Physicochemical and Biomimetic Properties in Drug Discovery

Chromatographic Techniques for Lead Optimization

Klara Valko

WILEY

PHYSICOCHEMICAL AND BIOMIMETIC PROPERTIES IN DRUG DISCOVERY

PHYSICOCHEMICAL AND BIOMIMETIC PROPERTIES IN DRUG DISCOVERY

Chromatographic Techniques for Lead Optimization

Klara Valko

CSC-Analytical Chemistry
Platform Technology & Sciences
GlaxoSmithKline, United Kingdom

and

Department of Pharmaceutical and Biological Chemistry
UCL, School of Pharmacy
London, United Kingdom

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Valko, Klara
Physicochemical and biomimetic properties in drug discovery : chromatographic techniques for lead optimization / Klara Valko, CSC-Analytical Chemistry, Platform Technology & Sciences, GlaxoSmithKline, United Kingdom, and Department of Pharmaceutical and Biological Chemistry, UCL, School of Pharmacy, London, United Kingdom.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-15212-6 (cloth)
1. Drug development–Statistical methods. 2. Chromatographic analysis. 3. Clinical trials–Statistical methods. 4. Pharmaceutical biotechnology. I. Title.
RM301.25.V35 2014
615.1′9–dc23

2013027975

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

During the past 10 years I have been teaching the Physchem/ADME (absorption, distribution, metabolism, and elimination) module for "Drug Discovery M.Sc." students at the School of Pharmacy, University of London (University College London since 2012). The module covers ADME and the underlying physicochemical properties of drugs. This book is intended to summarize the course material, providing detailed explanations of the physicochemical aspects of drug absorption and distribution *in vivo*. It is well recognized now that drug molecules occupy a relatively small physicochemical property space in comparison to the huge number of possible physicochemical entities. Lipophilicity, solubility, permeability, and the charge state of molecules are the most important properties that influence absorption and *in vivo* distribution. Thus, the measurement and calculation of lipophilicity, solubility, permeability, and the charge state are essential early on in the drug discovery process in order to select compounds for further studies. Compounds selected in this way have the best possible chance to make it to development and eventually help patients recover or at least improve their quality of life. The principles of the measurement of physicochemical and biomimetic properties are explained in more detail. Special emphasis is given to the chromatographic measurements of various physicochemical and biomimetic properties of drugs. It will be shown how to interpret the chromatographic retention data by putting them into various models for the estimation of *in vivo* distribution behavior of compounds. Other techniques are also mentioned, and further references are provided but not discussed in detail.

The chromatographic dynamic equilibrium process provides an excellent model to describe a compound's *in vivo* distribution between the moving plasma/blood compartment and the stationary tissue compartments. The reader will learn that chromatography is a powerful technique not only for the separation of closely related compounds but it is also very useful to measure a compound's interactions with various biomimetic stationary phases covered with alkyl chain, proteins, and phospholipids. The obtained binding data are suitable for deriving quantitative structure–physicochemical property relationships (QSPR) that can be used in drug design. Mathematical models can be constructed for the estimation of a compound's *in vivo* distribution. Active transporters and other specific binding of the

molecules to various types of proteins and phospholipids can, of course, modulate the estimated distribution, and these potential interactions need to be taken into account during the lead optimization process. When significant differences are observed between the *in vivo* measured and estimated distribution behavior of a compound based on physicochemical properties, it indicates to the drug discovery scientist that the compound undergoes some active transport mechanism. The active transport process can push the compound concentration from the thermodynamic equilibrium, but it requires constant energy investment from the body via biochemical processes.

This book contains previously published ideas, concepts, and methodologies. This literature review is far from comprehensive; it is rather critical and selective in order to reveal various ways to approach problems. However, the book provides simple descriptions and explanations of the essence of selected publications that usually help students understand the methods, results, and conclusions of scientific papers. A chapter dedicated to chromatography describes the basic principles and practical considerations that are needed to set up and run the measurements that are discussed in later chapters. Examples are shown how to use experimental data in various models of absorption, tissue binding, volume of distribution, and other *in vivo* ADME characteristics of the drugs. The property data obtained by fully automated chromatographic measurements can be used to select and prioritize compounds for further *in vitro* and *in vivo* studies at early stages of the drug discovery process. A special chapter is dedicated for the interpretation of the data, the interrelationship between the physicochemical properties, and some structure–property relationships in the hope that medicinal chemists can use the information for designing new potential drug molecules.

Although there are several books that cover similar topics, this book is unique in many ways. First, the chromatographic technique is normally described as an analytical separation method, and it is rarely discussed as a tool for property measurements of drug discovery compounds. There are hundreds of research papers and a few reviews that demonstrate the usefulness of the technique for such purposes but the general explanations together with a deeper understanding has not been published in a textbook yet. Sufficient detail is provided to enable the reproduction of chromatographic measurements in any laboratory equipped with HPLC. Second, this book is unique in the sense that it contains some new insight into the interdisciplinary knowledge needed for designing efficient drugs with minimal side effects. The intention is to reveal the relationships between various disciplines such as physical chemistry, analytical chemistry, biology, and pharmacokinetics/pharmacodynamics by the underlying basic thermodynamic laws.

