonists. Bioorg Med Chem Lett 16:2008–2012

79. Sisko JT, Tucker TJ, Bilodeau MT, Buser CA, Ciecko PA, Coll KE, Fernades C, Gibbs JB, Koester TJ, Kohl N, Lynch JJ, Mao X, McLoughlin D, Miller-Stein CM, Rodman LD, Rickert KW, Sepp-Lorrenzino L, Shipman JM, Thomas KA, Wong BK, Hartman GD (2006) Potent 2-[(pyrimidin-4-yl)amine]-1,3-thiazole-5-carbonitrile-based inhibitors of VEGFR-2 (KDR) kinase. Bioorg Med Chem Lett 16:1146–1150

80. Bhide RS, Cai ZW, Zhang YZ, Quian L, Wei D, Barbosa S, Lombardo LJ, Borzilleri RM, Zheng X, Wu LI, Barrish JC, Kim SH, Leavitt K, Mathur A, Leith L, Chao S, Wautlet B, Mortillo S, Jeyaseelan R, Kukral D, Hunt JT, Kamath A, Fura A, Vyas V, Marathe P, D'Arienzo C, Derbin G, Fargnoli J (2006) Discovery and preclinical studies of (R)-1-(4-(4-fluoro-2-methyl-1H-indol-5-yloxy)-5-methylpyrrolo[2,1-f][1,2,4]triazin-6-yloxy)propan-2-ol (BMS-540215), an in vivo active potent VEGFR-2 inhibitor. J Med Chem 49:2143–2146

81. Zgu B-Y, Jia ZJ, Zhang P, Su T, Huang W, Goldman E, Tumas D, Kadambi V, Eddy P, Sinha U, Scarborough RM, Song Y (2006) Inhibitory effect of carboxylic acid group on hERG binding. Bioorg Med Chem Lett 16:5507–5512

82. Rampe D, Wibble B, Brown AM, Dage RC (1993) Effects of terfenadine and its metabolites on a delayed rectifier K^+ channel cloned from human heart. Mol Pharmacol 44:1240–1245

83. Ghosh S, Elder A, Guo J, Mani U, Patane M, Carson K, Ye Q, Bennett R, Chi S, Jenkins T, Guan B, Kolbeck R, Smith S, Zhang C, LaRosa G, Jaffee B, Yang H, Eddy P, Lu C, Uttamsingh V, Horlick R, Harriman G, Flynn D (2006) Design, synthesis, and progress toward optimization of potent small molecule antagonists of cc chemokine receptor 8 (CCR8). J Med Chem 49:2669–2672

84. Edmondson SD, Mastracchio A, Duffy J, Eiermann GJ, He H, Ita I, Leiting B, Leone JF, Lyons KA, Makarewicz AM, Patel RA, Petrov A, Wu JK, Thornberry NA, Weber AE (2005) Discovery of potent and selective orally bioavailable β-substituted phenylalanine derived dipeptidyl peptidase IV inhibitors. Bioorg Med Chem Lett 15:3048–3052

85. Xu J, Mathvink R, He J, Park YJ, He H, Leiting B, Lyons KA, Marsilio F, Patel RA, Wu JK, Thornberry NA, Weber AE (2005) Discovery of potent and selective phenylalanine based dipeptidyl peptidase IV inhibitors. Bioorg Med Chem Lett 15:2533–2536

86. Hogberg T, Frimurer TM, Sasmal PK (2012) Melanin concentrating hormone receptor 1 (MCHR1) antagonists—still a viable approach for obesity treatment? Bioorg Med Chem Lett 22:6039–6047

87. Mihalic JT, Fan P, Chen X, Chen X, Fu Y, Motani A, Liang L, Lindstrom M, Tang L, Chen J-L, Jaen J, Dai K, Li L (2012) Discovery of a novel melanin concentrating hormone receptor 1 (MCHR1) antagonist with reduced hERG inhibition. Bioorg Med Chem Lett 22:3781–3785

88. Middleton DS, Maw GN, Challenger C, Jessiman A, Johnson PS, Million WA, Nichols CL, Price JA, Trevethick M (2006) Highly potent and selective zwitterionic agonists of the δ-opioid receptor, part 1. Bioorg Med Chem Lett 16:905–910

89. Gianotti M, Botta M, Brough S, Carletti R, Castiglioni E, Corti C, Dal-Cin M, Fratte SD, Korajac D, Lovric M, Merlo G, Mesic M, Pavone F, Piccoli L, Rast S, Roscic M, Sava A, Smehil M, Stasi L, Togninelli A, Wigglesworth M (2010) Novel Spirotetracyclic Zwitterionic Dual H1/5-HT2A receptor antagonists for the treatment of sleep disorders. J Med Chem 53:7778–7795

90. Sullivan S, Guilleminault C (2009) Emerging drugs for insomnia: new frontiers for old and novel targets. Expert Opin Emerg Drugs 14:411–422

91. McCauley JA, Theberge CR, Romano JJ, Billings SB, Anderson KD, Claremon DA, Freidinger RM, Bednar RA, Mosser SD, Gaul SL, Connolly TM, Condra CL, Xia M, Cunningham ME, Bednar B, Stump GL, Lynch JJ, Macaulay A, Wafford KA, Koblan KS, Liverton NJ (2004) NR2B-selective N-Methyl-D-aspartate antagonists: synthesis and evaluation of 5-substituted benzimidazoles. J Med Chem 47:2089–2096

92. Shu M, Loebach JL, Parker KA, Mills SG, Chapman KT, Shen DM, Malkowitz L, Gould SL, DeMartino JA, Siciliano SJ, DiSalvo J, Lyons K, Pivnichny JV, Kwei GY, Carella A, Carver G, Holmes K, Schleif WA, Danezeisen R, Hazuda D, Kessler J, Lineberger J, Miller MD, Emini EA (2004) Antagonists of the human CCR5 receptor containing 4-(pyrazolyl) piperidine side chains. Part 3: SAR studies on the benzylpyrazole segment. Bioorg Med Chem Lett 14:947–952

93. Bell IM, Gallicchio SN, Abrams M, Beshore DC, Buser CA, Culberson JC, Davide J, Ellis-Hutchings N, Fernandes C, Gibbs JB, Graham SL, Hartman GD, Heimbrook DC, Homnick CF, Huff JR, Hassahun K, Koblan KS, Kohl NE, Lobell RB, Lynch JJ, Miller PA, Omer CA, Rodrigues AD, Walsh ES, Williams TM (2001) Design and biological activity of (S)-4-(5-{[1-(3-Chlorobenzyl)-2-oxopyrrolidin-3-ylamino]methyl}imidazol-1-ylmethyl)benzonitrile, a 3-aminopyrrolidinone farnesyltransferase inhibitor with excellent cell potency. J Med Chem 44:2933–2949

94. Kang J, Reynolds WP, Chen XL, Ji J, Wang H, Rampe DE (2006) Mechanisms underlying the QT interval–prolonging effects of sevoflurane and its interactions with other QT-prolonging drugs. Anesthesiology 104:1015–1022

Drug-Induced Phospholipidosis: Prediction, Detection, and Mitigation Strategies

Umesh M Hanumegowda and Alicia Regueiro-Ren

Abstract In the fast-paced and resource-intensive process of the discovery and development of novel medicines, issues pertaining to safety often appear leading to a loss of valuable resource and time for the pharmaceutical industry. Drug-induced phospholipidosis (DIPL) is one such issue that often shows up late in the discovery process, especially after repeat-dose toxicity studies in animals. DIPL is long debated as to whether it is a manifestation of toxicity or just an adaptation response due to drug accumulation in a tissue. Irrespective of the argument on either side, the conservative approach is to avoid DIPL due to closely associated toxicities and pathological similarity with phospholipid storage disorders. Therefore, high importance is given to predict the potential of a novel drug compound/series and to identify/confirm their potential to induce phospholipidosis (PLD) to help steer structure-activity relationship (SAR) to avoid this potential early in discovery. Several drugs with similar physicochemical properties are known to cause PLD, and these are generally referred to as cationic amphiphilic drugs (CADs). Using these known CADs, several in silico, in vitro, and hybrid methods have been developed to predict the potential to induce PLD. Also, a few biomarkers have shown promise of being able to monitor for DIPL during an ongoing animal study without the need to confirm PLD in tissues from necropsy. Early prediction, detection, and mitigation strategies are of immense value in developing novel medicines without PLD-inducing potential.

Keywords Drug-induced phospholipidosis, Cationic amphiphilic drugs, lysosomes, toxicity, drug discovery

---

U.M. Hanumegowda
Discovery Toxicology, Bristol-Myers Squibb Research and Development, Wallingford, CT 06492, USA

A. Regueiro-Ren (✉)
Discovery Chemistry, Bristol-Myers Squibb Research and Development, Wallingford, CT 06492, USA
e-mail: Alicia.regueiroren@bms.com

Contents

Abbreviations

5-HT$_{1B}$	5-Hydroxytryptamine 1B
BMP	bis(monoacylglycero)phosphate
CAD	Cationic amphiphilic drug
Cell lines cited	HepG2, ARLJ301-3, CHO-K1, CHL/IU, J744A, HuH7, rat hepatocytes
Clog$P$	Calculated logarithm of the octanol–water partition coefficient
CMC	Critical micelle concentration
DIPL	Drug-induced phospholipidosis
DPP-IV	Dipeptidyl peptidase IV
Genes cited	MGC4171, NR0B2, INHBE, P8, SERPINA3, ASNS, C10, FLJ10055, FRCP1, AP1S1, TAGLN
H$_3$R	Histamine 3 receptor
hERG	Human ether-a-go-go-related gene
K_d	Dissociation constant
LBPA	Lyso-bis-phosphatidic acid
LDA	Linear discriminant analysis
NBD-PC	7-Nitrobenz-2-oxa-1,3-diazol-4-yl (NBD)-phosphatidylcholine
NBD-PE	7-Nitrobenz-2-oxa-1,3-diazol-4-yl (NBD)-phosphatidylethanolamine
NC	Net charge
PAG	Phenylacetylglycine
pKa	Logarithm of the acidic ionization constant
pKa-MB	Logarithm of the ionization constant-most basic
PLA2	Phospholipase A2
PLD	Phospholipidosis
QSAR	Quantitative structure-activity relationships
QT	Beginning of the QRS complex to the end of the T wave in the electrocardiogram
SAR	Structure-activity relationships
SMARTS	Smiles arbitrary target specification
Softwares/programs cited	ADMET, CAFCA, MC4PC, MDL-QSAR, Weka

TEM	Transmission electron microscopy
TRI	Triple reuptake inhibitor
Vd	Volume of distribution
$\Delta\Delta G_{AM}$	Free energy of amphiphilicity

1 Drug-Induced Phospholipidosis (DIPL)

As the term implies, DIPL is an excessive accumulation of phospholipids in cells due to drugs that are generally administered for longer periods of time. DIPL has been observed in animals more so than in humans for the basic reason that tissue evaluation in repeat-dose toxicity studies that are routinely done in support of safety packages for drugs is conducted at higher doses (relative to a lower efficacious human dose) that translates into a higher tissue burden. An example of DIPL is provided in Fig. 1. In general, DIPL is an adaptive response to drugs of particular characteristics that tend to accumulate in the lysosomal compartment of cells and, therefore, is dependent on dose and the duration of treatment with a drug, in addition to the physicochemical propensity of the drug itself. For this very reason, DIPL is often detected late in exploratory development during longer-term toxicity studies, which could translate into considerable loss of both investment and time. Although DIPL by itself is not considered a toxicity event, concomitant findings of toxicity, functional or histological, in the same tissue or other tissues have led to the consideration of DIPL as a potentially adverse event. Furthermore, the similarity of histological presentation of DIPL with those of lipid storage disorders (e.g., Niemann-Pick, Tay-Sachs) and the morbidities associated with those disorders have strengthened the conservative position of considering DIPL as potentially adverse. The FDA phospholipidosis (PLD) working group [1] has provided excellent examples of the possible correlation of PLD with other toxicities. For example, some of the PLD inducers such as amiodarone, erythromycin, and haloperidol are associated with QT (part of the electrocardiogram) prolongation, an antimalarial with neuropathy and myopathy, and perhexilene and coralgil with hepatotoxicity (refer to Table 1 for structures of the compounds). In particular, the overlap of pharmacophores with compounds inducing PLD and blockers of the cardiac hERG channel that results in QT prolongation with a potential to cause cardiac arrhythmia has been recognized [2]. Therefore, the FDA PLD working group's recommendation is to consider additional studies of QT prolongation, neurotoxicity, myopathy, and hepatotoxicity, if a compound is a PLD inducer. This further adds to the urgency of detecting PLD-inducing potential as early as possible in the drug discovery process in order to efficiently drive programs to develop safe drugs. In this chapter, we describe in brief with examples the mechanisms of DIPL, basic features of PLD inducers, methods to predict, detect, and monitor PLD, and medicinal chemistry strategies to mitigate PLD-inducing potential during early drug discovery.
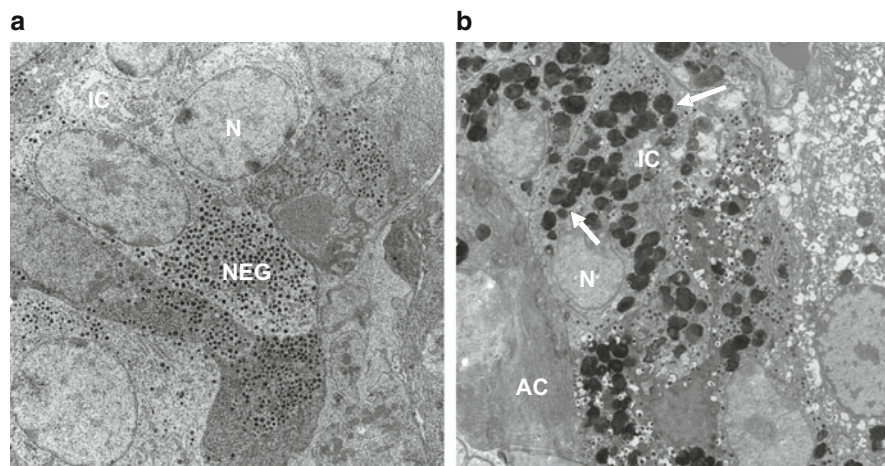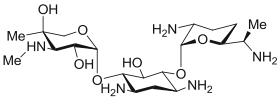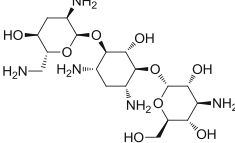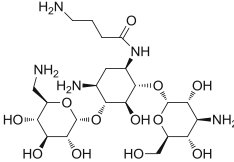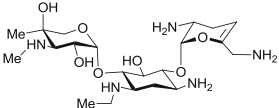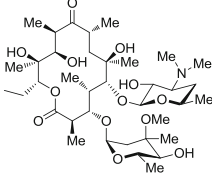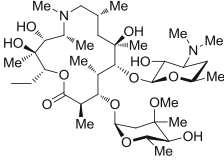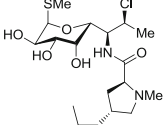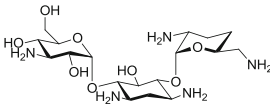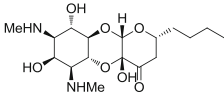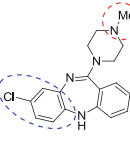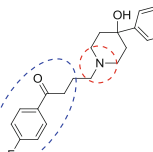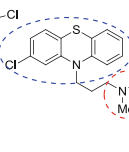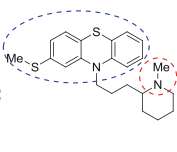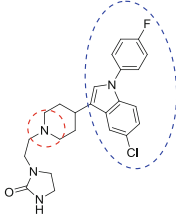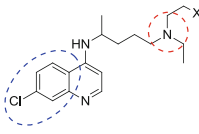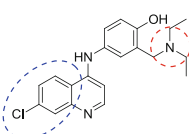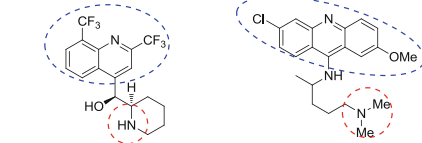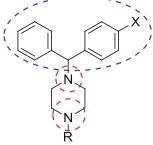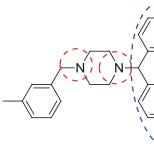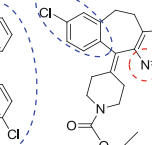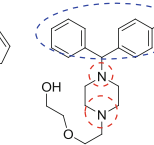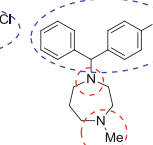
Fig. 1 Drug-induced phospholipidosis. DIPL in pancreas of rat as confirmed by transmission electron microscopy. Ultrastructural features consistent with PLD (lamellar bodies; arrows) were evident in pancreas with PLD (**b**) compared to control pancreas (**a**); *N* nucleus, *IC* islet cell, *AC* acinar cell, *NEG* neuroendocrine granules
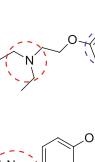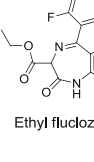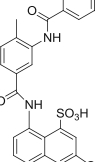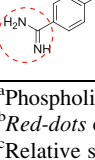
2 Mechanisms of DIPL
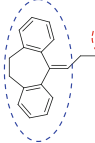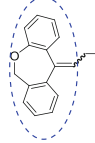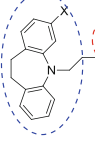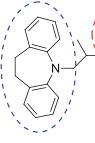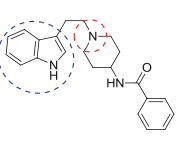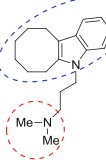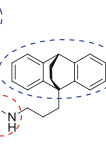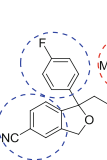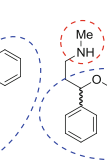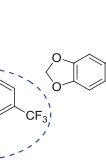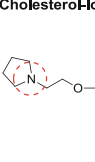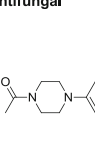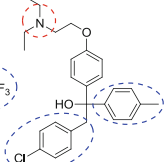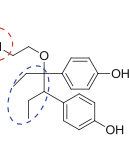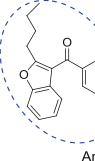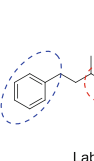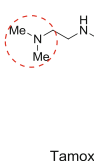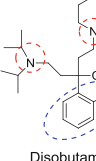
A variety of drugs spanning several therapeutic areas with diverse pharmacological activities are reported to induce PLD in animals and/or humans (Table 1). Although these drugs are structurally diverse, they generally share two common features: (1) a hydrophilic cationic amine side chain and (2) a hydrophobic core which is usually an aromatic and/or an aliphatic ring structure, hence termed cationic amphiphilic drugs (CAD) (Table 1). These features fulfill the two basic requirements for PLD induction by CADs: (1) entry into the cells and (2) enrichment in lysosomes [3]. In the most simplistic of terms, the hydrophobic features drive passage through biological membranes, while the cationic feature facilitates entrapment and accumulation in acidic lysosomes. There is no clear explanation on how this accumulation would lead to PLD, but is proposed that accumulation may alter the microenvironment for lysosomal catalytic enzymes, leading to decreased catabolism of phospholipids either by reduced breakdown of a drug-phospholipid complex by phospholipases or by direct inhibition of phospholipases, eventually leading to PLD. An exception to this rule is the class of antibiotics, exemplified by the aminoglycoside gentamicin, which is polycationic and highly hydrophilic and does not completely fit with the features of more typical PLD inducers. Extensive binding of gentamicin to membrane anionic phospholipids followed by internalization and sequestration in lysosomes triggers PLD [4]. A similar mechanism is proposed for PLD induced by the macrolide antibiotic azithromycin [5].

Table 1 Phospholipidosis-inducing drugs and their therapeutic class[a] with cationic and hydrophobic domains of selected PLD inducers[b]

Antibacterial

Gentamicin

Tobramycin

Amikacin

Netilmicin

Erythromycin

Azithromycin

Clindamycin

Dibekacin

Trospectomycin

Antipsychotics

Clozapine

Haloperidol

Chlorpromazine

Thioridazine

Sertindole

Antimalarials

X= H, Chloroquine
X= OH, Hydroxychloroquine

Amodiaquine

Mefloquine[c]

Quinacrine

Antihistamines

X= Cl, R= Me, Chlorcyclizine
X= Cl, R=H, Norchlorcyclizine
X= H, R = Me, Cyclizine

Meclizine

Loratadine

Homochlorcyclizine

Hydroxyzine

(continued)

Table 1 (continued)

Antidepressants

Amitriptyline   Doxepin

X= Cl, R= Me, Clomipramine
X= H, R= Me, Imipramine
X= H, R= H, Desipramine

Trimipramine   Indoramin

Iprindole   Maprotiline   Citalopram   Zimelidine   Fluoxetine   Paroxetine

Cholesterol-lowering

Boxidine   Triparanol   Coralgil

Antifungal

Ketoconazole

Analgesic

Clomipramine

Antianginal

Perhexiline

Anorectics

Chlorphentermine   Cloforex   Fenfluramine

Antiarrhythmics

Amiodarone   Disobutamide

Antihypertensive

Labetalol

Antiestrogen

Tamoxifen

Others

Tilorone   Ethyl fluclozepate

Pentamidine   Suramin

[a]Phospholipidosis in humans and/or animals
[b]*Red-dots* cationic domain, *Blue-dots* hydrophobic domain
[c]Relative stereochemistry
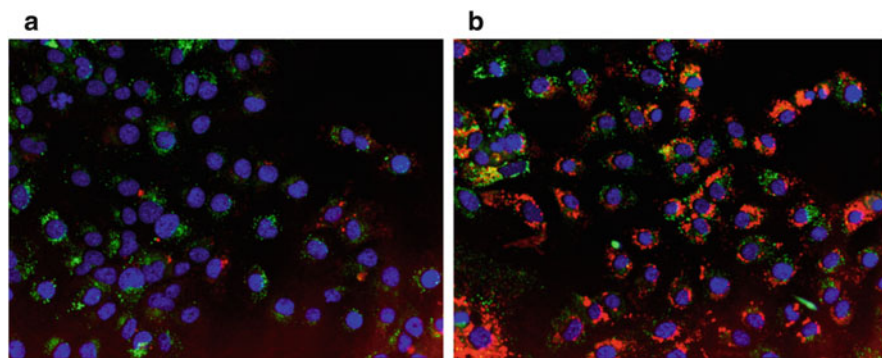
Fig. 2 Detection of PLD-inducing potential in vitro. PLD-inducing potential detected by differential staining of neutral lipids (*green*) and phospholipids (*red*) using LipidTox (from Invitrogen Corp); *blue* = nuclei. Huh7 cells were incubated with the lipid dyes as described by the kit in the absence (**a**) or presence (**b**) of a potent PLD inducer and later imaged using confocal microscopy. Increase in phospholipids, relative to neutral lipids, was evident with treatment with the PLD inducer (**b**)

3 General Physicochemical Properties of PLD Inducers

Several researchers have used various descriptors ranging from simple calculations of lipophilicity and basicity to complex electrotopological features and tools ranging from simple empirical observations to complex machine learning algorithms to understand the contribution of physicochemical attributes that predict the potential to induce PLD. All of these attempts have attested to cationic and amphiphilic properties as the hallmark features of PLD inducers. Although the empirical observation of these features of PLD inducers was made decades ago, no concerted efforts were made to use these attributes in predicting the potential for drugs to induce PLD until recently. Ploemen et al. first articulated the importance of the key physicochemical attributes of PLD inducers derived from simple software-driven calculations [6]. Since then researchers have made several modifications, extensions, and additions to the key attributes defined by Ploemen to improve upon the predictive capability. Further, advancements in data mining and analysis software and expansion of database of PLD inducers have facilitated an evaluation of additional molecular attributes in an effort to enhance the prediction capability. Additionally, several medium-throughput biochemical and cell-based methods have been developed to further enhance prediction based on experimental data. Taking a further step, the addition of experimental parameters of in vivo disposition such as metabolism and tissue distribution has further driven these prediction models much closer to predicting with high sensitivity the PLD-inducing potential in vivo. A few examples are of such models, and their prediction formulae are provided in Table 2, in the order of the complexity and the resource involved.

For a PLD inducer to exert its effect, it first needs to enter the cells. Entry into cells or passage across biological membranes is typically driven by the lipophilicity

Table 2 Formulae of selected prediction models[a]

Model (reference)	Prediction of PLD-inducing potential	
	Positive	Negative
In silico/Ploemen et al. [6]	$(pKa - MB)^2 + (clogP)^2 \geq 90$, provided $pKa \geq 8$ and $clogP \geq 1$	$(pKa - MB)^2 + (clogP)^2 < 90$, or $pKa < 8$, or $clogP < 1$
In silico/Modified Ploemen et al. [7]	$(pKa - MB)^2 + (clogP)^2 \geq 50$, provided $pKa \geq 6$ and $clogP \geq 2$	$(pKa - MB)^2 + (clogP)^2 < 50$, or $pKa < 6$, or $clogP < 2$
In silico/Tomizawa et al. [8]	Low if NC = 1 and $clogP < 1.61$ Medium if NC = 1 and $clogP \geq 1.61$ and <2.75 High if NC = 1 and $clogP \geq 2.75$ or NC > 1 and ≤ 2	NC < 1
In silico/Fischer et al. [9]	$pKa \geq 6.3$ but ≤ 11 and $\Delta\Delta G_{AM} \leq -6$ kJ/mol	$pKa < 6.3$ or > 11 and $\Delta\Delta G_{AM} > -6$ kJ/mol
In silico/Przyblak and Cronin (LDA-QSAR) [10]	$-4.58 + 0.963 \log P + 0.426$ pKa $+ 7.864$ (fifth order valence-corrected molecular connectivity) $+ 1.156$ (H-bond donors) $- 0.077$ (H-bond strength)	$-2.15 + 0.619 \log P + 0.226$ pKa $+ 18.077$ (fifth order valence-corrected molecular connectivity) $+ 0.436$ (H-bond donors) $+ 0.983$ (H-bond strength)
In vitro/Kasahara et al. [11]	Normalized NBD-PC/Hoechst33342 value ≥ 1.5	Normalized NBD-PC/Hoechst33342 value < 1.5
In vitro/Atienzar et al. [12]	PLD index (derivation of scores from normalized levels of gene expression) >1.5 at 50 μM	PLD index < 1.5
Hybrid/Hanumegowda et al. [13]	$(pKa\text{-most basic} \times clogP \times Vd) \geq 180$, provided $clogP \geq 2$	$(pKa\text{-most basic} \times clogP \times Vd) < 180$, or $clogP < 2$

[a]In order of increased complexity and resource involvement

of a compound. As noted in the previous section, the majority of PLD inducers have an aromatic and/or an aliphatic ring structure that provides hydrophobicity/lipophilicity to the compound. The lipophilicity of a compound can be calculated based on structural components by several computer software programs or measured experimentally by octanol–water partitioning. Several in silico models have clearly established lipophilicity as a criterion for a PLD inducer. A calculated logP (clogP) of at least 1 is required for a molecule to be a PLD inducer, with greater propensity for PLD induction with a higher clogP. As described earlier, an exception to this requirement is the class of antibiotics which are polycationic and less polar but enter cells by binding to membrane anionic phospholipids with subsequent internalization [3, 7, 14].

The primary cation of CADs is the basic amine which is unprotonated at higher physiological pH but becomes protonated upon diffusion into the lysosomal compartment, which is acidic and leads to its entrapment in this intracellular organelle. A temporary or focal increase of lysosomal pH as a result of the basic moieties of CADs is thought to lead to the malfunction of lysosomal enzymes. Trapped drug, presenting as a complex with phospholipids, makes bound phospholipids less susceptible to

degradation by phospholipases. Trapped drug itself can also bind to phospholipases and render them inactive. The eventual consequence of these events is the accumulation of undegraded phospholipids in the lysosomes. Similar to lipophilicity, several in silico models have clearly established basicity as a criterion for a PLD inducer. A calculated pK_a of at least 6 is required for a molecule to be a PLD inducer, with greater propensity for PLD induction associated with a higher pK_a [3, 7, 14].

4   Predicting PLD Based on In Silico Parameters

Despite a basic understanding of the physicochemical properties of PLD inducers, efforts to utilize these features to predict PLD-inducing potential were not made until recently. This was perhaps due to the lack of an extensive dataset of compounds that induce PLD and/or a lack of sharing of structures due to proprietary concerns. Academic labs have generally utilized what is available in the public domain while pharmaceutical companies have enriched the dataset with their own compounds. The PLD working group of the FDA has the largest dataset of PLD inducers originating from approved or failed INDs and NDAs. In this section we describe the models and criteria for predicting PLD inducers developed by several groups. High prediction accuracies have been reported by using complex model algorithms. Table 3 provides a representative list of in silico-based models/methods with the descriptors used for prediction, criteria, and statistics of model performance. It should, however, be noted that the predictive models used datasets that were not always similar, and, hence, the application of the reported models to different datasets may not give the same accuracies as originally reported.

One of the simplest models was described by Ploemen et al. [6] where they used only two descriptors, (1) clog$P$ and (2) pK_a, to represent the lipophilicity and basicity of compounds, respectively, to predict PLD-inducing potential. As both of these descriptors were calculated based on the structure by simple computer software programs, this method holds tremendous utility for screening compounds for PLD potential in the simplest way. Using a dataset of 41 compounds containing 29 PLD inducers, they were able to set criteria to correctly predict the majority of PLD inducers. Applying these criteria to a larger dataset of 201 compounds containing 85 PLD inducers, Pelletier et al. [7] found that the prediction performance of the Ploemen method was lower (sensitivity of 58%) and that it could be improved by modifying the criteria. With the modified criteria, the predictive sensitivity was improved to 79%. With the understanding that the subcellular target for PLD is lysosomes and the possible mechanism is entrapment of a basic drug in the lysosomes, Tomizawa et al. [8] modified the method of Ploemen to include net charge (NC) at the lysosomal pH instead of pK_a. Using clog$P$ and NC, a ratings criterion that rank ordered the PLD-inducing risk was developed. With this model, the predictive accuracy for a dataset of 63 compounds was 98%.

Fischer et al. [9] described the use of the free energy of amphiphilicity in predicting PLD. Using the measured free energy of amphiphilicity, an in-house

Table 3 In silico descriptor-based prediction models[a]

Method/reference	Descriptors/parameters	Prediction criteria	Dataset	Predictive statistics
ClogP-pKa based/Ploemen et al. [6]	clogP, pKa-most basic	Pos: $(pKa\text{-}MB)^2 + (clogP)^2 \geq 90$, provided $pKa \geq 8$ and $clogP \geq 1$; Neg: $pKa\text{-}MB)^2 + (clogP)^2 < 90$, or $pKa < 8$, or $clogP < 1$	41 (29 Pos, 12 Neg)	
ClogP-NC based/Tomizawa et al. [8]	clogP, NC	PLD risk: None if NC < 1; Low if NC = 1 and $clogP < 1.61$; Medium if NC = 1 and $clogP \geq 1.61$ and <2.75; High if NC = 1 and $clogP \geq 2.75$ or NC > 1 and ≤ 2	63	Prediction accuracy: 98%
Bayesian/Pelletier [7]	clogP, pKa, amphiphilic moment, basic and acidic centers, Scitegic proprietary structural fingerprint FCFP_4	Complex	201 (85 Pos, 116 Neg)	Specificity: 77%, Sensitivity: 92%, Concordance: 83%
Modified Ploemen/Pelletier et al. [7]	clogP, most basic pKa	Pos: $[(pKa - MB)^2 + (clogP)^2] \geq 50$, provided $pKa \geq 6$ and $clogP \geq 2$; Neg: $[pKa - MB)^2 + (clogP)^2] < 50$, or $pKa < 6$, or $clogP < 2$	201 (85 Pos, 116 Neg)	Specificity: 80%, Sensitivity: 79%, Concordance: 80%
MDL-QSAR/Kruhlak [15]	Molecular topological descriptors	Complex algorithms	583 (190 Pos, 393 Neg)	Specificity: 80%, Sensitivity: 76%, Concordance: 79%
MC4PC QSAR/Kruhlak [15]	Molecular fragment based	Complex rules	583 (190 Pos, 393 Neg)	Specificity: 92%, Sensitivity: 50%, Concordance: 78%
Combined-QSAR/Kruhlak et al. [15] 2008	MDL and MC4PC combined	Combined output	583 (190 Pos, 393 Neg)	Specificity: 79%, Sensitivity: 81%, Concordance: 79%

Method	Descriptor	Algorithm	Number of compounds	Results
Weka machine learning/ Ivanciuc [16]	Complex	Complex algorithms		Prediction accuracy: 97%
Machine learning/Lowe et al. [17]	Complex; E-Dragon, circular fingerprints	Complex algorithms	185 (102 Pos, 83 Neg)	Prediction accuracy: 82%
SMARTS/LDA-QSAR/ Przyblak and Cronin [10]	Complex	SMARTS patterns: Complex LDA-QSAR: Pos: $-4.58 + 0.963 \log P + 0.426\, pKa + 7.864$ (fifth order valence-corrected molecular connectivity) $+ 1.156$ (H-bond donors) $- 0.077$ (H-bond strength) Neg: $-2.15 + 0.619 \log P + 0.226\, pKa + 18.077$ (fifth order valence-corrected molecular connectivity) $+ 0.436$ (H-bond donors) $+ 0.983$ (H-bond strength)	450 (93 pos, 357 neg)	SMARTS Specificity: 88% Sensitivity: 85% Concordance: 87.5% LDA-QSAR Specificity: 71% Sensitivity: 87% Concordance: 74%
CAFCA-pKa/Fischer et al. [9]	Free energy of amphiphilicity ($\Delta\Delta G_{AM}$) and pKa	Pos: $pKa \geq 6.3$ but ≤ 11 and $\Delta\Delta G_{AM} \leq -6$ kJ/mol Neg: $pKa < 6.3$ or > 11 and $\Delta\Delta G_{AM} > -6$ kJ/mol	32	Concordance: 91%
QSAR/Choi et al. [18]	Complex; ADMET predictor platform Complex; symmetry platform	Artificial neural networks algorithm Combined probabilistic and similarity algorithm	379 (147 pos, 232 neg)	ADMET Predictor Specificity: 90% Sensitivity: 100% Concordance: 92% Symmetry platform Specificity: 92% Sensitivity: 83% Concordance: 88%

[a]In order of their publication year; *pos* positive, *neg* negative

proprietary program CAFCA (calculated free energy of amphiphilicity of small charged amphiphiles) was developed. Using this program, a correlation between PLD inducers in vitro and the prediction criteria for PLD inducers was established. This method predicted PLD inducers with a concordance of 91% for a set of 32 compounds.

Improvements in data mining and analysis tools have facilitated incorporation of several molecular attributes in PLD prediction models. Pelletier et al. [7] utilized several such molecular descriptors and statistical tools to describe a Bayesian model for predicting PLD inducers. Using a dataset of 201 compounds containing 85 PLD inducers, they were able to predict PLD inducers with 92% sensitivity and differentiate from non-inducers with 83% concordance. Ivanciuc [16] used Weka machine learning program to analyze a dataset with complex algorithms that was able to predict PLD inducers with 97% accuracy. Similarly, Lowe et al. [17] utilized machine learning tools and criteria similar to Ivanciuc to predict PLD inducers with an accuracy of 82% in a dataset of 185 compounds containing 101 known PLD inducers.

Kruhlak et al. [15] utilized a large dataset of 583 compounds containing 190 PLD inducers to develop prediction models. Using the commercial software MC4PC, they were able to identify 77 structural alerts, 23 of which were considered reliable indicators of PLD-inducing activity in conjunction with other physicochemical parameters. Structural alerts consisted of aromatic rings, amine fragments, and structures containing halogen-substituted aromatic rings, which overall fit into the defining features of CADs – a basic amine containing a hydrophilic side chain appended to a hydrophobic core. Using another commercially available software, MDL-QSAR, they identified 18 descriptors of high significance in distinguishing PLD inducers from non-inducers. The concordance of models developed using these programs was similar; however, the sensitivity of the model based on MC4PC was lower (50%) compared to MDL-QSAR (76%). Combining the output of both models, they were able to improve the sensitivity to 81% without loss of specificity. Similarly, Przybylak and Cronin [10] and Choi et al. [18] used various molecular descriptors and complex algorithms and software platforms to develop methods with prediction concordance up to 92% using large datasets of compounds.

5   Predicting PLD Based on Experimental Data

In parallel to in silico approaches as described above, several in vitro models ranging from biophysical interactions to cell-based imaging systems to hybrid methods were developed to predict PLD-inducing potential. Table 4 provides a representative list of in vitro and hybrid models/methods with criteria for prediction and statistics of model performance. It should be recognized that these methods also have their limitations as those of in silico methods for the simple reason that not all the PLD inducers in vitro are inducers in vivo, as in vivo disposition ranging from

Table 4 In vitro and hybrid-based prediction models[a]

Method/reference	Models/parameters	Prediction criteria	Dataset	Predictive statistics
Gene expression/ Sawada et al. [19]	HepG2; 17 genes (12 identified as relevant)	Increase or decrease of specific genes in a set of 12	30	NA
Gene expression/ Atienzar et al. [12]	HepG2; 11 genes (upregulated: MGC4171, NR0B2, INHBE, P8, SERPINA3, ASNS, C10, FLJ10055, FRCP1; downregulated: APIS1, TAGLN)	PLD index = Derivation of scores from normalized levels of gene expression. Pos: PLD Index > 1.5 at 50 μM, provided there is a dose–response relationship		
High-throughput screening/ Kasahara et al. [11]	Rat hepatocytes and cell lines: HepG2, ARLJ301-3, CHO-K1, CHL/IU, J744A	Fluorescent phospholipid (NBD-PC) and nuclear stain (Hoechst33342) uptake; correlation with confocal/electron microscopy. Pos: Normalized NBD-PC/Hoechst33342 value $\geq$ 1.5	24	Correlation coefficient of 0.812
Lipid staining/Nioi et al. [20]	HepG2; fluorescently labeled lipid staining using LipidTox kit from Invitrogen	Intensity of lipid staining as measured by fluorescent plate reader/confocal microscopy	24	Concordance: 100%
Biophysical/Vitovic [21]	Interaction with acidic phospholipid and impact of CMC	Effect on surface tension and CMC	53	NA
Vd based hybrid/ Hanumegowda et al. [13]	clog$P$, pKa-most basic, Vd	Pos: (pKa-most basic x clog$P$ x Vd) $\geq$180, provided clog$P$ $\geq$2. Neg: (pKa-most basic x clog$P$ x Vd) < 180, or clog$P$ <2	101 (51 Pos, 50 Neg)	Specificity: 94%; Sensitivity: 82%; Concordance: 88%
Biophysical, in silico/ Kuroda and Saito [22]	K_d from phospholipid containing vesicles, PLA2 inhibitory ratio, metabolic stability, pKa, clog$P$	Multivariate analysis	8 (5 pos, 3 neg)	NA
High content screening/van de Water et al. [23]	CHO-K1, HepG2 cell lines	Fluorescent phospholipid (NBD-PE) with counterstains for nucleus and cytoplasm; image analysis	56 (25 pos, 31 neg)	Sensitivity: 92% (CHO-K1), 88% (HepG2); Specificity: 87% (CHO-K1), 81% (HepG2)

[a]In order of their publication year

metabolism to pronounced dose-limiting tolerability and/or toxicity may limit the development or occurrence of PLD.

Vitovic et al. [21] developed a simple assay based on the interaction of a test compound with an acidic phospholipid and its influence on critical micelle concentration to determine the PLD-inducing potential. Kuroda and Saito [22] further added measured biophysical interaction, inhibition of phospholipase A_2 (PLA_2), and metabolic stability to in silico parameters to develop a multivariate analysis to predict PLD-inducing potential.

Numerous cell-based models are available to predict PLD-inducing potential. All of these methods are variants of staining with dyes, which are either general for lipids (such as nile red) or specific for the accumulation of fluorescent-conjugated phospholipids (such as 7-nitrobenz-2-oxa-1,3-diazol-4-yl-phosphatidylcholine (NBD-phosphatidylcholine, PC) or 7-nitrobenz-2-oxa-1,3-diazol-4-yl- phosphatidylethanolamine (NBD-phosphatidylethanolamine, PE)). These models have been developed in several cell lines and analyzed using various imaging platforms specifically for intensity of staining [11, 20, 23]. An example of detecting phospholipid accumulation in cells in vitro using a commercial kit is provided in Fig. 2. Sawada et al. tried a different approach using gene expression profiling in HepG2 cells to look for sensitive markers of PLD and identified 12 specific genes that were predictive of PLD induction based on a dataset of 30 compounds [19].

Recognizing that PLD in vivo occurs not only as a function of the inherent physicochemical properties of a compound, but also from the likely residence of a compound in tissues, Hanumegowda et al. [13] combined simple in silico parameters with the volume of distribution. In a dataset of 101 compounds containing 51 PLD inducers, they were able to differentiate PLD inducers from non-inducers with a concordance of 88%. This model highlights the importance of including a parameter related to the in vivo disposition of a compound to address the relevance to PLD in vivo, which is the eventual outcome of most concern.

6   Detecting and Monitoring PLD

As one might anticipate, methods for the detection of PLD have been known for several decades since the time of the first report using light microscopy [14, 24] although it took several additional years before these vacuolations were demonstrated to be characteristically composed of phospholipids. Since then, various techniques ranging from simple microscopy and lipid staining to electron microscopy and biomarkers in biofluids have been established to detect and/or monitor DIPL. While detection of DIPL is quite simply accomplished by evaluating collected tissues from animals or humans and visualization by light microscopy for vacuolation, staining for lipids using stains such as sudan black, nile red, and others, and immunohistochemical staining of markers such as lysosome-associated membrane protein-2 (LAMP-2) and adipophilin, a protein in the membranes of lipid droplets [25], confirmation of DIPL by electron microscopy for characteristic
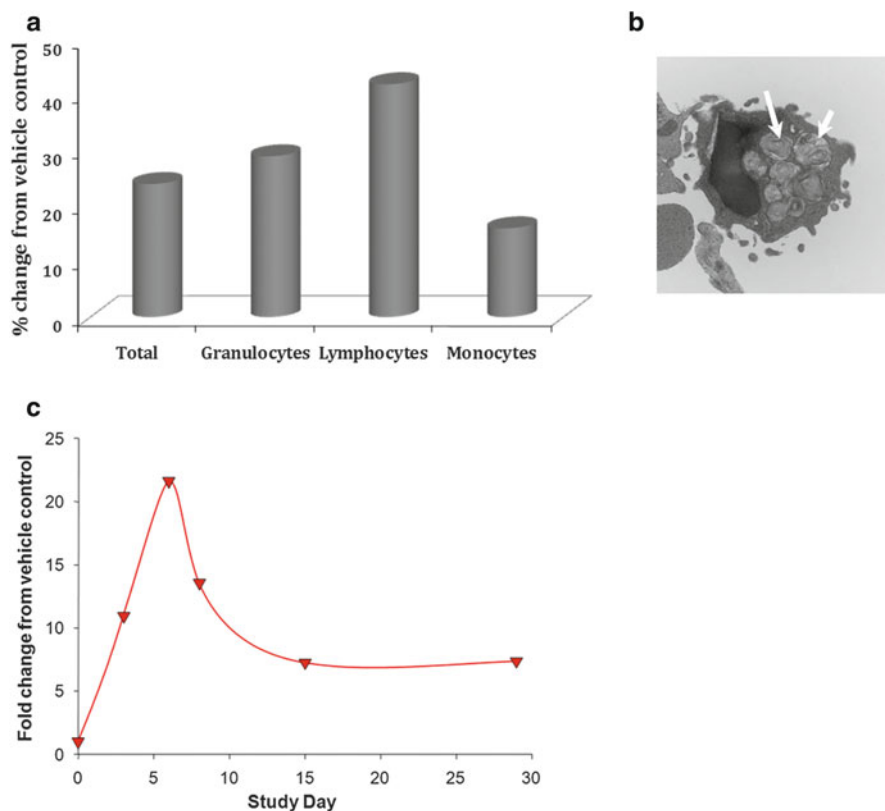
Fig. 3 Monitoring DIPL in nonterminal animal studies. DIPL monitored in peripheral blood leukocytes and urine of rat treated with a PLD inducer. Peripheral blood leukocytes (drawn of day 4 of the study) were stained by nile red staining and sorting cells by flow cytometry (**a**). Leukocytes, in general, and lymphocytes in particular had the highest intensity of nile red staining in rats treated with a PLD inducer. Lipid staining by nile red was confirmed to be PLD (lamellar bodies; *arrows*) by transmission electron microscopy of lymphocyte (**b**). Urine BMP (22:6) was increased in rats treated with PLD inducer especially during the early stages of the study (**c**)

concentric lamellar inclusion bodies is still the gold standard. Numerous other ways are being investigated with the intention of developing noninvasive or less invasive methods to detect and, moreover, monitor the progression of PLD. Further, correlation of animal to human DIPL is yet to be clearly established.

Since PLD is a process that develops over time, sometimes ranging from weeks to months, without apparent clinical manifestation, the discovery of PLD in tissues at the completion of a chronic animal toxicity study often ends up as a wasted investment for pharmaceutical companies. Hence, there is a greater emphasis to develop methods to monitor for the development of PLD in easily accessible tissues or biofluids at an early stage with less invasive methods. Lungs and lymphocytes (apart from liver) are the most common targets for DIPL. Naturally, several

methods to detect and monitor PLD in peripheral blood cells, especially lymphocytes, have gained considerable attention. Vacuolation by regular microscopy, staining by lipid stains such as nile red, and, ultimately, electron microscopy are ways of detecting and confirming DIPL. An example of monitoring of DIPL in a nonterminal study in rat peripheral leukocytes using nile red staining and confirmation by electron microscopy is provided in Fig. 3a, b. In human patients, amiodarone-induced PLD has been demonstrated by electron microscopy in leukocytes from blood [26].

Metabonomics has offered another sensitive methodology to monitor for DIPL. Two endogenous metabolites, phenylacetylglycine (PAG) and lyso-bis-phosphatidic acid (LBPA), also known as bis(monoacylglycero)phosphate (BMP), have been identified as promising biomarkers of DIPL. Urinary PAG in rodents is well correlated with DIPL; however, its utility is limited in humans since it is not produced. Nevertheless, from a drug development and hazard identification perspective, measuring urinary PAG in rodents offers a significant advantage in terms of monitoring for DIPL. LBPA is linked to lysosomal phospholipids and, therefore, is more appropriate as a biomarker for lysosomal storage disorders such as PLD. LBPA has also been shown to increase in several tissues in rats with PLD, including serum, which will be easier to monitor in an ongoing toxicity study [27]. An example of monitoring of DIPL in a non-terminal study in rat urine using BMP as a marker is provided in Fig. 3c. Similarly, in human patients, amiodarone-induced PLD has been demonstrated by an increase in levels of BMP and phosphatidylglycerol in alveolar macrophages collected by bronchoalveolar lavage [28].