Communication between scientists from different disciplines is hindered not only by differences in vocabulary but also many times by the use of different units. For example, the first measurements of activity of a compound is usually expressed as a pIC_{50}, which means a quantity of the compound expressed as the negative logarithm of the molar concentration that causes 50% of the maximum inhibition on a particular target or enzyme. Then we measure the solubility in millimolar or

milligram per milliliter units as the maximum soluble concentration of the drug. Experts in drug metabolism and pharmacokinetic (DMPK) measure the quantity of drug absorbed as a percentage, the plasma protein binding is measured as unbound fraction, and volume of distribution as liters per kilogram. Finally, the patient takes a quantity of drug (dose) expressed as an *X* mg tablets three times a day. In all of the above instances, we would like to express the amount or concentration of drug molecules. Using different units in various phases of the drug discovery may hinder our understanding. I do not think we can change this practice in the near future. However, at present, we can learn to convert the units and understand the relationships between various measurements. The mathematical and physical–chemical rules serve as links between disciplines in this book. It is demonstrated that mathematical approaches (paying special attention to the units of the measurements) help interdisciplinary thinking, converting, and translating knowledge between scientists working in different fields. In this way, we can maximize our understanding of the huge amount of data that are normally generated during the drug discovery process.

Chapter 12 presents examples, mostly using historical data of successful known drug molecules or published project examples from several pharmaceutical companies. At the end of each chapter, there is a short summary containing the conclusions and help revision for students taking examinations of this material. Finally, some typical examination questions are added at the end of each chapter to be able to test the reader's comprehension of the material. The answers are provided in Appendix A.

Some of the conclusions have been made on the basis of the author's personal view and experience in the subject. These do not always agree with the accepted and widely used theories in big pharmaceutical companies. Ideas for further research areas, unanswered questions, and hypotheses are presented *in italics* in order to encourage and motivate readers to form their own opinion. The book contains several questions that were raised by the students during the lessons and the answers still have to be found by the next generation of scientists. Therefore, I hope that the text raises the motivation and enthusiasm of scientists who are willing to engage themselves in drug discovery for a long period.

This book provides the essential experimental details for the determination of physicochemical and biomimetic properties that can be carried out in any analytical laboratories equipped with HPLC. Medicinal chemists will be able to understand these properties and the structure–property relationships described here and successfully use them in drug design and lead optimization. DMPK and ADME scientists will be able to interpret the results of their *in vitro* and *in vivo* experiments after thorough comprehension of the content, and it will help them in designing the necessary experiments to support the drug discovery process. I am aware that several drug discovery institutions and universities have already set up the chromatography-based determination of lipophilicity, protein binding, and phospholipid binding. I hope that with the help of this book, they will be able to use these properties in their full potential for the estimation of *in vivo* properties, thus reducing the need for animal experiments.

Finally, I would like to thank all those who have been of tremendous help during the preparation of this book. I am grateful to my past and present colleagues at GlaxoSmithKline, especially to Derek Reynolds, Chris Bevan, and Alan Hill with whom I gained experience in the field of physicochemistry and its application in drug discovery and who have supported me in many ways. I would like to thank my colleagues Shenaz Bunally, Elisabetta Chiarparin, and Paul Leeson, who read the manuscript and gave me excellent advice to improve it. I would like to acknowledge the help I have received from the colleagues at the School of Pharmacy, University College London, Professor David Thurston, Professor Simon Gibbons, Professor Anne Stevenson, Dr. Michael Munday, Dr. Mire Zloh, and Dr. Rosemary Smyth.

I have received the greatest inspiration for the preparation of this book from my students. Many of them are extremely talented, providing me with the hope that they will continue to improve drug discovery and will help millions of patients who desperately need cure for their disease. Drug Discovery M.Sc. students from year 2011 to 2012 have gone through the manuscript and provided excellent feedback where I need to explain the material in more detail or more clearly. I would like to thank especially Godfrey Mayoka, Samar Youshif, and Hajir Azam for carefully reading the manuscript. I have used and referenced the project works of Manju Kalyani and Andriana Rapti, who carefully investigated and validated the methodology for solubility determination and α-1-acid glycoprotein binding measurements during their M.Sc. studies.

I am very grateful for the well-known experts who provided valuable suggestions and corrections for specific chapters. Professor Krisztina Takács-Novák, School of Pharmacy, Semmelweis University, has reviewed chapters on solubility and pK_a; Dr. Alex Avdeef, In Adme, Ltd., has reviewed the chapter on permeability; Professor Michael Abraham, University College London; has reviewed the chapter on lipophilicity. They are recognized experts in these fields and their valuable corrections and suggestions are appreciated.

I would like to thank my family and friends for their emotional support and understanding, especially to my son Adam Valko who coped very well with his studies and his life on his own during the preparation of this book.

1

THE DRUG DISCOVERY PROCESS

The way we discover drugs now is very different from the process followed 30–50 years ago [1–3]. In the past, the drug discovery process was based on the structure of known active compounds. The structure of the active compound, which was either endogenous (e.g., acetylcholine, adrenaline, and steroid hormones) or a natural product (e.g., morphine, papaverine, cocaine, atropine, and digitalis glycosides), was identified and modified by a trial-and-error method. Scientists isolated the active material, revealed the structure and the mechanism of the pharmacological action, and based on this knowledge they would modify the chemical structures with the aim of improving the activity, decreasing the toxicity, and increasing the duration of action. Using this approach, they discovered, for example, a series of cholinesterase inhibitors, a series of adrenaline analogs for sympatomimetic or sympatolytic compounds, semisynthetic steroid hormones, and many more. Another example is pethidine containing the pharmacophores of morphine. The ethoxy derivative of papaverine called *No-Spa* has a longer half-life and is a stronger spasmolytic agent. Figure 1.1 shows a few analog-based drug molecules designed from natural compounds. There are numerous examples of naturally occurring molecules or their derivatives being used as drugs, such as warfarin (anticoagulant derived from dicumarol found in sweet clover, hirudin (anticoagulant from leeches), statins (a fungal metabolite that reduces plasma

Physicochemical and Biomimetic Properties in Drug Discovery: Chromatographic Techniques for Lead Optimization, First Edition. Klara Valko.