7 Chemical Strategies to Mitigate PLD-Inducing Potential

As discussed in the previous sections, the induction of PLD is mainly related to the structure of drug candidates and is independent of its targeted pharmacological activity. In most cases, DIPL is caused by CADs [29]. CADs are molecules that possess both a hydrophobic element containing an aromatic group and a hydrophilic moiety with at least one basic site. However, it is not well understood why some CADs induce PLD while others do not. Several medicinal chemistry strategies focusing on reducing the cationic amphiphilic character of drug candidates have been successfully applied to minimize the risk of DIPL [29, 30]. Representative examples of these strategies are highlighted in this section.

At Roche, a CAFCA program was used in two different projects to closely monitor the amphiphilicity of lead chemotypes [31]. Molecules with a free energy of amphiphilicity ($\Delta\Delta G_{AM}$) greater than −6 kJ/mol were prepared to lower the potential for PLD (Fig. 4). One molecule, the aminomethylpyrimidine **1**, was identified as a lead for a dipeptidyl peptidase-IV (DPP-IV) inhibitor project, which had an IC_{50} of 10 nM, good membrane permeability, solubility, and metabolic stability. However, **1** induced PLD in a concentration-dependent manner in cultured fibroblasts (2.5–20 μM). It also had a $\Delta\Delta G_{AM}$ of −6.6 kJ/mol. The Roche
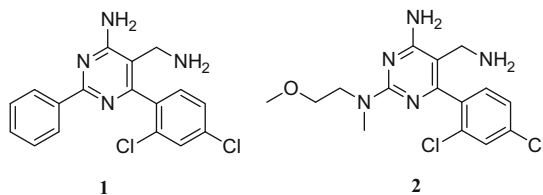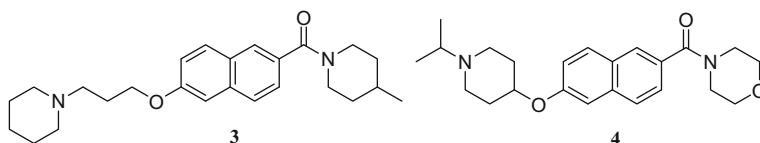
Fig. 4 Roche DPP-IV inhibitors

Fig. 5 Roche H$_3$ antagonists

team then focused on designing less amphiphilic analogues and found compound **2** with a $\Delta\Delta G_{AM} = -5.6$ kJ/mol having the same DPP-IV-inhibiting potency as **1**. Its free energy of amphiphilicity correlated with a lack of PLD induction in cultured fibroblasts at the highest concentration tested, 20 μM [32].

In a series of histamine 3 receptor (H$_3$R)-inverse agonists containing a naphthalene core (Fig. 5), **3** was identified as a lead with good potency and other favorable properties, including the absence of mutagenicity (genotoxicity and clastogenicity) or phototoxicity in vitro. Compound **3** induced PLD in cultured fibroblasts at concentrations of 2.5–20 μM with the extent dependent on concentration. The $\Delta\Delta G_{AM}$ for **3** was -10.4 kJ/mol. Modifications were introduced to the amine and amide side chains in order to generate the less amphiphilic compound **4**, which had a $\Delta\Delta G_{AM} = -4.27$ kJ/mol. Indeed, this compound showed no induction of PLD in cultured fibroblasts. Another approach to reduce amphiphilicity involved substitution of the naphthalene core with a less lipophilic quinoline. Those derivatives having the same substitution pattern at both amine and amide side chains showed reduced $\Delta\Delta G_{AM}$. Unfortunately, the potency at the H$_3$R was also reduced [33].

At Abbot, the H$_3$ receptor antagonist benzofuran **5** was shown to have excellent binding potency for both the human receptor, K$_i$ = 0.05 nM, and the rat variant, K$_i$ = 0.11 nM, (Fig. 6). However, its dibasic nature was largely responsible for the induction of PLD in vitro. It is well known that basic diamines containing an aromatic group have the propensity to bind to negatively charged phospholipid bilayers in membranes, causing inhibition of lysosomal phospholipases which leads to PLD. The Abbot group decided to reduce the basicity of one of the amines in **5** as a means of decreasing the overall cationic character. Since the pyrrolidine amine is critical for H$_3$ binding, the modifications were directed toward the primary amine. Either substitution with heterocycles, as indicated in compound **6**, or removal of the benzylamine with the heterocycle now moved closer to the benzofuran (compound **7**) proved to be successful approaches. Both **6** and **7** showed much lower potential for PLD induction in the cultured rat hepatocytes assay when compared to **5** [34].
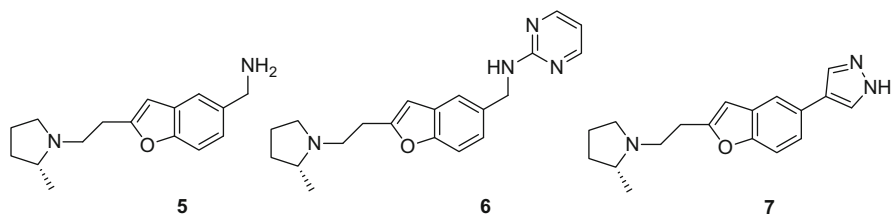
Fig. 6 Abbot H$_3$ antagonists
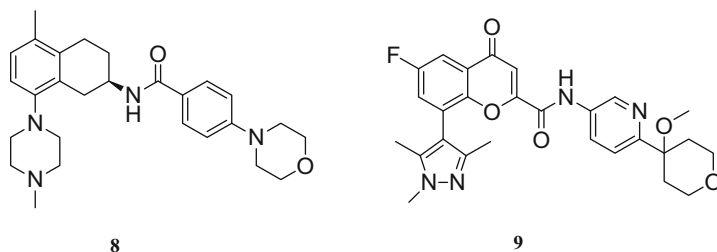
Fig. 7 AstraZeneca 5-HT$_{1B}$ antagonists
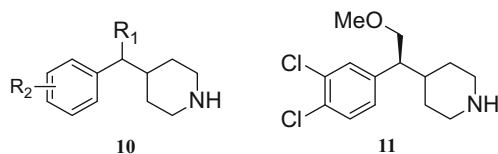
Fig. 8 Takeda triple reuptake inhibitors

Researchers at AstraZeneca disclosed the 5-hydroxytryptamine receptor 1B (5-HT$_{1B}$) antagonist **8** (Fig. 7) which had to be withdrawn from development due to toxicities linked to accumulation of the compound in vivo as well as PLD. Nevertheless, replacement of the basic piperazine with a pyrazole as well as the tetrahydronaphthalene core with a more polar chromene led to the more potent compound **9**. In an in vitro PLD assay where **8** showed 81% inhibition at 300 μM, **9** was inactive at the same concentration [30, 35].

A group of researchers at Takeda discovered a triple reuptake inhibitor (TRI) of the serotonin, norepinephrine, and dopamine transporters based on the piperidine scaffold **10** (Fig. 8). Analysis of existing monoamine reuptake inhibitors led the team to set up structural restrictions in order to maximize potency while minimizing liabilities, including PLD. Therefore, the MW limit was set at less than 300; the number of aromatic rings was restricted to one and the clog$P$ value at below 3.5. The lipophilicity could be tuned by appropriate selection of the side chain R1. Substituents R$_2$ on the aromatic ring had an important role in the potency. As a result of optimization and optical resolution, **11** was discovered as a TRI with low potential for PLD induction, as indicated by its PLD score (2.7) with NBD-PE and HepG2 cells [36].
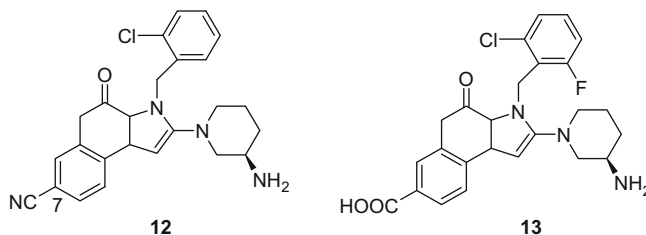
Fig. 9 Dainippon Sumitomo DPP-IV inhibitors

At Dainippon Sumitomo, researchers disclosed a series of potent DPP-IV inhibitors with a 3*H*-imidazo[4,5-*c*]quinolin-4(5*H*)-one skeleton (Fig. 9). Substitution at position 7 led to compounds such as **12**, which showed half-maximal DPP-IV inhibitory activity at a concentration of 6.3 nM. However, **12** also had a phospholipid index (PI) of 0.37 at 30 μM using isolated human hepatocytes and propranolol (PI = 1) as the reference agent. The introduction of a carboxylic acid at the same position led to **13**, a more potent DPP-IV inhibitor (IC_{50} = 0.48 nM) with a PI = 0.008 at 30 μM [37], and therefore low potential for PLD induction.

DIPL, as we understand and have discussed in this chapter, is a phenomenon that is generally adaptive in response to drug overload in a particular tissue, although in several instances it is incidentally related to concomitant injury in the same tissue where it occurs or in another organ system. Given the primary goal of developing safe drugs for humans and the conservative overall nature of drug development that targets conditions that require chronic treatment, it is a general understanding within the pharmaceutical industry and the regulatory agencies that the development of PLD-inducing drugs should be avoided. As is clearly understood, PLD-inducing potential is independent of the primary pharmacological activity of a molecule, although they tend to group into specific therapeutic classes (e.g., antipsychotic and antimalarial agents) because of the structural similarity of the drugs in those classes. For decades, CADs have been known to possess PLD-inducing potential and a concerted effort has been made to detect and predict the potential in an attempt to understand and develop medicinal chemistry strategies to mitigate PLD-inducing potential in promising chemical series. As demonstrated by the examples presented and discussed in this chapter, there are several approaches, both simple and complex, to predict, detect, or chemically mitigate PLD-inducing potential, although there is clearly no single universal solution. Every instance has to be treated on a case-by-case basis, keeping in mind the therapeutic area, duration of treatment, and unmet medical need, with a combination of multiple strategies in place as tiers to guide SAR and weed out the issue as the drug discovery progresses.

References

1. Willard J (2008) FDA Phospholipidosis Working Group, Preliminary results and developing opinions [abstract]. Society of Toxicological Pathology, 27th annual symposium, San Francisco, CA
2. Sun H, Xia M, Shahane SA, Jadhav A, Austin CP, Huang R (2013) Are hERG channel blockers also phospholipidosis inducers? Bioorg Med Chem Lett 23:4587–4590
3. Kodavanti UP, Mehendale HM (1990) Cationic amphiphilic drugs and phospholipid storage disorder. Pharmacol Rev 42:327–354
4. Laurent G, Kishore BK, Tulkens PM (1990) Aminoglycoside-induced renal phospholipidosis and nephrotoxicity. Biochem Pharmacol 40:2383–2392
5. Van Bambeke F, Montenez JP, Piret J, Tulkens PM, Courtoy PJ, Mingeot-Leclercq MP (1996) Interaction of the macrolide azithromycin with phospholipids. I. Inhibition of lysosomal phospholipase A1 activity. Eur J Pharmacol 314:203–214
6. Ploemen JP, Kelder J, Hafmans T, van de Sandt H, van Burgsteden J, Saleminki PJ, van Esch E (2004) Use of physicochemical calculation of pKa and CLogP to predict phospholipidosis-inducing potential: a case study with structurally related piperazines. Exp Toxicol Pathol 55:347–355
7. Pelletier DJ, Gehlhaar D, Tilloy-Ellul A, Johnson TO, Greene N (2007) Evaluation of a published in silico model and construction of a novel Bayesian model for predicting phospholipidosis inducing potential. J Chem Inf Model 47:1196–1205
8. Tomizawa K, Sugano K, Yamada H, Horii I (2006) Physicochemical and cell-based approach for early screening of phospholipidosis-inducing potential. J Toxicol Sci 31:315–324
9. Fischer H, Atzpodien EA, Csato M, Doessegger L, Lenz B, Schmitt G, Singer T (2012) In silico assay for assessing phospholipidosis potential of small drug like molecules: training, validation, and refinement using several data sets. J Med Chem 55:126–139
10. Przybylak KR, Cronin MTD (2011) In silico studies of the relationship between chemical structure and drug induced phospholipidosis. Mol Inform 30:415–429
11. Kasahara T, Tomita K, Murano H, Harada T, Tsubakimoto K, Ogihara T, Ohnishi S, Kakinuma C (2006) Establishment of an in vitro high-throughput screening assay for detecting phospholipidosis-inducing potential. Toxicol Sci 90:133–141
12. Atienzar F, Gerets H, Dufrane S, Tilmant K, Cornet M, Dhalluin S, Ruty B, Rose G, Canning M (2007) Determination of phospholipidosis potential based on gene expression analysis in HepG2 cells. Toxicol Sci 96:101–114
13. Hanumegowda U, Wenke G, Regueiro-Ren A, Yordanova R, Corradi JP, Adams SP (2010) Phospholipidosis as a function of basicity, lipophilicity, and volume of distribution of compounds. Chem Res Toxicol 23:749–755
14. Reasor MJ, Hastings KL, Ulrich RG (2006) Drug-induced phospholipidosis: issues and future directions. Expert Opin Drug Saf 5:567–583
15. Kruhlak NL, Choi SS, Contrera JF, Weaver JL, Willard JM, Hastings KL, Sancilio LF (2008) Development of a phospholipidosis database and predictive quantitative structure-activity relationship (QSAR) models. Toxicol Mech Methods 18:217–227
16. Ivanciuc O (2008) Weka machine learning for predicting the phospholipidosis inducing potential. Curr Top Med Chem 8:1691–1709
17. Lowe R, Glen RC, Mitchell JB (2010) Predicting phospholipidosis using machine learning. Mol Pharm 7:1708–1714
18. Choi SS, Kim JS, Valerio LG Jr, Sadrieh N (2013) In silico modeling to predict drug-induced phospholipidosis. Toxicol Appl Pharmacol 269:195–204
19. Sawada H, Takami K, Asahi SA (2005) Toxicogenomic approach to drug-induced phospholipidosis: analysis of its induction mechanism and establishment of a novel in vitro screening system. Toxicol Sci 83:282–292
20. Nioi P, Perry BK, Wang EJGYZ, Snyder RD (2007) In vitro detection of drug-induced phospholipidosis using gene expression and fluorescent phospholipid based methodologies. Toxicol Sci 99:162–173

21. Vitovic P, Alakoskela JM, Kinnunen PK (2008) Assessment of drug-lipid complex formation by a high-throughput Langmuir balance and correlation to phospholipidosis. J Med Chem 51:1842–1848
22. Kuroda Y, Saito M (2010) Prediction of phospholipidosis-inducing potential of drugs by in vitro biochemical and physicochemical assays followed by multivariate analysis. Toxicol In Vitro 24:661–668
23. van de Water FM, Havinga J, Ravesloot WT, Horbach GJ, Schoonen WG (2011) High content screening analysis of phospholipidosis: validation of a 96-well assay with CHO-K1 and HepG2 cells for the prediction of in vivo based phospholipidosis. Toxicol In Vitro 25:1870–1882
24. Nelson AA, Fitzhugh OG (1948) Chloroquine (SN-7618) pathologic changes observed in rats which for 2 years had been fed various proportions. Arch Pathol (Chic) 45:454–462
25. Obert LA, Sobocinski GP, Bobrowski WF, Metz AL, Rolsma MD, Altrogge DM, Dunstan RW (2007) An immunohistochemical approach to differentiate hepatic lipidosis from hepatic phospholipidosis in rats. Toxicol Pathol 35:728–734
26. Dake MD, Madison JM, Montgomery CK, Shellito JE, Hinchcliffe WA, Winkler ML, Bainton DF (1985) Electron microscopic demonstration of lysosomal inclusion bodies in lung, liver, lymph nodes, and blood leukocytes of patients with amiodarone pulmonary phospholipidosis. Am J Med 78:506–512
27. Monteith DK, Morgan RE, Halstead B (2006) In vitro assays and biomarkers for drug-induced phospholipidosis. Expert Opin Drug Metab Toxicol 2:687–696
28. Martin WJ III, Standing JE (1988) Amiodarone pulmonary toxicity: biochemical evidence for a cellular phospholipidosis in the bronchoalveolar lavage of human subjects. J Pharmacol Exp Ther 244:774–779
29. Ratcliffe AJ (2009) Medicinal chemistry strategies to minimize phospholipidosis. Curr Med Chem 16:2816–2823
30. Bernstein PR, Ciaccio P, Morelli J (2011) Drug-induced phospholipidosis. Ann Rep Med Chem 46:419–430
31. Fischer H, Kansy M, Bur D (2000) CAFCA: a novel tool for the calculation of amphiphilic properties of charged drug molecules. Chimia 54:640–645
32. Peters JU, Hunziker D, Fischer H, Kansy M, Weber S, Kritter S, Müller A, Wallier A, Ricklin F, Boehringer M, Poli SM, Csato M, Loeffler BM (2004) An aminomethylpyrimidine DPP-IV inhibitor with improved properties. Bioorg Med Chem Lett 14:3575–3578
33. Rodriguez Sarmiento RM, Nettekoven MH, Taylor S, Plancher JM, Richter H, Roche O (2009) Selective naphthalene H_3 receptor inverse agonists with reduced potential to induce phospholipidosis and their quinoline analogs. Bioorg Med Chem Lett 19:4495–4500
34. Sun M, Zhao C, Gfesser GA, Thiffault C, Miller TR, Marsh K, Wetter J, Curtis M, Faghih R, Esbenshade TA, Hancock AA, Cowart M (2005) Synthesis and SAR of 5-amino- and 5-(Aminomethyl)benzofuran histamine H_3 receptor antagonist with improved potency. J Med Chem 48:6482–6490
35. Nugiel DA, Krumrine JR, Hill DC, Damewood JR Jr, Bernstein PR, Sobotka-Briner DD, Liu JW, Zacco A, Pierson ME (2010) De novo design of a picomolar nonbasic 5-HT$_{1B}$ receptor antagonist. J Med Chem 53:1876–1880
36. Ishichi Y, Kimura E, Honda E, Yoshikawa M, Nakahata T, Terao Y, Suzuki A, Kawai T, Arakawa Y, Ohta H, Kanzaki N, Nakagawa H, Terauchi J (2013) Novel triple reuptake inhibitors with low risk of CAD associated liabilities: design, synthesis and biological activities of 4-[(1S)-1-(3,4-dichlorophenyl)-2-methoxyethyl]piperidine and related compounds. Bioorg Med Chem Lett 21:4600–4613
37. Ikuma Y, Hochigai H, Kimura H, Nunami N, Kobayashi T, Uchiyama K, Furuta Y, Sakai M, Horiguchi M, Masui Y, Okazaki K, Sato Y, Nakahira H (2012) Discovery of 3H-imidazo [4,5-c]quinolin-4(5H)-ones as potent and selective dipeptidyl peptidase IV (DPP-4) inhibitors. Bioorg Med Chem Lett 21:5864–5883

The Influence of Bioisosteres in Drug Design: Tactical Applications to Address Developability Problems

Nicholas A. Meanwell

Abstract The application of bioisosteres in drug discovery is a well-established design concept that has demonstrated utility as an approach to solving a range of problems that affect candidate optimization, progression, and durability. In this chapter, the application of isosteric substitution is explored in a fashion that focuses on the development of practical solutions to problems that are encountered in typical optimization campaigns. The role of bioisosteres to affect intrinsic potency and selectivity, influence conformation, solve problems associated with drug developability, including P-glycoprotein recognition, modulating basicity, solubility, and lipophilicity, and to address issues associated with metabolism and toxicity is used as the underlying theme to capture a spectrum of creative applications of structural emulation in the design of drug candidates.

Keywords Isostere, Bioisostere, Conformation, Carboxylic acid isosteres, Guanidine and Amidine isosteres, Phosphate isosteres, Drug-water isostere, Heterocyle isostere

Contents

N.A. Meanwell (✉)
Department of Discovery Chemistry, Bristol-Myers Squibb Research and Development,
5 Research Parkway, Wallingford, CT 06492, USA
e-mail: Nicholas.Meanwell@bms.com

Abbreviations

BACE	β-site amyloid precursor protein cleaving enzyme
CYP 450	Cytochrome P450
DFT	Density functional theory
FAAH	Fatty acid amide hydrolase
HCV	Hepatitis C virus
hERG	Human *ether a go-go*-related gene
hH-PGDS	Human prostaglandin D2 synthase
HIV-1	Human immunodeficiency virus-1
Hsp	Heat shock protein
KSP	Kinesin spindle protein
LE	Ligand efficiency
LELP	Lipophilicity-corrected ligand efficiency
LLE	Ligand-lipophilicity efficiency
NMR	Nuclear magnetic resonance
PAMPA	Parallel artificial membrane
PARP	Poly(ADP-ribose) polymerase
PGI₂	Prostacyclin
P-gp	P-glycoprotein
PPAR	Peroxisome proliferator-activated receptor
SAR	Structure–activity relationship

1   Introduction

Bioisosterism is a powerful concept that has found widespread application in drug design and continues to be an important tactical element in contemporary medicinal chemistry practices [1–10]. Bioisosterism, which evolved from the concept of shape

isosterism exhibited by simple structural elements, was originally recognized by Langmuir almost a century ago, and the first experiments that demonstrated the potential of bioisosterism in a practical setting were reported in 1932 and 1933 [1, 11–13]. In a series of experiments, Erlenmeyer and his colleagues showed that antibodies recognizing *ortho*-substituted tyrosine moieties in synthetic antigens, derived by reaction with diazonium ions, were not able to distinguish between a phenyl and thiophene ring or O, NH, and CH_2 in a linker element [12, 13]. The description of this phenomenon as bioisosterism is attributed to Harris Friedman who introduced the term in 1951 to define compounds demonstrating similar biological activities [14]. However, Friedman recognized that bioisosterism and isosterism were distinct concepts, an observation that anticipated contemporary drug design principles in which the utility of bioisosteres is frequently dependent on context and relies upon a less than exact structural or physicochemical mimicry for the manifestation of biological effect. The thoughtful deployment of a bioisostere offers potential value in drug design campaigns by providing an opportunity to probe the effect of steric size and shape, the modulation of dipole and electronic properties, lipophilicity and polarity, or pK_a on a biological response, which may be functional mimicry or antagonism of a biological regulator [1, 2, 10]. In addition to affecting potency and function, isosteres have demonstrated utility in addressing problems associated with pharmacokinetic and pharmaceutical properties, specificity, toxicity, and metabolic activation pathways in vivo in addition to being a source of novel intellectual property [1–10]. This chapter will summarize some of the more interesting tactical applications of bioisosteres in drug design where problems of the type commonly encountered by medicinal chemists have inspired innovative solutions that offer useful instruction. The vignettes are organized based on a specific property or problem under examination rather than the perhaps more traditional approach of cataloguing based on functional group mimicry.

2 Bioisosteres Designed to Enhance Drug Potency and Selectivity

2.1 *Fluorine as an Isostere of Hydrogen*

Fluorine has emerged as a remarkable element in drug design, and its applications continue to provide interesting effects on biological activity that can often be somewhat surprising, leading to its broad deployment by the medicinal chemistry community which, in turn, has promoted a deeper understanding of its properties [15–24]. The strategic introduction of fluorine to replace a hydrogen atom can markedly influence potency [25–30]. For example, the L-nucleoside analogue emtricitabine (**2**, FTC) is the 5-fluoro analogue of lamivudine (**1**, 3TC) and is consistently the more potent HIV-1 inhibitor in cell culture by four- to tenfold, which is reflected in the inhibition of HIV-1 reverse transcriptase by the respective triphosphate derivatives [25–27]. The F atom *ortho* to the piperazine and morpholine in the antibacterial agents, the quinolone-based DNA gyrase inhibitor **3** and the protein synthesis inhibitor linezolid **4**, respectively, is

of importance to both potency and efficacy in vivo while in indole-based HIV-1 attachment inhibitors, the 4-F substituent installed in **6** enhanced the potency of prototype **5** by more than 50-fold [15, 16, 28–30].

1 R = H: 3TC (lamivudine)
2 R = F: FTC (emtricitabine)

3

4: linezolid

5: R = H: EC$_{50}$ = 153 nM
6: R = F: EC$_{50}$ = 2.6 nM
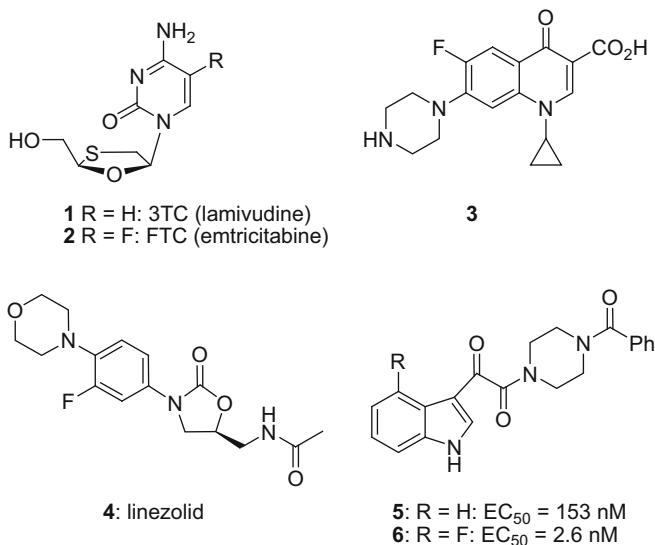
In a series of γ-secretase inhibitors that reduced Aβ protein production in a cell-based assay, the modest potency associated with the parent molecule **7** (evaluated in racemic form) was markedly improved by complete fluorination at C-5 although in this particular example gem-dimethyl substitution at C-5 also improved biological activity [31]. The introduction of a F atom at the *para*-position of the benzylated imide **10** enhanced thrombin inhibitory activity by over fivefold along with selectivity over the related serine protease trypsin by fourfold with **11** exhibiting a 67-fold preference for the coagulation cascade enzyme [17, 32]. An X-ray crystallographic structure of **11** bound to thrombin provided an understanding of the structure–activity relationship (SAR) observation with the F substituent proximal (2.4 Å) to the backbone C$_{\alpha}$-H atom of Asn$_{98}$ and aligned to establish an interaction with the C atom of the backbone amide C=O of Asn$_{98}$ at a distance of 3.5 Å and an angle of 96° to the plane of the C=O (Fig. 1) [17, 18, 32]. The F to C=O carbon interaction is the one with broader occurrence, and although this interaction is considered to be energetically modest, it appears to be sufficient to differentiate the potency between **10** and **11** by playing a supporting role to the primary interactions between the enzyme and the inhibitor [33]. In an analogous matched pair study, the thrombin inhibitory activity of the amidine **12** is embellished sixfold by F substitution to afford **13**, with an X-ray crystallographic co-crystal structure suggestive of a H-bonding or electrostatic interaction between the F atom and backbone NH of Gly$_{216}$ which are 3.47 Å apart, close to the upper distance limit for a H-bond [16]. The importance of this interaction was underscored from a comparison with the co-crystal of **12** and thrombin in which the aromatic ring adopted a significantly different conformation to that of **13** [16].

Fig. 1 Key interactions
of the 4-F substituent of
11 with Asn$_{98}$ of thrombin
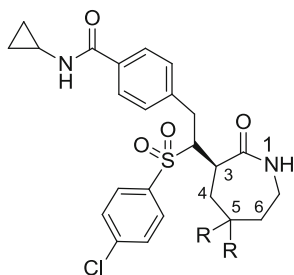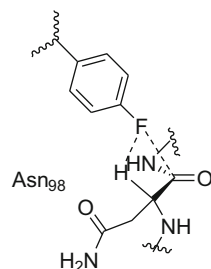
7: R = H (racemic): EC$_{50}$ = 170 nM
8: R = F: EC$_{50}$ = 4 nM
9: R = CH$_3$: EC$_{50}$ = 15 nM
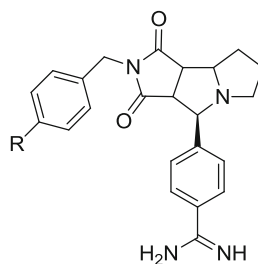
10: R = H: K_i = 310 nM
11: R = F: K_i = 57 nM
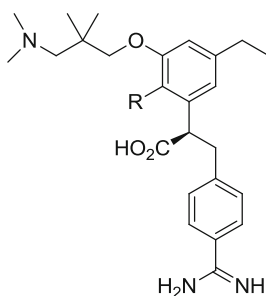
12: R = H: K_i = 1.6 μM
13: R = F: K_i = 0.26 μM

2.2 Carboxylic Acid Isosteres to Improve Potency

Carboxylic acid isosteres are a familiar part of the medicinal chemistry landscape
and have been deployed to address a number of challenges in drug design, including
positively influencing potency. A particularly compelling example that illustrates
the importance of carefully selecting the optimal isostere is provided by the SARs
that subtended the discovery of the potent angiotensin II receptor antagonist

losartan (**15**) and its analogues **14** and **16–20** [34–38]. A key insight was that the tetrazole moiety in **15** enhances potency over the carboxylic acid analogue EXP-7711 (**14**) by tenfold, attributed to the topological geometry in which the *2H* isomer depicted projects the acidic NH or negative charge 1.5 Å further from the aryl ring than does the carboxylic acid moiety [34, 39, 40]. Interestingly, as summarized by the comparison presented in Table 1, the position of the carboxylic acid moiety on the phenyl ring exerts minimal effect on potency, with **14** and **18** similarly active, while the *ortho*-tetrazole isomer **19** is over 50-fold more potent than the *meta*-substituted analogue **20** [38]. The carboxylic acid and tetrazole moieties are considered to bind in the same pocket in the angiotensin II receptor and interact with Lys_{199} [41]. However, this interaction does not appear to depend on the formation of a classic salt bridge since mutation of Lys_{199} to Gln exerted only a modest twofold effect on the binding of losartan (**15**) and a fivefold effect on the binding of EXP-7711 (**14**), data that are more consistent with a H-bonding interaction between drug and protein [41]. The poor activity of **20** may be a consequence of increased steric bulk compared to the carboxylic acid analogue **18**.

14
IC_{50} = 200 nM

15
IC_{50} = 19 nM

16
IC_{50} = 140 nM

17
IC_{50} = 6.2 nM
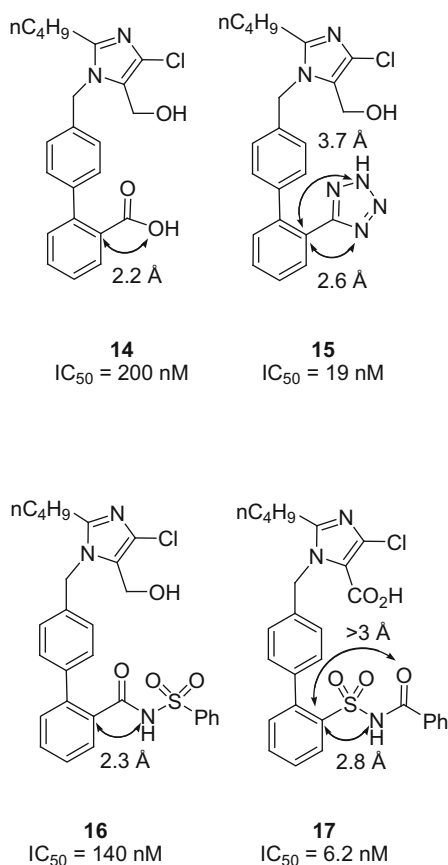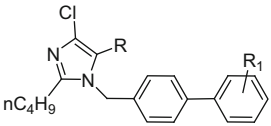
Table 1 Structure–activity relationships associated with the topology of the acidic moiety in angiotensin II receptor antagonists
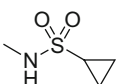
Compound no.	R	R_1	Inhibition of specific binding of [³H]-angiotensin II (2 nM) to rat adrenal cortical microsomes IC$_{50}$ (μM)
14	CH_2OH	$2\text{-}CO_2H$	0.20
18	CH_2OH	$3\text{-}CO_2H$	0.49
19	CHO	$2\text{-}CN_4H$	0.02
20	CHO	$3\text{-}CN_4H$	>1

The importance of topology and the associated geometric implications are also apparent in acylsulfonamide-based angiotensin II antagonists where this functionality functions suitably as an acid surrogate. The CONHSO$_2$Ph moiety of **16** is a 20-fold less potent ligand than the topologically reversed SO$_2$NHCOPh isomer **17**, an observation that is consistent with the longer C–S bond in **17** more effectively mimicking the topology of the tetrazole **15** by projecting the charge further from the biphenyl core than **16**, which more closely resembles the carboxylic acid **14** [34, 38, 42].

The discovery of the macrocyclic hepatitis C virus (HCV) NS3 protease inhibitor BILN-2061 (**21**) represented a milestone in HCV drug design by providing proof-of-concept for this mechanistic approach to the control of viremia in infected subjects. The disclosure of this series of tripeptide-based inhibitor stimulated considerable interest in further structural optimization, an endeavor that became of particular importance in the context of overcoming the cardiotoxicity that resulted in termination of this pioneering molecule [43–45]. The conversion of the carboxyl terminus to an acylsulfonamide, particularly the cyclopropyl acylsulfonamide moiety, afforded compounds with increased potency as a consequence of the cyclopropyl ring complementing the P1′ pocket and the sulfone oxygen atoms engaging the protein via H-bonding interactions with the catalytic elements of the enzyme. The significant advantage conferred by this structural modification can be readily assessed by comparing the potency of the matched pairs of acyclic inhibitors **22** and **23** and the P4-P2* macrocycles **24** and **25** [46–49].
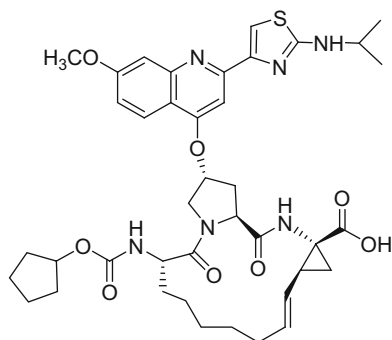
21
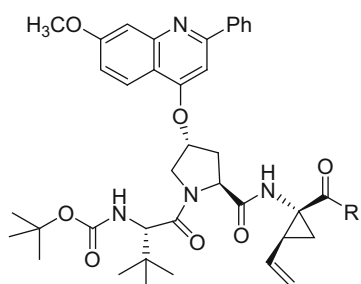
22: R = OH: IC$_{50}$ = 54 nM
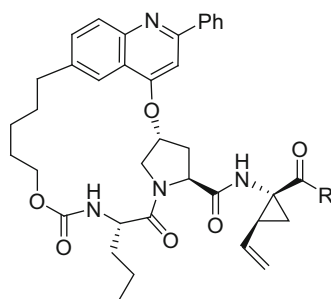
23: R = ▱   IC$_{50}$ = 1 nM
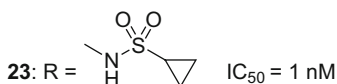
24: R = OH: K_i = 8.5 nM
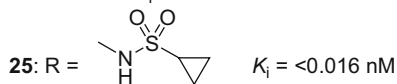
25: R = ▱ K_i = <0.016 nM

The optimization of inhibitors of the anti-apoptotic B-cell lymphoma 2 proteins Bcl-2, Bcl-x$_L$, and Bcl-w, for which the fragment leads biphenyl carboxylic acid **26** and phenol **27** were identified using a nuclear magnetic resonance (NMR) screen, involved tethering of the two molecules which bound weakly to proximal pockets in the Bcl-x$_L$ protein [50–52]. Initial attempts to link acid **26** to phenol **27** focused on substituting at the C atom *ortho* to the carboxylic acid moiety of **26**, a topology that was probed based on an analysis of NMR data [50, 51]. However, this approach was not initially successful in identifying hybrid molecules with increased affinity, attributed to poor geometric vectors that gave suboptimal binding. An acylsulfonamide was conceived to offer an improved vector while preserving the acidic moiety, of importance as a complement to Arg$_{139}$ of the Bcl protein, and a library of 120 molecules were prepared after establishing that the simple methyl acylsulfonamide exhibited comparable affinity for Bcl as **26** [51]. The acylsulfonamide **28** emerged from that exercise with further optimization to navitoclax (ABT-263, **30**) proceeding via the intermediacy of compound **29** [50–52].
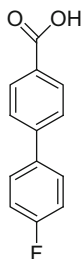
26
IC$_{50}$ FPA = 0.3 mM
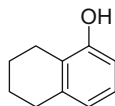
27
IC$_{50}$ FPA = 4.3 mM
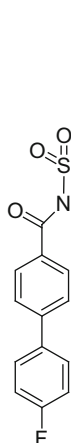
28
K_i FPA = 245 nM
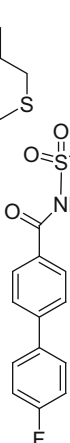
29
IC$_{50}$ FPA = 93 nM
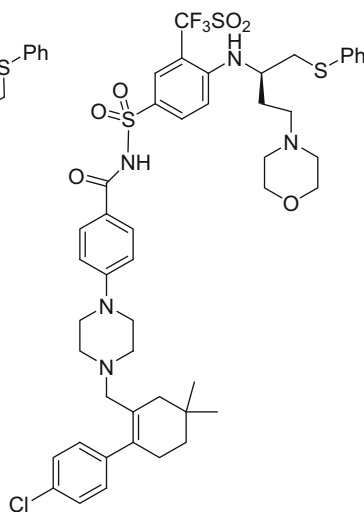
30: ABT-263
EC$_{50}$ = 5.9 nM

2.3 Applications of Heterocycles as Isosteres and Isosterism Between Heterocycles

Heterocycles find widespread application in drug design as important scaffolds for deploying functionality or as critical pharmacophoric elements that interact intimately with target proteins, a cue taken from Mother Nature who takes significant advantage of heterocycles. The immense utility of heterocycles resides in their typically facile synthetic accessibility combined with the versatility to project a range of vectors while the electronic properties of the ring can readily be modulated by the introduction of additional heteroatoms and substituent selection. Aspects of heterocycles that are of importance to the medicinal chemist revolve

around their ability to act as H-bond donors or acceptors, to influence the properties of substituents by withdrawing or donating charge and to engage in π–π and dipole–dipole interactions. Thus, the careful selection of a heterocycle for a specific application is frequently of considerable importance and there are many examples where the effects of substitution by rings offering a similar silhouette can be quite subtle in nature and not always well understood.

The H-bonding potential of a wide range of functionality, designated pK_{BHX}, has been determined based on the association of 4-fluorophenol with an acceptor in CCl_4 monitored at room temperature by the shift in the infrared stretching frequency of the OH bond [53–55]. On this scale, a stronger H-bond acceptor is associated with a higher pK_{BHX} value that shows dependence on several physical properties, including the position of the acceptor atom in the periodic table, polarizability, field and inductive/resonance effects, electronic resonance, and steric effects associated with substituents and their interaction with the acceptor atom. A particularly important aspect of H-bonding potential is that across a range of functionality it is not quantitatively related to basicity (pK_{BH}^+ or pKa) and there are several examples where a H-bonding interaction occurs preferentially at an atom of low basicity in the presence of a more basic acceptor [56–62]. For example, nicotine (**31**) protonates initially on the pyrrolidine nitrogen atom in H_2O, but H-bonding with 4-fluorophenol in CCl_4 occurs primarily at the pyridine nitrogen atom. For cotinine (**32**), the amide carbonyl is the primary site for H-bonding, which contrasts with basicity since the amide exhibits a pK_{BH}^+ value of −0.71 compared to 5.20 for the pyridine nitrogen atom where protonation occurs [56–62].


31   **32**
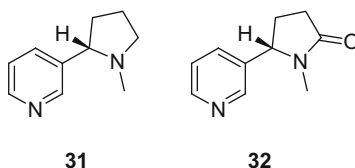
Table 2 presents a comparison of pK_{HBX} and pK_{BH}^+ (pKa) values for some common 5- and 6-membered ring heterocycles where within a homologous series there is some correlation between H-bonding potential (pK_{BH}^+) and pK_a values although isoxazole and pyridazine are notable exceptions. In the case of pyridazine, a heterocycle that has been proposed to be a privileged scaffold in medicinal chemistry, the high H-bonding potential for this poorly basic molecule is attributed to the α-effect in which unfavorable lone pair-lone pair interactions are relieved by one of the N atoms engaging in a H-bonding interaction, as depicted in Fig. 2 [63, 64].

The topological deployment of an oxazole heterocycle influenced the potency of blood platelet aggregation inhibition associated with the non-prostanoid prosta-cyclin (PGI$_2$) mimetics **33** and **34** [65–68]. The oxazole **33** is fivefold more potent than the isomer **34** which, in turn, is comparable to the *cis*-olefin **35** [65–68]. These data were interpreted in the context of the nitrogen atom of the central oxazole

Table 2 Comparison of pK_{HBX} and pK_{BH}^+ (pK_a) values for common 5- and 6-membered ring heterocycles

Heterocycle	pK_{BHX}	pK_{BH}^+(pK_a)
5-Membered heterocycles		
1-Methyl-imidazole	2.72	7.12
Imidazole	2.42	6.95
1-Methyl-pyrazole	1.84	2.06
Thiazole	1.37	2.52
Oxazole	1.30	0.8
Isoxazole	0.81	1.3
Furan	−0.40	–
6-Membered heterocycles		
Pyridine	1.86	5.20
Pyridazine	1.65	2.00
Pyrimidine	1.07	0.93
Pyrazine	0.92	0.37
Triazine	0.88	–

Fig. 2 H-bonding associated with the pyridazine heterocycle
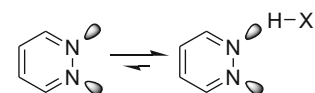
of **33** being ideally positioned to accept a H-bond from a donor within the PGI$_2$ receptor protein, consistent with earlier SAR observations. This interaction is not available to the isomer **34** since the O atom of an oxazole ring is a poorer H-bond acceptor or the olefin **35**, suggesting that the central oxazole ring of **34** is acting merely as a scaffold that is structurally analogous to **35** [65–68]. In the solid-state structures of **33** and **34**, although the molecules adopt an almost identical shape, the central oxazole rings accept H-bonds from the CO$_2$H moiety of an adjacent molecule in the unit cell, leading to a markedly different pattern of crystal packing between the two isomers [68]. Ab initio calculations indicate that a H-bond to the N atom of an oxazole is 10.4 kJ/mol or 2.48 kcal/mol more stable than that to the O atom (Fig. 3), reflected in the essentially exclusive occurrence of H-bonds only to the N atom of oxazoles observed in the Cambridge Structural Database (CSD) [68–70].

33: EC$_{50}$ = 27 nM   **34**: EC$_{50}$ = 160 nM   **35**: EC$_{50}$ = 180 nM
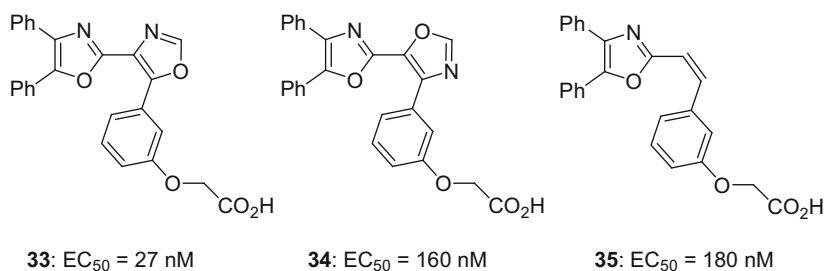
In a series of quinolone-based inhibitors of HCV NS5B polymerase, the 1,2,4-oxadiazole **36** exhibited potent enzyme inhibition, IC$_{50}$ = 41 nM, in part by

Fig. 3 Energetics of
H-bonding in oxazoles
based on ab initio
calculations

2.05 Å

-27.3 kJ/mole
-6.52 kcal/mole

2.03 Å

-16.9 kJ/mole
-4.04 kcal/mole
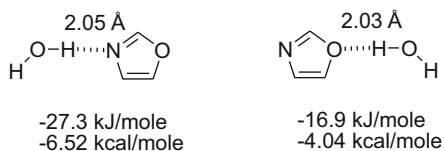
establishing a H-bond with the backbone NH of Ser$_{476}$, as determined by X-ray crystallographic analysis [71]. The topological isomer **37** is an order of magnitude weaker, consistent with the weaker potential for the O atom to engage in H-bonding despite projecting the benzyl moiety in a similar vector to **36**. In order for **37** to present the sterically somewhat hindered N-4 to the Ser$_{476}$ NH, the benzyl moiety would be required to adopt an alternate and unfavorable vector. Interestingly, the several other azole heterocycles **38–40** examined were relatively poor inhibitors of the enzyme, with **38** and **39** perhaps the most surprising given that they combine a strong H-bond acceptor with the preferred topology of the benzyl moiety that is found in **36** [71, 72].
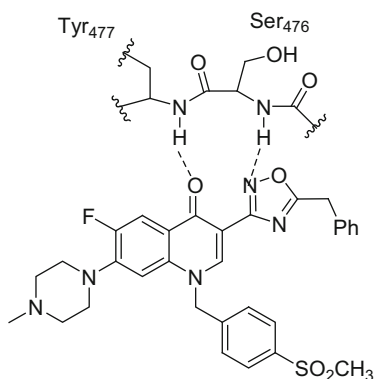
36: IC$_{50}$ HCV NS5B = 0.041 μM
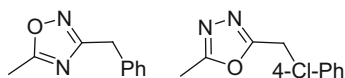
37: IC$_{50}$ = 0.73 μM   **38**: IC$_{50}$ =1.6 μM
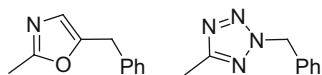
39: IC$_{50}$ =1.8 μM   **40**: IC$_{50}$ =>10 μM
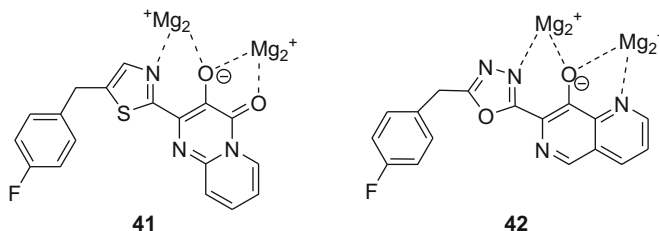
Fig. 4 Interactions between representative azole-substituted pyrido[1,2-*a*]pyrimidines and −1,6-naphthyridines and the catalytic Mg²⁺ atoms of HIV-1 integrase

Table 3 Structure–activity relationships for a series of azole-substituted pyrido[1,2-*a*]pyrimidine-based inhibitors of HIV-1 integrase

Compound no.	Heterocycle	IC$_{50}$ (nM)	Compound no.	Heterocycle	IC$_{50}$ (nM)
43		59	**47**		450
44		20	**48**		>10,000
45		45	**49**		–
46		310	**50**		225

The metal binding properties of azole rings were explored in the context of inhibitors of HIV-1 integrase and conducted with a chemotype in which the binding topology of the fluorobenzyl substituent controlled the presentation of the heteroatom to one of the catalytic magnesium atoms [73–77]. As depicted in Fig. 4, the positioning of the fluorobenzyl moieties in **41** and **42** dominates the binding conformation thereby dictating which azole heteroatom is presented to the active site metal atom, providing a sensitive assessment of metal binding potential, with SARs summarized in Table 3 [73–77]. The importance of Mg²⁺ coordination potential and its impact on potency is most evident in the comparison between the two 1,2,4-oxadiazoles **47** and **48** in which the O atom of **48** is an ineffective metal coordinator, an observation that is concordant with heteroatom H-bonding potential and was recapitulated in a series of 1,6-naphthyridine derivatives [73–75].
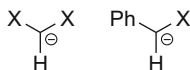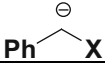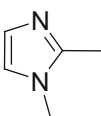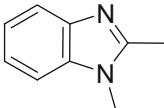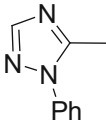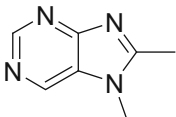
Fig. 5 Charge demand is defined as the fraction of π-charge transferred from a negatively charged trigonal carbon atom to the adjacent X group

Interestingly, in the pyrido[1,2-*a*]pyrimidine series, the thiazole **44** was selected as the basis for further optimization while in the 1,6-naphthyridine chemotype, the 1,3,4-oxadiazole ring that performs only modestly in this series (refer to compound **46**) was the most effective amide surrogate studied, presumably reflecting subtle differences in binding between the two structural classes [75–77].
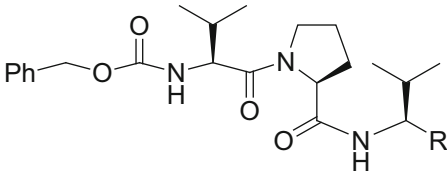
The electron-withdrawing properties of heterocyclic rings have been exploited as an important element in drug design with two aspects prominent: the acidifying effect of a heterocycle on a NH substituent, which enhances H-bonding donor properties and confers carboxylic acid mimicry in the case of sulfonamide antibacterial agents, and, secondly, the electron-withdrawing properties of a heterocycle ring that has been used to advantage in the design of mechanism-based inhibitors of serine-type proteases and hydrolases [78–81]. The site of attachment of substituents on a heterocyclic ring has been shown to be of importance to potency in all of these settings, interpretable based on the quantification of electron withdrawal. A scale of electron withdrawal has been formulated as the concept of charge demand, defined as the fraction of π-charge transferred from a negatively charged trigonal carbon atom to the adjacent X group, typically measured by ^{13}C-NMR chemical shifts of trigonal benzylic carbanions, as depicted in Fig. 5 [82–85]. The resonance electron-withdrawing capacity of common functional groups and a selection of heterocycles is summarized in Table 4, data that reveals interesting trends and emphasizes the importance of carefully selecting specific heterocycles for a particular application.