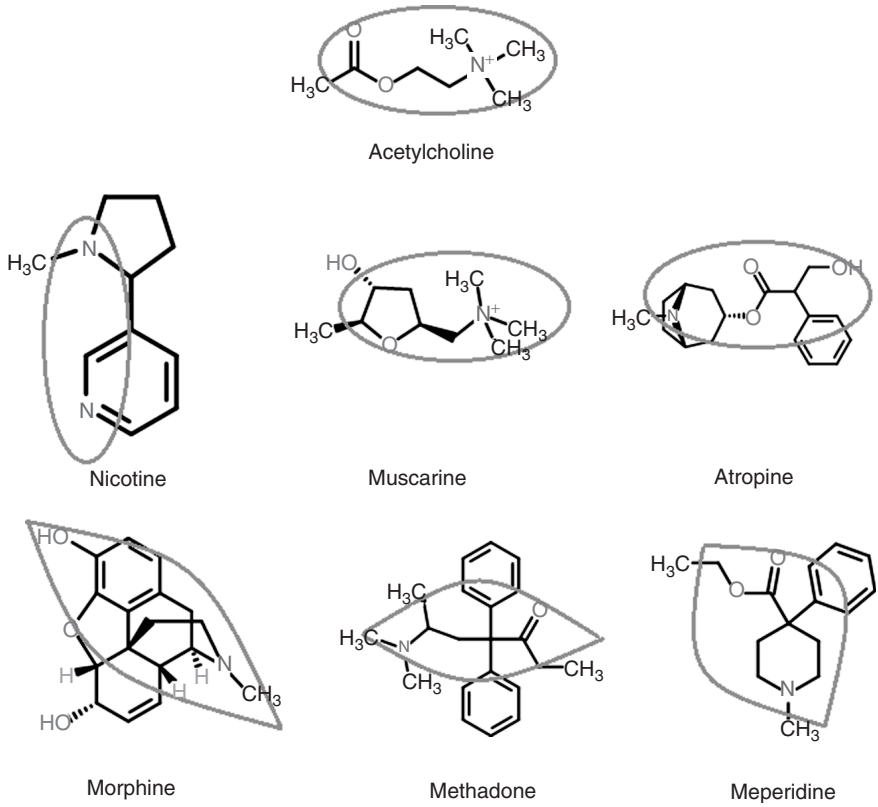

Figure 1.1 Drug molecules designed as analogs of natural active molecules containing the active pharmacophore.

cholesterol), vinca alkaloids (anticancer drugs isolated from periwinkle plant families), and antibiotics (derived from fungal metabolites).

Today, drug discovery utilizes recent advances in biochemistry and genetics. Drug discovery usually starts with discovering a so-called "target" that can be enzymes or receptors, such as G protein-coupled receptors (GPCR) and enzyme targets, such as kinases, ion channels, and hormone receptors.

The medicinal chemists and computational chemists work closely together to generate ideas about potential active molecules that would fit to the "target" molecule. They develop a so-called high throughput screening (HTS) method that enables the big pharmaceutical companies to screen large number of compounds against a particular target. Compounds that show activity in HTS are called *hits*. The activity of the "hits" is further evaluated by repeated experiments for obtaining the standard concentration–potency curve. Figure 1.2 shows a typical concentration–response/potency profile during hit generation and confirmation.

This phase of the drug discovery process is often called *hit generation*. Synthetic chemists then synthesize a whole library of molecules, making numerous

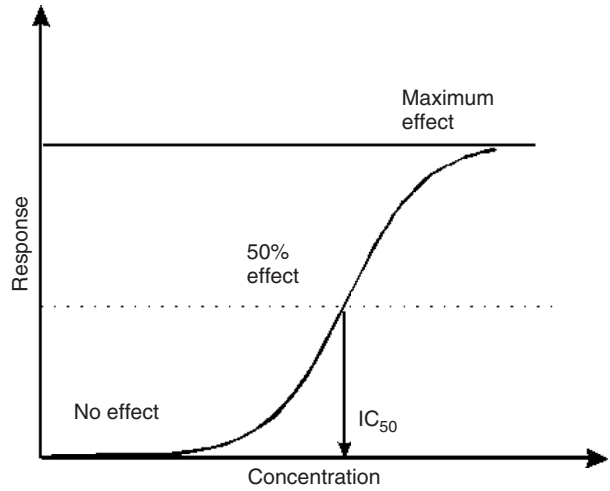

Figure 1.2 Typical concentration versus potency profile used to derive the pIC_{50} values that refer to a concentration of the compound producing 50% of the maximum measurable response in a particular screen.

combinations of substituents on an active "skeleton" or scaffold. This part of the process is often called *lead generation*. The "lead" molecule usually shows high activity on a particular target. It is advisable to test the selectivity of the lead molecule to avoid so-called promiscuous binders that produce a hit on a variety of targets; that is, the lead molecule should show activity preferably only on one target. The "lead" molecule also should show reasonable solubility and bioavailability during the first *in vivo* animal experiments.

Sometimes it is possible to get the three-dimensional structure of the targets using X-ray crystallography or nuclear magnetic resonance (NMR) and the binding sites are identified. Computational chemists, using various docking methods, find or design small molecules that fit to the active site of the target [4]. The primary role of a docking program is to identify biologically active compounds that may act as ligands to a target binding site based on the calculated intermolecular interaction energy. When the X-ray crystallographic structure is not available, there are various computational methods that can be used to assess the binding site of the target protein and design molecules that will bind to it.

The drug discovery process moves into the lead optimization phase, where molecules then undergo a series of structural modifications, which helps to establish a so-called SAR (structure–activity relationship) for the series. It allows the chemists to establish which part of the molecule is essential for activity. During this process, they often find much more active analogs than the original lead molecule. The activity measurements are repeated using cell-based assays or in the presence of plasma proteins. Usually that is the point where potential liabilities may be observed. The reduced cellular activity could be due to low permeability or solubility or strong nonspecific binding to cellular components.

This is the time when they approach the physicochemists to measure solubility, permeability, and protein binding in order to determine the cause of the loss of cellular activity. Usually the activity is not lost completely but only weakens. This is called *drop off* of the activity. The lead optimization process often takes a year or two, maybe even considerably longer. It includes various *in vitro* assays for the assessment of selectivity and *in vitro* microsomal stability. For a smaller number of potentially promising compounds, animal studies, bioavailability measurements, and pharmacokinetic studies are also carried out before candidate selection. At this stage, scientists try to establish the dose necessary for effective *in vivo* pharmacological activity. It is important to observe *in vivo* efficacy and receptor or target engagement assays to prove that the observed pharmacological activity is a consequence of the molecule binding to the target. The first toxicological studies are also performed at this stage.

Chemists utilize information from other disciplines to establish structure–property relationships and modify the structures accordingly. If successful, at the end of this process, the molecule that shows the desired potency, bioavailability, *in vivo* receptor occupancy, and acceptable pharmacokinetic properties will be declared as the "candidate" and will then enter a long development process. When a molecule is identified as a candidate, the research process ends and the selected molecule is passed over to development scientists, who then carry out preclinical studies, extensive toxicity studies using much higher doses than are needed for efficacy, formulation studies, large-scale synthesis, and preparation for the first clinical trial "first time in human." Usually, a "backup" compound is also selected in case the candidate molecule fails during the development process.

During this process, we also need a disease model to prove the concept that the compound acting on the selected target actually can cure the disease. It is called *proof of concept* and target validation. Figure 1.3 summarizes the major phases of the early drug discovery process. The high throughput chemistry is often replaced nowadays by lower throughput synthesis of smaller number of compounds as a designed library. The HTS campaign is very expensive; therefore, it is used when hit molecules are not available from other sources.

However, traditional HTS has not delivered on its promise of increasing the numbers and quality of new drugs entering clinical trials. The typical hit rate of a screening campaign is usually less than 1%. This lack of success is due in part to the complexity and the relatively large size of the compounds routinely being screened. On the basis of the mathematical probability, the ability of finding perfect binding, including four to six interactions at the binding site, is very low [5].

Recently, the so-called fragment approach has been used more often. It has been found that the probability of binding decreases rapidly as the complexity of the ligand increases. However, there is much higher probability of finding one or two binding interactions at a time and then mapping the other possible interactions at the binding site one by one. Therefore, instead of screening a very large number of molecules to find the one that shows affinity to the receptor, there is much higher probability of success in finding molecular fragments that have only one or two binding interactions. From the small, so-called "fragment" molecules, we can

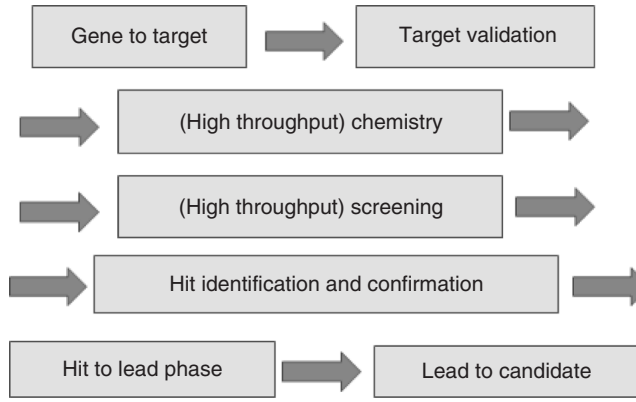

Figure 1.3 Major steps of the modern drug discovery process including both the biological and chemical aspects.

build bigger molecules by combining several fragments into a single molecule. In this way, we can increase the likelihood of finding a highly potent larger molecule. The effective use of the fragment-based approach is dependent on the use of biophysical techniques such as X-ray crystallography or NMR measurements to identify how and where the fragment binds to the target. Traditional bioassays used in HTS are generally unable to detect such small drug fragments because of their low potency binding to the protein target. The approach requires screening of molecules at higher concentration (millimolar and not micromolar or nanomolar) to be able to detect the weaker interactions. The fragment-based drug design approaches are adopted by many big and small pharmaceutical companies. The typical hit rate is usually an order of magnitude higher than the HTS hit rate (around 10%). The typical binding interactions between such small molecules and the target are hydrogen bonding (by donating or accepting proton), hydrophobic, and π–π interactions. The small molecule has to fit spatially well to the binding sites, so that the interacting molecular forces can produce energetically stable binding. Ideally, fragment molecules should allow chemists to attach other functional groups and grow the molecules. It means that the fragments should be chemically tractable for modification.