An important aspect associated with the design of serine protease inhibitors has focused on substrate mimetics that present the enzyme with an electrophilic carbonyl element. This acts as a decoy for the scissile amide bond that reacts with the catalytic serine hydroxyl to form a stable but unproductive tetrahedral intermediate [79]. The equilibrium in favor of the adduct is a function of the electrophilicity of the C=O moiety, with peptidic aldehydes an early vehicle that established the viability of this concept but which was subsequently refined by deploying trifluoromethyl ketones and α-ketoamides. Trifluoromethyl ketones provide an additional example of the beneficial effect on potency of replacing H with F since the corresponding methyl ketones are typically inactive [86]. α-Ketoamides are particularly interesting in inhibitor optimization because they provide an opportunity for the incorporation of structural elements designed to interact more extensively with the S′ pockets of an enzyme [87]. This design concept was further explored in the context of human neutrophil elastase (HNE) inhibitors that focused on tripeptidic, mechanism-based inhibitors that incorporated heterocycles as the carbonyl-activating element [88–94]. α-Ketoamides were included in this study for purpose of comparison and inhibitory potency was found to correlate nicely with

Table 4 Charge demand associated with functional groups and heterocycles

$$Ph \overset{\ominus}{\diagdown} X$$

Substituent	Charge demand c_X^{Ph}
P(O)(OEt)$_2$	0.26
SO·Ph	0.26
SO$_2$Ph	0.28
CN	0.28
Ph	0.29
CONMe$_2$	0.42
CO$_2$Me	0.40
4-Pyridyl	0.408
2-Pyridyl	0.411
3-pyridazinyl	0.417
2-Pyrimidinyl	0.430
Pyrazinyl	0.446
4-Pyrimidinyl	0.501
CO.Me	0.51
CO.Ph	0.56
2-Thiazolyl	0.380–0.413
2-Benzothiazolyl	0.457–0.471
2-Oxazolyl	0.346
2-Benzoxazolyl	0.424–0.436
(2-methylimidazol-1-yl structure)	0.283
(2-methylbenzimidazol-1-yl structure)	0.382
(5-methyl-1-phenyl-1,2,4-triazol structure)	0.411
(methylimidazopyrimidine structure)	0.536

the electrophilicity of the carbonyl moiety, as captured in the SARs summarized in Table 5 [89]. In addition to providing the potential for unique interactions with enzyme and the ability to probe interactions with S′ sites, the heterocycle moiety was also considered to offer the potential to sterically interfere with reductive metabolism of the carbonyl moiety [88–90, 94].

Table 5 SAR associated with a series of tripeptidic inhibitors of HNE incorporating activated carbonyl moieties to interact with the catalytic serine

R	HNE K_i (nM)
CO·CH$_3$	8,000
CHO	41
CO·CF$_3$	1.6
CO·Ph	16,000
CO-2-thienyl	4,300
CO-2-benzoxazole	3
CH(OH)-2-benzoxazole	21,000
CO-2-oxazole	28
CO-2-oxazolidine	0.6
CO-2-benzothiazole	25
CO-2-thiazole	270
CO-2-(1-Me)-imidazole	80,000
CO-2-(1-Me)-benzimidazole	12,000
CO-2-benzimidazole	5,600
CO-2-pyridine	22,000
CO-2-benzofuran	3,400

Studies with a matched pair of oxadiazole derivatives found that HNE inhibitory activity was sensitive to the topology of the heterocycle, a structure–activity relationship clearly rooted in the electron-withdrawing properties of the heterocycle [91, 92]. The 1,3,4-oxadiazole **51** is a potent inhibitor of HNE with a $K_i = 0.025$ nM, but the isomeric 1,2,4-oxadiazole **52** is 20-fold weaker, with a $K_i = 0.49$ nM [90]. In these inhibitors, the electron-withdrawing effect of the 1,3,4-oxadiazole moiety is similar to an oxazole, whereas the 1,2,4-oxadiazole ring performs more like an imidazole (Table 4). Based on this result, the 1,3,4-oxadiazole moiety was adopted as the carbonyl-activating moiety in more advanced SAR studies that ultimately led to the identification of ONO-6818 (**53**) as an orally bioavailable, non-peptidic inhibitor of HNE that was advanced into clinical trials [92].
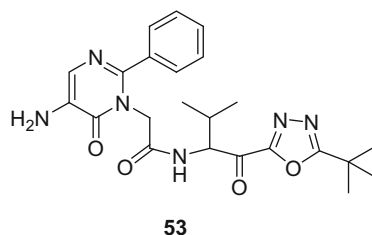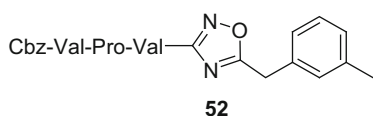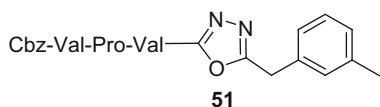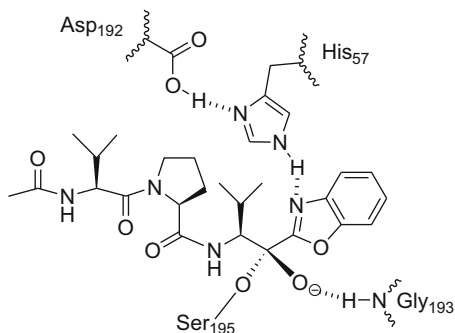
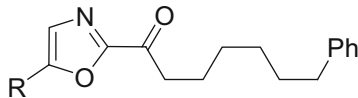**Fig. 6** Key interactions between a tripeptidic ketobenzoxazole derivative and HNE

The basic design principles were confirmed by an X-ray crystallographic analysis of a benzoxazole-based HNE inhibitor bound to the enzyme which revealed that the catalytic serine hydroxyl had reacted with the activated carbonyl moiety as anticipated to afford a tetrahedral intermediate [88]. In addition, the benzoxazole nitrogen atom engaged the NH of the imidazole of His$_{57}$, the residue that is part of the catalytic triad, in a H-bonding interaction, as summarized in Fig. 6 [88].

2-Keto-oxazole derivatives have featured prominently in the design of inhibitors of fatty acid amide hydrolase (FAAH), a serine hydrolase responsible for degrading endogenous lipid amides, including anandamide and related fatty acid amides that have been identified as neuronal modulators [80, 81, 95–103]. The electronic properties of the oxazole heterocycle were influenced by electron-donating and electron-withdrawing substituents introduced at C-5 of the ring that indirectly influenced the electrophilicity of the carbonyl moiety designed to react with the enzyme serine hydroxyl; a survey that provided the structure–potency relationships is compiled in Table 6. FAAH inhibitory activity was found to correlate with the Hammett σ constant associated with the substituent, with the plot of the data allowing the conclusion that the 5-CO$_2$H derivative bound as the anion while the 5-CHO and 5-CO.CF$_3$ analogues bound as their hydrated forms [82, 94–96, 104].
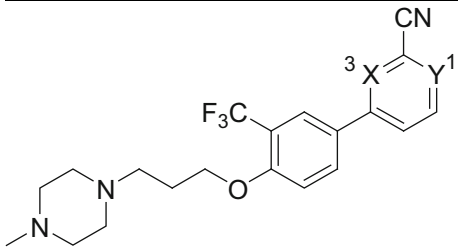
A particularly striking example of the importance of correctly deploying a heterocycle based on its electronic and H-bonding properties that relates closely to biological potency is illustrated by the series of 3 non-peptidic heteroaryl nitriles compiled in Table 7 that have been explored as inhibitors of the cysteine proteases cathepsins K and S [105]. These compounds depend on the electron-withdrawing properties of the heterocycle to activate the nitrile moiety toward addition of the catalytic cysteine thiol of the enzyme to form a stable, covalent but reversible imino thioether-based complex. The 2-cyanopyrimidine **54** is a potent and effective inhibitor of both cathepsins, but the high charge demand at this site (0.43, Table 4) is such that the nitrile moiety is indiscriminately reactive and forms adducts with thiols of microsomal proteins. The isomeric 2-cyanopyridines **55** and **56** are of lower intrinsic electrophilicity based on the reduced charge demand (0.41, Table 4) which would be expected to translate into weaker enzyme inhibition. However, **55** and **56** exhibited markedly different protease inhibition profiles that provided

Table 6 Structure–activity relationships associated with a series of ketooxazole-based inhibitors of FAAH

R	σ	FAAH K_i (nM)
H	0.00	48
CO_2H	0.00 (anion) 0.45 (acid)	30
CO_2CH_3	0.45	0.9
$CONH_2$	0.36	5
$CON(CH_3)_2$		2
$CO{\cdot}CH_3$	0.50	2
CHO	0.42 (C=O)	6
$CO{\cdot}CF_3$		3.5
CN	0.66	0.4
CH_3	−0.17	80
CF_3	0.54	0.8
I	0.18	3
Br	0.23	3
Cl	0.23	5
F	0.06	30
SCH_3	0.00	25

Table 7 Inhibitors of cathepsins S and K

Compound no.	X	Y	IC$_{50}$ (nM) Cathepsin S	Cathepsin K
54	N	N	1.3	13
55	N	CH	58	1,660
56	CH	N	Inactive	Inactive

unique insights into drug–target interactions. Pyridine **55** is 44- and 127-fold less potent toward cathepsins S and K, respectively, while the isomer **56** is inactive toward both cysteine proteases [105]. This observation was attributed to the concept that the pyrimidine N-3 is involved in cysteine thiol activation by acting as a general base to engage the thiol H atom, thereby catalyzing addition to the nitrile moiety, as depicted in Fig. 7. Replacing the pyrimidine N-3 atom with a C–H, as in pyridine **56**, not only removes this function but also introduces a negative steric interaction between the ring CH and the cysteine SH [105].
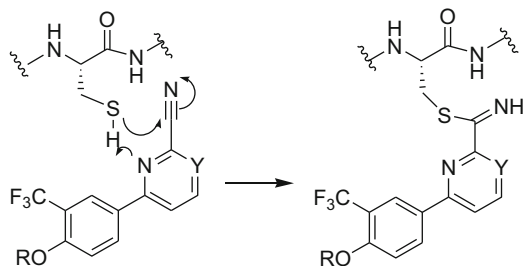
Fig. 7 Role of the pyrimidine N3 or pyridine N1 atom as a general base catalyst to facilitate addition of the catalytic cysteine thiol to the activated nitrile in cathepsin S and K inhibitors
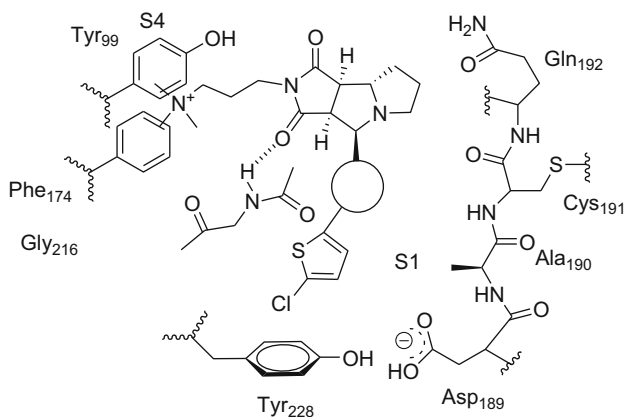

Fig. 8 Binding interactions of factor Xa inhibitors with structural elements in the active site

In a study inspired by observations of the marked potency differences between the two oxazole-based factor Xa inhibitors **57** and **58**, which represent a matched pair differing only by the topology of the oxazole rings, dipole–dipole interactions between the heterocycle ring and an amide moiety of the protein were analyzed [106, 107]. The oxazoles **57** and **58** differed in potency by over tenfold while the isoxazole **59** was the most potent with a $K_i = 9$ nM. X-ray co-crystal data indicated that these inhibitors were not engaging the enzyme via a H-bond donor–acceptor interaction, but the heterocycle rings were noted to be close to the amide bond between Cys_{191} and Gln_{192} with the planes of each almost parallel, leading to the suggestion of a dipole–dipole interaction, as illustrated in Fig. 8. Indeed, potency was shown to correlate with a favorable dipole–dipole interaction in the more active compound **57** that is mismatched in the less active isomer **58** (Fig. 9). The experimental dipole moment for oxazole is 1.7 D and for isoxazole is 3.0 D and gas phase calculations indicate that under optimal conditions the interaction energies between these heterocycles and N-methylacetamide, which has a dipole moment of 3.7 D, are −2.74 kcal/mol for oxazole and −2.95 kcal/mol for isoxazole.

57: K_i = 146 nM
favorable dipole interactions

58: K_i = 1620 nM
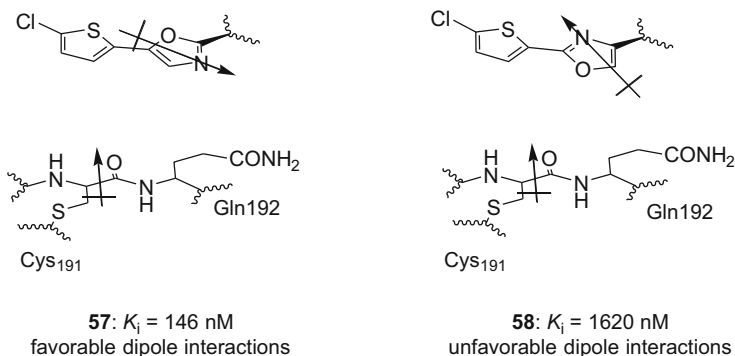unfavorable dipole interactions

Fig. 9 Dipole–dipole interactions between oxazole-based factor Xa inhibitors and the enzyme that are proposed to subtend the observed potency differences

Although the energies in an aqueous environment are expected to be lower based on solvation issues, these differences are sufficient to explain the differences in potency observed with the azoles **57–59**. A plot of the calculated energies versus dipole moment for a series of heterocycles interacting with N-methylacetamide exhibited a good correlation, $R^2 = 0.84$, while for benzene, pyrazine, and triazine, heterocyclic rings without a dipole moment, the interacting energies increase as the ring becomes more electron deficient. These studies suggested that potency can be optimized by the careful deployment of heterocycle rings in a fashion that aligns dipole moments in an antiparallel orientation compared to proximal amide moieties in a target protein. This effect can be enhanced by decreasing the π-electron density of the heteroarene in order to improve heteroarene-amide π interactions.

57: FXa K_i = 146 nM

58: FXa K_i = 1620 nM

59: FXa K_i = 9 nM
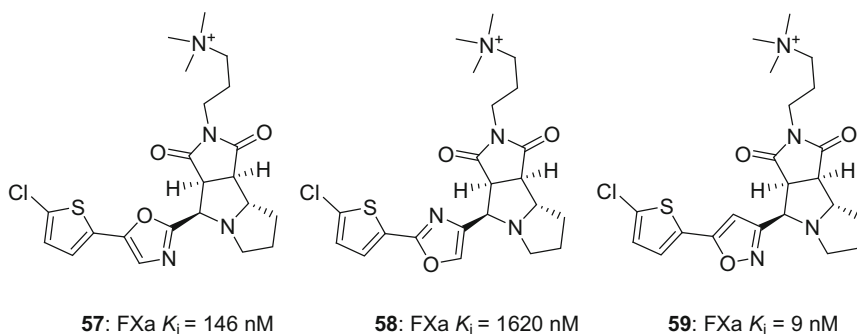
2.4   Isosteres of Drug and Water Molecules

Displacing a water molecule that mediates protein–ligand binding by introducing an appropriate structural element into the ligand to produce an isostere of the ligand–water complex can provide significant enhancements in potency by establishing new

Fig. 10 Cyclic HIV-1 protease inhibitors designed based on the premise of replacing a bound water molecule

$$60 \qquad 61$$

$$62$$
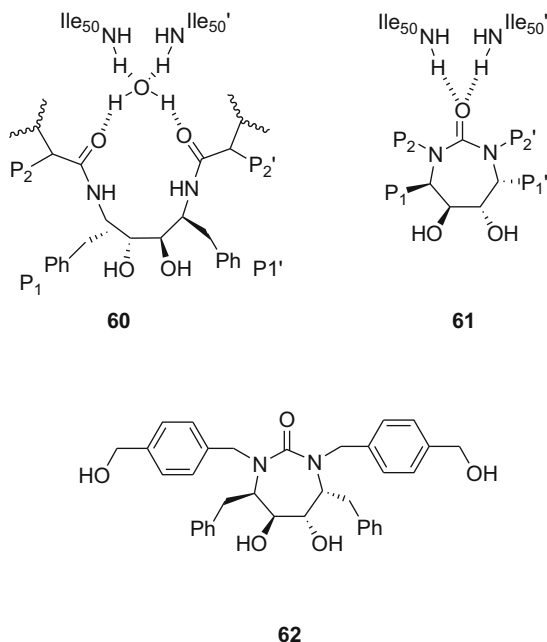
H-bonding interactions directly between the protein and ligand [108]. The increased potency arises from both increasing the enthalpic contribution to the thermodynamic signature of the association and taking advantage of the entropic energy gain that arises from releasing the bound water molecule into bulk solvent [108]. Although this can be a challenging enterprise since success depends on several factors, including the specific topology of the ligand-water-protein complex and the ability to effectively mimic the key interactions, several examples have been described where this approach has been used to considerable advantage [109–119]. The design of a series of cyclic urea-based inhibitors of human immunodeficiency virus-1 (HIV-1) protease provides a particularly compelling pioneering example of the concept that in this case is based on displacing the water molecule that mediates the interaction between the NHs of the 2 flap residues $Ile_{50}/Ile_{50'}$ and the P2/P2$'$ carbonyl O atoms of linear peptidomimetic inhibitors represented by **60**, as depicted in Fig. 10 [109–111]. A series of cyclic ureas exemplified by **61** were conceived based on the premise that the urea carbonyl moiety would displace the water molecule and establish H-bonds directly with the flap residues. As an additional benefit, the conformational constraint provided by the cyclic template would preorganize the inhibitor for recognition by the protease, providing an additional entropic advantage [109–111]. The urea **62** was identified from the initial phase of these studies as a potent HIV-1 protease inhibitor, $K_i = 0.27$ nM, that demonstrated antiviral activity in cell culture, $EC_{90} = 57$ nM [109].

Poly(ADP-ribose) polymerase-1 (PARP-1) is a DNA repair enzyme and inhibitors are potentially useful when combined with specific anticancer therapeutic
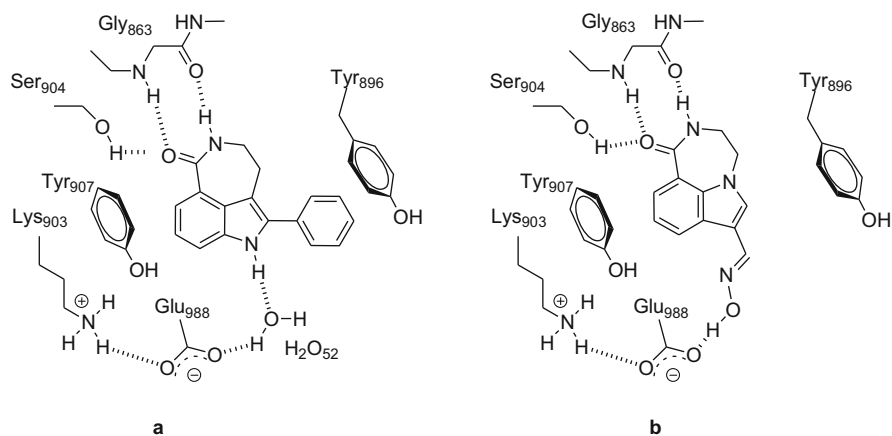
Fig. 11 Depiction of the key intermolecular interactions between chicken poly(ADP-ribose) polymerase-1 and indole-based inhibitors

Table 8 Inhibitor of human poly(ADP-ribose) polymerase-1 by 3,4-dihydro-[1,4]diazepino [6,7,1-hi]indol-1(2H)-ones

Compound no.	R_2	K_i vs human PARP (nM)
63	H	105
64	CH$_2$OH	79
65	(E)-CH=NOH	9.4
66	(E)-C=NOCH$_3$	809
67	(Z)-C=NOCH$_3$	121

agents by prolonging their antitumor activity [113–115]. Co-crystallization of the potent human PARP-1 inhibitor 5-phenyl-2,3,4,6-tetrahydro-1H-azepino[5,4,3-cd] indol-1-one, $K_i = 6$ nM, with chicken PARP revealed the presence of a water molecule, designated H$_2$O$_{52}$, that interfaced between the indole NH and Glu$_{988}$, as depicted in Fig. 11a [113]. A family of 3,4-dihydro-[1,4]diazepino[6,7,1-hi] indol-1(2H)-ones **63**–**67** (Table 8) was designed that relied upon the concept of topological inversion of the indole heterocycle to an isostere that allowed functionalization at C-3 with substituents capable of displacing H$_2$O$_{52}$ and directly engaging Glu$_{988}$ [114]. The (E)-carboxaldehyde oxime **65** depicted in Fig. 11b
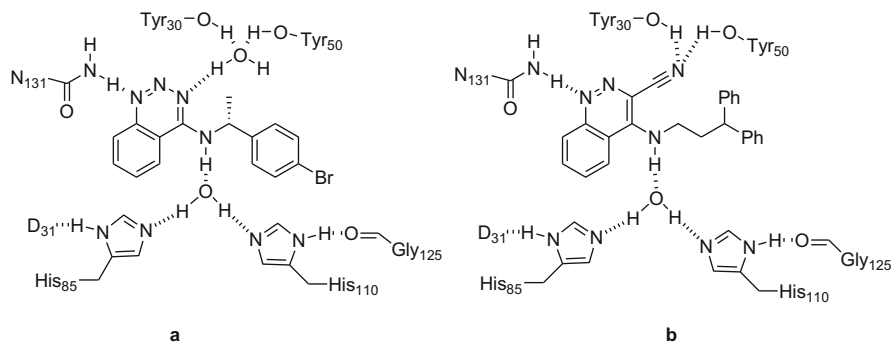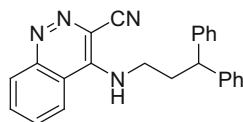
Fig. 12 Binding interactions between scytalone dehydratase and inhibitors

satisfied the design criteria and is characterized as a potent human PARP-1 inhibitor, $K_i = 9.4$ nM, that is over tenfold more potent than the prototype **63**. Indole **65** is severalfold more potent than the hydroxymethyl derivative **64** and the key structure-activity observations with the O-methylated derivatives **66** and **67** are consistent with the fundamental design hypothesis that the oxime moiety displaced H_2O_{52}. This was definitively confirmed by solving the co-crystal structure of oxime **65** with the enzyme which revealed the oxime OH engaging Glu_{988} as proposed, as captured in Fig. 11b. Interestingly, the potency-enhancing effect of introducing the oxime moiety was muted in a series of compounds that incorporated a C-6 aryl substituent [113–115].

The enzyme scytalone dehydratase catalyzes two steps in melanin biosynthesis in *Magnaporthe grisea*, a plant fungal pathogen [116]. The benzotriazine **68** is a potent scytalone dehydratase inhibitor and modeling of the analogue **72** in the active site of the enzyme (Fig. 12a) recognized its isosteric relationship with the salicylamide **71** which had been co-crystallized with the enzyme. A water molecule that played a critical role in mediating the interaction between the inhibitor and two enzyme residues, Tyr_{30} and Tyr_{50}, by H-bonding to the benzo-triazine moiety of **68** and the amide carbonyl of **72** was viewed as an opportunity for optimization of the scaffold. The removal of the key benzotriazine nitrogen atom that accepted a H-bond from the water molecule afforded the much weaker inhibitor cinnoline **69**, but the introduction of a nitrile moiety at the 3-position of the heterocycle provided a molecule, **70**, that is over 18,000-fold more potent than the progenitor [116]. The remarkable increase in potency for the addition of just two atoms represents a highly ligand-efficient modification that was also observed with an analogous quinazoline/quinoline matched pair [116, 121]. An X-ray co-crystal structure of **70** bound to the enzyme confirmed the modeling hypothesis and validated the design principle, as summarized in Fig. 12b [116].
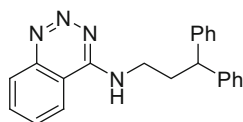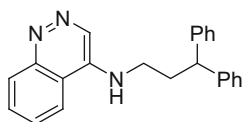
68, K_i = 0.22 nM   **69**, K_i = 140 nM   **70**, K_i = 0.0077 nM
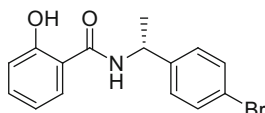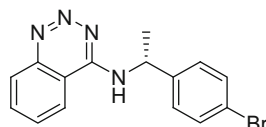

71   **72**

A similar strategy was adopted to enhance the potency of the triazine-based p38α MAP kinase inhibitor **73**, K_i = 3.7 nM, in which the key N atom of the triazine heterocycle was replaced by C-CN in order to displace a bound water molecule observed in an X-ray co-crystal structure [117]. This rationalization led to the design of the cyanopyrimidine **74** which, with a K_i = 0.057 nM, conferred a 60-fold enhancement of potency. X-ray crystallographic data obtained with a closely analogous compound demonstrated that the cyano moiety engaged the NH of Met_{109} in a H-bonding interaction, displacing the water molecule interfacing Met_{109} with the triazine N of **73**, as hypothesized during the design exercise.
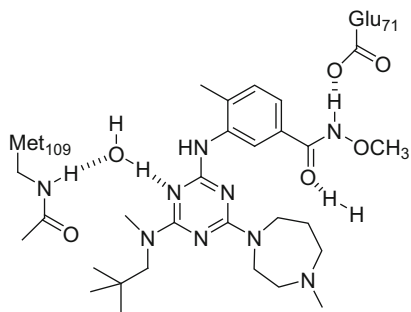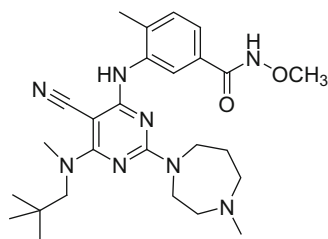
73   **74**

The displacement of a water molecule mediating the interaction between a series of heat shock protein 90 (Hsp90) inhibitors and the protein by the introduction of a strategically deployed cyano substituent contributed significantly to potency [119]. The pyrrolopyrimidine **75** was designed from the fragment screening lead **77** but was devoid of significant potency based on the results of a fluorescence polarization assay. This was attributed to the loss of a key H-bond between the imidazole
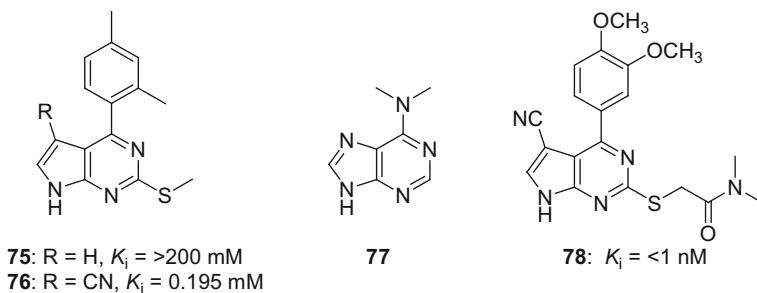
75: R = H, K_i = >200 mM
76: R = CN, K_i = 0.195 mM

77

78: K_i = <1 nM

Fig. 13 Principal drug–target interactions between pyrrolopyrimidine **78** and Hsp90

nitrogen atom of **77** and a water molecule that interfaced with Asn_{51} of Hsp90 and another water molecule in the active site. The introduction of the 3-cyano moiety to **75** gave the significantly more potent Hsp90 inhibitor **76** with the nitrile nitrogen atom engaging Asn_{51} directly, displacing the water molecule. Further optimization afforded the more potent **78** that engaged the Hsp90 protein principally via the interactions summarized in Fig. 13 [119].
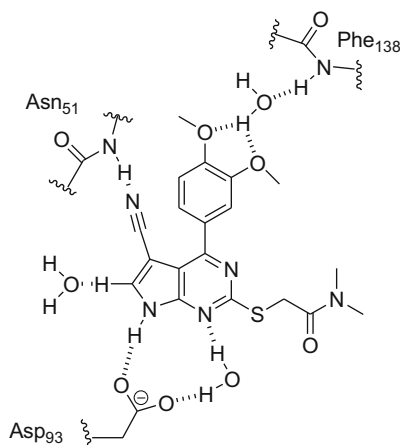
One of the difficulties associated with displacing bound water molecules is illustrated by observations with a series of human prostaglandin D2 synthase (hH-PGDS) inhibitors based on the prototype **79**, IC_{50} = 2.34 nM [120]. In the X-ray co-crystallographic structure of **79** with the enzyme, the isoquinoline nitrogen was observed to engage the hydrogen atom of a water molecule that interacted with Thr_{159} and Leu_{199} of hH-PGDS, leading to a study of the effect of incorporation of structural elements designed to displace the bridging water. The hydroxymethylated naphthalene derivative **80** and its amine analogue **81** were synthesized and confirmed by X-ray studies to perform as anticipated, displacing the water molecule and directly engaging Thr_{159} and Leu_{199} with the full complement of interactions. However, the potency of these compounds was significantly diminished compared

to **79** with IC$_{50}$s of 1,480 nM for **80** and 845 nM for **81**, representing 360- and 630-fold differences, respectively [120]. A potential explanation was suggested based on a closer analysis of the crystal structures which revealed that the topographical disposition of the hydroxyl and amino moieties of **80** and **81** was close to the plane of the naphthalene ring (21° and 27°, respectively), an energetically unfavorable arrangement that incurs allylic 1,3 strain [10, 122]. In addition, the dihedral angle between the phenyl and naphthalene rings was ~20° less than the 117° observed for **79**, data that taken together suggested that the increased energy associated with the imposed conformational constraints outweighed the entropic gains associated with displacing the water molecule [120].
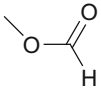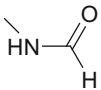

79   **80**: R = OH
              81: R = NH$_2$
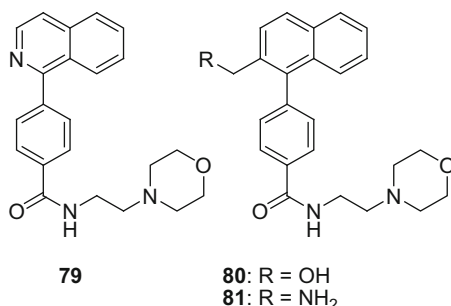
3 Bioisosteres to Modulate Conformation

3.1 *Fluorine as a Hydrogen Isostere*

The properties of fluorine make it an intriguing and useful isostere of a hydrogen atom, particularly in the context of alkanes where its unique properties have only recently begun to be examined in detail and more fully appreciated [16–21, 123–125]. Fluorine is the most electronegative of the atoms that forms covalent bonds with carbon and the polarization associated with the C–F bond affords a strong dipole moment, 1.85 D for CH$_3$F, that interacts with proximal functionality to influence conformational preferences in a fashion that can be exploited in drug design. These effects, the strength of which are dependent on the identity of the interacting substituent, are based on several underlying principles that include dipole–dipole interactions, attractive electrostatic effects, repulsion between fluorine and electronegative atoms, p orbital repulsion, and a hyperconjugative effect between an adjacent C–H bond and the low-lying C–F σ* orbital [21, 124–126]. Density functional theory (DFT) calculations have provided a relative ranking of the interaction energies between a fluorine atom and several elements and functional groups, with interactions with amine, alcohol, amide, and fluorine substituents of sufficient strength to be of practical utility in drug design exercises, data that are summarized in Table 9 [124–126].

Table 9 Energy differences between the *gauche* and *anti* conformers of α-substituted fluoroethanes based on DFT calculations

R	Δ Energy *gauche-anti* (kcal/mol) B3LYP	Δ Energy *gauche-anti* (kcal/mol) M05-2X	Preferred conformer	Predicted underlying effect
$-NH_3^+$	−6.65	−7.37	Strongly gauche	Electrostatic $F\delta^-$ and NH_3^+ δ^+
(formate ester)	−1.40	−2.18	Strongly gauche	Electrostatic C-F δ^- and C=O C δ^+
(formamide)	−1.00	−1.12	Strongly gauche	Electrostatic C-F δ^- and N-H δ^+
–F	−0.82	−0.66	Strongly gauche	σC(F)-H to σ*C-F
$-N_3$	−0.76	−1.21	Strongly gauche	Electrostatic C-F δ^- and central N δ^+
–CHNH	−0.25	−0.65	Strongly gauche	
$-CH_3$	−0.18	−0.35	Weakly gauche	
$-CHCH_2$	−0.01	−0.17	Weakly gauche	
$-C\equiv N$	0.64	−0.64	Strongly anti	p Orbital repulsion
–CHO	0.84	−1.20	Strongly anti	p Orbital repulsion and anti-parallel dipole: $C=O\delta^-\ldots\delta^+HCF$ δ^-
$-C\equiv CH$	0.98	−1.03	Strongly anti	p Orbital repulsion

1,2-Difluoroalkanes prefer a *gauche* conformation between the two fluorine atoms that is favored by 0.8 kcal/mol based upon a hyperconjugative interaction between an adjacent C–H bond donating into the low-lying C–F σ* orbital (Fig. 14a) [126, 127]. For 1,3-difluoroalkanes, the conformational preference depicted in Fig. 14b is favored by dipole–dipole interactions between the C–F bonds that is estimated to amount to a difference of 3.3 kcal/mol [128]. However, while these interactions have been shown to influence the conformational preferences of structurally simple fluoroalkane derivatives, they have yet to be fully exploited in drug design [126–130].

The conformation of an α-fluoro amide is influenced by the preference for the dipoles of the C–F and C=O moieties to align in an antiparallel fashion that in secondary amides can be reinforced by an electrostatic interaction between the electronegative F atom and amide NH [21, 131–136]. An interesting illustration of the utility of this effect is provided by the biological evaluation of the enantiomers of **83**, the α-fluorinated derivative of the vanilloid receptor agonist capsaicin (**82**) [136]. The *trans* conformer in which the C–F and C=O dipoles are aligned
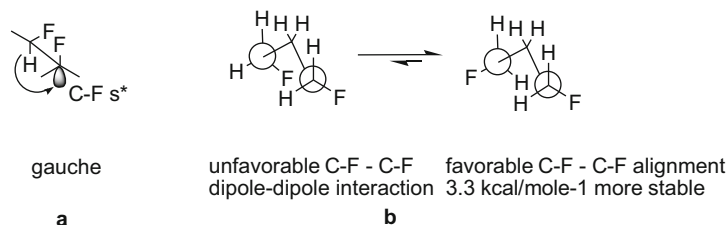
Fig. 14 Conformational preferences of 1,2- and 1,3-difluoroalkanes and the underlying interactions

in the same plane in an antiparallel fashion is calculated to be 6 kcal/mol more stable than the *gauche* conformer and 8 kcal/mol more stable than the *cis* conformer based on ab initio calculations, attributed to a favorable dipole alignment and an electrostatic interaction between the F and NH atoms, as depicted in Fig. 15 [131–136]. This conformational preference provided a potentially useful probe of the shape of the receptor based on the stereo-differentiation of the (*R*)- and (*S*)-enantiomers of **83** should the alkyl side chains project orthogonally to the plane of the amide moiety in the bound state, as depicted by the hashed bonds. However, the TRPVI receptor failed to discriminate between the two enantiomers of **83**, leading to the conclusion that the bound conformation of **82** is that in which the alkyl side chain vector projects in a plane similar to that of the amide moiety which for (*R*)- and (*S*)-**83** would be disposed in the parallel fashion depicted [136].

82: R = H, capsaicin
83: R = F, F-capsaicin
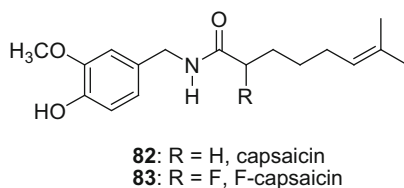
A fluorine atom deployed β- to the nitrogen atom of an amide also influences conformational preferences by favoring a *gauche* arrangement between these functionalities, a phenomenon established by X-ray crystallographic and NMR studies in the context of a series of β-peptide derivatives [126, 134, 135, 137–139]. For the difluoro succinamide esters **85** (*threo*) and **87** (*erythreo*) and the corresponding acids **84** (*threo*) and **86** (*erythreo*), solid-state structures revealed that all 4 molecules adopted a conformation that satisfied a vicinal F–F *gauche* interaction, a phenomenon that extended to solution based on an analysis $^3J_{HF}$ and $^3J_{HH}$ coupling constants in the NMR spectra [134, 135]. In the case of the *threo* isomers **84** and **85**, this conformational preference places the amide carbonyl moieties in an anti-periplanar arrangement while for the *erythreo* isomers, the amide carbonyls are disposed *gauche* [134, 135].
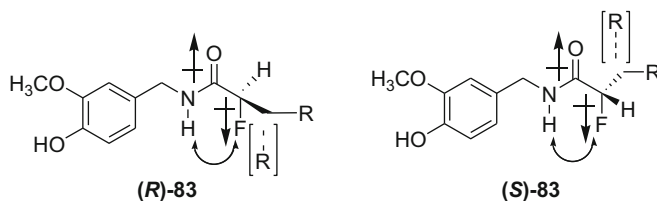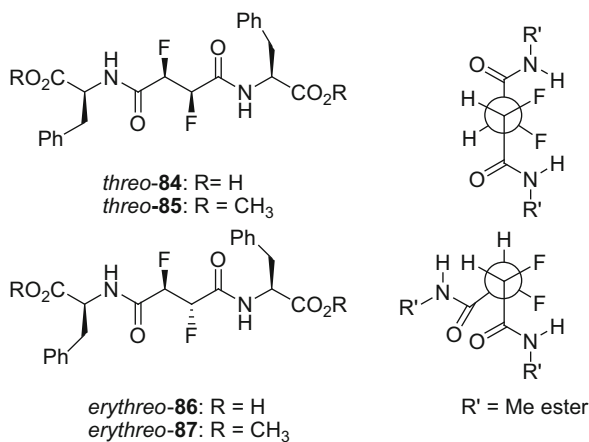
Fig. 15 Conformational analysis of the enantiomers of α-fluoro capsaicin (**83**)

threo-**84**: R= H
threo-**85**: R = CH$_3$

erythreo-**86**: R = H
erythreo-**87**: R = CH$_3$

R' = Me ester

For the GPR119 agonist **88**, EC$_{50}$ = 868 nM with intrinsic activity of 111% in a cAMP assay, isosteric replacement of the pro-(*S*) H atom β- to the amine moiety resulted in **89**, a compound with tenfold enhanced potency, EC$_{50}$ = 80 nM, and similar intrinsic agonistic properties, measured as 107% [140]. The F atom was introduced based on an appreciation of its potential to influence conformational preferences, a hypothesis confirmed by NMR analysis which revealed a population of the conformation in which the F and NH atoms are *gauche* amounting to 75% in CDCl$_3$ solution [140].

88: R = H
89: R = F
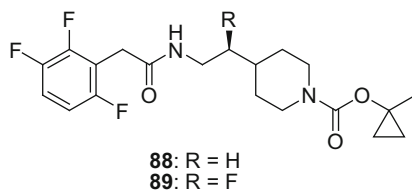
The *gauche* interaction between a fluorine atom and an amide influences the conformational preference of proline with the 4-(*R*)-F and 4-(*S*)-F derivatives **90** and **91**, respectively, favoring complementary conformers [124]. The incorporation of (*R*)- and (*S*)-4-F-Pro into collagen fibrils markedly affects the properties of the polymer in a fashion that has provided insight into the stabilizing effects of the
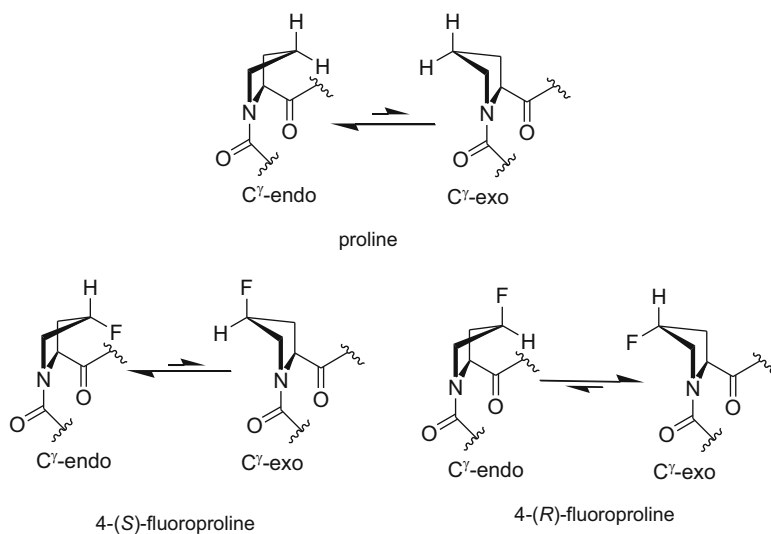
Fig. 16 Conformational preferences for proline, 4-(*S*)-fluoroproline, and 4-(*R*)-fluoroproline

4-(*R*)-hydroxyproline that occurs naturally in collagen [124, 141–146]. These studies have led to the suggestion that the stabilizing effect of 4-(*R*)-hydroxyproline on collagen, which is manifested as favoring a tightly wound helix structure, is the result of an inductive effect rather than a H-bonding interaction of the 4-OH substituent. The C$^\gamma$-endo conformation of proline is preferred by a modest 0.41 kcal/mol, leading to a 2:1 population of this conformer at room temperature while 4-(*R*)-hydroxyproline favors the C$^\gamma$-exo conformer by a calculated 0.48 kcal/ mol (Fig. 16) [147]. For 4-(*S*)-F proline (**91**), the C$^\gamma$-endo conformation is favored by 0.61 kcal/mol, attributed to stabilization by the *gauche* effect between fluorine and the amide moiety which leads to the F atom adopting an axial disposition, an arrangement that is observed in the single-crystal X-ray structure (Fig. 16) [147, 148]. In contrast, 4-(*R*)-fluoroproline (**92**) prefers the C$^\gamma$-exo conformation by 0.85 kcal/mol, with the *gauche* effect between the F and amide moieties favoring an equatorial disposition of the F atom [147]. The C$^\gamma$-exo conformation mimics that preferred by 4-(*R*)-hydroxyproline and provides an explanation for the structural similarity of collagen fragments that incorporate either 4-(*R*)-hydroxy- or 4-(*R*)-fluoroproline [149].

90
4-(*R*)-fluoroproline

91
4-(*S*)-fluoroproline
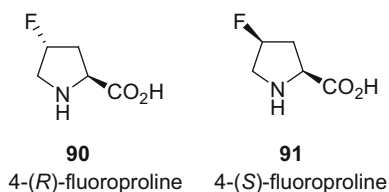
3-Fluoroproline also exhibits conformational bias as illustrated by studies with the diastereomeric esters **92** and **93** [150]. Both of these compounds crystallize with a *trans* configuration between the amide and ester moieties, as depicted in the structures of **92** and **93**, but the proline rings adopt quite different conformations. In each case, the F atom is axially disposed, favored by a stabilizing *gauche* interaction with the amide moiety such that the pyrrolidine ring of the (*R*)-isomer **92** adopts the C^{γ}-exo conformation while, in contrast, the (*S*)-isomer **93** favors the C^{γ}-endo arrangement [150]. In solution, the C^{γ}-endo conformation of **93** dominated based on the ≤ 1 Hz coupling constant between the α- and β-protons in the ^{1}H-NMR spectrum, an observation supported by calculations that indicate that this conformation should be preferred by a ratio of 97:3. For the 3-(*R*)-isomer **92**, unfavorable steric and electronic repulsive effects between the F atom and ester moiety led to a preference for the C^{γ}-exo conformation, populated to the extent of 69%. Notably, the conformational preferences of each of these 2 proline derivatives were expressed when they were incorporated into short peptide sequences [145, 150, 151].
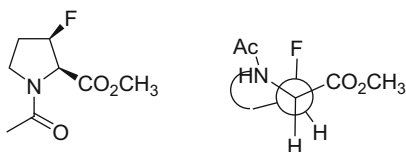

92: 3-(*R*)-fluoroproline
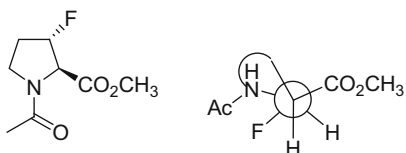
93: 3-(*S*)-fluoroproline

β-Amino fluoroalkanes exhibit conformational bias based on *gauche* interactions between the amine and F moieties, the energetics of which depend on the protonation state [21, 124, 152]. In the unprotonated form, a *gauche* relationship between the NH_2 and F of 2-fluoroethylamine is weakly favored by an energy estimated to be ~1 kcal/mol, an interaction considered to be a bridging H-bond. However, protonation results in a much stronger preference for a *gauche* relationship between the F atom and the charged amine that is estimated to be 5.8 kcal/mol and attributed to a favorable electrostatic (charge–dipole) interaction between the charged amine and the electronegative F atom [21, 124, 152]. These energetic preferences are such that cyclic amines experience considerable conformational bias, exemplified most effectively by the analysis of the protonated form of 3-fluoro-*N*-methyl-piperidine (**94**) as summarized in Fig. 17 [153–155]. Despite experiencing steric compression, the ring F atom of **94** overwhelmingly prefers an axial disposition, with conformer A in Fig. 17 representing the global minimum and favored to the extent of 95–96%,

Fig. 17 Conformational preference of *N*-methyl-3-fluoropiperidine (**94**) in its protonated form
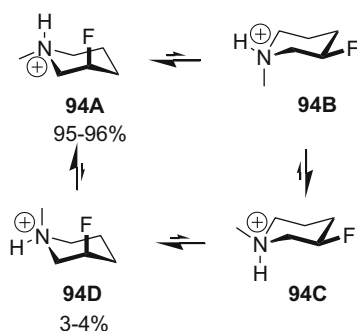
stabilized by an electrostatic interaction between the F and NH^+ moieties [155]. Conformer D contributes only 4–5% of the population at equilibrium, with a productive electrostatic effect compensating for unfavorable diaxial interactions between the F and CH_3 moieties.

The synthesis and evaluation of the two fluorinated enantiomers **96** and **97** of γ-aminobutyric acid (GABA, **95**) provided some insight into the conformation of the neurotransmitter when bound to cognate receptors or that recognized by the aminotransferase (Fig. 18) [156, 157]. Fluorination of GABA (**95**) increases the acidity of the carboxylic acid moiety and decreases the basicity of the amine but preserves the zwitterionic nature of the molecule at neutral pH. NMR analysis indicated that all molecules adopted an extended conformation in solution but with **96** and **97** further preferring an arrangement in which a favorable *gauche* interaction occurs between the fluorine and the NH_3^+ moiety [156]. The available conformations for each of the enantiomers **96** and **97** are captured in Fig. 18 with those that are disfavored based on the absence of a *gauche* effect noted [156]. Both **96** and **97** activated the cloned human $GABA_A$ receptor with comparable potency, although both were significantly less potent than the natural neurotransmitter **95**, results that are consistent with the bound form of the neurotransmitter being the extended conformation represented by B in Fig. 18 that is accessible to all three compounds [156, 157]. However, GABA aminotransferase was able to differentiate **96** and **97**, catalyzing elimination of HF from the latter with 20-fold higher efficiency than for **96**, an observation that suggested that the enzyme recognizes a conformation in which the NH_3^+ and CH_2CO_2- moieties are disposed in a *gauche* arrangement, as represented by conformation C in Fig. 18 [156, 157]. Both fluorinated derivatives were agonists at ρ1 and ρ2 $GABA_C$ receptors, with the (*R*)-isomer **97** tenfold weaker and the (*S*)-isomer **96** 20-fold weaker than the natural ligand **95** [158]. This was attributed to weaker electrostatic interactions due to the reduced basicity of the amine, observations that taken together suggested that GABA (**95**) binds to the $GABA_C$ receptor in the folded orientation represented by conformation C in Fig. 18 [158].

In an analogous fashion, the bound conformation of *N*-methyl-D-aspartate (**99**), an agonist mimetic of glutamic acid (Glu, **98**), at recombinant GluN2A and GluN2B receptors expressed in *Xenopus laevis* oocytes was probed through the synthesis
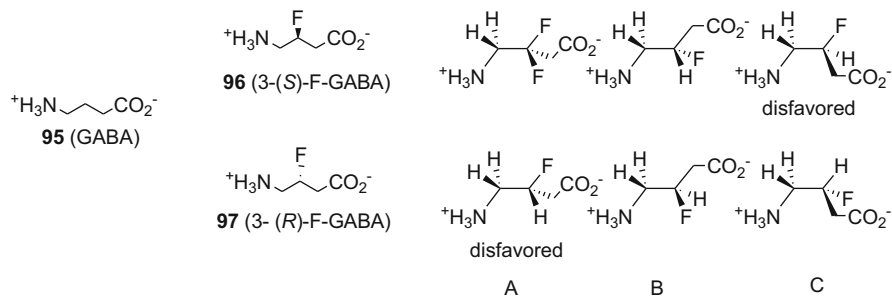
Fig. 18 Structures of GABA (**95**), the enantiomeric fluorinated derivatives **96** and **97**, and their preferred conformations

and evaluation of (2S,3S)-3F-NMDA (**100**) and (2S,3R)-3F-NMDA (**101**), the available conformations of which are depicted in Fig. 19 [159]. ^{1}H- and ^{19}F-NMR spectral data were consistent with **100** adopting conformation A in solution while DFT calculations suggested that **101** would prefer the conformation represented by B in Fig. 19, which was observed in the single-crystal X-ray structure. (2S,3R)-3F-NMDA (**101**) exhibited no significant effect in the biological assays while (2S,3S)-3F-NMDA (**98**) induced currents in oocytes expressing either GluN2A or GluN2B receptors, although the maximal responses were less than that observed with NMDA (**99**) [159]. These observations are consistent with only (2S,3S)-3F-NMDA (**100**) being able to adopt a conformation that mimics the receptor-bound form of NMDA (**99**), designated as A in Fig. 19, results that concur with the X-ray crystallographic structure of NMDA (**99**) bound to the GluN2D receptor that is highly homologous to GluN2A and GluN2B.

β-Fluoroethanol derivatives do not exhibit a strong conformational preference in the absence of an intramolecular interaction between the F and OH moieties which stabilizes a *gauche* arrangement by approximately 2 kcal/mol [152, 160, 161]. However, this effect is significantly reinforced by protonation of the alcohol which stabilizes the *gauche* conformer by an energy estimated at ~7 kcal/mol and which has been attributed to the combination of a stereoelectronic effect and an intramolecular H-bonding-type interaction. The HIV-1 protease inhibitor indinavir (**102**) and its epimer **103** presented an interesting opportunity to probe the consequences of introducing F-OH interactions on conformational disposition, probed with the congeners **104–107**, that could readily be equated with enzyme inhibitory activity, measured as the K_i values that are summarized in Fig. 20 [162]. In addition to influencing conformation, it was anticipated that the introduction of a F atom to indinavir (**102**) would affect the acidity of the OH, potentially introduce steric interactions between drug and target, and also affect solvation. In the indinavir series, the *syn,syn* analogue **104** exhibits potency comparable to the prototype **102** while the *anti,anti* isomer **106** is tenfold less active. The *syn,anti* epi-indinavir derivative **107** is eightfold more potent than progenitor **103**, but the *anti,syn* diastereomer **105** with a K_i of 5,900 nM is substantially (36-fold) less active [162]. The 2 most potent fluorinated compounds *syn,syn* (**104**) and *syn,anti* (**107**) were
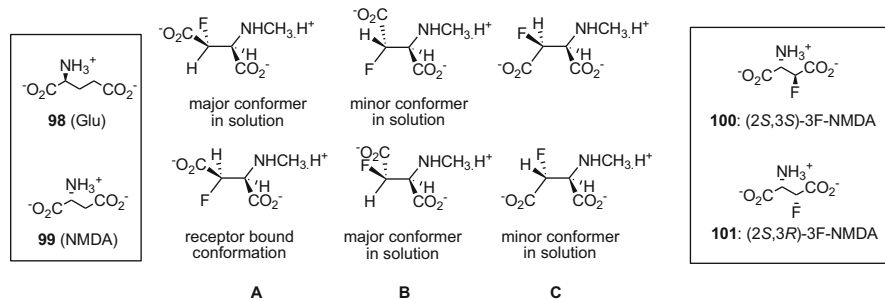
Fig. 19 Structures of glutamate (**98**) NMDA (**99**), the enantiomeric fluorinated derivatives **100** and **101**, and their preferred conformations
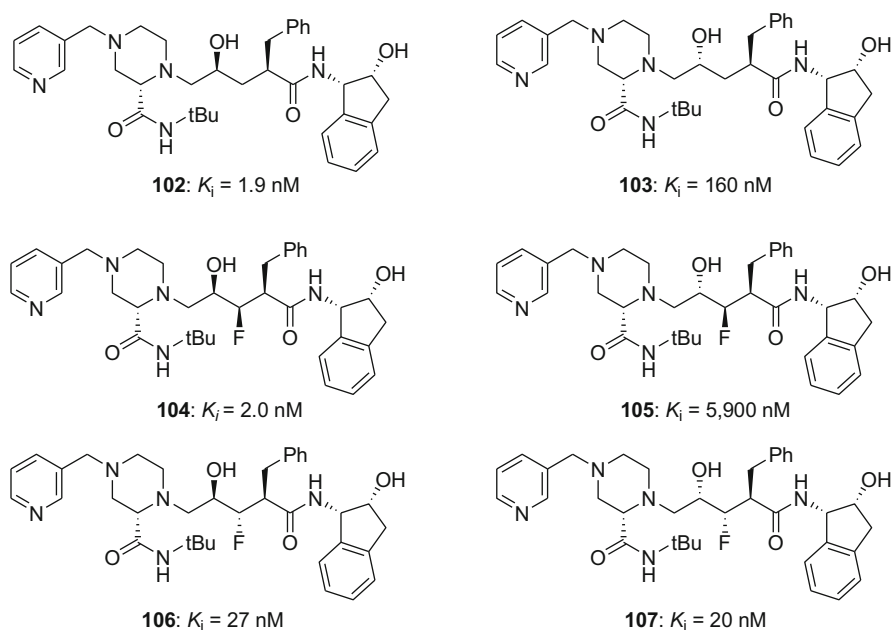

Fig. 20 The structures and HIV-1 protease inhibitory activity associated with indinavir (**102**), epi-indinavir (**103**), and the 4 fluorinated analogues **104–107**

both shown to adopt a fully extended conformation in solution based on ^{1}H-^{1}H and ^{1}H-^{19}F coupling constants, stabilized by a *gauche* relationship between the F and OH moieties, which is similar to that of indinavir (**102**) when bound to HIV-1 protease. However, the solution conformations of the less potent fluorinated diastereomers **105** and **106** were considerably more complex, sampling several additional populations and providing a potential explanation for the weaker protease inhibitory activity.

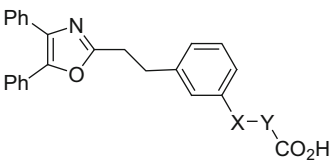## 3.2 The Conformation of Substituted Phenyl and Heteroaryl Derivatives and Isosterism

The topographical preferences of substituted aryl and heteroaryl rings are dependent upon both ring structure and the identity of the substituent atom [10, 163]. An illustration of the effect of substituent identity on biological activity is illustrated by the structure–activity relationships associated with the three non-prostanoid prostacyclin agonists **108–110** compiled in Table 10 which reveal a significant limitation on one of the simplest of the classic isosteric relationships, that between O and CH_2 [10, 65, 66]. The results have been interpreted based on the conformational preferences captured in Fig. 21 with the ether **108** and cinnamate **110** considered to be more potent based on the facility with which they are able to adopt an overall coplanar configuration with the phenyl ring, which contrasts with **109** in which the alkyl side chain projects orthogonally to the plane of the aromatic ring as a consequence of allylic 1,3-strain [10, 65, 66, 122, 163].
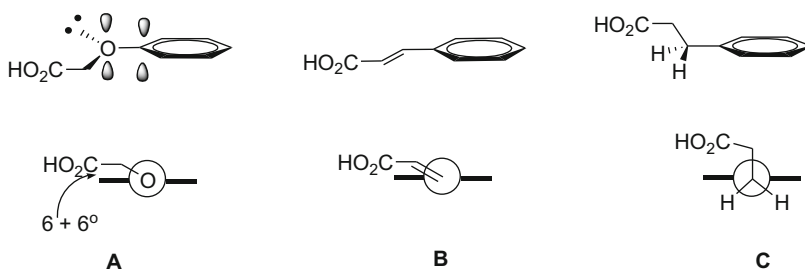
For the HIV-1 non-nucleoside reverse transcriptase inhibitor (NNRTI) **111**, replacing the OCH_2 moiety with a CH_2 linker afforded **112** which largely preserved antiviral activity, suggesting that in this context there is an isosteric relationship between the OCH_2 and CH_2 moieties [164]. This thesis was supported by the good alignment of the key elements of **112** when docked over the X-ray crystallographic structure of **111** bound to the enzyme and reflects the specific orientation of **111** in the enzyme in which the plane of the chlorophenyl and azaindazole rings approach orthogonality.

111: EC$_{95}$ WT = 8.4 nM   **112**: EC$_{95}$ WT = 22 nM
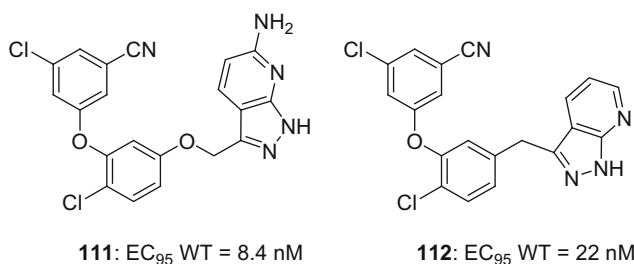
The conformational preference for arylmethyl ethers that are devoid of *ortho* substitution is one in which the MeO moiety is close to coplanarity with the phenyl ring and is rationalized on the basis of a rehybridization of the substituent to maximize electronic overlap between the oxygen lone pair and the π system [163]. However, in trifluoromethoxy- and difluoromethoxy-substituted benzenes, the substituent behaves more like an alkyl moiety, projecting orthogonally to the plane of the aromatic ring [165–169]. This conformation is calculated to be energetically favored by ~0.5 kcal/mol over the coplanar topography, which contrasts with the 3 kcal/mol preference for planarity calculated for anisole. An example where this effect appears to be manifested, at least in part, is provided by the cholesteryl ester transfer protein inhibitor **113**, IC$_{50}$ = 1.6 μM, which

Table 10 Platelet aggregation inhibition associated with a series of non-prostanoid prostacyclin agonists

| Compound no.	X	Y	Inhibition of platelet aggregation EC$_{50}$ (µM)
108	O	CH$_2$	1.2
109	CH$_2$	CH$_2$	16
110	tCH=CH		0.66

Fig. 21 Preferred conformations of phenoxyacetic acid, cinnamic acid, and phenylpropanoic acid

is eightfold less potent than the fluorinated analogue **114**, IC$_{50}$ = 0.2 µM, a structure–activity relationship attributed to both the effect of steric presentation of the substituent and a favorable drug–protein interaction [16, 166]. In this example, ab initio calculations indicate that while the ethoxy moiety in **113** is more stable when coplanar with the phenyl ring, the substituent in **114** prefers a vector that is closer to perpendicular.
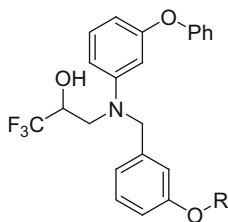
113: R = CH$_2$CH$_3$
114: R = CF$_2$CHF$_2$

Heteroaryl ethers, in which the substituent is attached to a carbon atom adjacent to a heteroatom, exhibit a topological preference based on lone pair-lone pair repulsive interactions, the energetics of which can be considerable and of both significance and utility in drug design [170–174]. The conformational preferences for a series of methoxy-substituted 5- and 6-membered heterocycles based on DFT calculations are summarized in Fig. 22 and reveal that, with the exception of furan and thiophene, the equilibrium is significantly biased for all cases examined [170]. This preference was used as a design principle to good effect in the identification of the factor Xa inhibitor ZK-807834 (**117**), a compound that has its origin in the 2,7-dibenzylidenecycloheptanone **115** [171–173]. The progenitor **115** gave a cause for concern based on the potential for photochemically induced olefin isomerization, prompting conception of the pyridine ether chemotype represented by **116** which was designed based on the conformational analysis summarized in Fig. 23. ZK-807834 (**117**) is a potent factor Xa inhibitor, $K_i = 0.11$ nM, and an X-ray co-crystal structure confirmed the predicted topology for the acyclic amidine-substituted phenyl ether although the ring with the cyclic amidine moiety adopted the less stable conformation [173].

115   **116**   **117**: ZK-807834
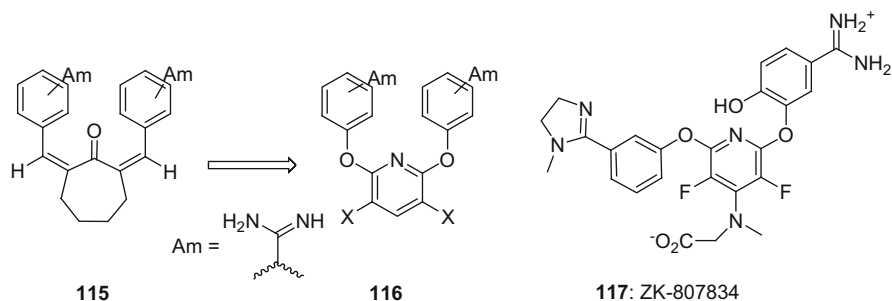
Two more recent examples where this phenomenon may be operative are provided by the corticotropin-releasing factor (CRF) antagonist **119**, which has potential for the treatment of depression and anxiety, and the GPR119 agonist **121** that promotes postprandial insulin secretion and may be useful for the treatment of diabetes [174, 175]. In both cases, a heteroaryl phenyl ether is employed to mimic the topology inherent to a fused heterocyclic ring system, the pyrrolo [2,3-*d*]pyrimidine **118** in the case of **119** and the pyrazolo[3,4-*d*]pyrimidine **120** in the case of **121**. In both examples, the favorable effect of the nonbonded lone pair-lone pair interactions may be augmented by intramolecular steric effects afforded by the heteroaryl CH_3 substituent that is manifested as allylic 1,3-strain [174, 175].

Fig. 22 Conformational
preferences or heteroaryl
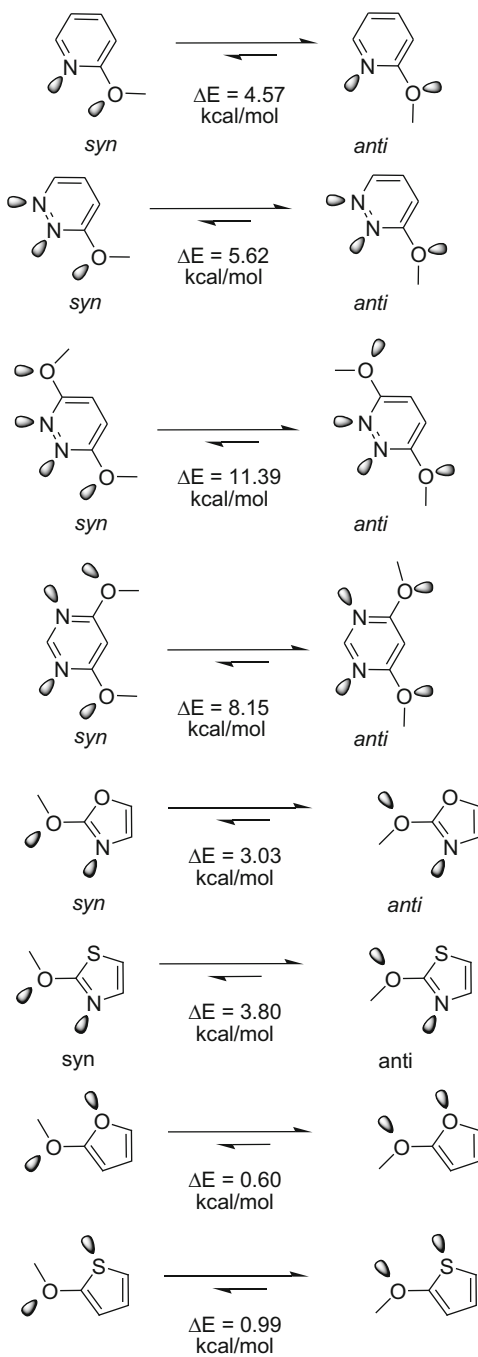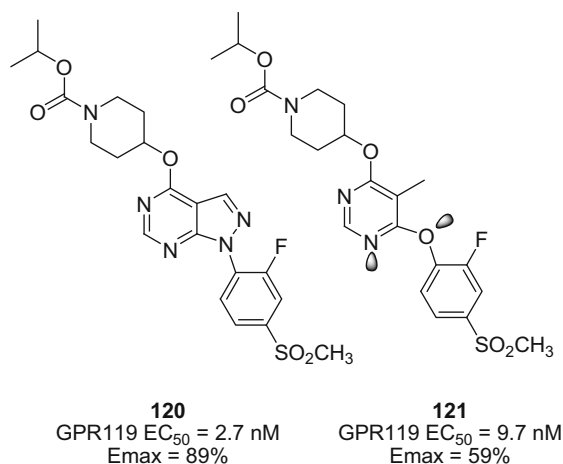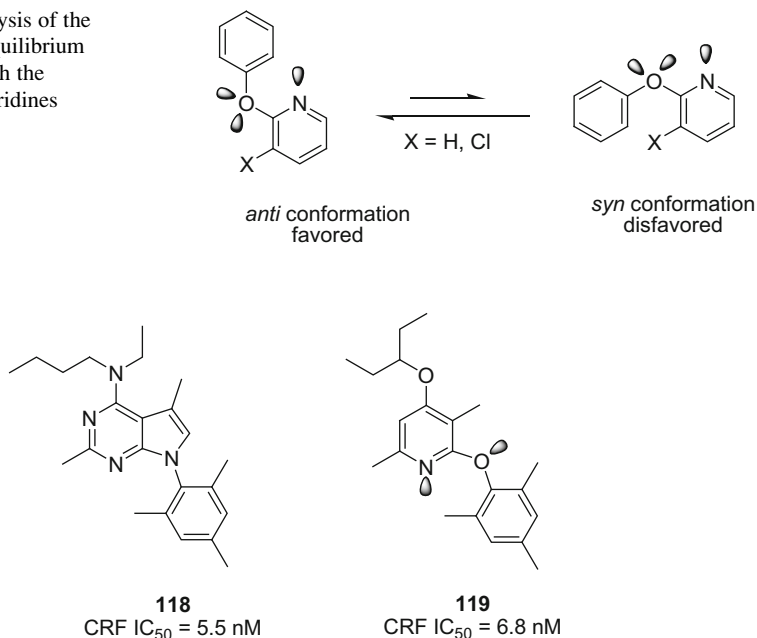ethers based on density
functional calculations

Fig. 23 Analysis of the topological equilibrium associated with the 2-phenoxy pyridines

anti conformation favored

X = H, Cl

syn conformation disfavored

118
CRF IC$_{50}$ = 5.5 nM

119
CRF IC$_{50}$ = 6.8 nM

120
GPR119 EC$_{50}$ = 2.7 nM
Emax = 89%

121
GPR119 EC$_{50}$ = 9.7 nM
Emax = 59%

The conformational bias provided by an alkyl substituent bound to a phenyl ring has been used to advantage in the design of the factor Xa inhibitor **123** (Table 11) which is based on the biphenyl prototype **122**, in which improved physical properties (lower molecular weight and cLog*P*) were sought [176]. In this example, a cyclopropyl substituent was designed to replace the phenyl ring distal from the methoxyphenyl moiety that engages the enzyme S1 sub-pocket. A careful analysis of phenylcyclopropane conformation indicated that for compounds with a benzylic

Table 11 Potency and physical parameters of drug–target interactions associated with the factor Xa inhibitors **122** and **123**

	122	**123**
K_i (nM)	0.3	0.035
Log$P$	5.94	4.99
MW	520.5	484.2
HAC	38	35
LE	0.34	0.41
LLE	3.58	5.47
LELP	17.47	12.35
Fsp^3	0.24	0.38

Fig. 24 Perpendicular and bisected conformations of phenylcyclopropane
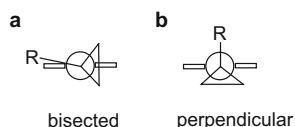
H atom, the cyclopropane ring adopts a conformation designated as bisected (conformation A in Fig. 24) that is stabilized by electronic effects associated with overlap of the cyclopropyl carbon-carbon bond orbitals with the π system. However, the introduction of a benzylic substituent alters the preference to favor a perpendicular conformation (conformation B in Fig. 24), 0.7 kcal/mol more stable for CH_3, that would more effectively mimic that of the biphenyl moiety of **122** and project the dimethylamine into the enzyme S4 pocket. Reduction to practice revealed that the phenylcyclopropane **123** is close to an order of magnitude more potent than **122**, attributed to optimized hydrophobic interactions with the S4 pocket and slightly reduced strain in the bound geometry, as summarized in Table 11 [176]. In addition, the phenylcyclopropane **123** exhibits a reduced cLog$P$ (1 $\log_{10}$) and a lower molecular weight that, when combined with the improved potency, affords significant improvements in ligand efficiency (LE), ligand-lipophilicity efficiency (LLE), lipophilicity-corrected ligand efficiency (LELP), and Fsp^3, the ratio of sp^3 C atoms to the total number of C atoms, all factors that are believed to be associated with increased drug durability in development [121, 177–182].

Fig. 25 Comparison of
vectors and dimensions of a
phenyl and bicyclo[1.1.1]
pentane moiety
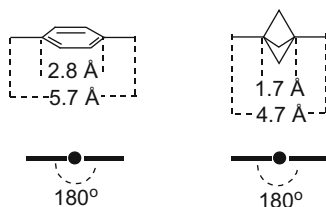
The bicyclo[1.1.1]pentane moiety is another isostere that mimics the conformation of a phenyl ring while simultaneously reducing lipophilicity and increasing Fsp^3. This moiety originally demonstrated value in the context of the glutamate antagonists **124–126** but more recently has found utility in inhibitors of the amyloid precursor protein processing enzyme γ-secretase **127** and **128** [183–186]. The presentation of the two key vectors by the bicyclo[1.1.1]pentane ring faithfully reproduces that of the *para*-substituents of a phenyl ring, but the dimensions are such that the distance between the substituents is ~1 Å shorter, as captured in Fig. 25 [186]. In the case of the glutamate antagonist **125**, a mimic of **124**, this could readily be compensated by deploying the larger tetrazole as a carboxylic acid isostere, although the result was less than impressive [183–185]. However, the propellane-based γ-secretase inhibitor **128** is an effective mimic of the phenyl analogue **127** without resort to additional structural adjustment and this compound offers improved solubility, membrane permeability, and metabolic stability in HLM than the progenitor, data summarized in Table 12 [186]. Moreover, **128** is less lipophilic, $ElogD = 3.8$ compared to 4.7 for **127**, LLE increases from 4.76 to 6.55, and the introduction of five additional sp^3 carbon atoms more than doubles Fsp^3 from 0.25 to 0.53 [186].
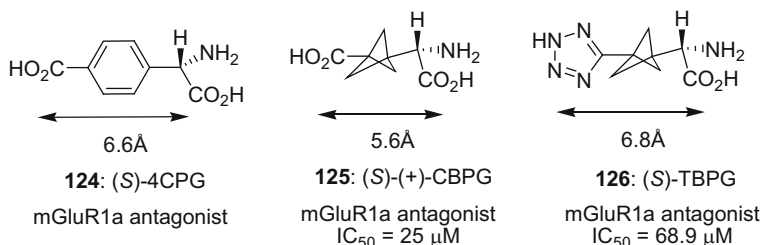

124: (*S*)-4CPG
mGluR1a antagonist

125: (*S*)-(+)-CBPG
mGluR1a antagonist
$IC_{50} = 25$ μM

126: (*S*)-TBPG
mGluR1a antagonist
$IC_{50} = 68.9$ μM

Table 12 Comparison of physical properties for the γ-secretase inhibitors **127** and **128**

Compound no.	Kinetic solubility, pH 6.5 (μM)	Thermodynamic solubility, pH 6.5 (μM)	Permeability: RRCK[a] P_{app} (A to B) [10^{-6} cm/s]	Human hepatocyte $CL_{int,app}$ (μL/min/million cells)	Elog$D$	LLE	LE	LELP	Fsp^3	Ar-sp^3
127	0.60	1.70	5.52	15.0	4.70	4.76	0.28	16.5	0.25	12
128	216	19.7	19.3	<3.80	3.80	6.55	0.30	12.5	0.53	1

[a]RRCK cells with low transporter activity were isolated from Madin–Darby canine kidney cells and were used to estimate intrinsic absorptive permeability

127
IC$_{50}$ (Aβ42) = 0.225 nM

128
IC$_{50}$ (Aβ42) = 0.178 nM
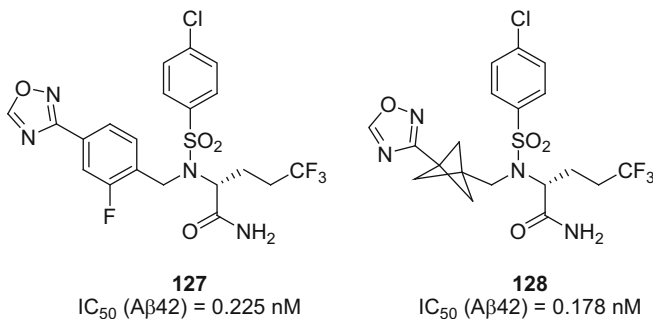
4 Bioisosteres to Modulate Drug Developability Properties

4.1 Isosteres to Modulate Permeability and P-Glycoprotein Recognition

The substitution of a H atom *ortho* to the anilide NH by a fluorine in the two closely related series of factor Xa inhibitors **129–132** and **133–135** resulted in improved Caco-2 cell permeability, as summarized in Table 13 [187, 188]. The observed increase in permeability for **130**, **132**, and **134** compared to the corresponding unsubstituted analogues **129**, **131**, and **133**, respectively, may be due to an effective masking of the H-bond donor properties of the anilide NH by association with the electronegative F atom in an electrostatic interaction [133, 189, 190]. In contrast, an *ortho* nitrile substituent in the aminobenzisoxazole chemotype afforded a compound **135** with significantly reduced permeability, presumably a function of increased acidity and H-bond donor capacity of the NH while geometrical constraints prevent the linear cyanide moiety from establishing an intramolecular H-bond.

The presence of a fluorine atom capable of engaging an amide or sulfonamide N–H in an electrostatic interaction has emerged as a common structural motif, particularly in kinase inhibitors, and may contribute to improved oral exposure, as illustrated by the matched pair comparisons between **136** and **137** and **138** and **139** that are compiled in Table 14 [191, 192].

Intramolecular H-bond capture by a pendent F atom was exploited to reduce P-glycoprotein (P-gp)-mediated brain efflux in a series of β-site amyloid precursor protein cleaving enzyme (BACE1) inhibitors [193–195]. The prototype BACE1 inhibitor **140** presented in Table 15 exhibited a large efflux ratio in cell lines expressing human or rat P-gp while analogues **141** and **142** show improved permeability properties, attributed to an intramolecular interaction between the F atom in the amide cap moiety and the NH, consistent with P-gp substrate recognition that relies upon protein–drug H-bonding [193–197].

Table 13 The effect of substituents *ortho* to an anilide on Caco-2 permeability in two series of factor Xa inhibitors
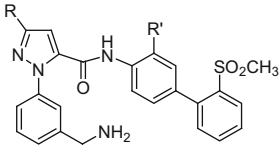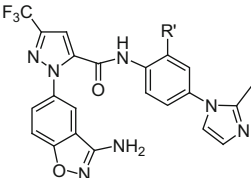

Compound no.	R	R'	Caco-2 permeability (cm/s)	Compound no.	R	Caco-2 permeability (cm/s)
129	CH_3	H	1.20×10^{-6}	**133**	H	0.82×10^{-6}
130	CH_3	F	3.14×10^{-6}	**134**	F	7.41×10^{-6}
131	CF_3	H	3.38×10^{-6}	**135**	CN	$<0.1 \times 10^{-6}$
132	CF_3	F	4.86×10^{-6}			

Table 14 Potency and oral bioavailability of Rho *kinase* (ROCK 1) inhibitors **136** and **137** and the *kinase* insert domain receptor (KDR) inhibitors **138** and **139**
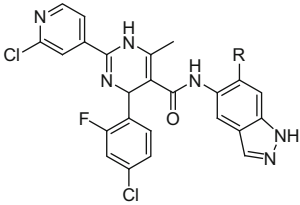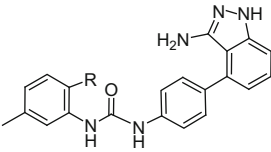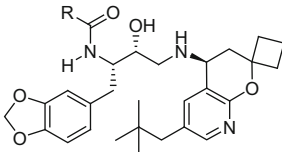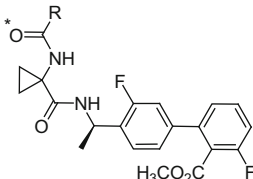

| Compound no.	R	ROCK 1 IC_{50} (nM)	F (%)	Compound no.	R	KDR IC_{50} (nM)	F (%)
136	H	4	7	**138**	H	3	38
137	F	7	49	**139**	F	4	100

Table 15 Caco-2 permeability and efflux ratios for the series of BACE1 inhibitors **140–142**

	Compound no.	R	BACE1 IC_{50} (nM)	Caco-2 P_{app} (10^{-6} cm/s)	Human efflux ratio	Rat efflux ratio
	140	CH_3	8	11	19	49
	141	CH_3 / F	28.9	16	2	4
	142	$2\text{-}FC_6H_4$	76.6	11	1	1

A similar tactical approach provided a solution to improved membrane in the series of CNS-penetrant bradykinin B1 antagonists **143–146** compiled in Table 16 which were explored as potential agents for the relief of pain [198]. The introduction of F atoms into the amide moiety resulted in improved passive permeability and reduced P-gp efflux ratios, attributed to a reduction in the strength of the

Table 16 Caco-2 permeability, efflux ratios, and H-bond strength for the series of bradykinin B1 antagonists **143–146**

| Compound no.	R	hBK1 K_i (nM)	Passive permeability P_{app} (nm/s)	P-gp efflux ratio	HBA (log) strength of C=O*
143	CH$_3$	0.93	210	8.6	2.12
144	CHF$_2$	0.40	310	3.2	1.63
145	CF$_3$	0.57	280	2.3	1.39
146	CF$_3$CF$_2$	1.6	310	1.4	1.35

Fig. 26 The *ortho*-fluorinated benzamide motif
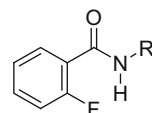
carbonyl marked with an asterisk (*) to act as a H-bond acceptor [198]. However, an intramolecular interaction between the F atoms and the NH may also be contributory.

The alternative topology depicted in Fig. 26 offers the potential for a similar electrostatic interaction between the *ortho*-F atom and the benzamide NH and this structural element is also prevalent in drug design [133, 189]. The potent tachykinin hNK$_2$ receptor antagonist **147** (Table 17) exhibited poor membrane permeability across a confluent Caco-2 cell layer, attributed to the polar amide NH moiety which was critical for potency and could not be modified by methylation or replaced by an isostere [199]. A halogen atom was introduced *ortho* to the amide carbonyl moiety to establish an intramolecular interaction with the NH, described as a H-bond with the fluoro derivative **149**, that resulted in improved permeability across a parallel artificial membrane (PAMPA) or a Caco-2 cell layer while a Cl substituent (**148**) was less effective but still superior to H. In this context, the *ortho*-F substituent in **149** performed somewhat similarly to the nitrogen atom of pyridine **151** which offers a more conventional H-bonding opportunity and markedly improved permeability in both assays compared to the analogous phenyl derivative **150** [199]. The *ortho*-F atom in the antiandrogen enzalutamide (**152**), which was approved for the treatment of castration-resistant prostate cancer by the FDA in August 2012, may contribute to its excellent pharmacokinetic profile [200, 201]. A similar motif is presented by ZD-9331 (**153**), a fluorinated methotrexate analogue which exhibits activity toward ovarian cancer cells resistant to classical thymidylate synthase inhibitors [202, 203].
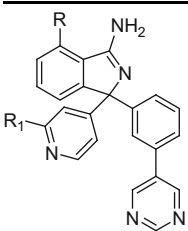
Table 17 Structure, human NK2 receptor binding affinity, PAMPA, and Caco-2 permeability for a series of phenyl alanine-based antagonists

| Compound no.	R	pK_i hNK$_2$	PAMPA P_{app} ($\times 10^{-6}$ cm/s)	Caco-2 P_{app} ($\times 10^{-6}$ cm/s)
147		9.63	ND	<1
148		8.40	0.16	4.21
149		8.49	0.66	13.80
150		8.34	1.60	9.34
151		7.57	7.23	17.61
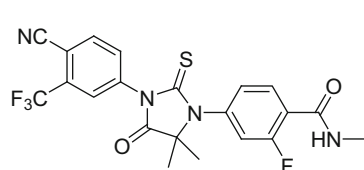
152 (enzalutamide, MDV-3100)
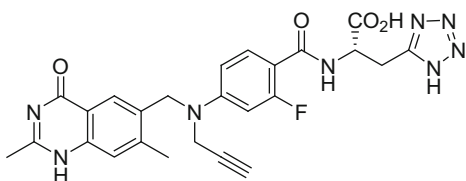
153 (ZD-9331)

Another productive F–NH relationship that exerts a beneficial effect on Caco-2 permeability and efflux ratio is provided by the series of cyclic benzamidine-based BACE1 inhibitors **154–157** (compiled in Table 18) [204]. The prototype compound **154** exhibited significant basicity, poor membrane permeability, and a high efflux ratio predictive of the molecule being a P-glycoprotein substrate which was anticipated to further contribute to reduced brain exposure. A key step toward solving the problem was the introduction of a fluorine atom *ortho* to the amidine moiety to afford **155**, a compound with a pK_a that is reduced by 1.3 units that exhibits improved Caco-2 permeability and a reduced efflux ratio while preserving BACE1 inhibitory activity. A similar effect was observed between the matched pairs **156** and **157** in which the F atom is believed to form a weak H-bond with the NH while calculations indicated that the solvation energy of the fluorinated derivative is less negative than for the H analogue, electronic and steric effects that shield the polar nitrogen atom which thus presents less than 2 H-bond donors to the environment [204].

Table 18 Caco-2 and pKa data associated with BACE1 inhibitors **154–157**

	Compound no.	R	R_1	BACE1 IC_{50} (nM)	Caco-2 P_{app} $(10^{-6}$ cm/s)	Caco-2 Efflux Ratio	pK_a
	154	H	H	500	3.4	12	8.4
	155	F	H	158	12	3.1	7.1
	156	H	CF_3	134	0.13	>10	–
	157	F	CF_3	241	39	0.6	6.9

4.2 Isosteres of Guanidines and Amidines

The high basicity associated with guanidine and amidine moieties limits the fraction of the more permeable unprotonated form that exists at physiological pH, providing an understanding for the generally poor permeability associated with molecules incorporating these structural elements [205, 206]. The prevalence of arginine as the P1 moiety of substrates of the serine protease enzymes that constitute the coagulation cascade has catalyzed the identification of guanidine and amidine surrogates as part of the effort to identify potent, selective, and orally bioavailable inhibitors that offer potential as antithrombotic agents [207–210]. The pK_a of a substituted guanidine moiety can be reduced significantly from the typical 13–14 range by modification to an acylguanidine, pK_a ~8, or an oxyguanidine, pK_a ~7–7.5, both of which represent isosteric replacements of a methylene moiety. However, these modifications have the potential to significantly affect molecular recognition that may be manifested as a reduction in potency, depending on the circumstance under study. Nevertheless, acylguanidines have been successfully deployed in a series of histamine H_2 agonists and NPY Y2 antagonists while an oxyguanidine moiety has proven to be an effective surrogate of arginine in thrombin inhibitors, with RWJ-671818 (**158**) a representative compound that was advanced into phase 1 clinical trials [211–218].
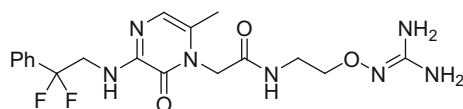
158 (RWJ-671818)

An extensive survey of benzamidine mimetics was conducted using the potent factor Xa inhibitor SN429 (**159**), $K_i = 13$ pM, as the basis for assessing the effect of this kind of structural variation on potency and oral bioavailability (Fig. 27) [219]. The study identified several neutral substituents that functioned as useful and effective amidine surrogates in this context, including the 3-chlorophenyl
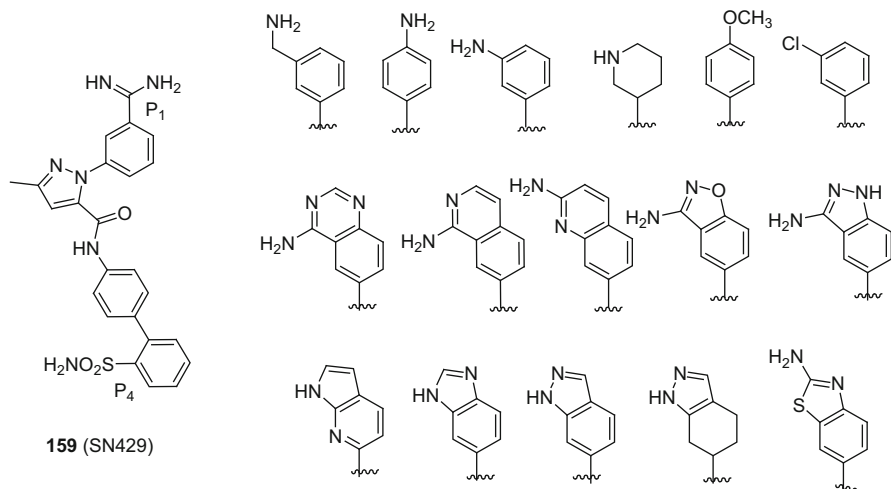
Fig. 27 The structure of the factor Xa inhibitor SN429 (**159**) and some of the benzamidine surrogates evaluated in an attempt to identify potent enzyme inhibitors with improved oral bioavailability

analogue which demonstrated a K_i of 37 nM, potency similar to that of the 3-aminophenyl derivative, which displays a K_i of 63 nM. However, both of these compounds are tenfold weaker than the more basic 3-aminomethyl compound, which exhibits a K_i of 2.7 nM, and are significantly less potent than the prototype **159**. However, the 4-methoxy analogue represented an acceptable compromise between potency ($K_i = 11$ nM) and pharmacokinetic properties and it is this P1 moiety that is found in the factor Xa inhibitor apixaban (**160**) which received marketing approval from the FDA in December 2012 as an agent for reducing the risk of blood clots and stroke in subjects experiencing atrial fibrillation that is not related to a heart valve problem [207, 220].

160 (apixaban)
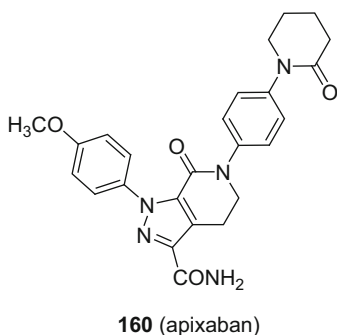
4.3 Isosteres of Phosphates and Phosphonates

Phosphates and phosphonates are highly acidic elements, and the low pK_a associated with these moieties provides a significant limit to membrane permeability and, hence, oral bioavailability [221–223]. Monofluoro- and difluoro-methylenebisphosphonic acids have been developed as useful chemically and enzymatically stable phosphate isosteres, but these rely upon preserving the inherently high acidity of the prototype and are not useful in the design of orally bioavailable drug candidates without resort to prodrug technology, which has been the most widely applied and successful tactic to deliver this kind of polar structural element [224–229]. Phosphate and phosphonate isosteres have been of particular interest in the design of nucleoside and nucleotide antiviral and anticancer agents and phosphatase inhibitors for which the design principles are based on substrate mimicry [230, 231]. However, the most notable success in the design of phosphate mimics has been accomplished in the arena of inhibitors of the strand transfer reaction catalyzed by HIV-1 integrase where the seminal identification of α,γ-diketo acids as phosphate transfer transition state mimics inspired the design of a wide range isosteres, summarized in Fig. 28, that are compatible with oral bioavailability [232–237].

The HIV-1 integrase inhibitors raltegravir (**161**), elvitegravir (**162**), and dolutegravir (**163**) have been licensed for marketing by the FDA for the treatment of HIV-1 infection [232–237].

161 (raltegravir)   **162** (elvitegravir)   **163** (dolutegravir)
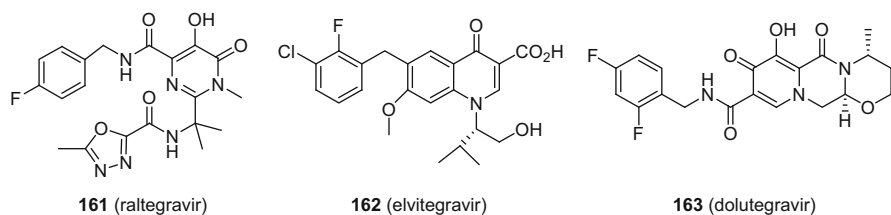
4.4 Isosteres to Modulate Basicity/Acidity and Solubility

The presence of a basic amine in a molecule can be associated with several problems including compound promiscuity, inhibition of the human *ether a go-go*-related gene type 1 (hERG) cardiac potassium channel, phospholipidosis, and recognition by P-glycoprotein [121, 238–246]. The basicity of an amine can readily be modulated by the introduction of proximal electron-withdrawing substituents or functionality, with the highly electronegative fluorine the most notable and one of the most widely utilized based on its metabolic stability and modest steric volume [247]. The pK_as of fluorinated ethylamines are summarized in
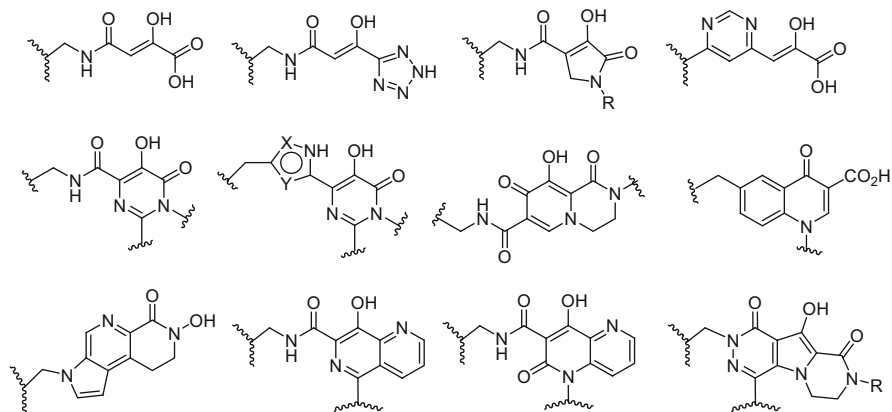
Fig. 28 A selection of HIV-1 integrase inhibiting motifs that mimic the transition state for phosphate transfer during integration of viral DNA into host chromosomal DNA

Amine	pK_a
$CH_3CH_2NH_3^+$	10.7
$CH_2FCH_2NH_3^+$	9.0
$CHF_2CH_2NH_3^+$	7.3
$CF_3CH_2NH_3^+$	5.7

Table 19 The effect of proximal F atoms on the basicity of ethylamine

Table 19 where the data indicate that the effect of introducing a F atom is additive in nature which allows the change in pK_a to be estimated with reasonable accuracy for aliphatic amines. Each fluorine atom introduced to a C atom β- to the amine reduces the pK_a by 1.7 units, an effect that is dependent on σ-transmission and therefore declines with increasing distance such that the effect at the γ-C is reduced to a shift of −0.7 units while a F at the δ-C reduces pK_a by 0.3 units, data that is reflected in Table 19.