The early drug discovery process sometimes produces several thousands of potent molecules, still only a few reach the market. The attrition is very high. Although this part of the drug discovery research is expensive, it costs only around 5% of the total cost of getting a drug molecule to the market, which is estimated to be around $800 million. Unfortunately, only 1 in 10 candidate molecules obtain approval from the Food and Drug Administration (FDA). This huge attrition of the compounds is due to so-called “developability” issues. If a candidate is to fail, the earlier the better, as it saves a lot of effort and investment for the company. The most common reasons for stopping the candidates from further development are disadvantageous pharmacokinetics, lack of efficacy in the clinic, toxicity,

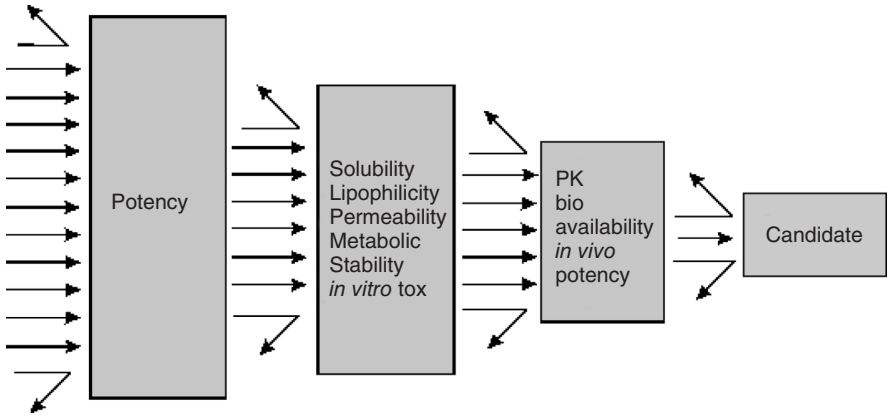

Figure 1.4 Developability considerations and physicochemical properties should be taken early in the drug discovery process.

and formulation problems. Sometimes development is stopped for commercial reasons, as other compounds by competitors appear on the market. The potential disadvantageous pharmacokinetics includes poor oral exposure, high clearance, drug–drug interaction, too short or too long half-life, etc. The physicochemical properties of the molecules have a crucial impact on these failures [6–8]; therefore, it is vital to include these measurements early in the drug discovery process. Figure 1.4 shows how the developability considerations and physicochemical properties are built into the candidate selection process.

For example, poor solubility of the discovery compounds may compromise the results of the screening. It may produce false-positive or false-negative potency values. Very lipophilic compounds often bind to many different types of proteins and to various nonspecific binding sites as well, thus reducing the actual available free concentration of the compound and shifting the pIC_{50} values. Ionized compounds can have various ionic interactions; they can easily change the physicochemical properties such as solubility and lipophilicity because of small changes in the pH of the environment. The most widely known description of optimum physicochemical property space of orally absorbed marketed drug molecules is the Lipinski rule of five [9].

1. Lipophilicity should be less than 5 (calculated by clog P).
2. Number of H-bond donor groups (–NH, –OH) should be less than 5.
3. Number of H-bond acceptor groups should be less than 10.
4. Molecular weight is less than 500 Da.
5. Molecules that are undergoing active transport are exceptions.

Since the introduction of the seminal work of Lipinski [9], a great number of publications have highlighted the importance of simple physicochemical properties such as size, lipophilicity, H-bond acidity and basicity, and polar surface area

that characterizes the drug-like property space. Thus, it was found that marketed drugs fall within a certain size, lipophilicity, and polarity range regardless of the pharmacological class [10–11]. As compounds progress through the lead optimization process, they tend to increase in size and lipophilicity, so-called "molecular obesity" [12], and fall outside the desired drug-like property space.

The aim of this book is to explain how to measure and use these physicochemical properties during the lead optimization process. We shall discuss how these properties can be used for compound selection and prioritization for further studies. At this stage, we need to measure solubility, lipophilicity, permeability, and other biomimetic properties of a large number of compounds in a short period of time. Compounds are usually not available in large quantities at this stage and often not available in a well-characterized solid form either. The measurements should be automated; compound dispensing and handling are being robotized. The physicochemical measurements are often pragmatic and limited to the experimental conditions used. Therefore, these measurements are quite different from the physicochemical measurements that are carried out on selected candidate molecules at the development stage.

SUMMARY

The drug discovery process has gone through major changes in the past decade. The discovery of defined human molecular targets based on genomics provides the possibility to develop HTS methods. In spite of finding numerous potent molecules, the number of drug approvals has not increased proportionally as drug candidates often fail at later stages in the drug development process. The failure is most often due to the lack of efficacy, poor bioavailability, unfavorable pharmacokinetics, and toxic side effects. The undesirable outcome in the clinic is often due to the physicochemical properties of the molecules, such as size, lipophilicity, permeability, and solubility. The scope of this book is to investigate the measurement and calculation of these properties at early stages of the drug discovery process, reveal general structure/property relationships, and suggest appropriate structural modifications during the lead optimization. The aim is to provide insight and explain why these physicochemical properties are important to take into consideration for reducing late-stage attrition of potent candidate molecules.

QUESTION FOR REVIEW

Q1.1 What are the major steps of the drug discovery process?

REFERENCES

1. Drug Discovery and Development, Ed. Rang, H. P. (2006) Elsevier, London.

2. Kerns, E. H., Di, L. (2008) Drug-like Properties: Concepts, Structure Design and Methods. Academic Press, Elsevier, Amsterdam.
3. Analogue Based Drug Discovery II, Eds Fischer, J., Ganellin, C. R. (2010) Wiley-VCH, Weinheim.
4. Balius, T. E., Mukherjee, S., Rizzo, R. C. (2011) Implementation and evaluation of a docking-rescoring method using molecular footprint comparisons. *Journal of Computational Chemistry*, *32*, 2273–2289.
5. Leach, A. R., Hann, M. M. (2011) Molecular complexity and fragment-based drug discovery: Ten years on. *Current Opinion in Chemical Biology*, *15*, 1–8.
6. Valkó, K., Reynolds, D. P. (2005) High throughput physico-chemical and *in vitro* ADMET screening: A role in pharmaceutical profiling, *American Journal of Drug Delivery*, *3*, 83–100.
7. Van de Waterbeemd, H. (2009) Improving compound quality through *in vitro* and *in silico* physicochemical profiling, *Chemistry and Biodiversity*, *6*, 1760–1766.
8. Gleeson, M. P., Hersey, A., Montanari, D., Overington, J. (2011) Probing the links between *in vitro* potency, ADMET and physicochemical parameters. *Nature Reviews; Drug Discovery*, *10*, 197–208.
9. Lipinski, C. A., Lombardo, F. B., Dominy, W., Feeney, P. J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *23*, 3–25.
10. Lipinski, C. A. (2004) Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today: Technologies*, *1*, 337–341.
11. Vieth, M., Siegel, M. G., Higgs, R. E., Watson, I. A., Robertson, D. H., Savin, K. A., Durst, G. L., Hipskind, P. A. (2004) Characteristic physical properties and structural fragments of marketed oral drugs. *Journal of Medicinal Chemistry*, *47*, 224–232.
12. Hann, M. M. (2011) Molecular obesity, potency and other addictions in drug discovery. *Medicinal Chemistry Communications*, *2*, 349–355.

2

DRUG-LIKENESS AND PHYSICOCHEMICAL PROPERTY SPACE OF KNOWN DRUGS

There has been a lot of interest in finding common features in drug molecules that are used for the treatment of various diseases. The potent molecules usually contain a so-called pharmacophore group that is responsible for the receptor binding, enabling it to produce the desired pharmacological activity. However, it has long been recognized that it is not enough to make an active molecule in order to produce a successful drug. Many highly potent compounds failed to reach the market because of poor absorption, poor distribution, safety concerns, and so on. On the basis of the study by Kola and Landis [1], only around 5% of selected candidate molecules reach patients as drugs. A pioneering study about the desirable properties of drugs was published by Lipinski et al. [2]. The Lipinski "rule of five" has been mentioned in Chapter 1. The rule has been established on the basis of analysis of the physicochemical properties of orally administered drugs. Their publication has been cited over 10,000 times, and many new modified rules have been proposed [3–11].

Other physicochemical properties have also been considered in subsequent publications such as the polar surface area, number of rotatable bonds, and number of heavy atoms. More recently, the importance of the number of aromatic rings has been emphasized [12–14]. Increasing the number of aromatic rings in molecules increases the chance of poor solubility. Lipinski [15] introduced the term *drugability*, emphasizing that the chemical space, where the drug molecules can be

Physicochemical and Biomimetic Properties in Drug Discovery: Chromatographic Techniques for Lead Optimization, First Edition. Klara Valko.

found, occupies only a small region of the chemical space of all possible types of molecules. Van de Waterbeemd et al. [16] introduced the new phase of drug discovery, where the property-based design replaces the structure-based design of new drug molecules. Kerns and Di [17] published a book in 2008 in which they discuss the drug-like properties, their measurements, and applications in the drug discovery. The essence of these publications is that drug molecules occupy a reasonably well-defined physicochemical property space. It was found that size, lipophilicity, solubility, permeability, polar surface area, and H-bond acidity and basicity play an important role in the description of the physicochemical property space. Gleeson et al. [18] published various *in silico* models for estimating absorption, plasma protein binding, metabolism, volume of distribution, central nervous system (CNS) penetration, and so on. These models highlighted the most important physicochemical properties that influence these absorption, distribution, metabolism, and elimination (ADME) characteristics. Lipophilicity and size were the two most dominant descriptors in estimating almost all of the ADME properties. For oral absorption, the compound should be easily dissolved in biorelevant media and should be permeable through the intestinal wall. Solubility and permeability can be measured *in vitro*. However, measurements are not sufficient for designing drug molecules; we need to understand how these properties relate to each other and also how they relate to the chemical structure of the molecules. Both solubility and permeability can be related to compound lipophilicity or hydrophobicity. The more lipophilic the compound, the less likely it is to be soluble in water. On the other hand, the molecule has to be reasonably lipophilic to be able to go through the lipophilic membrane bilayers. Hydrophilic compounds might contain H-bond donor and acceptor groups that can form hydrogen bonds in crystal form, which would reduce solubility and dissolution rate due to the increased crystal lattice energy. Solubility and permeability are both dependent on lipophilicity in the opposite direction and we need to find the optimum lipophilicity in order to achieve acceptable solubility and permeability. Also, when introducing a solubilizing group into a molecule, we increase the size of the molecule too, which can have a detrimental effect on the permeability. Figure 2.1 illustrates a hypothetical multidimensional property space projected onto a two-dimensional space using the first two principal components as axes. We need to be in the right chemical property space to achieve good potency, good absorption, good safety profile, good metabolic stability, and so on. We do it by making compounds and measuring its properties. Figure 2.1 shows a desirable situation when there is a small space that overlaps with all the required property spaces. This is usually not known to the researcher at the start of the chemistry program. Figure 2.2 demonstrates a hypothetical situation when potent compounds are outside the desired property space for good absorption and metabolic stability. Therefore, where possible, it is very important to start a research program with several series of leads and try to map the boundaries of the desired property space. Several research programs have ended without success as, for example, all the soluble and potent compounds were metabolically unstable, and only the insoluble potent compounds had the desired metabolic stability. In such cases, it is better to search for a new chemical starting point using different scaffolds. An in-depth