An insightful example of the application of the effect of F to reduce amine basicity is provided by the design of inhibitors of the motor protein kinesin spindle protein (KSP) explored as a potential therapy for the treatment of taxane-refractory solid tumors [248]. The efficacy of this series of compounds was restricted by P-gp efflux which was determined experimentally to be minimized by adjusting the pK_a of the amine to a range between 6.5 and 8.0. The *N*-cyclopropyl derivative **164** represented an initial solution to this problem but was associated with time-dependent cytochrome P450 inhibition, a known liability of this structural element [249]. The β-fluoroethyl derivative **165** also satisfied the pK_a requirement but was found to be *N*-dealkylated in vitro and in vivo to release fluoroacetaldehyde which was oxidized to fluoroacetic acid, a toxin that is metabolized in vivo to a potent inhibitor of aconitase, an enzyme in the tricarboxylic acid cycle [139, 250]. This problem was solved by deploying the fluorine atom in the piperidine ring β- to the amine, an arrangement that produced the *cis* derivative **166**, in which the F atom is axially disposed as confirmed by single-crystal X-ray analysis, and the *trans*

Fig. 29 Isosterism between an amide and trifluoroethylamine and difluoroethylamine moieties
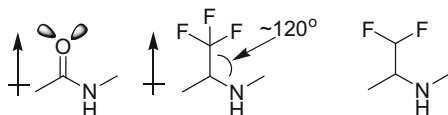
analogue **167**, where the F atom adopts an equatorial disposition. The effect of this structural modification on the basicity of the piperidine was dependent on the stereochemical disposition of the F atom with the pK_a of **167** determined to be 6.6 while **168** was more basic with a pK_a of 7.6 and it was this compound, designated as MK-0371, that was selected for clinical evaluation [248].
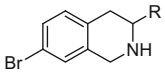
164: R = cyclopropyl
165: R = CH_2CH_2F

166 (MK-0371)

167

pK_a = 7.6

pK_a = 6.6
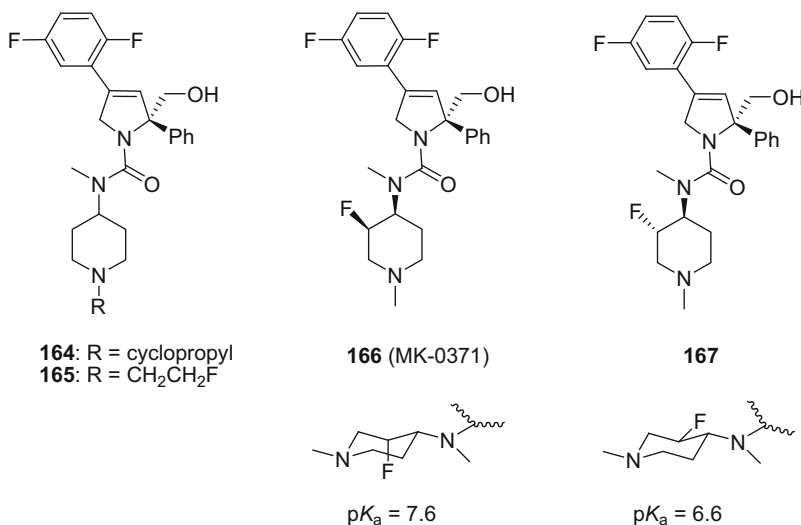
The electron-withdrawing properties of the CF_3 and CHF_2 deployed β- to an amine reduce basicity to an extent that these functionalities have found application as amide mimics, an isosteric relationship furthered by the similarity of dipoles and the geometry of the N-C-CF_3 and N-C-CF_2, which approximates the 120° associated with an amide moiety, as illustrated in Fig. 29 [251–254]. Although this bioisostere was originally conceived to replace the amide bonds of peptide derivatives, it has begun to find a similar application in drug design, with the most prominent example being the cathepsin K inhibitor odanacatib (**170**) which has completed phase 3 clinical trials for the treatment of osteoporosis. Cathepsin K is a lysosomal cysteine protease in osteoclasts that is responsible for bone degradation during remodeling and inhibitors prevent bone resorption. L-006235 (**168**), in which the nitrile moiety is presented to the enzyme as an electrophile to react reversibly with the catalytic cysteine thiol, emerged as a refined cathepsin K inhibitor that exhibited good pharmacokinetic properties. However, this compound has poor selectivity for cathepsin K versus the analogues enzymes cathepsin B, L, and S due to its strongly basic nature, which promoted accumulation in the acidic

environment of lysosomes [255]. L-873724 (**169**) offered good enzyme inhibitory selectivity but poor PK, with further optimization leading to odanacatib (**170**) which solved both problems by installing a F atom in the Ile residue to block hydroxylation while the cyclopropyl moiety reduced the propensity for amide hydrolysis [255, 256]. While odanacatib (**170**) has been quite successful clinically, with the phase 3 trial halted early due to the observation of good efficacy and safety, the pharmaceutical properties of this molecule are less than ideal [257–260]. Odanacatib (**170**) is highly crystalline and exhibits low aqueous solubility, properties that contribute to the ≤10% bioavailability across preclinical species after dosing the drug as a suspension in methocel. In an effort to overcome the dissolution-limited bioavailability, modification of the CF_3 moiety to a CHF_2 was explored as a means of improving the physical properties by increasing the pK_a of the amine [260]. This modification afforded **171** which preserved cathepsin K inhibitory potency and selectivity while decreasing log D by 3 units, attributed to the increased basicity, with the result that bioavailability from a 1% methocel suspension was improved fourfold from 6% with odanacatib (**170**) to 23% with **171** in the dog and 8–22% in the rat. The increase in basicity with **171** facilitated salt formation with strong acids like HCl although, interestingly, salts offered no advantage over the neutral form in rat PK studies [260].
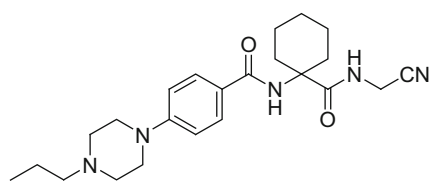
168: L-006235
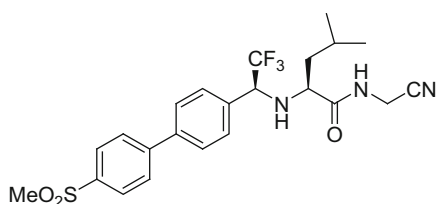
169: L-873724
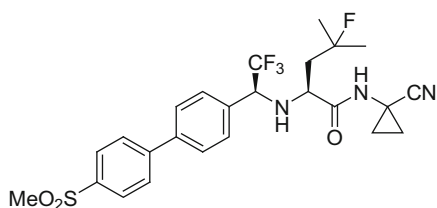
170: odanacatib (MK-0822)
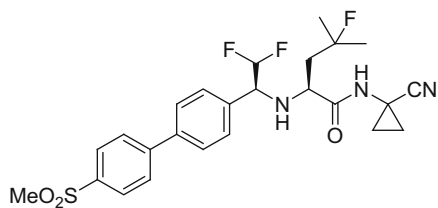
171

In a series of tetrahydroisoquinoline-based inhibitors of phenylethanolamine *N*-methyltransferase (PNMT), the enzyme that catalyzes the final step in epinephrine biosynthesis by methylating norepinephrine using *S*-adenosyl L-methionine as the cofactor, binding of these compounds to the α_2 adrenoreceptor was a significant issue. The prototypical 3-methyl derivative **172** is only modestly selective (Table 20) [261]. In an effort to address this problem, the effect of modulating the basicity of the amine by the successive introduction of F atoms into the

Table 20 Calculated basicity and inhibitory potency of a 3-substituted tetrahydroisoquinolines toward human PNMT and the α_2 adrenoreceptor

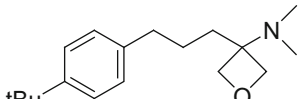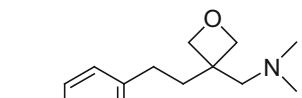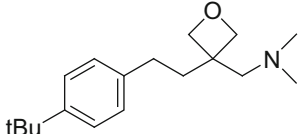| Compound no.	R	Calculated pK_a	K_i PNMT (μM)	K_i for α_2 adrenoreceptor (μM)	Selectivity: α_2/PNMT
172	CH$_3$	9.29	0.017	1.1	65
173	CH$_2$F	7.77	0.023	6.4	280
174	CHF$_2$	6.12	0.094	230	2,400
175	CF$_3$	4.33	3.2	>1,000	>310
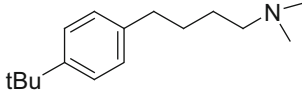
3-methyl group was explored, an approach that proved to be both productive and informative. The monofluoro analogue **173** exhibited similar affinity for the enzyme as **172** while α_2 binding declined only modestly and this molecule was calculated to be less basic than the prototype by 1.5 pK_a units. The trifluoromethyl homologue **175** exhibited poor potency in both assays and is poorly basic; however, the difluoromethylated compound **174** provided the optimal balance of properties, retaining good affinity for PNMT while reducing α_2 adrenoreceptor binding by ~200-fold and providing the first example of fluorination of an amine modulating target selectivity [261].

A detailed analysis of the properties of 3-substituted oxetane rings and applications in the context of broader functionality have established this heterocycle as an advantageous structural element when utilized in a fashion that takes advantage of its electron-withdrawing properties and topographical isosterism with the carbonyl functionality [262–265]. These properties allow the oxetane ring to be deployed as a ketone or amide mimetic depending on the structural context with the electron-withdrawing effects reducing the basicity of an amine in a fashion that is strictly dependent on proximity [262]. As is evident from the homologous series presented in Table 21, the amino oxetane **177** is the weakest base while the aminomethyl and aminoethyl homologues **178** and **179**, respectively, exhibit progressively increased basicity which shows a correlation with solubility, with the exception of the parent molecule **176** which is poorly soluble despite being the most basic [262]. It is the weak basicity associated with **177** coupled with the structural mimicry of the carbonyl moiety by the oxetane ring which adopts a planar topography that has led to this motif being considered as a useful isostere of an amide functionality (Fig. 30).

4.5 Isosteres to Modulate Lipophilicity and sp^2 Atom Count

The oxetane ring has also been proposed as a useful replacement for the gem-dimethyl moiety that provides a similar vectorial presentation of the 2 substituents (Thorpe-Ingold effect) and a similar size while reducing the lipophilic

Table 21 Basicity and solubility of a series of oxetane-substituted phenylbutyl amines

Compound no.	Structure	pK_a of N	Solubility (mg/mL)
176		9.9	<1
177		7.2	57
178		8.0	25
179		9.2	4,100

Fig. 30 Topographical similarity between the carbonyl and oxetane moieties

burden [121, 262–266]. This concept is particularly useful in an era where there is a considerable focus on the role of physical properties in drug design that has fostered the belief that contemporary practices relies far too heavily on lipophilicity to derive potency, a strategy that is based on taking advantage of entropic rather than enthalpic contributions to binding affinity [121, 177, 267–270]. The rising appreciation of the potential problems associated with this approach to drug design, referred to as molecular obesity, is beginning to be manifested in the description of strategies and tactics to identify scaffolds that replace sp^2 carbon centers by sp^3-based motifs [121, 181, 182, 238, 271–275]. One example of this is provided by studies of oxytocin antagonists of which **180** was the prototype, a compound potent receptor affinity, $K_i = 6$ nM, but poor aqueous solubility of 6 μg/mL and a low Fsp^3 of 0.18 [276]. The effect of replacing the pyrazine ring with piperidine, pyrrolidine, and azetidine was examined from which the azetidine **181** emerged as a preferred compound that bound to the oxytocin receptor with a K_i of 9.5 nM but which exhibited tenfold improved aqueous solubility of 59 μg/mL and a much increased Fsp^3 of 0.46 [276].
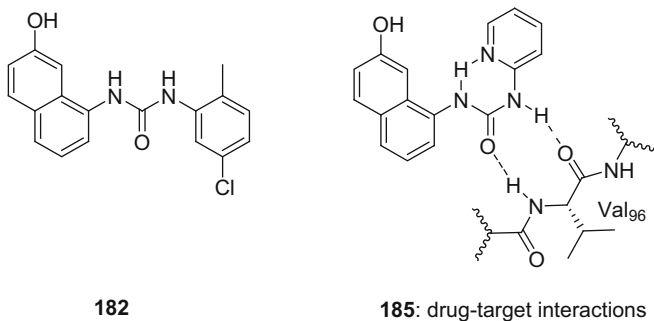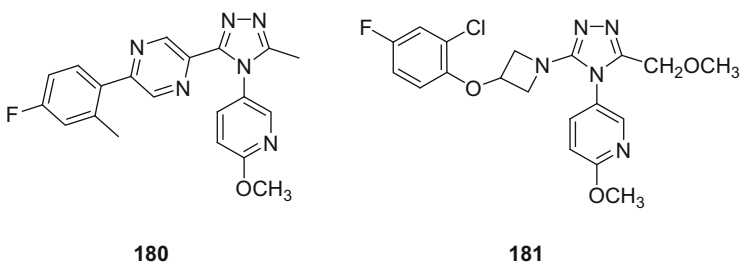
Fig. 31 Structure of the Cdk4 kinase inhibitor lead **182** and drug-target interactions for the analogue **185**

The establishment of an intramolecular H-bonding interaction can mimic the topology of an aromatic or heterocyclic ring in a fashion that reduces the dependence on sp^2 carbon centers, and this isosterism-based concept has found application in kinase inhibitor design. The Cdk4 inhibitor **182** (Fig. 31) was identified as a weak lead inhibitor, IC$_{50}$ = 44 μM, derived by structure-based scaffold generation using a model of Cdk4 and a proprietary modeling program [277]. The SAR associated with the homologous series of pyridine derivatives **183–185** is summarized in Table 22 and clearly reveals the importance of the topological deployment of the N atom, attributed to the formation of an intramolecular H-bond between the 2-pyridyl N atom of **185** and the distal urea NH that favors a *cis* amide topology. This hypothesis was subsequently confirmed by a co-crystal of the inhibitor with Cdk2 (Fig. 31) and the effect could be recapitulated with the thiazole **186** although this compound is threefold less potent than **185** [277].

PD-166285 (**187**) is a broad-acting kinase inhibitor that was used as a vehicle to explore the concept of replacing the pyridone ring with a mimic based on an intramolecular H-bond to favor the preferred topology [278, 279]. Ab initio calculations predicted that the *cis* urea is favored by 0.5 kcal/mol in H$_2$O and 3.2 kcal/mol in the gas phase, a contention supported by a search of the pyrimidinyl urea substructure in the CSD where seven molecules were found, all of which

Table 22 SAR associated with urea-based Cdk4 kinase inhibitors

| Compound no.	R	IC$_{50}$ (μM)
183	4-Pyridyl	110
184	3-Pyridyl	340
185	2-Pyridyl	7.6
186	2-Thiazolyl	23
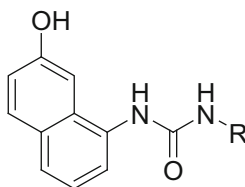
Fig. 32 Preferred conformation of 1-methyl-3-(pyrimidin-4-yl)urea
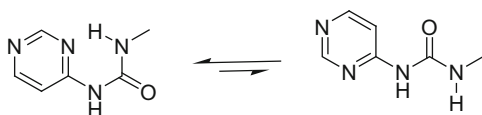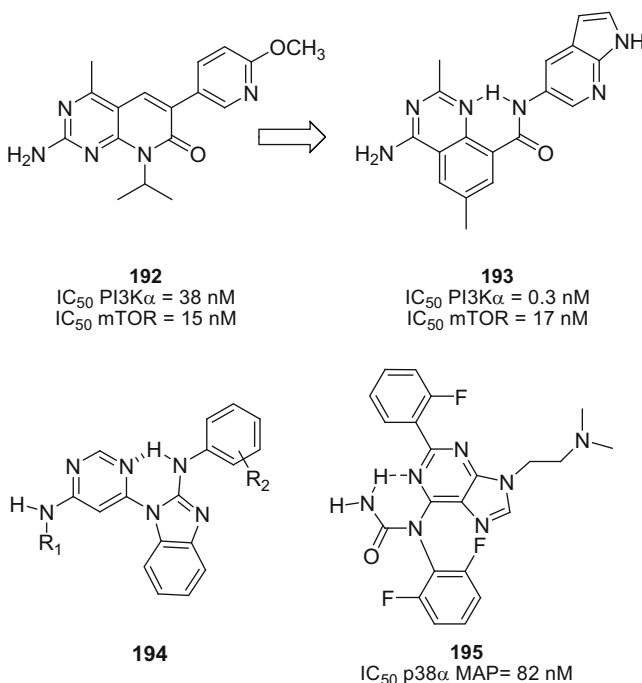

exhibited an intramolecular H-bond (Fig. 32). The urea **188** demonstrated broad-based kinase inhibitory activity with potency comparable to **187** and the derivative NVP-BGJ398 (**189**) was advanced into phase 1 clinical trials [278, 279].


187

188

189
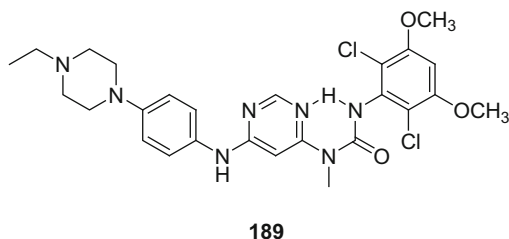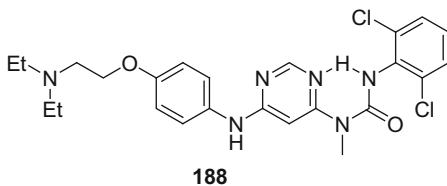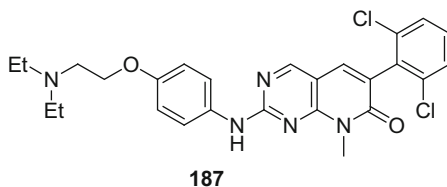
Another successful example of the application of this concept is provided by the anthranilamide **190** which mimics the phthalazine **191**, discovered as an inhibitor of the fibroblast growth factor receptor family of receptor tyrosine kinases by high-throughput screening [280]. The *anti* topology of the *para*-chlorophenyl ring depicted is calculated by *ab initio* methods to be 3.1 kcal/mol more stable than the *cis* conformer. This facilitated the design concept of replacing the pyridazine ring by an intramolecularly H-bonded moiety which led to compounds with potent kinase inhibitory properties [280].

190

191
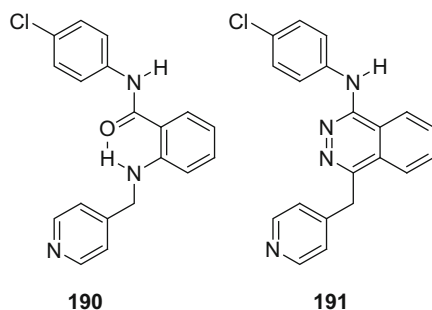
Other examples from the kinase inhibitor arena where this concept is well represented include the dual PI3Kα/mTOR inhibitor **193** which is derived from **192**, Lck inhibitors **194**, and the p38α MAP kinase inhibitor **195** [281–283]. The presence of intramolecular H-bonds in molecules of this type in an aqueous environment has recently been verified by NMR [284, 285].

192
IC$_{50}$ PI3Kα = 38 nM
IC$_{50}$ mTOR = 15 nM

193
IC$_{50}$ PI3Kα = 0.3 nM
IC$_{50}$ mTOR = 17 nM

194

195
IC$_{50}$ p38α MAP= 82 nM

An interesting application of the isosteric interplay between rings and H-bonded surrogates is provided by the design of the B-Raf^{V600E} kinase inhibitor **200** which has its origins in the benzamide derivative **196** [286]. The lead amide **196** exhibits poor solubility, 3–9 μg/mL, and a high melting point of 226°C and offers only a very weakly basic aminopyridine moiety, pK_a = 0.7, as a vehicle for salt formation, leading to a dependence on an amorphous dispersion-based formulation to deliver adequate oral exposure. The potency of **196** was improved markedly by modifying the aminopyridine moiety to the pyrazolopyridine found in **198** and both compounds were used as vehicles to explore the design of urea-based inhibitors that relied upon intramolecular H-bonding to correctly orient the pharmacophoric elements. This was successful in the context of the ureas **198** and **199**, but chemical instability led to further structural refinement with optimization ultimately focusing on the reverse amide series represented by **200**, a series with improved physical properties and robust antitumor activity in a B-Raf^{V600E} mouse xenograft model [286].
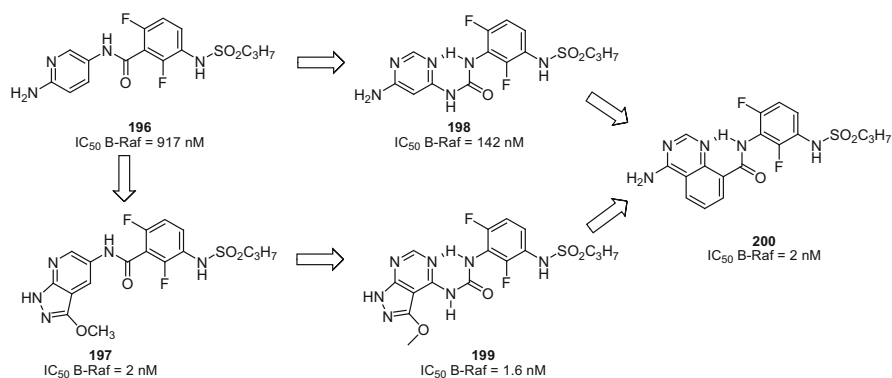

The sulfonamide moiety was developed as a useful phenol isostere in the context of β-adrenergic antagonists with advantage since the high polar surface area and H-bonding properties of this moiety frequently restrict blood–brain barrier penetration as a consequence of reduced permeability and recognition by P-gp [1–10, 196, 197, 287–292]. These physical properties have been exploited in restricting drug molecules to peripheral tissues, but the increased polarity may also reduce oral bioavailability due to poor membrane permeability and there are occasions where more lipophilic moieties are desirable [293, 294]. Although the 2,2-difluoro-2-(pyridin-2-yl)ethanamine moiety in the thrombin inhibitor **202** was introduced to interfere with benzylic oxidation by cyctochrome P450 (CYP 450), this moiety may, in essence, be functioning as a sulfonamide isostere based on analogy with the prototype **201** [216–218, 295–297]. In this context, not recognized in the design process, the two electronegative F atoms may be viewed as nonpolar mimics of the sulfone oxygen atoms that also reduce the basicity of the nitrogen atom and enhance the H-bonding donating properties of the NH to more effectively mimic the sulfonamide. The extra atom introduced with the phenethyl moiety may

Table 23 Structure and properties associated with the dual D2-receptor/β2-adrenoceptor agonists **203** and **204**

201
Ki = 4 nM

202
Ki = 1.2 nM

Compound no.	**203**	**204**
X	SO_2	CF_2
Log $D_{7.4}$	2.51	3.26
β2 $p[A]_{50}$	8.09	7.77
Intrinsic agonist activity	0.51	0.49
β2 duration (min)	146	>180
pK_a of the N atom	<8	>8

compensate to some extent for the longer bonds associated with the C–S bonds in a sulfonamide. However, the difluoroalkyl moiety has also been invoked as a heteroaryl ether isostere with the caveat that the geometry and topology of this moiety differs from that of an oxygen-based substituent; indeed, the difluoroalkyl substituent may more effectively mimic a fluoroalkyl ether which adopts an orthogonal disposition to the phenyl ring [298]. Interestingly, **202** also incorporates a fluorobenzene moiety to mimic the pyridone with a close interaction between the F atom and the thrombin peptide backbone NH observed in the co-crystal, indicative of isosterism between C–F and C=O bonds.
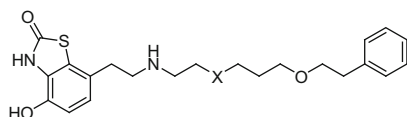

Another example where a difluoromethyl mimics a sulfone can be found in the matched pair of long-acting dual D2-receptor/β2-adrenoceptor agonists **203** and **204** captured in Table 23 [299]. The affinity of **203** and **204** for the β2-adrenergic receptor and the intrinsic agonistic properties are very similar, but the log $D_{7.4}$ and pK_a of the nitrogen atom for the 2 molecules differ, with the difluoromethyl analogue **204** more lipophilic but more basic. The design of **204** was based on the development of a model that identified secondary amines with a $pK_a > 8.0$ and a log $D_{7.4} > 2$ as molecules possessing an ultra-long duration of action [299].

5 Isosteres to Address Metabolism and Toxicity

5.1 Substitution of Hydrogen by Deuterium

The substitution of H by D is the most conservative form of bioisosteric replacement and this tactic has attracted considerable attention recently as a practical approach to drug design. The underlying premise relies upon the effect of deuterium substitution to influence pharmacokinetic properties when deployed in a strategic fashion that takes advantage of the kinetic isotope effect (KIE) associated with the heavier atom [300–304]. The differences between the isotopes are small but measurable, with deuterium 0.140 cm^3/mol per atom smaller than hydrogen, the C–D bond is 0.005 Å shorter than a C–H bond, and the log P_{oct} of D is 0.006 units less lipophilic, an effect that can be measured in per-deuterated alkanes and which has facilitated the chromatographic separation of enantiomers that are based solely on H/D isotopic substitution [305–307]. Deuteration slightly increases the basicity of proximal amines and reduces the acidity of phenols and carboxylic acid derivatives [308–311]. Following the discovery of the isotope, deuteration of drug molecules was quickly adopted as a useful structural modification for the study of metabolic pathways and the origins of toxicity, with the KIE playing the key role in this application [300, 312–314]. However, the study of the deuteration of drug candidates as an approach to improving pharmacokinetic properties by reducing the rate of metabolic modification at sites where H atom abstraction determines the rate of reaction has been adopted only recently, although the first demonstration of the practicality of the approach was described in 1961 [315]. The KIE for a D for H substitution is dependent on the circumstances but usually ranges from one- to sevenfold with calculations suggesting a seven- to tenfold difference; however, in specific cases, much higher isotope effects have been measured [316, 317]. D for H substitution has been shown to translate into meaningful changes in the clinical pharmacokinetic profiles of drugs, exemplified most effectively by CTP-347 (**205**), a dideuterated analogue of the antidepressant paroxetine, and SD-254 (**206**), a per-deuterated derivative of the dual serotonin/norepinephrine reuptake inhibitor venlafaxine that also finds utility for its antidepressant properties [303, 318, 319]. In CPT-347 (**205**), deuterium is introduced at the methylenedioxy moiety, the site of metabolic modification by CYP 2D6, which leads to mechanism-based inhibition of the enzyme through the intermediacy of a carbene-based metabolite, resulting in drug accumulation on repeat dosing due to autoinhibition and precipitating drug–drug interactions [303, 320–322]. CPT-347 (**205**) does not exhibit significant mechanism-based inhibition of CYP 2D6 in vitro and a phase 1 clinical study demonstrated that subjects dosed with the deuterated drug were able to metabolize the CYP 2D6 substrate dextromethorphan more effectively when compared to historical data for paroxetine [303]. In a phase 1 clinical trial in normal healthy volunteers, SD-254 (**206**) was reported to be metabolized more slowly than venlafaxine which is O-demethylated by CYP 2D6 [303].
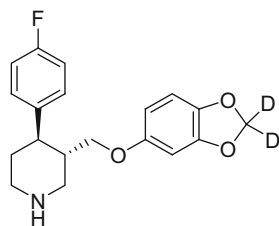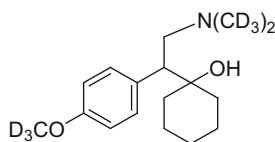
205: CTP-347   **206**: SD-254

However, selective deuteration at a site of metabolism does not always lead to enhanced pharmacokinetic properties, illustrated by a study of the tyrosine kinase inhibitor imatinib (**207**) for which N-demethylation to the less active piperazine **208** is a major metabolic pathway [323–325]. The tri-deuterio analogue **209** exhibited increased stability toward N-demethylation in rat and human liver microsomes, as anticipated, but intravenous administration of the compound to rats was not associated with increased exposure and the rate of demethylation of **209** was similar to **207**, attributed to a relatively low rate of demethylation of the parent in rat liver microsomes and the low rate clearance of both compounds in rats [323].
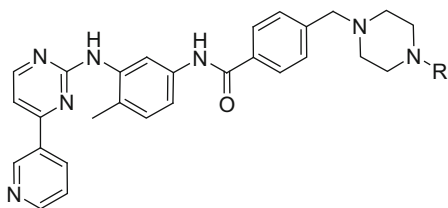

207: R = CH$_3$
208: R = H
209: R = CD$_3$

An alternative application of deuterium incorporation that led to improved pharmacokinetic properties is illustrated by the HCV NS3 protease inhibitor telaprevir (**210**) which suffers facile racemization at the P1 residue at higher pH and in human plasma [326]. The (R)-diastereomer is the major metabolite of telaprevir in vivo but is 30-fold weaker than the (S)-isomer which contributes to the need for this drug to be dosed on a TID schedule. Installing a deuterium atom at the configurationally labile carbon atom gave a compound that inhibited HCV NS3 protease with a $K_i = 20$ nM, approximately twofold more potent than the telaprevir (**210**), and reduced the propensity for racemization in human plasma: the deuterated compound produced only 10% of the epimer over 1 h of incubation compared to 35% with the protio homologue. Although the stability of the deuterio derivative in rat plasma was lower, this compound exhibited a 13% increased AUC following oral administration to rats when compared to telaprevir (**210**) [326].
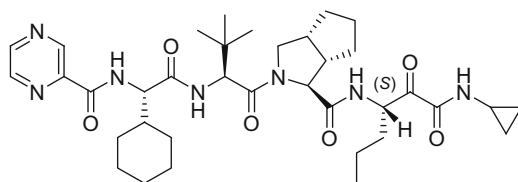
210: telaprevir

Deuteration at metabolically labile sites can redirect metabolism away from a toxic species, as in the case of the HIV-1 non-nucleoside reverse transcriptase inhibitor efavirenz (**211**), or enhance bioactivation of prodrugs, as illustrated by the platelet aggregation inhibitor clopidogrel (**214**) [327, 328]. In rats, the metabolism of efavirenz (**211**) is complex, with the initial step hydroxylation of the aromatic ring to afford a phenol that is a substrate for sulfation. Subsequent hydroxylation at the propargylic position affords a cyclopropylcarbinol that facilitates the addition of glutathione to the alkyne to afford an adduct from which the glutamate is subsequently cleaved, a process blocked by acivicin, to afford **213**, which was determined to be the source of kidney toxicity seen uniquely in rats. As part of this investigation, deuteration at the propargylic position was designed to slow the hydroxylation step and reduce the amount of **213** produced, a successful enterprise since **212** exhibits less severe nephrotoxicity of lower frequency [327].

211: R = H, efavirenz
212: R = D

213
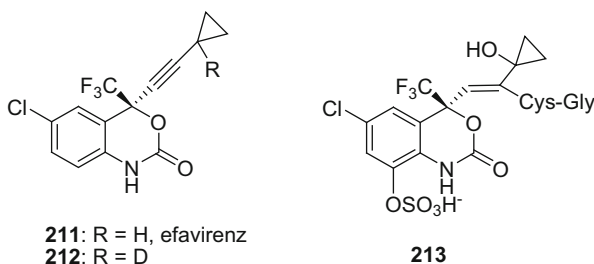
Clopidogrel (**214**) requires metabolic activation in order to express its antithrombotic effects, a process that is initiated by oxidation of the thiophene ring by CYP 450 enzymes to produce the thiolactone **215** [328, 329]. Thiolactone **215** is further oxidized to the acylated sulfoxide **216**, a species that is readily hydrolyzed to the highly reactive sulfenic acid **217**, considered to be the ultimate active principle of the drug that reacts covalently with and blocks the P2Y$_{12}$ receptor that is activated by adenosine diphosphate (ADP) [328, 329]. However, the metabolism of clopidogrel (**214**) into the active species is limited, complicated by alternative pathways that involve cleavage of the ester moiety to afford an inactive acid and CYP-mediated oxidation of the piperidine ring [328]. In an effort to direct metabolic activation toward the thiophene ring and enhance the antiplatelet activity of clopidogrel (**214**) in vivo, the effect of deuteration of the piperidine was examined. The d_6 derivative **218** exhibited similar metabolic

stability to clopidogrel (**214**) in rat and human liver microsomes but improved conversion to the thiolactone **215**, an effect that manifested in vivo in rats as enhanced ex vivo inhibition of ADP-induced platelet aggregation following oral administration of **218** [328].

214   **215**

216   **217**

218
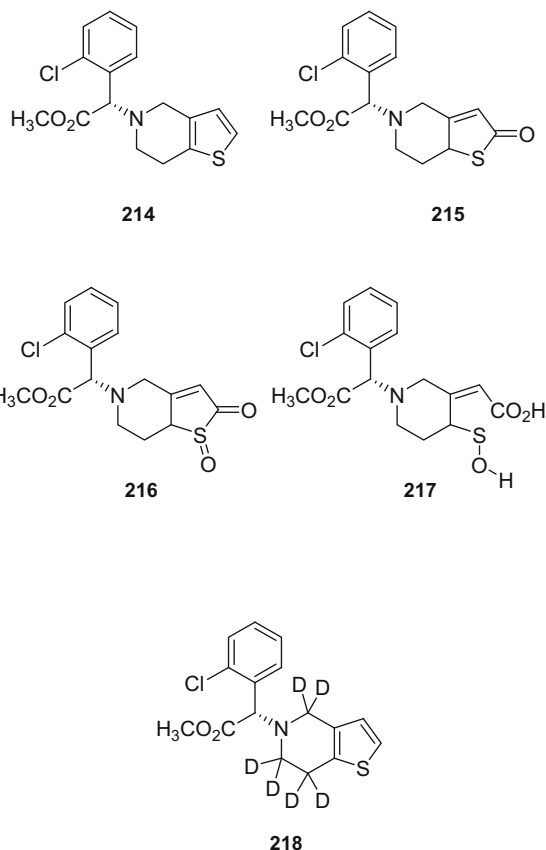
5.2 Substitution of Carbon by Silicon

The replacement of carbon atoms by silicon has attracted attention as a drug design principle that may offer advantage in specific circumstances where the unique properties of silicon can be exploited [330, 331]. A particularly compelling example of the application of deploying silicon in a tactical sense is provided by studies with the antipsychotic agent haloperidol (**219**), a dopamine D_2 antagonist, in which the replacement of the carbinol carbon by silicon was examined as a means of eliminating a problematic metabolic pathway [332, 333]. Haloperidol (**219**) is metabolized, in part, by dehydration of the piperidinol which affords the
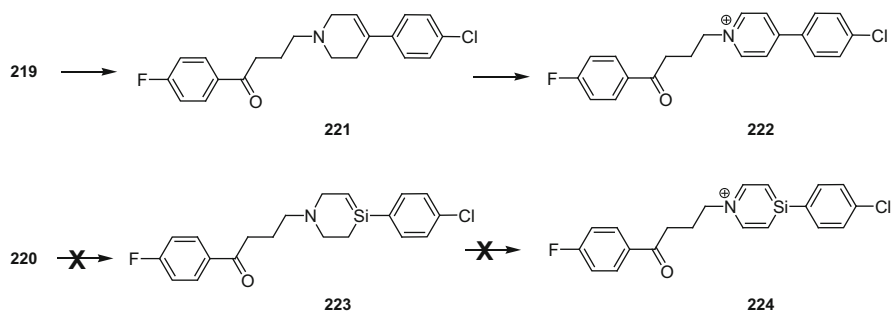
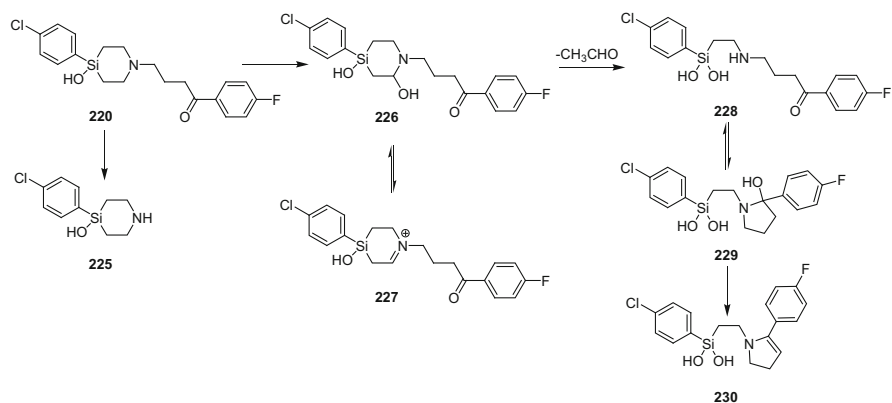**Fig. 33** Metabolism of haloperidol (**219**) to a pyridinium

**Fig. 34** Metabolism of sila-haloperidol (**220**)

tetrahydropiperidine **221**, a precursor to the pyridinium **222** which is neurotoxic, as summarized in Fig. 33. Sila-haloperidol (**220**) exhibits biological properties that are quite similar to the progenitor **219**, but this compound cannot be metabolized by dehydration to **223** and further oxidized to **224** due to the inherent instability of carbon-silicon double bonds [332]. Indeed, the metabolism of sila-haloperidol (**220**) follows more conventional pathways associated with the piperidine moiety in which oxidative metabolism occurs adjacent to the N atom leading to dealkylation of the fluorophenyl-containing side chain to afford **225** and ring hydroxylation that produces **228** and **230** (Fig. 34) [334].
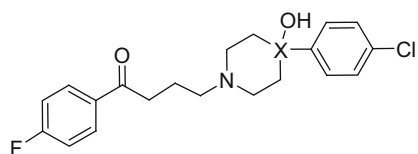
219: X = C - haloperidol
220: X = Si - sila-haloperidol

5.3   Isosteres of Alkyl Groups and Alkylene Moieties

tert-Butyl moieties are susceptible to metabolic oxidation of the methyl groups and this pathway can be a source of poor pharmacokinetic properties. Replacing a single methyl group with CF_3 in the context of the NK1 receptor antagonist **231** and the vanilloid receptor antagonist **233** afforded **232** and **234**, respectively, as compounds that demonstrate enhanced metabolic stability in human liver microsomes compared to the prototypes [335]. A further refinement of this approach identified the trifluoromethylcyclopropyl as a metabolically more stable *tert*-butyl isostere, exemplified in the type II 5α-reductase inhibitor finasteride (**235**) where this substitution provided **236** which exhibits a half-life in human liver microsomes of 114 min, a substantial improvement on the 63-min half-life measured for the progenitor [336]. The potent respiratory syncytial virus fusion inhibitor **237** exhibits poor stability in human liver microsomes with a $t_{1/2}$ of just 4 min, similar to the isopropyl- and cyclobutyl-substituted homologues while the cyclopropyl derivative **238** uniquely addressed this deficiency exhibiting a tenfold improved $t_{1/2}$ of 39 min [337].

231: R = CH_3
232: R = CF_3

233: R = CH_3
234: R = CF_3

235: R = tBu
236: R =

237: R = tBu
238: R = cyclopropyl
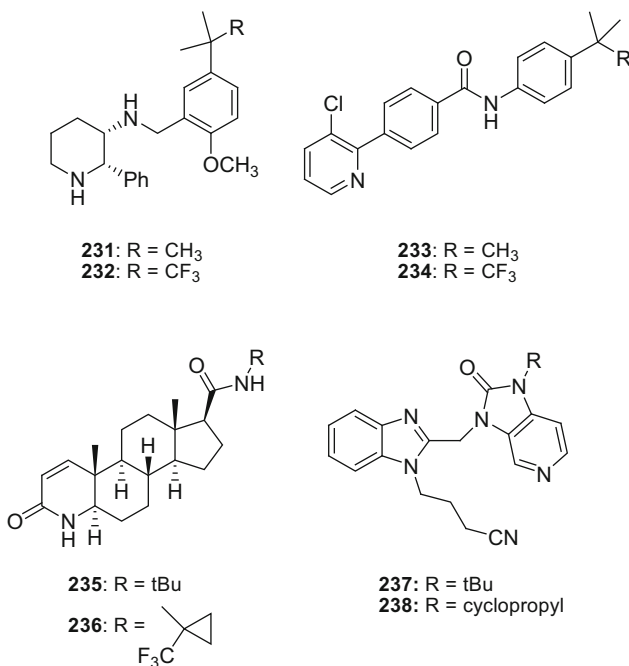
Alkanoic acid derivatives can be metabolized by β-oxidation that can lead to poor pharmacokinetic properties in vivo, a problem encountered with iloprost (**240**), an analogue of prostacyclin (**239**) in which the chemically labile bicyclic enol ether oxygen atom is replaced with a CH_2 isostere [338]. β-Oxidation of fatty acids is a degradative pathway that proceeds mechanistically as depicted in Fig. 35
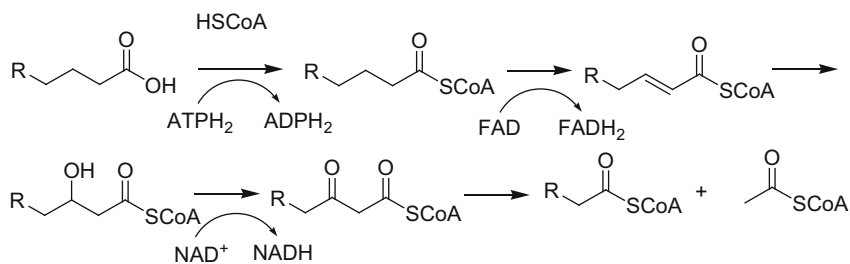
Fig. 35 The β-oxidation pathway for alkanoic carboxylic acids

and involves the initial formation of a CoA ester as a substrate for oxidation to an
α,β-unsaturated CoA ester that is further metabolized to a CoA ester with two fewer
carbon atoms [339, 340]. The essential formation of the α,β-unsaturated CoA ester
provides opportunity for rational structural modification of xenobiotic carboxylic
acids to block this pathway. The introduction of germinal substitution at the α- or
β-positions and replacing the β-carbon with a heteroatom are effective tactics that
prevent olefin formation and it was the latter that was successfully applied to
iloprost (**240**) leading to the design of cicaprost (**241**), a compound with longer-
lasting hypotensive effects in rats following oral administration [338].

239: prostacyclin   **240**: iloprost   **242**: cicaprost
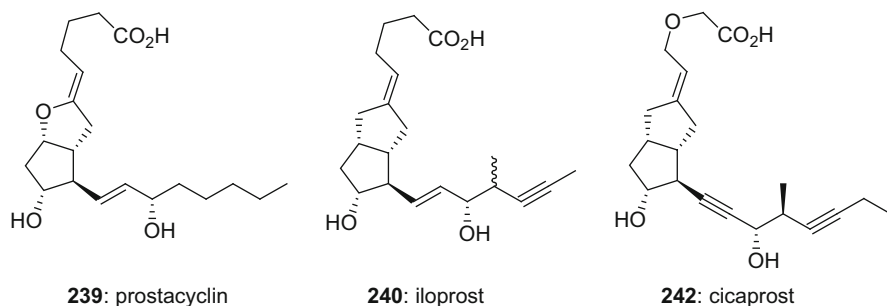
5.4   Isosteres of Carboxylic Acids to Reduce Reactive
Metabolite Formation

Acyl-CoA esters have been shown to be inherently electrophilic acylating agents,
providing further impetus to replace the carboxylic acid moiety with an isostere that
cannot engage in this metabolic pathway [10, 341–345]. Carboxylic acids are also
readily conjugated with glucuronic acid in vivo to afford acyl glucuronides that can
undergo sequential rearrangement of the acyl moiety to the adjacent hydroxyl
substituents on the pyranose ring, a process summarized in Fig. 36. This becomes
problematic when the pyran ring opens to reveal an unnatural aldehyde which can
react with amine moieties of proteins, lysine, for example, to generate an imine
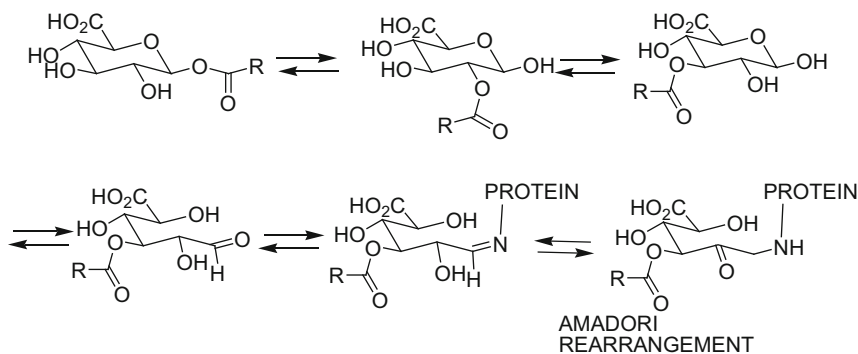
Fig. 36 Rearrangement of acyl glucuronides and exposure of an electrophilic aldehyde moiety

[346, 347]. The imine intermediate can be trapped with $NaB^3[H]_4$, but this functionality can also undergo an Amadori rearrangement to afford the corresponding aminomethyl ketone, resulting in an irreversible protein modification that may create a hapten. Insights have been gleaned into the properties of a carboxylic acid that promote this rearrangement which has been shown to be more facile for electron-deficient acids but slowed by bulky substituents at the α-carbon atom [348–350]. Guidelines have been developed that attempt to equate the propensity of rearrangement with the potential for idiosyncratic toxicity, with the half-life of the acyl glucuronide in buffer that separated safe from problematic acids estimated to be 3.6 h [351]. The level of mechanistic understanding underlying the potential toxicity of carboxylic acid provides a rationale for structural modification to avoid this metabolic pathway by deploying an acid isostere incapable of undergoing rearrangement after glucuronidation. Figure 37 captures the range of acidic functionality that has been examined to address problems across a wide variety of medicinal chemistry campaigns [10, 344, 345, 352, 353].