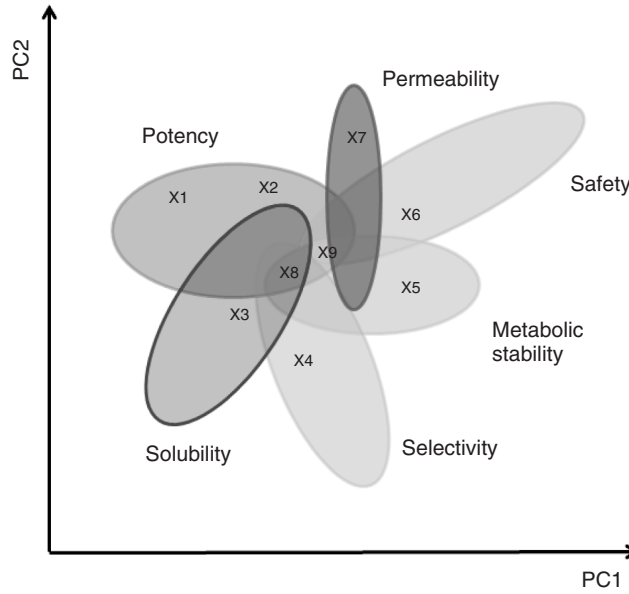

Figure 2.1 Illustration of the small overlapping property space (plotted using the first two principal components PC1 and PC2) that covers all the required important attributes of a drug molecule.

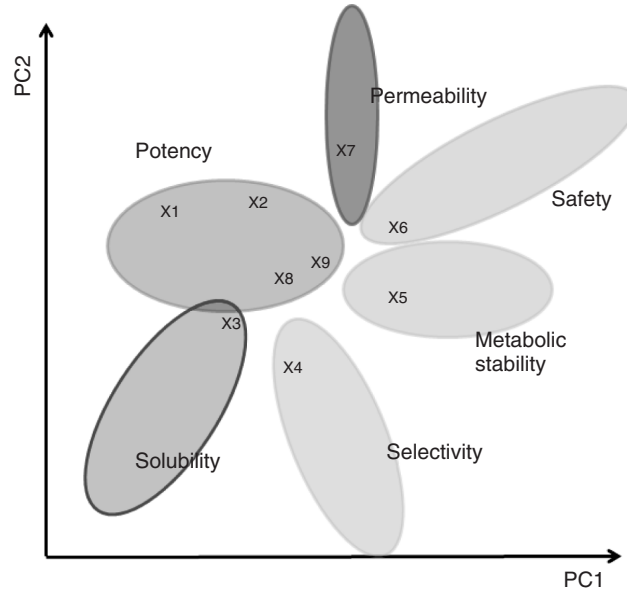

Figure 2.2 Illustration of the lack of overlapping property space (plotted using the first two principal components PC1 and PC2) that fulfills the requirement of a drug molecule.

lead optimization program should start only when a high probability exists that the optimization process is possible. As the number of newly discovered targets is increasing and no previous example of potent drug molecule exists, there is no guarantee that it is possible to find molecules that have the overlapping desired property space of a drug.

There is a debate in literature as to make molecules that are outside the boundaries of the physicochemical property space of known drug molecules [11]. The analysis of lipophilicity data suggests that only a very narrow lipophilicity range (from 1 to 3 log D, logarithmic value of the octanol/water distribution coefficient at pH 7.4) provides a high probability of finding a good drug molecule. The paper discusses in detail what the risks are when a research molecule is outside of the optimum lipophilicity range. There are examples of successful drug molecules that do not obey the Lipinski "rule of five." There are always exceptions to the general rules. As has been shown by Bade et al. [10], many natural products, for example, erythromycin, digitalis glycosides, do not fit the rule of five as the molecular weight is over 1000 Da and the number of H-bond donor and acceptor groups are also over the limit and still they are orally absorbed (most often by active transport). In drug discovery, we need to have a sense of urgency not only due to considerations of the cost but also the benefit of finding new drug molecules that can help patients live longer and healthier lives. In modern drug discovery, we need to focus on the probability of the success and need to increase our knowledge and understanding of why some molecules fail during the development process.