5.5 Isosteres of Thiols and Alcohols

The thiol moiety is relatively uncommon in drugs and drug candidates due to potential problems arising from metabolic lability or chemical reactivity, with the angiotensin-converting enzyme inhibitor captopril (**243**) and penicillamine (**244**), a chelator used to promote heavy metal excretion, the most prominent exceptions [354, 355].
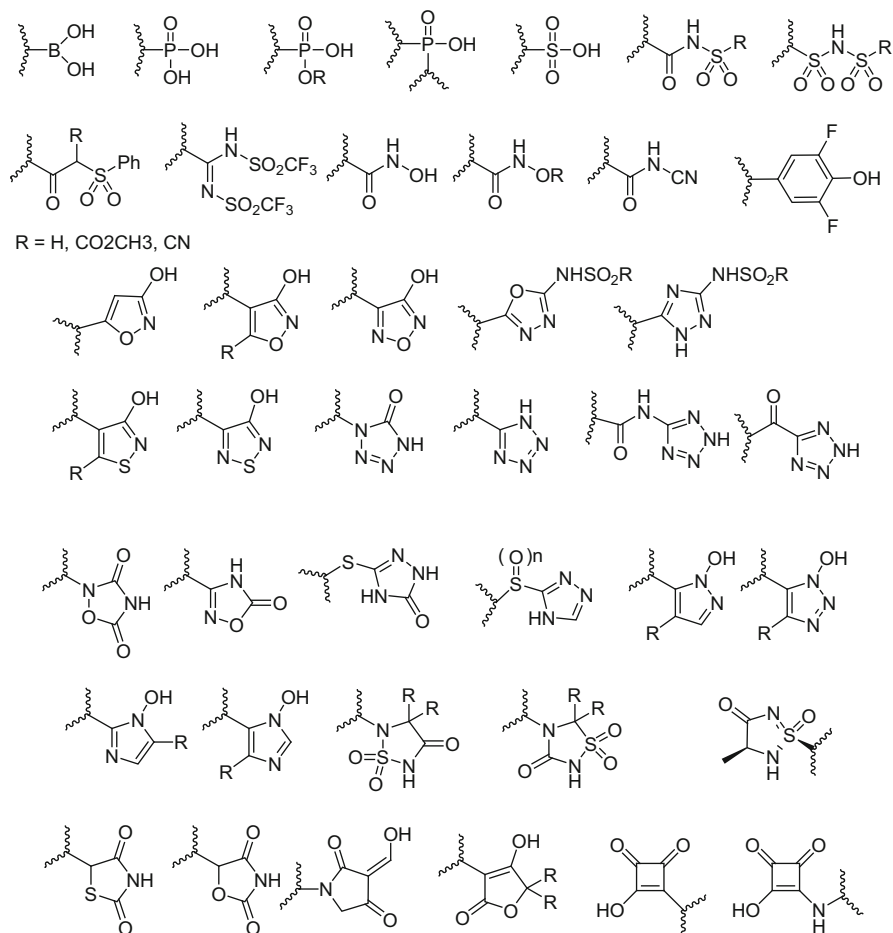

243   **244**
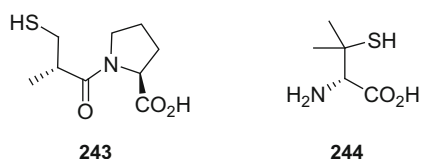
Fig. 37 A sampling of carboxylic acid isosteres

A useful although somewhat underutilized isostere of a thiol is the CHF_2 moiety which has been applied in an elegant fashion to the design of HCV NS3 protease inhibitors where the P1 cysteine moiety of hexapeptide, substrate-based inhibitors, was considered to be a liability [356]. Capitalizing on earlier observations of the intramolecular H-bonding properties of the CHF_2 moiety in **248**, this element was examined as a thiol surrogate in the context of the potent NS3 inhibitor **245** which exhibited a $K_i = 40$ nM [356, 357]. Replacing the P1 SH with CH_3 led to an almost 20-fold erosion of potency and **246** inhibited the enzyme with a $K_i = 700$ nM. However, a CHF_2 terminus at P1 proved to be very effective, fully restoring the potency of **247** to that of the cysteine progenitor, with a $K_i = 30$ nM [356]. This amino acid, referred to as difluoro-Abu, was conceived to be suitable for this context after an insightful analysis of its properties which suggested considerable potential to mimic the cysteine. Mimicry between a SH and a CHF_2 was based on

several observations that included the similarity of the van der Waals surfaces of the two structural elements, 46.7 Å for HCF_2CH_3 compared to 47.1 Å for $HSCH_3$ and electrostatic potential maps, which revealed that the negative potential around the sulfur lone pairs of electrons was similar to that of the two fluorine atoms while there was positive potential around both the SH and CF_2H hydrogen atoms, suggestive of H-bonding capability. Indeed, an X-ray co-crystal structure of a related inhibitor revealed that the CF_2H moiety donated a H-bond to the C=O of Lys_{136} while one of the fluorine atoms was close enough to the C-4-hydrogen atom of Phe_{154} to suggest the presence of a weak C–H to F H-bond [356].
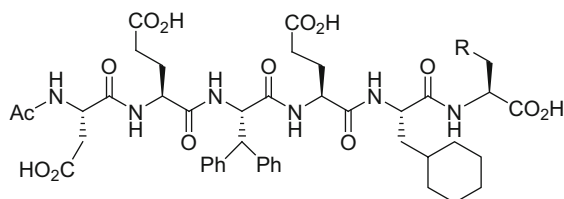
245: R = SH
246: R = CH_3
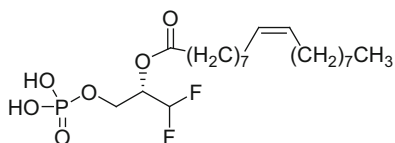**247**: R = CHF_2
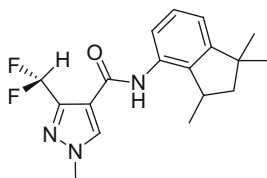
**248**
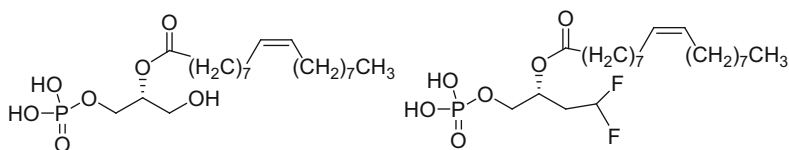
249

250

251

Lysophosphatidic acid (**249**) is a mitogen that has been shown to interact with four G protein-coupled receptors designated LPA 1–4 in addition to being an agonist at the peroxisome proliferator-activated receptor γ (PPARγ), a nuclear hormone receptor, and antagonists offer potential in a range of clinical conditions including the treatment of liver and lung fibrosis, several cancers, and neuropathic pain [358, 359]. The acyl moiety in lysophosphatidic acid (**249**) can migrate to the primary alcohol, stimulating the design of analogues that are functionally incapable of entering this rearrangement pathway [360, 361]. In this context the CHF_2 moieties of **250** and **251** were conceived as potential mimics of the OH moiety of **249** that would be chemically stable. Neither compound was recognized by LPA receptors 1, 2, or 3, with both agonistic and antagonistic properties evaluated, but **250** was found to stimulate luciferase production in CV-1 cells transfected with luciferase under the control of a PPARγ-responsive element [360, 361].

## 5.6   Substitution of C–H by N in Phenyl Rings and C by N in Dihydropyridines

The substitution of the C–H moiety of a phenyl ring by nitrogen is a useful and well-established tactic for mitigating metabolic activation to chemically reactive and potentially toxic species that has found broad-based application and is captured in synoptic form by several illustrative examples [362–367]. The 4,7-dimethoxy indole derivative **252** is a potent HIV-1 attachment inhibitor, but the observation of O-demethylation as a metabolic pathway in HLM gave rise to the concern that the chemically electrophilic quinone **253** could be formed in vivo [362]. The 4,7-dimethoxy-6-azaindole **254** analogue largely preserved the antiviral activity of **252** but, in contrast, would afford the far less electrophilic **255** upon dealkylative metabolism, and it was this compound, designated BMS-488043, that was advanced into proof-of-concept clinical studies [362]. In addition, the mild basicity associated with the azaindole ring of **254** led to improved aqueous solubility.
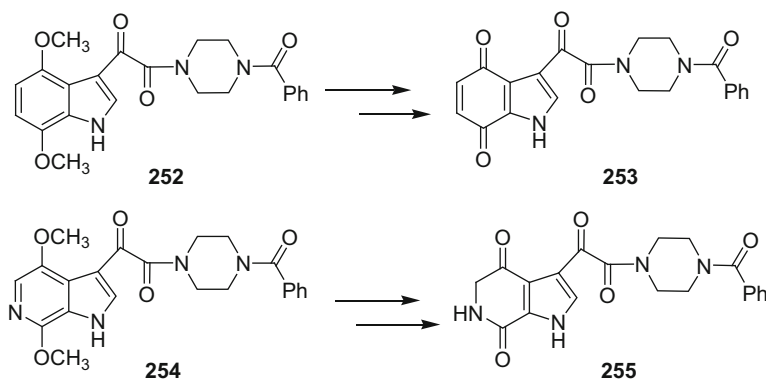
The phenol **256** is a short-acting calcium-sensing receptor antagonist that was examined for its potential as a treatment for osteoporosis [363, 364]. However, CYP 3A4-mediated oxidation of the phenol ring of **256** afforded the catechol **257** which underwent further oxidation to the corresponding *ortho* quinone, a metabolite identified as the source of GSH adducts in both human and rat liver microsomes. The strategic introduction of a nitrogen atom into the phenol ring of **256** to give **258** was anticipated to reduce the rate of metabolic activation of the catechol based on quantum chemistry calculations which indicated that oxidation of the aza-catechol derived from **258** to the corresponding quinone was energetically less favorable than for **257**. This was confirmed experimentally with a 50-fold reduction in the formation of GSH adducts compared to **256** when **258** was incubated in liver microsomes [363, 364].

256: R = H
257: R = OH

258
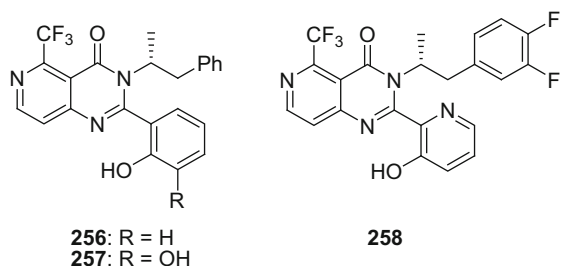
The CRF antagonist **259** was found to undergo metabolic activation following dosing to rats followed by trapping of the reactive species by GSH which afforded adducts amounting to 25% of the dose, with **260** identified as the major species produced in vivo [365–367]. This metabolic process was postulated to proceed via iminoquinone formation, providing a rationale for studying the effect of replacing the electron-rich phenyl ring with a pyridine, an approach that, after additional optimization, led to the identification of **261** as a compound with targeted biological and pharmacokinetic properties. As part of the structural refinement of **259**, the pyrazinone chlorine substituent was replaced with an electron-withdrawing nitrile moiety in order to reduce metabolic activation of the olefin while the methoxy was introduced into the side chain to redirect metabolism to an alternative site that would provide benign products [365–367].

GS = glutathione

259   **260**   **261**
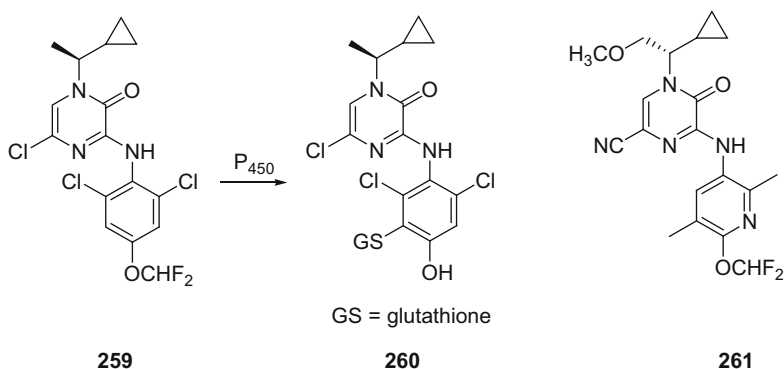
Another example of the beneficial effect of replacing a C atom with N to interfere with metabolism is provided by the bioisosteric replacement of the dihydropyridine (DHP) ring of the Ca^{2+} channel blocker class of smooth muscle relaxant exemplified by nifedipine (**262**, R=*ortho*-NO$_2$) with a pyrimidinone heterocycle [368–371]. The dihydropyridine ring is subject to rapid first-pass oxidative metabolism in vivo to give the corresponding pyridine **263** which is an inactive metabolite. This led to the design of the pyrimidinone **264** as a probe of the idea that this ring system would be an effective substitute of the DHP ring. The design concept was based on the premise that the amide NH of **264** would mimic the H-bond donor properties of the dihydropyridine NH of **262** while the second nitrogen atom of the ring would provide a site for the introduction of the important carbonyl moiety, initially examined in the context of the methoxycarbonyl derivative **264**, as well as providing resistance toward facile oxidation of the heterocycle. However, as an acylated urea derivative, **264** suffered from chemical instability, necessitating replacement by the more robust ureido moiety found in **265**. In this analogue, the orientation of the exocyclic carbonyl group is as depicted in **265**, favored by both dipole–dipole interactions and an intramolecular H-bond that projects the R substituent in a vector compatible with vasodilatory activity [368–371].
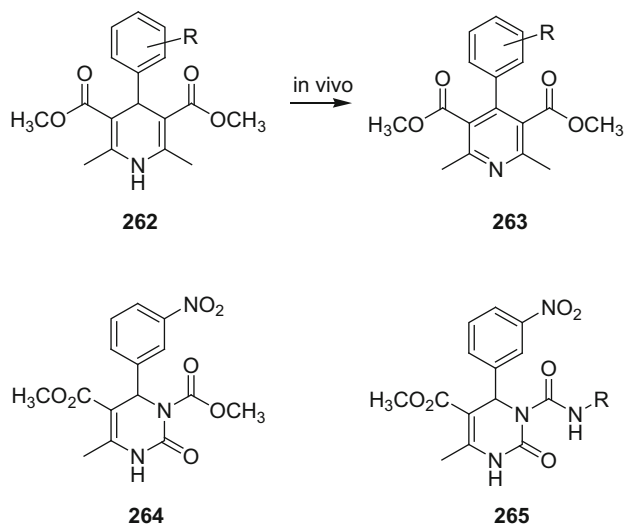

5.7 Isosteres of Heterocycles to Reduce Metabolic Activation

The potent bradykinin B$_1$ antagonist **266**, K_i = 11.8 nM, developed as a potential treatment for pain, was found to produce glutathione adducts when incubated in rat and human liver microsomal preparations, a metabolic pathway also observed

in vivo following oral administration of the drug to rats with GSH adducts derived from the drug identified in bile [372]. The site of GSH adduction was determined to be the pyridine ring, characterized as **268** and hypothesized to arise either from the diiminoquinone **267** or the epoxide **270**, while the pyridine N-oxide **269** was determined to be a minor contributory pathway.
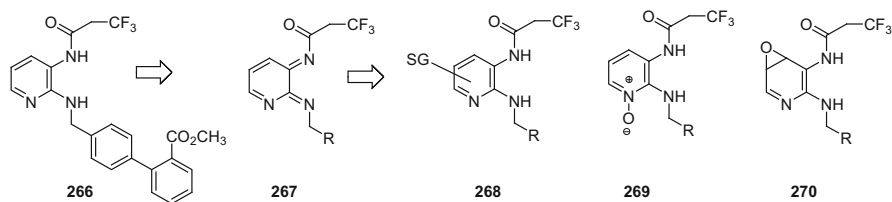
The design of a suitable isostere of the diaminopyridine ring was based upon the assumption that both NHs were of importance while the pyridine N atom was considered to function as a H-bond acceptor, an analysis that anticipated the simple glycine derivative **271** as the most rudimentary mimic. In order to influence the conformation of **271** in a fashion that allowed mimicry of the topology of the *ortho*-disposed substituents of **266**, the gem-dimethyl derivative **272** was prepared, but this compound expressed only modest affinity for the bradykinin B_1 receptor, $K_i = 3.5\ \mu M$ [372]. Further refinement to the cyclopropyl analogue **273** led to a more than 50-fold increase in potency, attributed to conformational constraint due to π–π hyperconjugation between the cyclopropyl C–C bond and the amide C=O that favors the two conformations depicted in Fig. 38, with that represented by A mimicking the topology associated with **266**. In addition, it was noted that the 116° bond angle associate with the cyclopropane ring substituents more closely matches the 120° vectors projected by the pyridine ring.
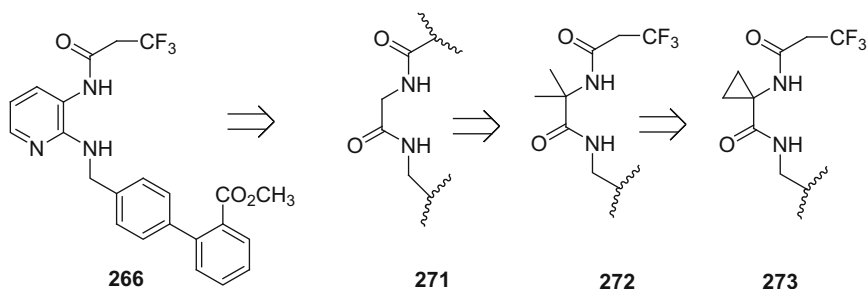
Amide and ester isosteres are susceptible to protease and/or esterase-mediated degradation in vivo while alkyl esters can be degraded by oxidative dealkylation by CYP 450 enzymes to afford the corresponding carboxylic acids, providing an impetus to identify surrogates with resistance to metabolic modification [373–377]. Pioneering studies in this area were focused on the design of heterocycles as ester replacements in the context of benzodiazepine derivatives of general
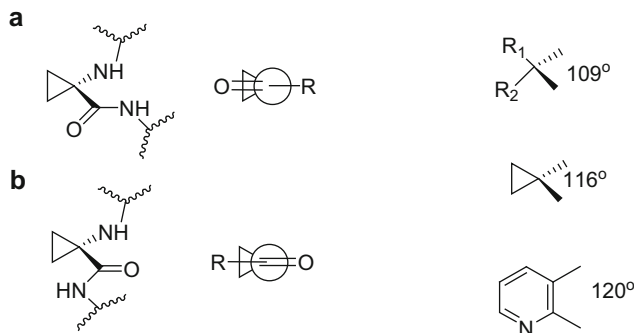
Fig. 38 Conformational topology and bond angles of spirocyclopropyl glycinamides
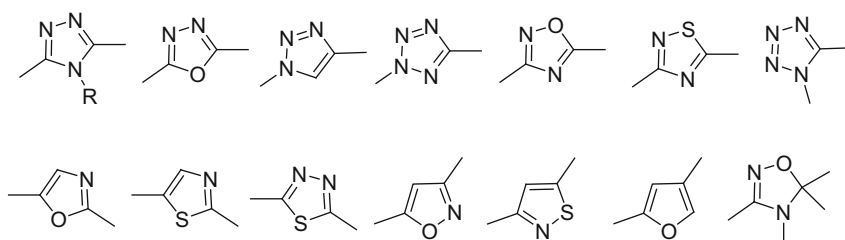
Fig. 39 Azole isosteres of esters and amides

structure **274** and **275** and muscarinic agonists based on the naturally occurring arecoline (**276**) [378–380]. Azole heterocycles are the most common amide and ester replacements that have been explored, captured in synoptic fashion in Fig. 39. This tactic remains an approach of contemporary interest although there can be marked differences between heterocycles in the ability to emulate carbonyl-based functionality dependent on context that may be attributed to subtle effects associated with the underlying electronic properties of a heterocycle that are not always understood [2, 3, 9, 10, 71, 381–387].

<div align="center">

274   **275**   **276**

</div>
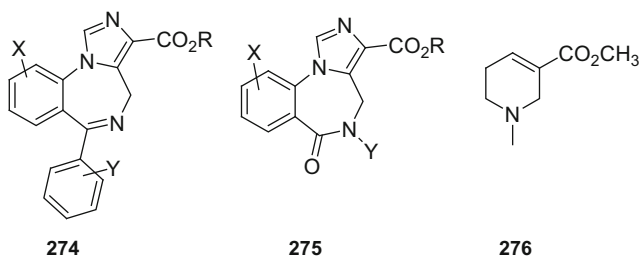
Fig. 40 Intermolecular interactions of nitro groups from the Cambridge Structural Database

5.8 Isosteres of the NO₂ Moiety

The unique properties of the nitro group have often made this a difficult structural motif to emulate by isosteric replacement [388, 389]. For example, the n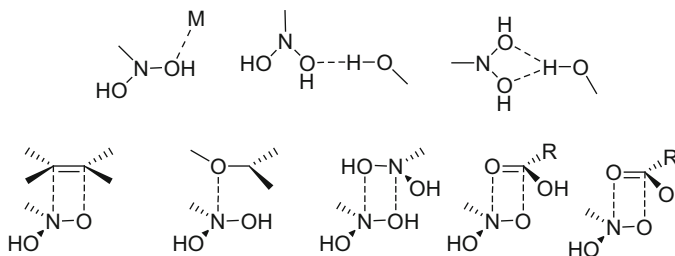itrophenyl moiety of **277**, an inhibitor of the protein–protein interaction between murine double minute 2 (MDM2) and p53, offers the optimal potency from an extensive survey of potential substitutes [389]. These observations were suggested to be a function of a lipophilic and strongly directional interaction between the nitro group and the MDM2 protein that is also dependent on the electron-withdrawing properties. Where the latter effect is of importance for drug–target interactions, several electron-withdrawing functional groups can substitute for the nitro, including pyridine, pyridine N-oxide, sulfonyl, trifluoromethyl, amide, and a carboxylic acid moiety that is considered to be an exact isostere [104, 390–395]. Intermolecular interactions for the nitro moiety as collected from the Cambridge Structural Database of small molecules are summarized in Fig. 40 [387].

277
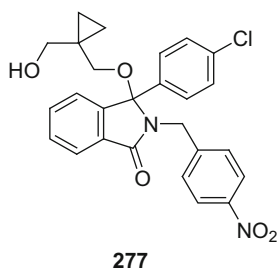
6 Epilogue

The design of bioisosteres is a powerful and effective concept in drug design that has been applied to solve a wide range of problems encountered in drug discovery campaigns. This chapter has focused on the practical utility of applying bioisosteric

substitution to solve contemporary developability challenges that are encountered almost routinely by the practicing medicinal chemistry. The creativity and ingenuity of medicinal chemists is quite clear from this synopsis which captures many insightful and elegant examples of drug design. The most successful and creative designs are frequently based on a careful analysis of not only the chemistry underlying a particular problem but also a deep and detailed understanding of the physicochemical properties of the atoms and structural elements that are selected to address an issue. While medicinal chemists have developed a broad toolbox of useful structural elements to draw upon to emulate a range of commonly encountered functionalities, it is anticipated that the challenges in drug design that lie ahead will continue to provide a stimulus for creativity that will expand the range of bioisosteric replacements.

References

1. Burger A (1991) Isosterism and bioisosterism in drug design. Progress Drug Res 37:288–362
2. Patani GA, LaVoie EJ (1996) Bioisosterism: a rational approach in drug design. Chem Rev 96:3147–3176
3. Thornber CW (1979) Isosterism and molecular modification in drug design. Chem Soc Rev 8:563–580
4. Lipinski CA (1986) Bioisosterism in drug design. Ann Rep Med Chem 21:283–291
5. Wermuth CG, Bourguignon J-J (2003) In: Wermuth CG (ed) The practice of medicinal chemistry, 2nd edn. Chaps 12–16, Academic, London, pp 181–310. ISBN: 0-12-744640-0
6. Olsen PH (2001) The use of bioisosteric groups in lead optimization. Curr Opin Drug Discov Devel 4:471–478
7. Sheridan RP (2002) The most common chemical replacements in drug-like compounds. J Chem Inf Comput Sci 42:103–108
8. Wermuth CG (2005) Similarity in drugs: reflections on analogue design. Drug Discov Today 11:348–354
9. Lima LM, Barriero EJ (2005) Bioisosterism: a useful strategy for molecular modification in drug design. Curr Med Chem 12:23–49
10. Meanwell NA (2011) Synopsis of some recent tactical applications of bioisosteres in drug design. J Med Chem 54:2529–2591
11. Langmuir I (1919) Isomorphism, isosterism and covalence. J Am Chem Soc 41:1543–1559
12. Erlenmeyer H, Berger E (1932) Studies on the significance of structure of antigens for the production and the specificity of antibodies. Biochem Z 252:22–36
13. Erlenmeyer H, Berger E, Leo M (1933) Relationship between the structure of antigens and the specificity of antibodies. Helv Chim Acta 16:733–738
14. Friedman HL (1951) Influence of isosteric replacements upon biological activity. NASNRS 206:295–358
15. Bégué J-P, Bonnet-Delpon D (2008) Bioorganic and medicinal chemistry of fluorine. Wiley, Hoboken
16. Böhm H-J, Banner D, Bendels S, Kansy M, Kuhn B, Müller K, Obst-Sander U, Stahl M (2004) Fluorine in medicinal chemistry. ChemBioChem 5:637–643
17. Müller K, Faeh C, Diederich F (2007) Fluorine in pharmaceuticals: looking beyond intuition. Science 317:1881–1886
18. Hagman WK (2008) The many roles for fluorine in medicinal chemistry. J Med Chem 51:4359–4369

19. Shah P, Westwell AD (2007) The role of fluorine in medicinal chemistry. J Enzyme Inhib Med Chem 22:527–540
20. Purser S, Moore PR, Swallow S, Gouverneur V (2008) Fluorine in medicinal chemistry. Chem Soc Rev 37:320–330
21. O'Hagan D (2008) Understanding organofluorine chemistry. An introduction to the C–F bond. Chem Soc Rev 37:308–319
22. O'Hagan D (2010) Fluorine in health care: organofluorine containing blockbuster drugs. J Fluorine Chem 131:1071–1081
23. Hodgetts KJ, Combs KJ, Elder AM, Harriman GC (2010) The role of fluorine in the discovery and optimization of CNS agents: modulation of drug-like properties. Ann Rep Med Chem 45:429–448
24. Smart BE (2001) Fluorine substituent effects (on bioactivity). J Fluorine Chem 109:3–11
25. Schinazi RF, McMillan A, Cannon D, Mathis R, Loyd RM, Peck A, Sommadossi J-P, St. Clair M, Wilson J, Furman PA, Painter G, Choi W-B, Liotta DC (1992) Selective inhibition of human immunodeficiency viruses by racemates and enantiomers of cis-5-fluoro-1-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine. Antimicrob Agents Chemother 36:2423–2431
26. Wang LH, Begley J, St. Claire RL, III, Harris J, Wakeford C, Rousseau FS (2004) Pharmacokinetic and pharmacodynamic characteristics of emtricitabine support its once daily dosing for the treatment of HIV infection. AIDS Res Human Retroviruses 20:1173–1182
27. Feng JY, Shi J, Schinazi RF, Anderson KS (1999) Mechanistic studies show that (−) FTC-TP is a better inhibitor of HIV-1 reverse transcriptase than 3TC-TP. FASEB J 13:1511–1517
28. Domagala JM, Hanna LD, Heifetz CL, Hutt MP, Mich TF, Sanchez JP, Solomon M (1986) New structure–activity relationships of the quinolone antibacterials using the target enzyme. The development and application of a DNA gyrase assay. J Med Chem 29:394–404
29. Barbachyn MR, Ford CW (2003) Oxazolidinone structure–activity relationships leading to linezolid. Angew Chem Int Ed 42:2010–2023
30. Meanwell NA, Wallace OB, Fang H, Wang H, Deshpande M, Wang T, Yin Z, Zhang Z, Pearce BC, James J, Yeung K-S, Qiu Z, Wright JJK, Yang Z, Zadjura L, Tweedie DL, Yeola S, Zhao F, Ranadive S, Robinson BA, Gong Y-F, Wang H-GH, Blair WS, Shi P-Y, Colonno RJ, P-f L (2009) Inhibitors of HIV-1 attachment. Part 2: an initial survey of indole substitution patterns. Bioorg Med Chem Lett 19:1977–1981
31. Kitas EA, Galley G, Jakob-Roetne R, Flohr A, Wostl W, Mauser H, Alker AM, Czech C, Ozmen L, David-Pierson P, Reinhardt D, Jacobsen H (2008) Substituted 2-oxo-azepane derivatives are potent, orally active γ-secretase inhibitors. Bioorg Med Chem Lett 18:304–308
32. Olsen JA, Banner DW, Seiler P, Wagner B, Tschopp T, Obst-Sander U, Kansy M, Müller K, Diederich F (2004) Fluorine interactions at the thrombin active site: protein backbone fragments H-Cα-C=O comprise a favorable C-F environment and interactions of C-F with electrophiles. ChemBioChem 5:666–675
33. Olsen JA, Banner DW, Seiler P, Sander UO, D'Arcy A, Stihle M, Muller K, Diederich F (2003) A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: evidence for C-F•••C=O interactions. Angew Chem Int Ed 42:2507–2511
34. Carini DJ, Christ DD, Duncia JV, Pierce ME (1989) The discovery and development of angiotensin II antagonists. Pharmaceutical Biotechnol 11:29–56
35. Duncia JV, Carini DJ, Chiu AT, Johnson AL, Price WA, Wong PC, Wexler RR, Timmermans PBMWM (1992) The discovery of DuP 753, a potent, orally active nonpeptide angiotensin II receptor antagonist. Med Res Rev 12:149–191
36. Duncia JV, Chiu AT, Carini DJ, Gregory GB, Johnson AL, Price WA, Wells GJ, Wong PC, Calabrese JC, Timmermans PBMWM (1990) The discovery of potent nonpeptide angiotensin II receptor antagonists: a new class of potent antihypertensives. J Med Chem 33:1312–1329

37. Carini DJ, Duncia JV, Johnson AL, Chiu AT, Price WA, Wong PC, Timmermans PBMWM (1990) Nonpeptide angiotensin II receptor antagonists: *N*-[(benzyloxy)benzyl]imidazoles and related compounds as potent antihypertensives. J Med Chem 33:1330–1336

38. Carini DJ, Duncia JV, Aldrich PE, Chiu AT, Johnson AL, Pierce ME, Price WA, Santella JB III, Wells GJ, Wexler RR, Wong PC, Yoo S-E, Timmermans PBMWM (1991) Nonpeptide angiotensin II receptor antagonists: the discovery of a series of *N*-(biphenylylmethyl)imidazoles as potent, orally active antihypertensives. J Med Chem 34:2525–2547

39. Herr RJ (2002) 5-Substituted-1*H*-tetrazoles as carboxylic acid isosteres: medicinal chemistry and synthetic methods. Bioorg Med Chem 10:3379–3393

40. Allen FH, Groom CR, Liebeschuetz JW, Bardwell DA, Olsson TSG, Wood PA (2012) The hydrogen bond environments of 1*H*-tetrazole and tetrazolate rings: the structural basis for tetrazole−carboxylic acid bioisosterism. J Chem Inf Model 52:857–866

41. Noda K, Saad Y, Kinoshita A, Boyle TP, Graham RM, Husain A, Karnik SS (1995) Tetrazole and carboxylate groups of angiotensin receptor antagonists bind to the same subsite by different mechanisms. J Biol Chem 270:2284–2289

42. Naylor EM, Chakravarty PK, Costello CA, Chang RS, Chen T-B, Faust KA, Lotti VJ, Kivlighn SD, Zingaro GJ, Siegl PKS, Wong PC, Carini DJ, Wexler RR, Patchett AA, Greenlee WJ (1993) Potent imidazole angiotensin II antagonists: acyl sulfonamides and acyl sulfamides as tetrazole replacements. Bioorg Med Chem Lett 3:69–74

43. Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, Bonneau P, Boes M, Cameron DR, Cartier M, Cordingley MG, Faucher A-M, Goudreau N, Kawai SH, Kukolj G, Lagace L, LaPlante SR, Narjes H, Poupart M-A, Rancourt J, Sentjens R,E, St. George R, Simoneau B, Steinmann G, Thibeault D, Tsantrizos YS, Weldon SM, Yong C-L, Llinas-Brunet M (2003) An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. Nature 426:186–189

44. Llinas-Brunet M, Bailey MD, Bolger G, Brochu C, Faucher A-M, Ferland JM, Garneau M, Ghiro E, Gorys V, Grand-Maitre C, Halmos T, Lapeyre-Paquette N, Liard F, Poirier M, Rheaume M, Tsantrizos YS, Lamarre D (2004) Structure-activity study on a novel series of macrocyclic inhibitors of the hepatitis C virus NS3 protease leading to the discovery of BILN 2061. J Med Chem 47:1605–1608

45. Stoltz JH, Stern JO, Huang Q, Seidler RW, Pack FD, Knight BL (2011) A twenty-eight-day mechanistic time course study in the rhesus monkey with hepatitis C virus protease inhibitor BILN 2061. Toxicol Pathol 39:496–501

46. Campbell JA, Good A (2002) Preparation of tripeptides as hepatitis C inhibitors. PCT Int. Appl. WO 2002/060,926

47. Scola PM, Sun L-Q, Chen J, Wang AX, Sit S-Y, Chen Y, D'Andrea SV, Zheng Z, Sin N, Venables BL, Cocuzza A, Bilder D, Carini D, Johnson B, Good AC, Rajamani R, Klei HE, Friborg J, Barry D, Levine S, Chen C, Sheaffer A, Hernandez D, Falk P, Yu F, Zhai G, Knipe JO, Mosure K, Shu Y-Z, Phillip T, Arora VK, Loy J, Adams S, Schartman R, Browning M, Levesque PC, Li D, Zhu JL, Sun H, Pilcher G, Bounous D, Lange RW, Pasquinelli C, Eley T, Colonno R, Meanwell NA, McPhee F (2010) Discovery of BMS-650032, an NS3 protease inhibitor for the treatment of Hepatitis C. In: 239th ACS national meeting and exposition, March 21–25, 2010, San Francisco. Abstract: MEDI-38

48. Liverton NJ, Holloway MK, McCauley JA, Rudd MT, Butcher JW, Carroll SS, DiMuzio J, Fandozzi C, Gilbert KF, Mao S-S, McIntyre CJ, Nguyen KT, Romano JJ, Stahlhut M, Wan B-L, Olsen DB, Vacca JP (2008) Molecular modeling based approach to potent P2-P4 macrocyclic inhibitors of hepatitis C NS3/4A protease. J Am Chem Soc 130:4607–4609

49. Cummings MD, Lindberg J, Lin T-I, de Kock H, Lenz O, Lilja E, Fellander S, Baraznenok V, Nystrom S, Nilsson M, Vrang L, Edlund M, Rosenquist A, Samuelsson B, Raboisson P, Simmen K (2010) Induced-fit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target. Angew Chemie Int Ed 49:1652–1655

50. Wendt MD (2008) Discovery of ABT-263, a Bcl-family protein inhibitor: observations on targeting a large protein–protein interaction. Exp Opin Drug Discov 3:1123–1143

51. Petros AM, Dinges J, Augeri DA, Baumeister SA, Betebenner DA, Bures MG, Elmore SW, Hajduk PJ, Joseph MK, Landis SK, Nettesheim DG, Rosenberg SH, Shen W, Thomas S, Wang X, Zanze I, Zhang H, Fesik SW (2006) Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. J Med Chem 49:656–663
52. Park C-M, Bruncko M, Adickes J, Bauch J, Ding H, Kunzer A, Marsh KC, Nimmer P, Shoemaker AR, Song X, Tahir SK, Tse C, Wang X, Wendt MD, Yang X, Zhang H, Fesik SW, Rosenberg SH, Elmore SW (2008) Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. J Med Chem 51:6902–6915
53. Laurence C, Berthelot M (2000) Observations on the strength of hydrogen bonding. Perspect Drug Discov Des 18:39–60
54. Laurence C, Brameld KA, Graton J, Le Questel J-Y, Renault E (2009) The pK$_{BHX}$ database: toward a better understanding of hydrogen-bond basicity for medicinal chemists. J Med Chem 52:4073–4086
55. Abraham MH, Duce PP, Prior DV (1989) Hydrogen bonding. Part 9. Solute proton donor and proton acceptor scales for use in drug design. J Chem Soc Perkin Trans 2 1355–1375
56. Graton J, Berthelot M, Gal J-F, Laurence C, Lebreton J, Le Questel J-Y, Maria P-C, Robins R (2003) The nicotinic pharmacophore: thermodynamics of the hydrogen-bonding complexation of nicotine, nornicotine, and models. J Org Chem 68:8208–8221
57. Arnaud V, Le Questel J-Y, Mathé-Allainmat M, Lebreton J, Berthelot M (2004) Multiple hydrogen-bond accepting capacities of polybasic molecules: the case of cotinine. J Phys Chem A 108:10740–10748
58. Locati A, Berthelot M, Evain M, Lebreton J, Le Questel J-Y, Mathé-Allainmat M, Planchat A, Renault E, Graton J (2007) The exceptional hydrogen-bond properties of neutral and protonated lobeline. J Phys Chem A 111:6397–6405
59. Arnaud V, Berthelot M, Evain M, Graton J, Le Questel J-Y (2007) Hydrogen-bond interactions of nicotine and acetylcholine salts: a combined crystallographic, spectroscopic, thermodynamic and theoretical study. Chem Eur J 13:1499–1510
60. Arnaud V, Berthelot M, Le Questel J-Y (2005) Hydrogen-bond accepting strength of protonated nicotine. J Phys Chem A 109:3767–3770
61. Arnaud V, Berthelot M, Felpin F-X, Lebreton J, Le Questel J-Y, Graton J (2009) Hydrogen-bond accepting strength of five-membered N-heterocycles: the case of substituted phenylpyrrolines and myosmines. Eur J Org Chem 4939–4948
62. Graton J, Berthelot M, Gal J-F, Girard S, Laurence C, Lebreton J, Le Questel J-Y, Maria P-C, Naus P (2002) Site of protonation of nicotine and nornicotine in the gas phase: pyridine or pyrrolidine nitrogen? J Am Chem Soc 124:10552–10562
63. Wermuth CG (2012) Are pyridazines privileged structures? Med Chem Commun 2:935–941
64. Taft RW, Anvia F, Taagpera M, Catalán J, Elguero J (1986) Electrostatic proximity effects in the relative basicities and acidities of pyrazole, imidazole, pyridazine, and pyrimidine. J Am Chem Soc 108:3237–3239
65. Meanwell NA, Romine JL, Seiler SM (1994) Non-prostanoid prostacyclin mimetics. Drugs Future 19:361–385
66. Meanwell NA, Rosenfeld MJ, Trehan AK, Wright JJK, Brassard CL, Buchanan JO, Federici ME, Fleming JS, Gamberdella M, Zavoico GB, Seiler SM (1992) Nonprostanoid prostacyclin mimetics. 2. 4,5-Diphenyloxazole derivatives. J Med Chem 35:3483–3497
67. Meanwell NA, Rosenfeld MJ, Wright JJK, Brassard CL, Buchanan JO, Federici ME, Fleming JS, Gamberdella M, Hartl KS, Zavoico GB, Seiler SM (1993) Nonprostanoid prostacyclin mimetics. 4. Derivatives of 2-[3-[2-(4,5-diphenyl-2-oxazolyl)ethyl]phenoxy]acetic acid substituted α- to the oxazole ring. J Med Chem 36:3871–3883
68. Meanwell NA, Romine JL, Rosenfeld MJ, Martin SW, Trehan AK, Wright JJK, Malley MF, Gougoutas JZ, Brassard CL, Buchanan JO, Federici ME, Fleming JS, Gamberdella M, Hartl KS, Zavoico GB, Seiler SM (1992) Nonprostanoid prostacyclin mimetics. 5. Structure–activity relationships associated with [3-[4-(4,5-diphenyl-2-oxazolyl)-5-oxazolyl]phenoxy] acetic acid. J Med Chem 36:3884–3903

69. Böhm H-J, Klebe G, Brode S, Hesse U (1996) Oxygen and nitrogen in competitive situations: which is the hydrogen-bond acceptor? Chem Eur J 2:1509–1513
70. Nobeli I, Price SL, Lommerse JPM, Taylor R (1997) Hydrogen bonding properties of oxygen and nitrogen acceptors in aromatic heterocycles. J Comput Chem 18:2060–2074
71. Kumar DV, Rai R, Brameld KA, Riggs J, Somoza JR, Rajagopalan R, Janc JW, Xia YM, Ton TL, Hu H, Lehoux I, Ho JD, Young WB, Hart B, Green MJ (2012) 3-Heterocyclyl quinolone inhibitors of the HCV NS5B polymerase. Bioorg Med Chem Lett 22:300–304
72. Boström J, Hogner A, Llinàs A, Wellner E, Plowright AT (2012) Oxadiazoles in medicinal chemistry. J Med Chem 55:1817–1830
73. Jones ED, Vandegraaff N, Le G, Choi N, Issa W, MacFarlane K, Thienthong N, Winfield LJ, Coates JAV, Lu L, Li X, Feng X, Yu C, Rhodes DI, Deadman JJ (2010) Design of a series of bicyclic HIV-1 integrase inhibitors. Part 1: selection of the scaffold. Bioorg Med Chem Lett 20:5913–5917
74. Le G, Vandegraaff N, Rhodes DI, Jones ED, Coates JAV, Thienthong N, Winfield LJ, Lu L, Li X, Yu C, Feng X, Deadman JJ (2010) Design of a series of bicyclic HIV-1 integrase inhibitors. Part 2: azoles: effective metal chelators. Bioorg Med Chem Lett 20:5909–5912
75. Le G, Vandegraaff N, Rhodes DI, Jones ED, Coates JAV, Lu L, Li X, Yu C, Feng X, Deadman JJ (2010) Discovery of potent HIV integrase inhibitors active against raltegravir resistant viruses. Bioorg Med Chem Lett 20:5013–5018
76. Johns BA, Weatherhead JG, Allen SH, Thompson JB, Garvey EP, Foster SA, Jeffrey JL, Miller WH (2009) The use of oxadiazole and triazole substituted naphthyridines as HIV-1 integrase inhibitors. Part 1: establishing the pharmacophore. Bioorg Med Chem Lett 19:1802–1806
77. Johns BA, Weatherhead JG, Allen SH, Thompson JB, Garvey EP, Foster SA, Jeffrey JL, Miller WH (2009) 1,3,4-Oxadiazole substituted naphthyridines as HIV-1 integrase inhibitors. Part 2: SAR of the C5 position. Bioorg Med Chem Lett 19:1807–1810
78. Kalgutkar AS, Jones R, Sawant A (2010) Sulfonamide as an essential functional group in drug design. In: Smith DA (ed) Metabolism, pharmacokinetics and toxicity of functional groups. RSC drug discovery series 1. Chap 5, Royal Society of Chemistry, London, pp 210–274
79. Hedstrom L (2002) Serine protease mechanism and specificity. Chem Rev 102:4501–4523
80. Long JZ, Cravatt BF (2011) The metabolic serine hydrolases and their functions in mammalian physiology and disease. Chem Rev 111:6022–6063
81. Otrubova K, Boger DL (2012) α-Ketoheterocycle-based inhibitors of fatty acid amide hydrolase (FAAH). ACS Chem Neurosci 3:340–348
82. Bradamante S, Pagani GA (1996) Benzyl and heteroarylmethyl carbanions: structure and substituent effects. Adv Carbanion Chem 2:189–263
83. Abbotto A, Bradamante S, Pagani GA (1993) Charge mapping in carbanions. Weak charge demand of the cyano group as assessed from a ^{13}C-NMR study of carbanions of α-activated acetonitriles and phenylacetonitriles: breakdown of a myth. J Org Chem 58:449–455
84. Abbotto A, Bradamante S, Pagani GA (1996) Diheteroarylmethanes. 5. E-Z isomerism of carbanions substituted by 1,3-azoles: ^{13}C and ^{15}N ð-charge/shift relationships as source for mapping charge and ranking the electron-withdrawing power of heterocycles. J Org Chem 61:1761–1769
85. Abbotto A, Bradamante S, Facchetti A, Pagani GA (1999) Diheteroarylmethanes. 8. Mapping charge and electron-withdrawing power of the 1,2,4-triazol-5-yl substituent. J Org Chem 64:6756–6763
86. Imperiali B, Abeles RH (1986) Inhibition of serine proteases by peptidyl fluoromethyl ketones. Biochemistry 25:3760–3767
87. Imperiali B, Abeles RH (1987) Extended binding inhibitors of chymotrypsin that interact with leaving group subsites S1′–S3′. Biochemistry 26:4474–4477
88. Edwards PD, Meyer EF Jr, Vijayalakshmi J, Tuthill PA, Andisik DA, Gomes B, Strimplerg A (1992) Elastase inhibitors, the peptidyl α-ketobenzoxazoles, and the X-ray crystal structure of

the covalent complex between porcine pancreatic elastase and Ac-Ala-Pro-Val-2-benzoxazole. J Am Chem Soc 114:1854–1863

89. Edwards PD, Wolanin DJ, Andisik DA, David MW (1995) Peptidyl α-ketoheterocyclic inhibitors of human neutrophil elastase. 2. Effect of varying the heterocyclic ring on in vitro potency. J Med Chem 38:76–85

90. Edwards P, Zottola MA, Davis M, Williams J, Tuthill PA (1995) Peptidyl α-ketoheterocyclic inhibitors of human neutrophil elastase. 3. In vitro and in vivo potency of a series of peptidyl α-ketobenzoxazoles. J Med Chem 38:3972–3982

91. Ohmoto K, Yamamoto T, Horiuchi T, Imanishi H, Odagaki Y, Kawabata K, Sekioka T, Hirota Y, Matsuoka S, Nakai H, Toda M, Cheronis JC, Spruce LW, Gyorkos A, Wieczorek M (2000) Design and synthesis of new orally active nonpeptidic inhibitors of human neutrophil elastase. J Med Chem 43:4927–4929

92. Ohmoto K, Yamamoto T, Okuma M, Horiuchi T, Imanishi H, Odagaki Y, Kawabata K, Sekioka T, Hirota Y, Matsuoka S, Nakai H, Toda M, Cheronis JC, Spruce LW, Gyorkos A, Wieczorek M (2001) Development of orally active nonpeptidic inhibitors of human neutrophil elastase. J Med Chem 44:1268–1285