These rules and boundaries have to be used with caution. Intuition and experience can also increase the probability of success of a research program. The conclusion from the studies of the physicochemical properties of drugs, research compounds, and failed drug candidates is that we can define physicochemical property boundaries into which the majority of known drug molecules can fit. The most important physicochemical properties that we can measure are the lipophilicity, solubility, permeability, and acid/base character. Size is also important and is usually described by molecular weight or calculated molar refraction (cMR). These measurable properties are not independent of each other; so, it is difficult to alter only one property at a time. We also need to understand the effect of the conditions of the measurements (pH, solvents, etc.).

The following chapters provide deeper insight into these important physicochemical properties and how we can measure them using chromatographic techniques. It is demonstrated through examples that describe how we can apply measured data to develop general and local models to estimate a compound's *in vivo* behavior in the early stages of the drug discovery process and increase the probability of success.

SUMMARY

Several physicochemical properties of known drug molecules have been analyzed. It has been found that absorption, distribution, metabolism, and other

pharmacokinetic parameters, such as volume of distribution and half-life can be related to the physicochemical properties of drugs. The majority of known drug molecules fit into a narrow range of physicochemical properties that are termed *drug-like* properties. When drug discovery compounds show these "drug-like" properties, they have a higher probability of becoming successful drugs. Therefore, it is important to identify, understand, and measure these properties early in the drug discovery process. The most important physicochemical properties that show correlations to the absorption, distribution, and pharmacokinetic properties are lipophilicity, solubility, permeability, and acid/base character. The size of the molecule is also very important and can be easily and precisely characterized by the molecular weight or calculated molar refractivity. These physicochemical parameters are not independent. The measured values are highly dependent on the conditions of the measurements. Therefore, it is very important to gain a better understanding of how to measure them and how to relate the measured values to the measurement conditions and to the chemical structure.

QUESTIONS FOR REVIEW

Q2.1 What are the most important physicochemical properties that can be related to the ADME properties?

Q2.2 What is "drug-likeness?"

Q2.3 Why is there a need to measure solubility, lipophilicity, and permeability in the early stages of the drug discovery process?

Q2.4 Which physicochemical properties are the most important in terms of drug-like property space?

REFERENCES

1. Kola, I., Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery*, *3*, 711–716.
2. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *46*, 3–26.
3. Oprea, T. I., Allu, T., Fara, D., Rad, R., Ostopovici, L., Bologa, C. (2007) Lead-like, drug-like or "Pub-like": How different are they? *Journal of Computer Aided Molecular Design*, *21*, 113–119.
4. Gleeson, M. P. (2008) Generation of a set of simple, interpretable ADME rules of thumb. *Journal of Medicinal Chemistry*, *51*, 817–834.
5. Hou, T. J., Wang, J. M., Zhang, W., Xu, X. J. (2007) ADME evaluation in drug discovery. 6. Can the oral bioavailability in humans be effectively predicted by simple molecular property-based rules? *Journal of Chemical Information and Computational Sciences*, *47*, 460–463.
6. Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., Kopple, K. D. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, *45*, 2615–2623.

7. Leeson, P. D., Empfield, J. R. (2010) Reducing the risk of drug attrition associated with physicochemical properties. *Annual Reports in Medicinal Chemistry*, *45*, 393–407.
8. Hughes, J. D., Blagg, J., Price, D. A., Bailey, S., DeCrescenzo, G. A., Devraj, R. V., Ellsworth, E., Fobian, Y. M., Gibbs, M. E., Gilles, R. W., Greene, N., Huang, E., Krieger-Burke, T., Loesel, J., Wager, T., Whiteley, L., Zhang, Y. (2008) Physicochemical drug properties associated with *in vivo* toxicological outcomes. *Bioorganic and Medicinal Chemistry Letters*, *18*, 4872–4875.
9. Keserü, Gy., Makara, G. (2008) The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews Drug Discovery*, *8*, 2003–212.
10. Bade, R., Chan, H.-F., Reynisson, J. (2010) Characteristics of known drug space. Natural products, their derivatives and synthetic drugs. *European Journal of Medicinal Chemistry*, *45*, 5646–5652.
11. Waring, M. J. (2010) Lipophilicity in drug discovery. *Expert Opinion in Drug Discovery*, *5(3)*, 235–248.
12. Lovering, F., Bikker, J., Humblet, C. (2009) Escape from flatland: Increasing saturation as an approach to improving clinical success. *Journal of Medicinal Chemistry*, *52*(*21*), 6752–6756.
13. Ritchie, T. J., MacDonald, S. J. F. (2009) The impact of aromatic ring count on compound developability—Are too many aromatic rings a liability in drug design? *Drug Discovery Today*, *14*, 1011–1020.
14. Hill, A. P. Young, R. J. (2010) Getting physical in drug discovery: A contemporary perspective on solubility and hydrophobicity. *Drug Discovery Today*, *15*, 648–655.
15. Lipinski, C. A. (2000) Drug like properties and the causes of poor solubility and permeability. *Journal of Pharmacological and Toxicological Methods*, *44*, 235–249.
16. Van de Waterbeemd, H., Smith, D. A., Beaumont, K., Walker, D. K. (2001) Property-based design: Optimization of drug absorption and pharmacokinetics, *Journal of Medicinal Chemistry*, *44*, 1313–1333.
17. Kerns, E. H., Di, L. (2008) Drug-like Properties: Concepts, Structure Design and Methods. Elsevier, Academic Press, San Diego, California, USA.
18. Gleeson, M. P., Hersey, A., Hannongbua, S. (2011) In-silico ADME models: A general assessment of their utility in drug discovery applications. *Current Topics in Medicinal Chemistry*, *11*, 358–381.