93. Costanzo MJ, Almond HR Jr, Hecker LR, Schott MR, Yabut SC, Zhang H-C, Andrade-Gordon P, Corcoran TW, Giardino EC, Kauffman JA, Lewis JM, de Garavilla L, Haertlein BJ, Maryanoff BE (2005) In-depth study of tripeptide-based α-ketoheterocycles as inhibitors of thrombin. Effective utilization of the S1' subsite and its implications to structure-based drug design. J Med Chem 48:1984–2008

94. Maryanoff BE, Costanzo MJ (2008) Inhibitors of proteases and amide hydrolases that employ α-ketoheterocycles as a key enabling functionality. Bioorg Med Chem 16:1562–1595

95. Romero FA, Hwang I, Boger DL (2006) Delineation of a fundamental α-ketoheterocycle substituent effect for use in the design of enzyme inhibitors. J Am Chem Soc 128:14004–14005

96. Mileni M, Garfunkle J, DeMartino JK, Cravatt BF, Boger DL, Stevens RC (2009) Binding and inactivation mechanism of a humanized fatty acid amide hydrolase by α-ketoheterocycle inhibitors revealed from cocrystal structures. J Am Chem Soc 131:10497–10506

97. Boger DL, Miyauchi H, Du W, Hardouin C, Fecik RA, Cheng H, Hwang I, Hedrick MP, Leung D, Acevedo O, Guimarães CRW, Jorgensen WL, Cravatt BF (2005) Discovery of a potent, selective, and efficacious class of reversible α-ketoheterocycle inhibitors of fatty acid amide hydrolase effective as analgesics. J Med Chem 48:1849–1856

98. Hardouin C, Kelso MJ, Romero FA, Rayl TJ, Leung D, Hwang I, Cravatt BF, Boger DL (2007) Structure–activity relationships of α-ketooxazole inhibitors of fatty acid amide hydrolase. J Med Chem 50:3359–3368

99. DeMartino JK, Garfunkle J, Hochstatter DG, Cravatt BF, Boger DL (2008) Exploration of a fundamental substituent effect of α-ketoheterocycle enzyme inhibitors: potent and selective inhibitors of fatty acid amide hydrolase. Bioorg Med Chem Lett 18:5842–5846

100. Garfunkle J, Ezzili C, Ray TJ, Hochstatter DG, Hwang I, Boger DL (2008) Optimization of the central heterocycle of α-ketoheterocycle inhibitors of fatty acid amide hydrolase. J Med Chem 51:4392–4403

101. Mileni M, Garfunkle J, Ezzili C, Kimball FS, Cravatt BF, Stevens RC, Boger DL (2009) X-ray crystallographic analysis of α-ketoheterocycle inhibitors bound to a humanized variant of fatty acid amide hydrolase. J Med Chem 53:230–240

102. Mileni M, Garfunkle J, Ezzili C, Cravatt BF, Stevens RC, Boger DL (2011) Fluoride-mediated capture of a noncovalent bound state of a reversible covalent enzyme inhibitor: X-ray crystallographic analysis of an exceptionally potent α-ketoheterocycle inhibitor of fatty acid amide hydrolase. J Am Chem Soc 133:4092–4100

103. Otrubova K, Ezzili C, Boger DL (2011) The discovery and development of inhibitors of fatty acid amide hydrolase (FAAH). Bioorg Med Chem Lett 21:4674–4685

104. Hansch C, Leo A, Unger SH, Kim KH, Nikaitani D, Lien EJ (1973) Aromatic substituent constants for structure-activity correlations. J Med Chem 16:1207–1216

105. Cai J, Fradera X, van Zeeland M, Dempster M, Cameron KS, Bennett DJ, Robinson J, Popplestone L, Baugh M, Westwood P, Bruin J, Hamilton W, Kinghorn E, Long C, Uitdehaag JCM (2010) 4-(3-Trifluoromethylphenyl)-pyrimidine-2-carbonitrile as cathepsin S inhibitors: N3, not N1 is critically important. Bioorg Med Chem Lett 20:4507–4510

106. Salonen LM, Holland MC, Kaib PSJ, Haap W, Benz J, Mary J-L, Kuster O, Schweizer WB, Banner DW, Diederich F (2012) Molecular recognition at the active site of factor Xa: cation–π interactions, stacking on planar peptide surfaces, and replacement of structural water. Chem Eur J 18:213–222

107. Harder M, Kuhn B, Diederich F (2013) Efficient stacking on protein amide fragments. ChemMedChem 8:397–404

108. Michel J, Tirado-Rives J, Jorgensen WL (2009) Energetics of displacing water molecules from protein binding sites: consequences for ligand optimization. J Am Chem Soc 131:15403–15411

109. Lam PYS, Jadhav PK, Eyermann CJ, Hodge CN, Ru Y, Bacheler LT, Meek JL, Otto MJ, Rayner MM, Wong YN, Chang C-H, Weber PC, Jackson DA, Sharpe TR, Erikson-Viitanen S (1994) Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science 263:380–384

110. Lam PYS, Ru Y, Jadhav PK, Aldrich PE, DeLucca GV, Eyermann CJ, Chang C-H, Emmett G, Holler ER, Daneker WF, Li L, Confalone PN, McHugh RJ, Han Q, Li R, Markwalder JA, Seitz SP, Sharpe TR, Bacheler LT, Rayner MM, Klabe RM, Shum L, Winslow DL, Kornhauser DM, Jackson DA, Erickson-Viitanen S, Hodge CN (1996) Cyclic HIV protease inhibitors: synthesis, conformational analysis, P2/P2' structure–activity relationship, and molecular recognition of cyclic ureas. J Med Chem 39:3514–3525

111. De Lucca GV, Erickson-Viitanen S, Lam PYS (1997) Cyclic HIV protease inhibitors capable of displacing the active site structural water molecule. Drug Discov Today 2:6–18

112. Nalam MNL, Peeters A, Jonckers TMH, Dierynck I, Schiffer CA (2007) Crystal structure of lysine sulfonamide inhibitor reveals the displacement of the conserved flap water molecule in human immunodeficiency virus type 1 protease. J Virol 81:9512–9518

113. Canan Koch SS, Thoresen LH, Tikhe JG, Maegley KA, Almassy RJ, Li J, Yu X-H, Zook SE, Kumpf RA, Zhang C, Boritzki TJ, Mansour RN, Zhang KE, Ekker A, Calabrese CR, Curtin NJ, Kyle S, Thomas HD, Wang L-Z, Calvert AH, Golding BT, Griffin RJ, Newell DR, Webber SE, Hostomsky Z (2002) Novel tricyclic poly(ADP-ribose) polymerase-1 inhibitors with potent anticancer chemopotentiating activity: design, synthesis, and X-ray cocrystal structure. J Med Chem 45:4961–4974

114. Tikhe JG, Webber SE, Hostomsky Z, Maegley KA, Ekkers A, Li J, Yu X-H, Almassy RJ, Kumpf RA, Boritzki TJ, Zhang C, Calabrese CR, Curtin NJ, Kyle S, Thomas HD, Wang L-Z, Calvert AH, Golding BT, Griffin RJ, Newell DR (2004) Design, synthesis, and evaluation of 3,4-dihydro-2*H*-[1,4]diazepino[6,7,1-*hi*]indol-1-ones as inhibitors of poly(ADP-Ribose) polymerase. J Med Chem 47:5467–5481

115. García-Sosa AT, Firth-Clark S, Mancera RL (2005) Including tightly-bound water molecules in de novo drug design. Exemplification through in silico generation of poly(ADP-ribose) polymerase inhibitors. J Chem Inf Model 45:624–633

116. Chen JM, Xu SL, Wawrzak Z, Basarab GS, Jordan DB (1998) Structure-based design of potent inhibitors of scytalone dehydratase: displacement of a water molecule from the active site. Biochemistry 37:17735–17744

117. Liu C, Wrobleski ST, Lin J, Ahmed G, Metzger A, Wityak J, Gillooly KM, Shuster DJ, McIntyre KW, Pitt S, Shen DR, Zhang RF, Zhang H, Doweyko AM, Diller D, Henderson I, Barrish JC, Dodd JH, Schieven GL, Leftheris K (2005) 5-Cyanopyrimidine derivatives as a novel class of potent, selective, and orally active inhibitors of p38α MAP kinase. J Med Chem 48:6261–6270

118. Wissner A, Berger DM, Boschelli DH, Floyd MB Jr, Greenberger LM, Gruber BC, Johnson BD, Mamuya N, Nilakantan R, Reich MF, Shen R, Tsou H-R, Upeslacis E, Wang YF, Wu B, Ye F, Zhang N (2000) 4-Anilino-6,7-dialkoxyquinoline-3-carbonitrile inhibitors of epidermal

growth factor receptor kinase and their bioisosteric relationship to the 4-anilino-6,7-dialkoxy-quinazoline inhibitors. J Med Chem 43:3244–3256

119. Davies NGM, Browne H, Davis B, Drysdale MJ, Foloppe N, Geoffrey S, Gibbons B, Hart T, Hubbard R, Rugaard Jensen M, Mansell H, Massey A, Matassova N, Moore JD, Murray J, Pratt R, Ray S, Robertson A, Roughley SD, Schoepfer J, Scriven K, Simmonite H, Stokes S, Surgenor A, Webb P, Wood M, Wright L, Brough P (2012) Targeting conserved water molecules: design of 4-aryl-5-cyanopyrrolo[2,3-d]pyrimidine Hsp90 inhibitors using fragment-based screening and structure-based optimization. Bioorg Med Chem 20:6770–6789

120. Trujillo JI, Kiefer JR, Huang W, Day JE, Moon J, Jerome GM, Bono CP, Kornmeier CM, Williams ML, Kuhn C, Rennie GR, Wynne TA, Carron CP, Thorarensen A (2012) Investigation of the binding pocket of human hematopoietic prostaglandin (PG) D2 synthase (hH-PGDS): a tale of two waters. Bioorg Med Chem Lett 22:2795–3799

121. Meanwell NA (2011) Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. Chem Res Toxicol 24:1420–1456

122. Johnson F (1968) Allylic strain in six membered rings. Acc Chem Res 68:375–413

123. Yamazaki T, Taguchi T, Ojima I (2009) Unique properties of fluorine and their relevance to medicinal chemistry and chemical biology. In: Ojima I (ed) Fluorine in medicinal chemistry and chemical biology. Chap 1, Wiley, Chichester, UK pp 1–46

124. Hunter L (2010) The C–F bond as a conformational tool in organic and biological chemistry. Beilstein J Org Chem 6. doi:10.3762/bjoc.6.38

125. Zimmer LE, Sparr C, Gilmour R (2011) Fluorine conformational effects in organocatalysis: an emerging strategy for molecular design. Angew Chem Int Ed 50:11860–11871

126. Buissonneaud DY, van Mourik T, O'Hagan D (2010) A DFT study on the origin of the fluorine gauche effect in substituted fluoroethanes. Tetrahedron 66:2196–2202

127. O'Hagan D (2012) Organofluorine chemistry: synthesis and conformation of vicinal fluoromethylene motifs. J Org Chem 77:3689–3699

128. Wu D, Tian A, Sun H (1998) Conformational properties of 1,3-difluoropropane. J Chem Phys A 102:9901–9905

129. Tavasli M, O'Hagan D, Pearson C, Petty MC (2002) The fluorine gauche effect. Langmuir isotherms report the relative conformational stability of (±)-erythro- and (±)-threo-9,10-difluorostearic acids. Chem Commun 1226–1227

130. Hunter L, Kirsch P, Slawin AMZ, O'Hagan D (2009) Synthesis and structure of stereoisomeric multivicinal hexafluoroalkanes. Angew Chem Int Ed 48:5457–5460

131. Banks JW, Batsanov AS, Howard JAK, O'Hagan D, Rzepa H, Martin-Santamaria S (1999) The preferred conformation of α-fluoroamides. J Chem Soc Perkin 2:2409–2411

132. Briggs CRS, O'Hagan D, Howard JAK, Yulfi DS (2003) The C–F bond as a tool in the conformational control of amides. J Fluorine Chem 119:9–13

133. Dalvit D, Vulpetti A (2012) Intermolecular and intramolecular hydrogen bonds involving fluorine atoms: implications for recognition, selectivity, and chemical properties. ChemMedChem 7:262–272

134. Schüler M, O'Hagan D, Slawin AMZ (2005) The vicinal F-C-C-F moiety as a tool for influencing peptide conformation. Chem Commun 4324–4326

135. O'Hagan D, Rzepa HS, Schüler M, Slawin AMZ (2006) The vicinal difluoro motif: the synthesis and conformation of erythreo- and threo-diastereomers of 1,2-difluorodiphenylethanes, 2,3-difluorosuccinic acids and their derivatives. Beilstein J Org Chem 2. doi:10.1186/1860-5397-2-19

136. Winkler M, Moraux T, Khairy HA, Scott RH, Slawin AMZ, O'Hagan D (2009) Synthesis and vanilloid receptor (TRPV1) activity of the enantiomers of α-fluorinated capsaicin. ChemBioChem 10:823–828

137. Peddie V, Butcher RJ, Robinson WT, Wilce MCJ, Traore DAK, Abell AD (2012) Synthesis and conformation of fluorinated β-peptidic compounds. Chem Eur J 18:6655–6662

138. O'Hagan D, Bilton C, Howard JAK, Knight L, Tozer DJ (2001) The preferred conformation of N-β-fluoroethylamides. Observation of the fluorine amide gauche effect. J Chem Soc Perkin Trans 2:605–607
139. O'Hagan D, Rzepa HS (1997) Some influences of fluorine in bioorganic chemistry. Chem Commun 645–652
140. Mascitti V, Stevens BD, Choi C, McClure KF, Guimarães CRW, Farley KA, Munchhof MJ, Robinson RP, Futatsugi K, Lavergen SY, Lefker BA, Cornelius P, Bonin PD, Kalgutkar AS, Sharma R, Chen Y (2011) Design and evaluation of a 2(2,3,6-trifluorophenyl)acetamide derivative as an agonist of the GPR119 receptor. Bioorg Med Chem Lett 21:1306–1309
141. Holmgren SK, Taylor KM, Bretscher LE, Raines RT (1998) Code for collagen's stability deciphered. Nature 392:666–667
142. Bretscher LE, Jenkins CL, Taylor KM, DeRider ML, Raines RT (2001) Conformational stability of collagen relies on a stereoelectronic effect. J Am Chem Soc 123:777–778
143. Hodges JA, Raines RT (2003) Stereoelectronic effects on collagen stability: the dichotomy of 4-fluoroproline diastereomers. J Am Chem Soc 125:9262–9263
144. Doi M, Nishi Y, Uchiyama S, Nisiuchi Y, Nakazawa T, Ohkubo T, Kobayasji Y (2003) Characterization of collagen model peptides containing 4-fluoroproline: (4(S)-fluoroproline-Pro-Gly)$_{10}$ forms a triple helix, but (4(R)-fluoroproline-Pro-Gly)$_{10}$ does not. J Am Chem Soc 125:9922–9923
145. Hodges JA, Raines RT (2005) Stereoelectronic and steric effects in the collagen triple helix: toward a code for strand association. J Am Chem Soc 127:15923–15932
146. Shoulders MD, Kamer KJ, Raines RT (2009) Origin of the stability conferred upon collagen by fluorination. Bioorg Med Chem Lett 19:3859–3862
147. DeRider ML, Wilkins SJ, Waddell MJ, Bretscher LE, Weinhold F, Raines RT, Markley JL (2002) Collagen stability: insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. J Am Chem Soc 124:2497–2505
148. Doi M, Nishi Y, Kiritoshi N, Iwata T, Nago M, Nakano H, Uchiyama S, Nakazawa T, Wakamiya T, Kbayashi Y (2002) Simple and efficient syntheses of Boc- and Fmoc-protected 4(R)- and 4(S)-fluoroproline solely from 4(R)-hydroxyproline. Tetrahedron 58:8453–8459
149. Park S, Radmer RJ, Klein TE, Pande VS (2005) A new set of molecular mechanics parameters for hydroxyproline and its use in molecular dynamics simulations of collagen-like peptides. J Comput Chem 26:1612–1616
150. Kim W, Hardcastle KL, Conticello VP (2006) Fluoroproline flip-flop: regiochemcial reversal of a stereoelectronic effect on peptide and protein structures. Angew Chem Int Ed 45:8141–8814
151. Kitamoto T, Ozawa T, Abe M, Marubayashi S, Yamazaki T (2008) Incorporation of fluoroprolines to proctolin: study on the effect of a fluorine atom towards peptidic conformation. J Fluorine Chem 129:286–293
152. Briggs CRS, Allen MJ, O'Hagan D, Tozer DJ, Slawin AMZ, Goeta AE, Howard JAK (2004) The observation of a large gauche preference when 2-fluoroethylamine and 2-fluoroethanol become protonated. Org Biomol Chem 2:732–740
153. Lankin DC, Chandrakumar NS, Rao SN, Spangler DP, Snyder JP (1993) Protonated 3-fluoropiperidines: an unusual fluoro directing effect and a test for quantitative theories of solvation. J Am Chem Soc 115:3356–3357
154. Snyder JP, Chandrakumar NS, Sato H, Lankin DC (2000) The unexpected diaxial orientation of cis-3,5-difluoropiperidine in water: a potent CF–NH charge-dipole effect. J Am Chem Soc 122:544–545
155. Sun A, Lankin DC, Harcastle K, Snyder JP (2005) 3-Fluoropiperidines and N-methyl-3-fluoropiperidinium salts: the persistence of axial fluorine. Chem Eur J 11:1579–1591
156. Deniau G, Slawin AMZ, Lebl T, Chorki F, Issberner JP, van Mourik T, Heygate JM, Lambert JJ, Etherington L-A, Sillar KT, O'Hagan D (2007) Synthesis, conformation and biological

evaluation of the enantiomers of 3-fluoro-γ-aminobutyric acid ((*R*)- and (*S*)-3F-GABA): an analogue of the neurotransmitter GABA. ChemBioChem 8:2265–2274

157. Clift MD, Ji H, Deniau GP, O'Hagan D, Silverman RB (2007) Enantiomers of 4-amino-3-fluorobutanoic acid as substrates for γ-aminobutyric acid aminotransferase. Conformational probes for GABA binding. Biochemistry 46:13819–13828

158. Yamamoto I, Deniau GP, Gavande N, Chebib M, Johnston GAR, O'Hagan D (2011) Agonist responses of (*R*)- and (*S*)-3-fluoro-γ-aminobutyric acids suggest an enantiomeric fold for GABA binding to GABA$_C$ receptors. Chem Commun 47:7956–7958

159. Chia PW, Livesey MR, Slawin AMZ, van Mourik T, Wyllie DJA, O'Hagan D (2012) 3-Fluoro-*N*-methyl-D-asaprtic acid (3F-NMDA) stereoisomers as conformational probes for exploring agonist binding at NMDA receptors. Chem Eur J 18:8813–8819

160. Abraham RJ, Chambers EJ, Thomas AW (1994) Conformational analysis. Part 22. An NMR and theoretical investigation of the *gauche* effect in fluoroethanols. J Chem Soc Perkin Trans 2:949–955

161. Abraham RJ, Smith TAD, Thomas AW (1996) Conformational analysis. Part 28. OH-F hydrogen bonding and the conformation of trans-2-fluorocyclohexanol. J Chem Soc Perkin Trans 2:1949–1955

162. Myers AG, Barbay JK, Zhong B (2001) Asymmetric synthesis of chiral organofluorine compounds: use of nonracemic fluoroiodoacetic acid as a practical electrophile and its application to the synthesis of monofluoro hydroxyethylene dipeptide isosteres within a novel series of HIV protease inhibitors. J Am Chem Soc 123:7207–7219

163. Brameld KA, Kuhn B, Reuter DC, Stahl M (2008) Small molecule conformational preferences derived from crystal structure data. A medicinal chemistry focused analysis. J Chem Inf Model 48:1–24

164. Gomez R, Jolly S, Williams T, Tucker T, Tynebor R, Vacca J, McGaughey G, Lai M-T, Felock P, Munshi V, DeStefano D, Touch S, Miller M, Yan Y, Sanchez R, Liang Y, Paton B, Wan B-L, Anthony N (2011) Design and synthesis of pyridone inhibitors of non-nucleoside reverse transcriptase. Bioorg Med Chem Lett 21:7344–7350

165. Leroux F, Jeschke P, Schlosser M (2005) α-Fluorinated ethers, thioethers, and amines: anomerically biased species. Chem Rev 105:827–856

166. Massa MA, Spangler DP, Durley RC, Hickory BS, Connolly DT, Witherbee BJ, Smith ME, Sikorski JA (2001) Novel heteroaryl replacements of aromatic 3-tetrafluoroethoxy substituents in trifluoro-3-(tertiaryamino)-2-propanols as potent inhibitors of cholesterol ester transfer protein. Bioorg Med Chem Lett 11:1625–1628

167. Horne DB, Bartberger MD, Kaller MR, Monenschein H, Zhong W, Hitchcock SA (2009) Synthesis and conformational analysis of α, α-difluoroalkyl heteroaryl ethers. Tetrahedron Lett 50:5452–5455

168. Anderson GM III, Kollman PA, Domelsmith LN, Houk KN (1979) Methoxy group nonplanarity in o-dimethoxybenzenes. Simple predictive models for conformations and rotational barriers in alkoxyaromatics. J Am Chem Soc 101:2344–2352

169. Klocker J, Karpfen A, Wolschann P (2003) On the structure and torsional potential of trifluoromethoxybenzene: an ab initio and density functional study. Chem Phys Lett 367:566–575

170. Chien RJ, Corey EJ (2010) Strong conformational preferences of heteroaromatic ethers and electron pair repulsion. Org Lett 12:132–135

171. Phillips GB, Buckman BO, Davey DD, Eagen KA, Guilford WJ, Hinchman J, Ho E, Koovakkat S, Liang A, Light DR, Mohan R, Ng HP, Post JM, Shaw KJ, Smith D, Subramanyam B, Sullivan ME, Trinh L, Vergona R, Walters J, White K, Whitlow M, Wu S, Xu W, Morrissey MM (1998) Discovery of *N*-[2-[5-[amino(imino)methyl]-2-hydroxyphenoxy]-3, 5-difluoro-6-[3-(4, 5-dihydro-1-methyl-1*H*-imidazol-2-yl)phenoxy]pyridin-4-yl]-*N*-methylglycine (ZK-807834): a potent, selective, and orally active inhibitor of the blood coagulation enzyme factor Xa. J Med Chem 41:3557–3562

172. Phillips G, Davey DD, Eagen KA, Koovakkat SK, Liang A, Ng HP, Pinkerton M, Trinh L, Whitlow M, Beatty AM, Morrissey MM (1999) Design, synthesis, and activity of 2,6-diphenoxypyridine-derived factor Xa inhibitors. J Med Chem 42:1749–1756

173. Adler M, Davey DD, Phillips GB, Kim S-H, Jancarik J, Rumennik G, Light DR, Whitlow M (2000) Preparation, characterization, and the crystal structure of the inhibitor ZK-807834 (CI-1031) complexed with factor Xa. Biochemistry 39:12534–12542

174. Chen YL, Mansbach RS, Winter SM, Brooks E, Collins J, Michael L. Corman ML, Dunaiskis AR, Faraci WS, Gallaschun RJ, Schmidt A, Schulz DW (1997) Synthesis and oral efficacy of a 4-(butylethylamino)pyrrolo[2,3-d]pyrimidine: a centrally active corticotropin-releasing factor-1 receptor antagonist. J Med Chem 40:1749–1754

175. Semple G, Lehmann J, Wong A, Ren A, Bruce M, Shin Y-J, Sage CR, Morgan M, Chen W-C, Sebring K, Chu Z-L, Leonard JN, Al-Shamma H, Grottick AJ, Du F, Liang Y, Demarest K, Jones RM (2012) Discovery of a second generation agonist of the orphan G-protein coupled receptor GPR119 with an improved profile. Bioorg Med Chem Lett 22:1750–1755

176. Qiao JX, Cheney DL, Alexander RS, Smallwood AM, King SR, He K, Rendina AR, Luettgen JM, Knabb RM, Wexler RR, Lam PYS (2008) Achieving structural diversity using the perpendicular conformation of *alpha*-substituted phenylcyclopropanes to mimic the bioactive conformation of *ortho*-substituted biphenyl P4 moieties: discovery of novel, highly potent inhibitors of factor Xa. Bioorg Med Chem Lett 18:4118–4123

177. Leeson PD, Springthorpe B (2007) The influence of drug-like concepts on decision making in medicinal chemistry. Nat Rev Drug Discov 6:881–890

178. Tarcsay A, Nyíri K, Keserű GM (2012) Impact of lipophilic efficiency on compound quality. J Med Chem 55:1252–1260

179. Keserű GM, Makara GM (2009) The influence of lead discovery strategies on the properties of drug candidates. Nat Rev Drug Discov 8:203–212

180. Hann M, Keserű GM (2012) Finding the sweet spot: the role of nature and nurture in medicinal chemistry. Nat Rev Drug Discov 11:355–365

181. Lovering F, Bikker J, Humblet C (2009) Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 52:6752–6756

182. Lovering F (2013) Escape from flatland 2: complexity and promiscuity. Med Chem Commun 4:515–519

183. Pellicciari R, Raimondo M, Marinozzi M, Natalini B, Costantino G, Thomsen C (1996) (*S*)-(+)-2-(3'-Carboxybicyclo[1.1.1]pentyl)glycine, a structurally new group 1 metabotropic glutamate receptor antagonist. J Med Chem 39:2874–2876

184. Pellicciari R, Costantino G, Giovagnoni E, Mattoli L, Brabet I, Pin J-P (1998) Synthesis and preliminary evaluation of (*S*)-2-(4'-carboxycubyl)glycine, a new selective mGluR1 antagonist. Bioorg Med Chem Lett 8:1569–1574

185. Costantino G, Maltoni K, Marinozzi M, Camaioni E, Prezeau L, Pin J-P, Pellicciari R (2001) Synthesis and biological evaluation of 2-(3'-(1*H*-tetrazol-5-yl)bicyclo[1.1.1]pent-1-yl)glycine (S-TBPG), a novel mGlu1 receptor antagonist. Bioorg Med Chem 9:221–227

186. Stepan AF, Subramanyam C, Efremov IV, Dutra JK, O'Sullivan TJ, DiRico KJ, McDonald WS, Won A, Dorff PH, Nolan CE, Becker SL, Pustilnik LR, Riddell DR, Kauffman GW, Kormos BL, Zhang L, Lu Y, Capetta SH, Green ME, Karki K, Sibley E, Atchison KP, Hallgren AJ, Oborski CE, Robshaw AE, Sneed B, O'Donnell CJ (2012) Application of the bicyclo[1.1.1]pentane motif as a nonclassical phenyl ring bioisostere in the design of a potent and orally active γ-secretase inhibitor. J Med Chem 55:3414–3424

187. Pinto DJP, Orwat MJ, Wang S, Fevig JM, Quan ML, Amparo E, Cacciola J, Rossi KA, Alexander RS, Smallwood AM, Luettgen JM, Liang L, Aungst BJ, Wright MR, Knabb RM, Wong PC, Wexler RR, Lam PYS (2001) Discovery of 1-[3-(aminomethyl)phenyl]-*N*-[3-fluoro-2'-(methylsulfonyl)-[1,1'-biphenyl]-4-yl]-3-(trifluoromethyl)-1*H*-pyrazole-5-carboxamide (DPC423), a highly potent, selective, and orally bioavailable inhibitor of blood coagulation factor Xa. J Med Chem 44:566–578

188. Quan ML, Lam PYS, Han Q, Pinto DJP, He MY, Li R, Ellis CD, Clark CG, Teleha CA, Sun J-H, Alexander RS, Bai S, Luettgen JM, Knabb RM, Wong PC, Wexler RR (2005) Discovery of 1-(3′-aminobenzisoxazol-5′-yl)-3-trifluoromethyl-N-[2-fluoro-4-[(2′-dimethylaminomethyl) imidazol-1-yl]phenyl]-1H-pyrazole-5-carboxyamide hydrochloride (razaxaban), a highly potent, selective, and orally bioavailable factor Xa inhibitor. J Med Chem 48:1729–1744

189. Nayak SK, Reddy MK, Guru Row TN, Chopra D (2011) Role of hetero-halogen (F⋯X, X = Cl, Br, and I) or homo-halogen (X⋯X, X = F, Cl, Br, and I) interactions in substituted benzanilides. Crystal Growth Des 11:1578–1596

190. Chopra D, Row TNG (2008) Evaluation of the interchangeability of C–H and C–F groups: insights from crystal packing in a series of isomeric fluorinated benzanilides. Cryst Eng Comm 10:54–67

191. Sehon CA, Wang GZ, Viet AQ, Goodman KB, Dowdell SE, Elkins PA, Semus SF, Evans C, Jolivette LJ, Kirkpatrick RB, Dul E, Khandekar SS, Yi T, Wright LL, Smith GK, Behm DJ, Bentley R, Doe CP, Hu E, Lee D (2008) Potent, selective and orally bioavailable dihydropyrimidine inhibitors of rho kinase (ROCK1) as potential therapeutic agents for cardiovascular diseases. J Med Chem 51:6631–6634

192. Dai Y, Hartandi K, Ji Z, Ahmed AA, Albert DH, Bauch JL, Bouska JJ, Bousquet PF, Cunha GA, Glaser KB, Harris CM, Hickman D, Guo J, Li J, Marcotte PA, Marsh KC, Moskey MD, Martin RL, Olson AM, Osterling DJ, Pease LL, Soni NB, Stewart KD, Stoll VS, Tapang P, Reuter DR, Davidsen SK, Michaelides MR (2007) Discovery of N-(4-(3-amino-1H-indazol-4-yl)phenyl)-N′-(2-fluoro-5-methylphenyl)urea (ABT-869), a 3-aminoindazole-based orally active multitargeted receptor tyrosine kinase inhibitor. J Med Chem 50:1584–1597

193. Weiss MM, Williamson T, Babu-Khan S, Bartberger MD, Brown J, Chen K, Cheng Y, Citron M, Croghan MD, Dineen TA, Esmay J, Graceffa RF, Harried SS, Hickman D, Hitchcock SA, Horne DB, Huang H, Imbeah-Ampiah R, Judd T, Kaller MR, Kreiman CR, La DS, Li V. Lopez P, Louie S, Monenschein H, Nguyen TT, Pennington LD, Rattan C, San Miguel T, Sickmier EA, Wahl RC, Wen PH, Wood S, Xue Q, Yang BH, Patel VF, Zhong W (2012) Design and preparation of a potent series of hydroxyethylamine containing β-secretase inhibitors that demonstrate robust reduction of central β-amyloid. J Med Chem 55:9009–9024

194. Dineen TA, Weiss MM, Williamson T, Acton P, Babu-Khan S, Bartberger MD, Brown J, Chen K, Cheng Y, Citron M, Croghan MD, Dunn RT, Esmay J, Graceffa RF, Harried SS, Hickman D, Hitchcock SA, Horne DB, Huang H, Imbeah-Ampiah R, Judd T, Kaller MR, Kreiman CR, La DS, Li V, Lopez P, Louie S, Monenschein H, Nguyen TT, Pennington LD, San Miguel T, Sickmier EA, Vargas HM, Wahl RC, Wen PH, Whittington DA, Wood S, Xue Q, Yang BH, Patel VF, Zhong W (2012) Design and synthesis of potent, orally efficacious hydroxyethylamine derived β-site amyloid precursor protein cleaving enzyme (BACE1) inhibitors. J Med Chem 55:9025–9044

195. Kaller MR, Harried SS, Albrecht B, Amarante P, Babu-Khan S, Bartberger MD, Brown J, Brown R, Chen K, Cheng Y, Citron M, Croghan MD, Graceffa R, Hickman D, Judd T, Kriemen C, La D, Li V, Lopez P, Luo Y, Masse C, Monenschein H, Nguyen T, Pennington LD, Miguel TS, Sickmier EA, Wahl RC, Weiss MM, Wen PH, Williamson T, Wood S, Xue M, Yang B, Zhang J, Patel V, Zhong W, Hitchcock S (2012) A potent and orally efficacious, hydroxyethylamine-based inhibitor of β-secretase. ACS Med Chem Lett 3:886–891

196. Hitchcock SA (2012) Structural modifications that alter the P-glycoprotein efflux properties of compounds. J Med Chem 55:4877–4895

197. Desai PV, Raub TJ, Blanco M-J (2012) How hydrogen bonds impact P-glycoprotein transport and permeability. Bioorg Med Chem Lett 22:6540–6548

198. Kuduk SD, Di Marco CN, Chang RK, Wood MR, Schirripa KM, Kim JJ, Wai JMC, DiPardo RM, Murphy KL, Ransom RW, Harrell CM, Reiss DR, Holahan MA, Cook J, Hess JF, Sain N, Urban MO, Tang C, Prueksaritanont T, Pettibone DJ, Bock MG (2007) Development of orally bioavailable and CNS penetrant biphenylaminocyclopropane carboxamide brady-kinin B1 receptor antagonists. J Med Chem 50:272–282

199. Ettorre E, D'Andrea P, Mauro S, Porcelloni M, Rossi C, Altamura M, Catalioto RM, Giuliani S, Maggi CA, Fattori D (2011) hNK$_2$ receptor antagonists. The use of intramolecular hydrogen bonding to increase solubility and membrane permeability. Bioorg Med Chem Lett 21:1807–1809

200. Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V, Wongvipat J, Smith-Jones PM, Yoo D, Kwon A, Wasielewska T, Welsbie D, Chen C, Higano CS, Beer TM, Hung DT, Scher HI, Jung ME, Sawyers CL (2009) Development of a second-generation antiandrogen for treatment of advanced prostate cancer. Science 324:787–790

201. Jung ME, Ouk S, Yoo D, Sawyers CL, Chen C, Tran C, Wongvipat J (2010) Structure–activity relationship for thiohydantoin androgen receptor antagonists for castration-resistant prostate cancer (CRPC). J Med Chem 53:2779–2796

202. Marsham PR, Wardleworth JM, Boyle FT, Hennequin LF, Kimbell R, Brown M, Jackman AL (1999) Design and synthesis of potent non-polyglutamatable quinazoline antifolate thymidylate synthase inhibitors. J Med Chem 42:3809–3820

203. Niculesco-Duvaz I (2000) ZD-9331 (Astra-Zeneca). Curr Opin Investig Drugs 1:141–149

204. Swahn B-M, Kolmodin K, Karlstroem S, von Berg S, Soederman P, Holenz J, Berg S, Lindstroem J, Sundstroem M, Turek D, Kihlstroem J, Slivo C, Andersson L, Pyring D, Rotticci D, Oehberg L, Kers A, Bogar K, Bergh M, Olsson L-L, Janson J, Eketjaell S, Georgievska B, Jeppsson F, Faelting J (2012) Design and synthesis of β-site amyloid precursor protein cleaving enzyme (BACE1) inhibitors with in vivo brain reduction of β-amyloid peptides. J Med Chem 55:9346–9361

205. Peterlin-Mašič L, Kikelj D (2001) Arginine mimetics. Tetrahedron 57:7073–7105

206. Peterlin-Mašič L (2006) Arginine mimetic structure in biologically active antagonists and inhibitors. Curr Med Chem 13:3627–3648

207. Pinto DJP, Smallheer JM, Cheney DL, Knabb RM, Wexler RR (2010) Factor Xa inhibitors: next generation antithrombotic agents. J Med Chem 53:6243–6274

208. Danalev D (2012) Inhibitors of serine proteinases from blood coagulation cascade – view on current developments. Mini Rev Med Chem 12:721–730

209. Nar H (2012) The role of structural information in the discovery of direct thrombin and factor Xa inhibitors. Trends Pharmacol Sci 33:279–288

210. Srivastave S, Goswami LN, Dikshit DK (2005) Progress in the design of low molecular weight thrombin inhibitors. Med Res Rev 25:66–92

211. Ghorai P, Kraus A, Keller M, Gotte C, Igel P, Schneider E, Schnell D, Bernhardt G, Dove S, Zabel M, Elz S, Seifert R, Buschauer A (2008) Acylguanidines as bioisosteres of guanidines: N^G-acylated Imidazolylpropylguanidines, a new class of histamine H$_2$ receptor agonists. J Med Chem 51:7193–7204

212. Kraus A, Ghorai P, Birnkammer T, Schnell D, Elz S, Seifert R, Dove S, Bernhardt G, Buschauer A (2009) N^G-acylated aminothiazolylpropylguanidines as potent and selective histamine H$_2$ receptor agonists. ChemMedChem 4:232–240

213. Pluym N, Brennauer A, Keller M, Ziemek R, Pop N, Bernhardt G, Buschauer A (2011) Application of the guanidine–acylguanidine bioisosteric approach to argininamide-type NPY Y2 receptor antagonists. ChemMedChem 6:1727–1738

214. Tomczuk B, Lu T, Soll RM, Fedde C, Wang A, Murphy L, Crysler C, Dasgupta M, Eisennagel S, Spurlino J, Bone R (2003) Oxyguanidines: application to non-peptidic phenyl-based thrombin inhibitors. Bioorg Med Chem Lett 13:1495–1498

215. Lu T, Markotan T, Coppo F, Tomczuk B, Crysler C, Eisennagel S, Spurlino J, Gremminger L, Soll RM, Giardino EC, Bone R (2004) Oxyguanidines. Part 2: discovery of a novel orally active thrombin inhibitor through structure-based drug design and parallel synthesis. Bioorg Med Chem Lett 14:3727–3731

216. Lee L, Kreutter KD, Pan W, Crysler C, Spurlino J, Player MR, Tomczuk B, Lu T (2007) 2-(2-Chloro-6-fluorophenyl)acetamides as potent thrombin inhibitors. Bioorg Med Chem Lett 17:6266–6269

217. Kreutter KD, Lu T, Lee L, Giardino EC, Patel S, Huang H, Xu G, Fitzgerald M, Haertlein BJ, Mohan V, Crysler C, Eisennagel S, Dasgupta M, McMillan M, Spurlino JC, Huebert ND, Maryanoff BE, Tomczuk BE, Damiano BP, Player MR (2008) Orally efficacious thrombin inhibitors with cyanofluorophenylacetamide as the P2 motif. Bioorg Med Chem Lett 18:2865–2870
218. Lu T, Markotan T, Ballentine SK, Giardino EC, Spurlino J, Brown K, Maryanoff BE, Tomczuk BE, Damiano BP, Shukla U, End D, Andrade-Gordon P, Bone RF, Player MR (2010) Discovery and clinical evaluation of 1-{N-[2-(amidinoaminooxy)ethyl]amino} carbonylmethyl-6-methyl-3-[2,2-difluoro-2-phenylethylamino]pyrazinone (RWJ-671818), a thrombin inhibitor with an oxyguanidine P1 motif. J Med Chem 53:1843–1856
219. Lam PYS, Clark CG, Li R, Pinto DJP, Orwat MJ, Galemmo RA, Fevig JM, Teleha CA, Alexander RS, Smallwood AM, Rossi KA, Wright MR, Bai SA, He K, Luettgen JM, Wong PC, Knabb RM, Wexler RR (2003) Structure-based design of novel guanidine/benzamidine mimics: potent and orally bioavailable factor Xa inhibitors as novel anticoagulants. J Med Chem 46:4405–4418
220. Pinto DJP, Orwat MJ, Koch S, Rossi KA, Alexander RS, Smallwood A, Wong PC, Rendina AR, Luettgen JM, Knabb RM, He K, Xin B, Wexler RR, Lam PYS (2007) Discovery of 1-(4-methoxyphenyl)-7-oxo-6-[4-(2-oxo-1-piperidinyl)phenyl]-4,5,6,7-tetrahydro-1H-pyrazolo [3,4-c]pyridine-3-carboxamide (apixaban, BMS-562247), a highly potent, selective, efficacious, and orally bioavailable inhibitor of blood coagulation factor Xa. J Med Chem 50:5339–5356
221. Rye CS, Baell JB (2005) Phosphate isosteres in medicinal chemistry. Curr Med Chem 12:3127–3141
222. Romanenko VD, Kukhar VP (2006) Fluorinated phosphonates: synthesis and biomedical application. Chem Rev 106:3868–3935
223. Elliott TS, Slowey A, Yeb Y (2012) Conway SJ (2012) The use of phosphate bioisosteres in medicinal chemistry and chemical biology. Med Chem Commun 3:735–751
224. Blackburn GM, Kent DE (1981) A novel synthesis of α- and γ-fluoroalkylphosphonates. JCS Chem Comm 511–513
225. Blackburn GM, England DA, Kolkmann F (1981) Monofluoro- and difluoro-methylenebisphosphonic acids : isopolar analogues of pyrophosphoric acid. JCS Chem Comm 930–932
226. Smyth MS, Ford H Jr, Burke TR Jr (1992) A general method for the preparation of benzylic α, α-difluorophosphonic acids; non-hydrolyzable mimetics of phosphotyrosine. Tet Lett 33:4137–4140
227. Hecker SJ, Erion MD (2008) Prodrugs of phosphates and phosphonates. J Med Chem 51:2328–2345
228. Hostetler KY (2009) Alkoxyalkyl prodrugs of acyclic nucleoside phosphonates enhance oral antiviral activity and reduce toxicity: current state of the art. Antiviral Res 82:A84–A98
229. Pertusati F, Serpi M, McGuigan C (2012) Medicinal chemistry of nucleoside phosphonate prodrugs for antiviral therapy. Antiviral Chem Chemother 22:181–203
230. Combs AP (2007) Structure-based drug design of new leads for phosphatase research. IDrugs 10:112–115
231. Combs AP (2010) Recent advances in the discovery of competitive protein tyrosine phosphatase 1B inhibitors for the treatment of diabetes, obesity, and cancer. J Med Chem 53:2333–2344
232. Savarino A (2006) A historical sketch of the discovery and development of HIV-1 integrase inhibitors. Exp Opin Investig Drugs 15:1507–1522
233. Johns BA, Svolto AC (2008) Advances in two-metal chelation inhibitors of HIV integrase. Exp Opin Ther Patents 18:1225–1237
234. Ramkumar K, Serrao E, Odde S, Neamati N (2010) HIV-1 integrase inhibitors: 2007–2008 update. Med Res Rev 30:890–954

235. Pendri A, Meanwell NA, Peese KM, Walker MA (2011) New first and second generation inhibitors of human immunodeficiency virus-1 integrase. Exp Opin Ther Patents 21:1173–1189

236. Ingale KB, Bhatia MS (2011) HIV-1 integrase inhibitors: a review of their chemical development. Antiviral Chem Chemother 22:95–105

237. Beare KD, Coster MJ, Rutledge PJ (2012) Diketoacid inhibitors of HIV-1 integrase: from L-708,906 to raltegravir and beyond. Curr Med Chem 19:1177–1192

238. Leeson PD, St-Gallay SA, Wenlock MC (2011) Impact of ion class and time on oral drug molecular properties. Med Chem Commun 2:91–105

239. Manallack DT, Prankerd RJ, Yuriev E, Oprea TI, Chalmers DK (2013) The significance of acid/base properties in drug discovery. Chem Soc Rev 42:485–496

240. Azzaoui K, Hamon J, Faller B, Whitebread S, Jacoby E, Bender A, Jenkins JL, Urban L (2007) Modeling promiscuity based on in vitro safety pharmacology profiling data. ChemMedChem 2:874–880

241. Peters J-E, Schnider P, Mattei P, Kansy M (2009) Pharmacological promiscuity: dependence on compound properties and target specificity in a set of recent Roche compounds. ChemMedChem 4:680–686

242. Waring MJ, Johnstone C (2007) A quantitative assessment of hERG liability as a function of lipophilicity. Bioorg Med Chem Lett 17:1759–1764

243. Jamieson C, Moir EM, Rankovic Z, Wishart G (2006) Medicinal chemistry of hERG optimizations: highlights and hang-ups. J Med Chem 49:5029–5046

244. Hanumegowda UM, Wenke G, Regueiro-Ren A, Yordanova R, Corradi JP, Adams SP (2010) Phospholipidosis as a function of basicity, lipophilicity, and volume of distribution of compounds. Chem Res Toxicol 23:749–755

245. Ratcliffe AJ (2009) Medicinal chemistry strategies to minimize phospholipidosis. Curr Med Chem 16:2816–2823

246. Bernstein PR, Ciaccio P, Morelli J (2011) Drug-induced phospholipidosis. Ann Rep Med Chem 46:419–430

247. Morgenthaler M, Schweizer E, Hoffman-Röder A, Benini F, Martin RE, Jaeschke G, Wagner B, Fischer H, Bendels S, Zimmeli D, Scheider J, Diedeich F, Kansy M, Müller K (2007) Predicting and tuning physicochemical properties in led optimization: amine-basicities. ChemMedChem 2:1100–1115

248. Cox CD, Coleman PJ, Breslin MJ, Whitman DB, Garbaccio RM, Fraley ME, Buser CA, Walsh ES, Hamilton K, Schaber MD, Lobell RB, Tao W, Davide JP, Diehl RE, Abrams MT, South VJ, Huber HE, Torrent M, Prueksaritanont T, Li C, Slaughter DE, Mahan E, Fernandez-Metzler C, Yan Y, Kuo LC, Kohl NE, Hartman GD (2008) Kinesin spindle protein (KSP) inhibitors. 9. Discovery of (2S)-4-(2,5-difluorophenyl)-N-[(3R,4S)-3-fluoro-1-methylpiperidin-4-yl]-2-(hydroxymethyl)-N-methyl-2-phenyl-2,5-dihydro-1H-pyrrole-1-carboxamide (MK-0731) for the treatment of taxane-refractory cancer. J Med Chem 51:4239–4252

249. Kalgutkar AS, Obach RS, Maurer TS (2007) Mechanism-based inactivation of cytochrome P450 enzymes: chemical mechanisms, structure–activity relationships and relationship to clinical drug–drug interactions and idiosyncratic adverse drug reactions. Curr Drug Metab 8:407–447

250. Goncharov NV, Jenkins RO, Radilov AS (2006) Toxicology of fluoroacetate: a review, with possible directions for therapy research. J Appl Toxicol 26:148–161

251. Sani M, Volonterio A, Zanda M (2007) The trifluoroethylamine function as peptide bond replacement. ChemMedChem 2:1693–1700

252. Zanda M (2004) Trifluoromethyl group: an effective xenobiotic function for peptide backbone modification. New J Chem 28:1401–1411

253. Molteni M, Pesenti C, Sani M, Volonterio A, Zanda M (2004) Fluorinated peptidomimetics: synthesis, conformational and biological features. J Fluorine Chem 125:1735–1743

254. Molteni M, Bellucci MC, Bigotti S, Mazzini S, Volonterio A, Zanda M (2009) Ψ[CH(CF$_3$)NH]Gly-peptides: synthesis and conformation analysis. Org Biomol Chem 7:2286–2296

255. Black WC, Bayly C, Davis DE, Desmarais S, Falgueyret J-P, Leger S, Li CS, Masse F, McKay DJ, Palmer JT, Percival MD, Robichaud J, Tsou N, Zamboni R (2005) Trifluoroethylamines as amide isosteres in inhibitors of cathepsin K. Bioorg Med Chem Lett 15:4741–4744

256. Gauthier JY, Chauret N, Cromlish W, Desmarais S, Duong LT, Falgueyret J-P, Kimmel DB, Lamontagne S, Leger S, LeRiche T, Li CS, Masse F, McKay DJ, Nicoll-Griffith DA, Oballa RM, Palmer JT, Percival MD, Riendeau D, Robichaud J, Rodan GA, Rodan SB, Seto C, Therien M, Truong V-L, Venuti MC, Wesolowski G, Young RN, Zamboni R, Black WC (2008) The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. Bioorg Med Chem Lett 18:923–928

257. Langdahl B, Binkley N, Bone H, Gilchrist N, Resch H, Rodriguez Portales J, Denker A, Lombardi A, Le Bailly De Tilleghem C, Da Silva C, Rosenberg E, Leung A (2012) Odanacatib in the treatment of postmenopausal women with low bone mineral density: five years of continued therapy in a phase 2 study. J Bone Mineral Res 27:2251–2258

258. Brixen K, Chapurlat R, Cheung AM, Keaveny TM, Fuers T, Engelke K, Recker R, Dardzinski B, Verbruggen N, Ather S, Rosenberg E, de Papp AE (2013) Bone density, turnover, and estimated strength in postmenopausal women treated with odanacatib: a randomized trial. J Clin Endocrinol Metab 98:571–580

259. Ng KW (2012) Potential role of odanacatib in the treatment of osteoporosis. Clin Intervent Aging 7:235–247

260. Isabel E, Mellon C, Boyd MJ, Chauret N, Deschênes D, Desmarais S, Falgueyret J-P, Gauthier JY, Khougaz K, Lau CK, Léger S, Levorse DA, Li CS, Massé F, Percival MD, Roy B, Scheigetz J, Thérien M, Truong VL, Wesolowski G, Young RN, Zamboni R, Black WC (2011) Difluoroethylamines as an amide isostere in inhibitors of cathepsin K. Bioorg Med Chem Lett 21:920–923

261. Grunewald GL, Seim MR, Lu J, Makboul M, Criscione KR (2006) Application of the Goldilocks effect to the design of potent and selective inhibitors of phenylethanolamine N-methyltransferase: balancing pK_a and steric effects in the optimization of 3-methyl-1,2,3,4-tetrahydroisoquinoline inhibitors by α-fluorination. J Med Chem 49:2939–2952

262. Wuitschik G, Rogers-Evans M, Müller K, Fischer H, Wagner B, Schuler F, Polonchuk L, Carreira EM (2006) Oxetanes as promising modules in drug discovery. Angew Chem Int Ed 45:7736–7739

263. Wuitschik G, Rogers-Evans M, Buckl A, Bernasconi M, Marki M, Godel T, Fischer H, Wagner B, Parrilla I, Schuler F, Schneider J, Alker A, Schweizer WB, Müller K, Carreira EM (2008) Spirocyclic oxetanes: synthesis and properties. Angew Chem Int Ed 47:4512–4515

264. Wuitschik G, Carreira EM, Wagner B, Fischer H, Parrilla I, Schuler F, Rogers-Evans M, Müller K (2010) Oxetanes in drug discovery: structural and synthetic insights. J Med Chem 53:3227–3246

265. Burkhard JA, Wuitschik G, Rogers-Evans M, Müller K, Carreira EM (2010) Oxetanes as versatile elements in drug discovery and synthesis. Angew Chem Int Ed 49:9052–9067

266. Jung ME, Piizzi G (2005) gem-Disubstituent effect: theoretical basis and synthetic applications. Chem Rev 105:1735–1766

267. Freire E (2008) Do enthalpy and entropy distinguish first in class from best in class? Drug Discov Today 13:869–874

268. Ladbury JE, Klebe G, Freire E (2010) Adding calorimetric data to decision making in lead discovery: a hot tip. Nat Rev Drug Discov 9:23–27

269. Lafont V, Armstrong AA, Ohtaka H, Kiso Y, Amzel LM, Freire E (2007) Compensating enthalpic and entropic changes hinder binding affinity optimization. Chem Biol Drug Des 69:413–422

270. Freire E (2009) A thermodynamic approach to the affinity optimization of drug candidates. Chem Biol Drug Des 74:468–472

271. Hann MM (2011) Molecular obesity, potency and other addictions in drug discovery. Med Chem Commun 2:349–355

272. Tarcsay A, Keserű GM (2013) Contributions of molecular properties to drug promiscuity. J Med Chem 56:1789–1795

273. Ritchie TR, Macdonald SJF (2009) The impact of aromatic ring count on compound developability – are too many aromatic rings a liability in drug design? Drug Discov Today 14:1011–1020

274. Ritchie TR, Macdonald SJF, Young RJ, Pickett SD (2011) The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. Drug Discov Today 16:164–171

275. Ritchie TJ, Macdonald SJF, Peace S, Pickett SD, Luscombe CN (2012) The developability of heteroaromatic and heteroaliphatic rings – do some have a better pedigree as potential drug molecules than others? Med Chem Commun 3:1062–1069

276. Brown A, Brown TB, Calabrese A, Ellis D, Puhalo N, Ralph M, Watson L (2010) Triazole oxytocin antagonists: identification of an aryloxyazetidine replacement for a biaryl substituent. Bioorg Med Chem Lett 20:516–520

277. Honma T, Hayashi K, Aoyama T, Hashimoto N, Machida T, Fukasawa K, Iwama T, Ikeura C, Ikuta M, Suzuki-Takahashi I, Iwasawa Y, Hayama T, Nishimura S, Morishima H (2001) Structure-based generation of a new class of potent Cdk4 inhibitors: new de novo design strategy and library design. J Med Chem 44:4615–4627

278. Furet P, Caravatti G, Guagnano V, Lang M. Meyer T, Schoepfer J (2008) Entry into a new class of protein kinase inhibitors by pseudo ring design. Bioorg Med Chem Lett 18:897–900

279. Guagnano V, Furet P, Spanka C, Bordas V, Le Douget M, Stamm C, Brueggen J, Jensen MR, Schnell C, Schmid H, Wartmann M, Berghausen J, Drueckes P, Zimmerlin A, Bussiere D, Murray J, Graus Porta D (2012) Discovery of 3-(2,6-dichloro-3,5-dimethoxy-phenyl)-1-{6-[4-(4-ethyl-piperazin-1-yl)-phenylamino]-pyrimidin-4-yl}-1-methyl-urea (NVP-BGJ398), a potent and selective inhibitor of the fibroblast growth factor receptor family of receptor tyrosine kinase. J Med Chem 54:7066–7083

280. Furet P, Bold G, Hofmann F, Manley P, Meyer T, Altmann K-H (2003) Identification of a new chemical class of potent angiogenesis inhibitors based on conformational considerations and database searching. Bioorg Med Chem Lett 13:2967–2971

281. Liu KKC, Huang X, Bagrodia S, Chen JH, Greasley S, Cheng H, Sun S, Knighton D, Rodgers C, Kristina Rafidi K, Zou A, Xiao J, Yan S (2011) Quinazolines with intra-molecular hydrogen bonding scaffold (iMHBS) as PI3K/mTOR dual inhibitors. Bioorg Med Chem Lett 21:1270–1274

282. Zhang G, Ren P, Gray NS, Sim T, Liu Y, Wang X, Che J, TianS-S SML, Spalding TA, Romeo R, Iskandar M, Chow D, Seidel HM, Karanewsky DS, He Y (2008) Discovery of pyrimidine benzimidazoles as Lck inhibitors: Part I. Bioorg Med Chem Lett 18:5618–5621

283. Wan Z, Boehm JC, Bower MJ, Kassis S, Lee JC, Zhao B, Adams JL (2003) N-Phenyl-N-purin-6-yl ureas: the design and synthesis of P38α MAP kinase inhibitors. Bioorg Med Chem Lett 13:1191–1194

284. Jansma A, Zhang Q, Li B, Ding Q, Uno T, Bursulaya B, Liu Y, Furet P, Gray NS, Geierstanger BH (2007) Verification of a designed intramolecular hydrogen bond in a drug scaffold by nuclear magnetic resonance spectroscopy. J Med Chem 50:5875–5877

285. Kuhn B, Mohr P, Stahl M (2010) Intramolecular hydrogen bonding in medicinal chemistry. J Med Chem 53:2601–2611

286. Mathieu S, Gradl SN, Ren L, Wen Z, Aliagas I, Gunzner-Toste J, Lee W, Pulk R, Zhao G, Alicke B, Boggs JW, Buckmelter AJ, Choo EF, Dinkel V, Gloor SL, Gould SE, Hansen JD, Hastings G, Hatzivassiliou G, Laird ER, Moreno D, Ran Y, Voegtli WC, Wenglowsky S, Grina J, Rudolph J (2012) Potent and selective aminopyrimidine-based B-Raf inhibitors with favorable physicochemical and pharmacokinetic properties. J Med Chem 55:2869–2881

287. Larsen AA, Lish PM (1964) A new bioisostere: alkylsulphonamidophenethanolamines. Nature 203:1283–1284

288. Hitchcock SA, Pennington LD (2006) Structure–brain exposure relationships. J Med Chem 49:7559–7583

289. Nirogi RVS, Daulatabad AV, Parandhama G, Mohammad S, Sastri KR, Shinde AK, Dubey PK (2010) Synthesis and pharmacological evaluation of aryl aminosulfonamide derivatives as potent 5-HT$_6$ receptor antagonists. Bioorg Med Chem Lett 20:4440–4443

290. Ahmed M, Briggs MA, Bromidge SM, Buck T, Campbell L, Deeks NJ, Garner A, Gordon L, Hamprecht DW, Holland V, Johnson CN, Medhurst AD, Mitchell DJ, Moss SF, Powles J, Seal JT, Stean TO, Stemp G, Thompson M, Trail B, Upton N, Winborn K, Witty DR (2005) Bicyclic heteroarylpiperazines as selective brain penetrant 5-HT$_6$ receptor antagonists. Bioorg Med Chem Lett 15:4867–4871

291. Zhao S-H, Berger J, Clark RD, Sethofer SG, Krauss NE, Brothers JM, Martin RS, Misner DL, Schwab AL (2007) 3,4-Dihydro-2H-benzo[1,4]oxazine derivatives as 5-HT$_6$ receptor antagonists. Bioorg Med Chem Lett 17:3504–3507

292. Nirogi R, Shinde A, Daulatabad A, Kambhampati R, Gudla P, Shaik M, Gampa M, Balasubramaniam S, Gangadasari P, Reballi V, Badange R, Bojja K, Subramanian R, Bhyrapuneni G, Muddana N, Jayarajan P (2012) Design, synthesis, and pharmacological evaluation of piperidin-4-yl amino aryl sulfonamides: novel, potent, selective, orally active, and brain penetrant 5-HT$_6$ receptor antagonists. J Med Chem 55:9255–9269

293. Odan M, Ishizuka N, Hiramatsu Y, Inagaki M, Hashizume H, Fujii Y, Mitsumori S, Morioka Y, Soga M, Deguchi M, Yasui K, Arimura A (2012) CB 1/2 dual agonists with 3-carbamoyl 2-pyridone derivatives as antipruritics: reduction of CNS side effects by introducing polar functional groups. Bioorg Med Chem Lett 22:2894–2897

294. Fulp A, Bortoff K, Zhang Y, Seltzman H, Snyder R, Maitra R (2011) Towards rational design of cannabinoid receptor 1 (CB1) antagonists for peripheral selectivity. Bioorg Med Chem Lett 21:5711–5714

295. Burgey CS, Robinson KA, Lyle TA, Sanderson PEJ, Lewis SD, Lucas BJ, Krueger JA, Singh R, Miller-Stein C, White RB, Wong B, Lyle EA, Williams PD, Coburn CA, Dorsey BD, Barrow JC, Stranieri MT, Holahan MA, Sitko GR, Cook JJ, McMasters DR, McDonough CM, Sanders WM, Wallace AA, Clayton FC, Bohn D, Leonard YM, Detwiler TJ Jr, Lynch JJ Jr, Yan Y, Chen Z, Kuo L, Gardell SJ, Shafer JA, Vacca JP (2003) Metabolism-directed optimization of 3-aminopyrazinone acetamide thrombin inhibitors. development of an orally bioavailable series containing P1 and P3 pyridines. J Med Chem 46:461–473

296. Reiner JE, Siev DV, Araldi G-L, Cui JJ, Ho JZ, Reddy KM, Mamedova L, Vu PH, Lee K-SS, Minami NK, Gibson TS, Anderson SM, Bradbury AE, Nolan TG, Semple JE (2002) Non-covalent thrombin inhibitors featuring P3-heterocycles with P1-monocyclic arginine surrogates. Bioorg Med Chem Lett 12:1203–1208

297. Cui JJ, Araldi G-L, Reiner JE, Reddy KM, Kemp SJ, Ho JZ, Siev DV, Mamedova L, Gibson TS, Gaudette JA, Minami NK, Anderson SM, Bradbury AE, Nolan TG, Semple JE (2002) Non-covalent thrombin inhibitors featuring P3-heterocycles with P1-bicyclic arginine surrogates. Bioorg Med Chem Lett 12:2925–2930

298. Zhou Q, Ruffoni A, Gianatassio R, Fujiwara Y, Sella E, Shabat D, Baran PS (2013) Direct synthesis of fluorinated heteroarylether bioisosteres. Angew Chem Int Ed 52:3949–3952

299. Austin RP, Barton P, Bonnert RV, Brown RC, Cage PA, Cheshire DR, Davis AM, Dougall IG, Ince F, Pairaudeau G, Young A (2003) QSAR and the rational design of long-acting dual D2-receptor/β2-adrenoceptor agonists. J Med Chem 46:3210–3220

300. Blake MI, Crespi HL, Katz JJ (1975) Studies with deuterated drugs. J Pharm Sci 64:367–391

301. Yarnell AT (2009) Heavy-hydrogen drugs turn heads, again. Firms seek to improve drug candidates by selective deuterium substitution. Chem Eng News 87:36–39

302. Shao L, Hewitt MC (2010) The kinetic isotope effect in the search for deuterated drugs. Drug News Perspect 23:398–404

303. Harbeson SL, Tung RD (2011) Deuterium in drug discovery and development. Ann Rep Med Chem 46:403–417

304. Katsnelson A (2013) Heavy drugs draw heavy interest from pharma backers. Nature Med 19:656

305. El Tayar N, van de Waterbeemd H, Gryllaki M, Testa B, Trager WF (1984) The lipophilicity of deuterium atoms. A comparison of shake-flask and HPLC methods. Int J Pharmaceutics 19:271–281

306. Turowski M, Yamakawa N, Meller J, Kimata K, Ikegami T, Hosoya K, Tanaka N, Thornton ER (2003) Deuterium isotope effects on hydrophobic interactions: the importance of dispersion interactions in the hydrophobic phase. J Am Chem Soc 125:13836–13849

307. Kimata K, Hosoya K, Araki T, Tanaka N (1997) Direct chromatographic separation of racemates on the basis of isotopic chirality. Anal Chem 69:2610–2612

308. Perrin CL, Ohta BK, Kuperman J (2003) β-Deuterium isotope effects on amine basicity, "inductive" and stereochemical. J Am Chem Soc 125:15008–15009

309. Perrin CL, Ohta BK, Kuperman J, Liberman J, Erdélyi M (2005) Stereochemistry of β-deuterium isotope effects on amine basicity. J Am Chem Soc 127:9641–9647

310. Perrin CL, Dong Y (2008) Nonadditivity of secondary deuterium isotope effects on basicity of triethylamine. J Am Chem Soc 130:11143–11148

311. Perrin CL, Dong Y (2007) Secondary deuterium isotope effects on the acidity of carboxylic acids and phenols. J Am Chem Soc 129:4490–4497

312. Foster AB (1984) Deuterium isotope effects in studies of drug metabolism. Trends Pharmacol Sci 5:524–527

313. Nelson SD, Trager WF (2003) The use of deuterium isotope effects to probe the active site properties, mechanism of cytochrome P450-catalyzed reactions, and mechanisms of metabolically dependent toxicity. Drug Metab Disp 31:1481–1498

314. Mutlib AE (2008) Application of stable isotope-labeled compounds in metabolism and in metabolism-mediated toxicity studies. Chem Res Toxicol 21:1672–1689

315. Belleau B, Burba J, Pindell M, Reiffenstein J (1961) Effect of deuterium substitution in sympathomimetic amine on adrenergic responses. Science 133:102–104

316. Westheimer FH (1961) The magnitude of the primary kinetic isotope effect for compounds of hydrogen and deuterium. Chem Rev 61:265–273

317. Peng S, van der Donk WA (2003) An unusual isotope effect on substrate inhibition in the oxidation of arachidonic acid by lipoxygenase. J Am Chem Soc 125:8988–8989

318. Gant TG, Sarshar S (2008) Substituted phenethylamines with serotonergic and/or norepinephrinergic activity. US Patent 7,456, 317, November 25, 2008

319. Tung R (2007) Novel benzo-[D] [1,3]-dioxol derivatives. World Patent Application WO 2007/016,431, February 8th, 2007

320. Murray M (2000) Mechanisms of inhibitory and regulatory effects of methylenedioxyphenyl compounds on cytochrome P450-dependent drug oxidation. Curr Drug Metab 1:67–84

321. Bertelsen KM, Venkatakrishnan K, von Moltke LL, Obach RS, Greenblatt DJ (2003) Apparent mechanism-based inhibition of human CYP 2D6 in vitro by paroxetine: comparison with fluoxetine and quinidine. Drug Metab Disp 31:289–293

322. Venkatakrishnan K, Obach RS (2005) In vitro-in vivo extrapolation of CYP2D6 inactivation by paroxetine: prediction of nonstationary pharmacokinetics and drug interaction magnitude. Drug Metab Disp 33:845–852

323. Manley PW, Blasco F, Mestan J, Aichholz R (2013) The kinetic deuterium isotope effect as applied to metabolic deactivation of imatinib to the des-methyl metabolite, CGP74588. Bioorg Med Chem 21:3231–3239

324. Shao L, Abolin C, Hewitt MC, Koch P, Varney M (2006) Derivatives of tramadol for increased duration of effect. Bioorg Med Chem Lett 16:691–694

325. Velthuisen EJ, Baughman TM, Johns BA, Temelkoff DP, Weatherhead JG (2013) Synthesis and pharmacokinetic profile of highly deuterated brecanavir analogs. Eur J Med Chem 63:202–212

326. Maltais F, Jung YC, Chen M, Tanoury J, Perni RB, Mani N, Laitinen L, Huang H, Liao S, Gao H, Tsao H, Block E, Ma C, Shawgo RS, Town C, Brummel CL, Howe D, Pazhanisamy S, Raybuck S, Namchuk M, Bennani YL (2009) In vitro and in vivo isotope

effects with hepatitis C protease inhibitors: enhanced plasma exposure of deuterated telaprevir versus telaprevir in rats. J Med Chem 52:7993–8001

327. Mutlib AE, Gerson RJ, Meunier PC, Haley PJ, Chen H, Gan LS, Davies MH, Gemzik B, Christ DD, Krahn DF, Markwalder JA, Seitz SP, Robertson RT, Miwa GT (2000) The species-dependent metabolism of efavirenz produces a nephrotoxic glutathione conjugate in rats. Toxicol Appl Pharmacol 169:102–113

328. Zhu Y, Zhou J, Jiao B (2013) Deuterated clopidogrel analogues as a new generation of antiplatelet agents. ACS Med Chem Lett 4:349–352

329. Laizure SC, Parker RB (2010) A comparison of the metabolism of clopidogrel and prasugrel. Expert Opin Drug Metab Toxicol 6:1417–1424

330. Bains W, Tacke R (2003) Silicon chemistry as a novel source of chemical diversity in drug design. Curr Opin Drug Discov Dev 6:526–543

331. Franz AK, Wilson SO (2013) Organosilicon molecules with medicinal applications. J Med Chem 56:388–405

332. Tacke R, Popp F, Müller B, Theis B, Burschka C, Hamacher A, Kassack MU, Schepmann D, Wünsch B, Jurva U, Wellner E (2008) Sila-haloperidol, a silicon analogue of the dopamine (D$_2$) receptor antagonist haloperidol: synthesis, pharmacological properties, and metabolic fate. ChemMedChem 3:152–164

333. Subramanyam B, Woolf T, Castagnoli N Jr (1991) Studies on the *in-vitro* conversion of haloperidol to a potentially neurotoxic pyridinium metabolite. Chem Res Toxicol 4:123–128

334. Johansson T, Weidolf L, Popp F, Tacke R, Jurva U (2010) In vitro metabolism of haloperidol and sila-haloperidol: new metabolic pathways resulting from carbon/silicon exchange. Drug Metab Disp 38:73–83

335. Tanaka H, Shishido Y (2007) Synthesis of aromatic compounds containing a 1,1-dialkyl-2-trifluoromethyl group, a bioisostere of the tert-alkyl moiety. Bioorg Med Chem Lett 17:6079–6085

336. Barnes-Seeman D, Jain M, Bell L, Ferreira S, Cohen S, Chen X-H, Amin J, Snodgrass B, Hatsis P (2013) Metabolically stable tert-butyl replacement. ACS Med Chem Lett 4:514–516

337. Yu K-L, Sin N, Civiello RL, Wang XA, Combrink KD, Gulgeze HB, Venables BL, Wright JJK, Dalterio RA, Zadjura L, Marino A, Dando S, D'Arienzo C, Kadow KF, Cianci CW, Li Z, Clarke J, Genovesi EV, Medina I, Lamb L, Colonno RJ, Yang Z, Krystal M, Meanwell NA (2007) Respiratory syncytial virus fusion inhibitors. Part 4: optimization for oral bioavailability. Bioorg Med Chem Lett 17:895–901

338. Skuballa W, Schillinger E, C-St S, Vorbrüggen H (1986) Synthesis of a new chemically and metabolically stable prostacyclin analogue with high and long-lasting oral activity. J Med Chem 29:313–315

339. Kim J-JP, Battaile KP (2002) Burning fat: the structural basis of fatty acid β-oxidation. Curr Opin Struct Biol 12:721–728

340. Świzdor A, Panek A, Milecka-Tronina N, Kołek T (2012) Biotransformations utilizing β-oxidation cycle reactions in the synthesis of natural compounds and medicines. Int J Mol Sci 13:16514–16543

341. Li C, Benet LZ, Grillo MP (2002) Studies on the chemical reactivity of 2-phenylpropionic acid 1-*O*-acyl glucuronide and *S*-acyl-CoA thioester metabolites. Chem Res Toxicol 5:1309–1317

342. Grillo MP (2011) Drug-*S*-acyl-glutathione thioesters: synthesis, bioanalytical properties, chemical reactivity, biological formation and degradation. Curr Drug Metab 12:229–244

343. Darnell M, Weidolf L (2013) Metabolism of xenobiotic carboxylic acids: focus on coenzyme A conjugation, reactivity, and interference with lipid metabolism. Chem Res Toxicol. doi:10.1021/tx400183y

344. Ballatore C, Huryn DM, Smith AB III (2013) Carboxylic acid (bio)isosteres in drug design. ChemMedChem 8:385–395

345. Hamada Y, Kiso Y (2012) The application of bioisosteres in drug design for novel drug discovery: focusing on acid protease inhibitors. Expert Opin Drug Discov 7:903–922

346. Stachulski AV, Harding JR, Lindon JC, Maggs JL, Park BK, Wilson ID (2006) Acyl glucuronides: biological activity, chemical reactivity, and chemical synthesis. J Med Chem 49:6931–6945
347. Regan SL, Maggs JL, Hammond TG, Lambert C, Williams DP, Park BK (2010) Acyl glucuronides: the good, the bad and the ugly. Biopharm Drug Dispos 31:367–395
348. Baba A, Yoshioka T (2009) Structure–activity relationships for degradation reaction of 1-β-O-acyl glucuronides: kinetic description and prediction of intrinsic electrophilic reactivity under physiological conditions. Chem Res Toxicol 22:158–172
349. Baba A, Yoshioka T (2009) Structure–activity relationships for degradation reaction of 1-β-O-acyl glucuronides. Part 2: Electronic and steric descriptors predicting the reactivity of 1-β-O-acyl glucuronides derived from benzoic acids. Chem Res Toxicol 22:1559–1569
350. Baba A, Yoshioka T (2009) Structure–activity relationships for degradation reaction of 1-β-O-acyl glucuronides. Part 3: Electronic and steric descriptors predicting the reactivity of aralkyl carboxylic acid 1-β-O-acyl glucuronides. Chem Res Toxicol 22:1998–2008
351. Sawamura R, Okudaira N, Watanabe K, Murai T, Kobayashi Y, Tachibana M, Ohnuki T, Masuda K, Honma H, Kurihara A, Okazaki O (2010) Predictability of idiosyncratic drug toxicity risk for carboxylic acid-containing drugs based on the chemical stability of acyl glucuronide. Drug Metab Disp 38:1857–1864
352. Krogsgaard-Larsen P, Ebert B, Lund TM, Briiuner-Osbome H, Slok FA, Johansen TN, Brehm L, Madsen U (1996) Design of excitatory amino acid receptor agonists, partial agonists and antagonists: ibotenic acid as a key lead structure. Eur J Med Chem 31:515–537
353. Pemberton N, Graden H, Evertsson E, Bratt E, Lepistö M, Johannesson P, Svensson PH (2012) Synthesis and functionalization of cyclic sulfonimidamides: a novel chiral heterocyclic carboxylic acid bioisostere. ACS Med Chem Lett 3:574–578
354. Petrillo E Jr, Ondetti M (1982) Angiotensin-converting enzyme inhibitors: medicinal chemistry and biological actions. Med Res Rev 2:1–41
355. Ishak R, Abbas O (2013) Penicillamine revisited: historic overview and review of the clinical uses and cutaneous adverse effects. Am J Clin Dermatol 14:223–233
356. Narjes F, Koehler KF, Koch U, Gerlach B, Colarusso S, Steinkühler C, Brunetti M, Altamura S, De Francesco R, Matassa VG (2002) A designed P1 cysteine mimetic for covalent and non-covalent inhibitors of HCV NS3 protease. Bioorg Med Chem Lett 12:701–704
357. Erickson JA, McLoughlin JI (1995) Hydrogen bond donor properties of the difluoromethyl group. J Org Chem 60:1626–1631
358. Albers HMHG, Ovaa H (2012) Chemical evolution of autotaxin inhibitors. Chem Rev 112:2593–2603
359. Mutoh T, Rivera R, Chun J (2012) Insights into the pharmacological relevance of lysophospholipid receptors. Br J Pharmacol 165:829–844
360. Xu Y, Qian L, Pontsler AV, McIntyre TM, Prestwich GD (2004) Synthesis of difluoromethyl substituted lysophosphatidic acid analogues. Tetrahedron 60:43–49
361. Xu Y, Prestwich GD (2002) Concise synthesis of acyl migration-blocked 1,1-difluorinated analogues of lysophosphatidic acid. J Org Chem 67:7158–7161
362. Wang T, Yin Z, Zhang Z, Bender JA, Yang Z, Johnson G, Yang Z, Zadjura LM, D'Arienzo CJ, DiGiugno PD, Gesenberg C, Yamanaka GA, Gong Y-F, Ho H-T, Fang H, Zhou N, McAuliffe BV, Eggers BJ, Fan L, Nowicka-Sans B, Dicker IB, Gao Q, Colonno RJ, Lin P-F, Meanwell NA, Kadow JF (2009) Inhibitors of human immunodeficiency virus type 1 (HIV-1) attachment. 5. An evolution from indole to azaindoles leading to the discovery of 1-(4-benzoylpiperazin-1-yl)-2-(4,7-dimethoxy-1H-pyrrolo[2,3-c]pyridin-3-yl)ethane-1,2-dione (BMS-488043), a drug candidate that demonstrates antiviral activity in HIV-1-infected subjects. J Med Chem 52:7778–7787
363. Kalgutkar AS, Griffith DA, Ryder T, Sun H, Miao Z, Bauman JN, Didiuk MT, Frederick KS, Zhao SX, Prakash C, Soglia JR, Bagley SW, Bechle BM, Kelley RM, Dirico K, Zawistoski M, Li J, Oliver R, Guzman-Perez A, Liu KKC, Walker DP, Benbow JW, Morris J (2010)

Discovery tactics to mitigate toxicity risks due to reactive metabolite formation with 2-(2-hydroxyaryl)-5-(trifluoromethyl)pyrido[4,3-*d*]pyrimidin-4(3*H*)-one derivatives, potent calcium-sensing receptor antagonists and clinical candidate(s) for the treatment of osteoporosis. Chem Res Toxicol 23:1115–1126

364. Southers JA, Bauman JN, Price DA, Humphries PS, Balan G, Sagal JF, Maurer TS, Zhang Y, Oliver R, Herr M, Healy DR, Li M, Kapinos B, Fate GD, Riccardi KA, Paralkar VM, Brown TA, Kalgutkar AS (2010) Metabolism-guided design of short-acting calcium-sensing receptor antagonists. ACS Med Chem Lett 1:219–223

365. Hartz RA, Ahuja VT, Zhuo X, Mattson RJ, Denhart DJ, Deskus JA, Vrudhula VM, Pan S, Ditta JL, Shu Y-Z, Grace JE, Lentz KA, Lelas S, Li Y-W Molski TF, Krishnananthan S, Wong H, Qian-Cutrone J, Schartman R, Denton R, Lodge NJ, Zaczek R, Macor JE, Bronson JJ (2009) A strategy to minimize reactive metabolite formation: discovery of (*S*)-4-(1-cyclopropyl-2-methoxyethyl)-6-[6-(difluoromethoxy)-2,5-dimethylpyridin-3-ylamino]-5-oxo-4,5-dihydropyrazine-2-carbonitrile as a potent, orally bioavailable corticotropin-releasing factor-1 receptor antagonist. J Med Chem 52:7653–7658

366. Hartz RA, Ahuja VT, Schmitz WD, Molski TF, Mattson GK, Lodge NJ, Bronson JJ, Macor JE (2010) Synthesis and structure–activity relationships of N^3-pyridylpyrazinones as corticotropin-releasing factor-1 (CRF1) receptor antagonists. Bioorg Med Chem Lett 20:1890–1894

367. Zhuo X, Hartz RA, Bronson JJ, Wong H, Ahuja VT, Vrudhula VM, Leet JE, Huang S, Macor JE, Shu Y-Z (2010) Comparative biotransformation of pyrazinone-containing corticotropin-releasing factor receptor-1 antagonists: minimizing the reactive metabolite formation. Drug Metab Disp 38:5–15

368. Atwal KS, Rovnyak GC, Schwartz J, Moreland S, Hedberg A, Gougoutas JZ, Malley MM, Floyd DM (1990) Dihydropyrimidine calcium channel blockers: 2-heterosubstituted 4-aryl-1,4-dihydro-6-methyl-5-pyrimidinecarboxyaliccid esters as potent mimics of dihydropyridines. J Med Chem 33:1510–1515

369. Atwal KS, Rovnyak GC, Kimball SD, Floyd DM, Moreland S, Swanson BN, Gougoutas JZ, Schwartz J, Smillie KM, Malley MF (1990) Dihydropyrimidine calcium channel blockers. 2. 3-Substituted-4-aryl-1,4-dihydro-6-methyl-5-pyrimidinecarboxyalciicd esters as potent mimics of dihydropyridines. J Med Chem 33:2629–2635

370. Atwal KS, Swanson BN, Unger SE, Floyd DM, Moreland S, Hedberg A, O'Reilly BC (1991) Dihydropyrimidine calcium channel blockers. 3. 3-Carbamoyl-4-aryl-1,2,3,4-tetrahydro-6-methyl-5-pyrimidinecarboxylic acid esters as orally effective antihypertensive agents. J Med Chem 34:806–811

371. Rovnyak GC, Atwal KS, Hedberg A, Kimball SD, Moreland S, Gougoutas JZ, O'Reilly BC, Schwartz J, Malley MF (1992) Dihydropyrimidine calcium channel blockers. 4. Basic 3-substituted-4-aryl-1,4-dihydropyrimidine-5-carboxylic acid esters. Potent antihypertensive agents. J Med Chem 35:3254–3263

372. Wood MR, Schirripa KM. Kim JJ, Wan B-L, Murphy KL, Ransom RW, Chang RSL, Tang C, Prueksaritanont T, Detwiler TJ, Hettrick LA, Landis ER, Leonard YM, Krueger JA, Lewis SD, Pettibone DJ, Freidinger RM, Bock MG (2006) Cyclopropylamino acid amide as a pharmacophoric replacement for 2,3-diaminopyridine. Application to the design of novel bradykinin B1 receptor antagonists. J Med Chem 49:1231–1234

373. Testa B, Pedretti A, Vistoli G (2012) Reactions and enzymes in the metabolism of drugs and other xenobiotics. Drug Discov Today 17:549–560

374. Testa B, Krämer SD (2007) The biochemistry of drug metabolism – an introduction Part 3. Reactions of hydrolysis and their enzymes. Chem Biodivers 4:2031–2122

375. Testa B, Krämer SD (2007) The biochemistry of drug metabolism – an introduction Part 2. Redox reactions and their enzymes. Chem Biodivers 4:257–405

376. Guengerich FP (2001) Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. Chem Res Toxicol 14:611–650

377. Peng HM, Raner GM, Vaz ADN, Coon MJ (1995) Oxidative cleavage of esters and amides to carbonyl products by cytochrome P450. Arch Biochem Biophys 318:333–339

378. Watjen F, Baker R, Engelstoff M, Herbert R, MacLeod A, Knight A, Merchant K, Moseley J, Saunders J, Swain CJ, Wong E, Springer JP (1989) Novel benzodiazepine receptor partial agonists: oxadiazolylimidazobenzodiazepines. J Med Chem 32:2282–2291

379. Saunders J, Cassidy M, Freedman SB, Harley EA, Iversen LL, Kneen C, MacLeod AM, Merchant KJ, Snow RJ, Baker R (1990) Novel quinuclidine-based ligands for the muscarinic cholinergic receptor. J Med Chem 33:1128–1138

380. Krogsgaard-Larsen P, Falch E, Sauerberg P, Freedman SB, Lembøl HL, Meier E (1988) Bioisosteres of arecoline as novel CNS-active muscarinic agonists. Trends Pharmacol Sci Suppl:69–74

381. Kim J, Ok T, Park C, So W, Jo M, Kim Y, Seo M, Lee D, Jo S, Ko Y, Choi I, Park Y, Yoon J, Kyeong Ju M, Ahn JY, Kim J, Han S-J, Kim T-H, Cechetto J, Nam J, Liuzzi M, Sommer P, No Z (2012) A novel 3,4-dihydropyrimidin-2($1H$)-one: HIV-1 replication inhibitors with improved metabolic stability. Bioorg Med Chem Lett 22:2522–2526

382. Rajapakse HA, Nantermet PG, Selnick HG, Barrow JC, McGaughey GB, Munshi S, Lindsley SR, Young MB, Ngo PL, Holloway MK, Lai M-T, Espeseth AS, Shi X-P, Colussi D, Pietrak B, Crouthamel M-C, Tugusheva K, Huang Q, Xu M, Simon AJ, Kuo L, Hazuda DJ, Graham S, Vacca JP (2010) SAR of tertiary carbinamine derived BACE1 inhibitors: role of aspartate ligand amine pK_a in enzyme inhibition. Bioorg Med Chem Lett 20:1885–1889

383. Feil SC, Hamilton S, Krippner GY, Lin B, Luttick A, McConnell DB, Nearn R, Parker MW, Ryan J, Stanislawski PC, Tucker SP, Watson KG, Morton CJ (2012) An orally available 3-ethoxybenzisoxazole capsid binder with clinical activity against human rhinovirus. ACS Med Chem Lett 3:303–307

384. Sun Z-Y, Asberom T, Bara T, Bennett C, Burnett D, Chu I, Clader J, Cohen-Williams M, Cole D, Czarniecki M, Durkin J, Gallo G, Greenlee W, Josien H, Huang X, Hyde L, Jones N, Kazakevich I, Li H, Liu X, Lee J, MacCoss M, Mandal MB, McCracken T, Nomeir A, Mazzola R, Palani A, Parker EM, Pissarnitski DA, Qin J, Song L, Terracina G, Vicarel M, Voigt J, Xu R, Zhang L, Zhang Q, Zhao Z, Zhu X, Zhu Z (2012) Cyclic hydroxyamidines as amide isosteres: discovery of oxadiazolines and oxadiazines as potent and highly efficacious γ-secretase modulators in vivo. J Med Chem 55:489–502

385. Valverde IE, Bauman A, Kluba CA, Vomstein S, Walter MA, Mindt TL (2013) 1,2,3-Triazoles as amide bond mimics: triazole scan yields protease-resistant peptidomimetics for tumor targeting. Angew Chem Int Ed 52:8957–8960

386. Xia G, You X, Liu L, Liu H, Wang J, Shi Y, Li P, Xiong B, Liu X, Shen J (2013) Design, synthesis and SAR of piperidyl-oxadiazoles as 11β-hydroxysteroid dehydrogenase 1 inhibitors. Eur J Med Chem 62:1–10

387. Staben ST, Blaquiere N, Tsui V, Kolesnikov A, Do S, Bradley EK, Dotson J, Goldsmith R, Heffron TP, Lesnick J, Lewis C, Murray J, Nonomiya J, Olivero AG, Pang J, Rouge L, Salphati L, Wei BQ, Wiesmann C, Wu P (2013) Cis-Amide isosteric replacement in thienobenzoxepin inhibitors of PI3-kinase. Bioorg Med Chem Lett 23:897–901

388. Taylor R, Mullaley A, Mullier GW (1990) Use of crystallographic data in searching for isosteric replacements: composite crystal-field environments of nitro and carbonyl groups. Pestic Sci 29:197–213

389. Hardcastle IR, Liu J, Valeur E, Watson A, Ahmed SU, Blackburn TJ, Bennaceur K, Clegg W, Drummond C, Endicott JA, Golding BT, Griffin RJ, Gruber J, Haggerty K, Harrington RW, Hutton C, Kemp S, Lu X, McDonnell JM, Newell DR, Noble MEM, Payne SL, Revill CH, Riedinger C, Xu Q, Lunec J (2011) Isoindolinone inhibitors of the murine double minute 2 (MDM2)-p53 protein–protein interaction: structure–activity studies leading to improved potency. J Med Chem 54:1233–1243

390. Firestine SM, Davisson VJ (1993) A tight binding inhibitor of 5-aminoimidazole ribonucleotide carboxylase. J Med Chem 36:3484–3486

391. Thorson JS, Chapman E, Schultz PG (1995) Analysis of hydrogen bonding strengths in proteins using unnatural amino acids. J Am Chem Soc 117:9361–9362
392. Alston TA, Porter DJT, Bright HJ (1983) Enzyme inhibition by nitro and nitroso compounds. Acc Chem Res 16:418–424
393. Julémont F, de Leval X, Michaux C, Damas J, Charlier C, Durant F, Pirotte B, Dogné J-M (2002) Spectral and crystallographic study of pyridinic analogues of nimesulide: determination of the active form of methanesulfonamides as COX-2 selective inhibitors. J Med Chem 45:5182–5185
394. Renard J-F, Arslan D, Garbacki N, Pirotte B, de Leval X (2009) Pyridine analogues of nimesulide: design, synthesis, and in vitro and in vivo pharmacological evaluation as promising cyclooxygenase 1 and 2 inhibitors. J Med Chem 52:5864–5871
395. Gougoutas JZ, Ojala WH, Malley MF (1982) 3,4-*bis*(4-Chlorophenyl)sydnone. Cryst Struct Comm 11:1731–1736

Erratum to: Improving Solubility via Structural Modification

Michael A. Walker

Erratum to: Top Med Chem
   DOI: 10.1007/7355_2013_32

1. On page 69, Michael A. Walker's affiliation was incorrect. The correct information is given below:

   Department of Discovery Chemistry, Bristol Myers-Squibb, 5 Research Pkwy, Wallingford, CT 06492, USA

2. On page 105, reference [64] had incorrect information. The correct information is given below:

   64. Zhang L, Zhu H, Mathiowetz A, Gao H et al (2011) Deep understanding of structure–solubility relationship for a diverse set of organic compounds using matched molecular pairs. Bioorg Med Chem 19:5763–5770

The online version of the original chapter can be found under
DOI: 10.1007/7355_2013_32

M.A. Walker
Department of Discovery Chemistry, Bristol Myers-Squibb, 5 Research Pkwy, Wallingford, CT 06492, USA
e-mail: michael.a.walker@bms.com

Erratum to: Drug Discovery vs hERG

Derek J. Leishman and Zoran Rankovic

Erratum to: Top Med Chem
   DOI: 10.1007/7355_2014_38

1. On page 227, 4th line below Figure 1:
     Replace "CMP" with "CPMP"
2. On page 233, 1st line below heading 2.2 In Vitro and In Vivo Assays:
     Replace "CHMP" with "CPMP"
3. On page 236, 10th line below subhead 2.2.3 Effects of hERG Block in In Vitro Tissue Assays:
     Replace reference citation "[4]" with "[3]"

The online version of the original chapter can be found under
DOI: 10.1007/7355_2014_38

D.J. Leishman and Z. Rankovic (✉)
Eli Lilly, 893 South Delaware Street, Indianapolis, IN 46285, USA
e-mail: rankovic_zoran@lilly.com

Erratum to: Drug-Induced Phospholipidosis: Prediction, Detection, and Mitigation Strategies

Umesh M. Hanumegowda and Alicia Regueiro-Ren

Erratum to: Top Med Chem
DOI: 10.1007/7355_2013_34

On page 281, references [36] and [37] had incorrect information. The correct information is given below:

36. Ishichi Y, Kimura E, Honda E, Yoshikawa M, Nakahata T, Terao Y, Suzuki A, Kawai T, Arakawa Y, Ohta H, Kanzaki N, Nakagawa H, Terauchi J (2013) Novel triple reuptake inhibitors with low risk of CAD associated liabilities: design, synthesis and biological activities of 4-[(1S)-1-(3,4-dichlorophenyl)-2-methoxyethyl]piperidine and related compounds. Bioorg Med Chem 21:4600–4613

37. Ikuma Y, Hochigai H, Kimura H, Nunami N, Kobayashi T, Uchiyama K, Furuta Y, Sakai M, Horiguchi M, Masui Y, Okazaki K, Sato Y, Nakahira H (2012) Discovery of 3H-imidazo [4,5-c]quinolin-4(5H)-ones as potent and selective dipeptidyl peptidase IV (DPP-4) inhibitors. Bioorg Med Chem 20:5864–5883

The online version of the original chapter can be found under
DOI: 10.1007/7355_2013_34

U.M. Hanumegowda
Discovery Toxicology, Bristol-Myers Squibb Research and Development, Wallingford, CT 06492, USA

A. Regueiro-Ren (✉)
Discovery Chemistry, Bristol-Myers Squibb Research and Development, Wallingford, CT 06492, USA
e-mail: Alicia.regueiroren@bms.com

Erratum to: The Influence of Bioisosteres in Drug Design: Tactical Applications to Address Developability Problems

Nicholas A. Meanwell

Erratum to: Top Med Chem
   DOI: 10.1007/7355_2013_29

1. On page 289, last paragraph from the bottom, last line from the top of the paragraph:
   Replace reference citations "[46–49]" with "[10, 46–49]"
2. On page 307, the legend for Figure 13 was incorrectly placed. It needs to be moved to the bottom of the next figure on the same page.
3. On page 316, figure 19 was incorrect. The correct version is below:
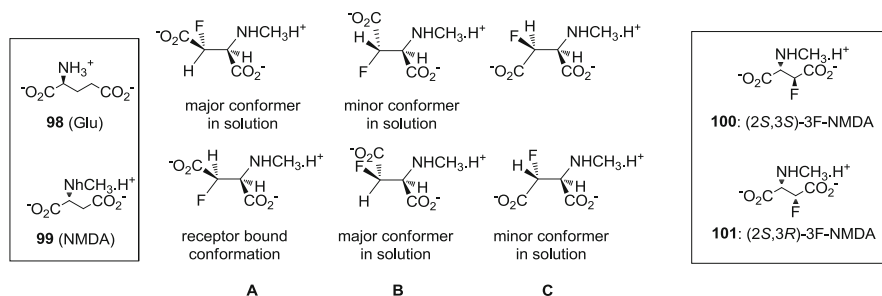
N.A. Meanwell (✉)

Department of Discovery Chemistry, Bristol-Myers Squibb Research and Development,
5 Research Parkway, Wallingford, CT 06492, USA
e-mail: Nicholas.Meanwell@bms.com

4. On page 322, 2nd line below Fig. 24 to match Fig. 24:
   Replace "conformation A" with "conformation a"
5. On page 322, 5nd line below Fig. 24 to match Fig. 24:
   Replace "conformation B" with "conformation b"
6. On page 333, 4th line below Fig. 29:
   Replace "**168**" with "**166**"
7. On pages 360, 372, 377, 379, 387, references [42], [247], [343], [368] and [387] had incorrect information. The correct information is given below:

   42. Naylor EM, Chakravarty PK, Costello CA, Chang RS, Chen T-B, Faust KA, Lotti VJ, Kivlighn SD, Zingaro GJ, Siegl PKS, Wong PC, Carini DJ, Wexler RR, Patchett AA, Greenlee WJ (1994) Potent imidazole angiotensin II antagonists: acyl sulfonamides and acyl sulfamides as tetrazole replacements. Bioorg Med Chem Lett 4:69–74

   247. Morgenthaler M, Schweizer E, Hoffman-Röder A, Benini F, Martin RE, Jaeschke G, Wagner B, Fischer H, Bendels S, Zimmeli D, Scheider J, Diedeich F, Kansy M, Müller K (2007) Predicting and tuning physicochemical properties in lead optimization: amine-basicities. ChemMedChem 2:1100–1115

   343. Darnell M, Weidolf L (2013) Metabolism of xenobiotic carboxylic acids: focus on coenzyme A conjugation, reactivity, and interference with lipid metabolism. Chem Res Toxicol 26:1139–1155

   368. Atwal KS, Rovnyak GC, Schwartz J, Moreland S, Hedberg A, Gougoutas JZ, Malley MM, Floyd DM (1990) Dihydropyrimidine calcium channel blockers: 2-heterosubstituted 4-aryl-1,4-dihydro-6-methyl-5-pyrimidinecarboxylic acid esters as potent mimics of dihydropyridines. J Med Chem 33:1510–1515

   387. Staben ST, Blaquiere N, Tsui V, Kolesnikov A, Do S, Bradley EK, Dotson J, Goldsmith R, Heffron TP, Lesnick J, Lewis C, Murray J, Nonomiya J, Olivero AG, Pang J, Rouge L, Salphati L, Wei BQ, Wiesmann C, Wu P (2013) cis-Amide isosteric replacement in thienobenzoxepin inhibitors of PI3-kinase. Bioorg Med Chem Lett 23:897–901

Index