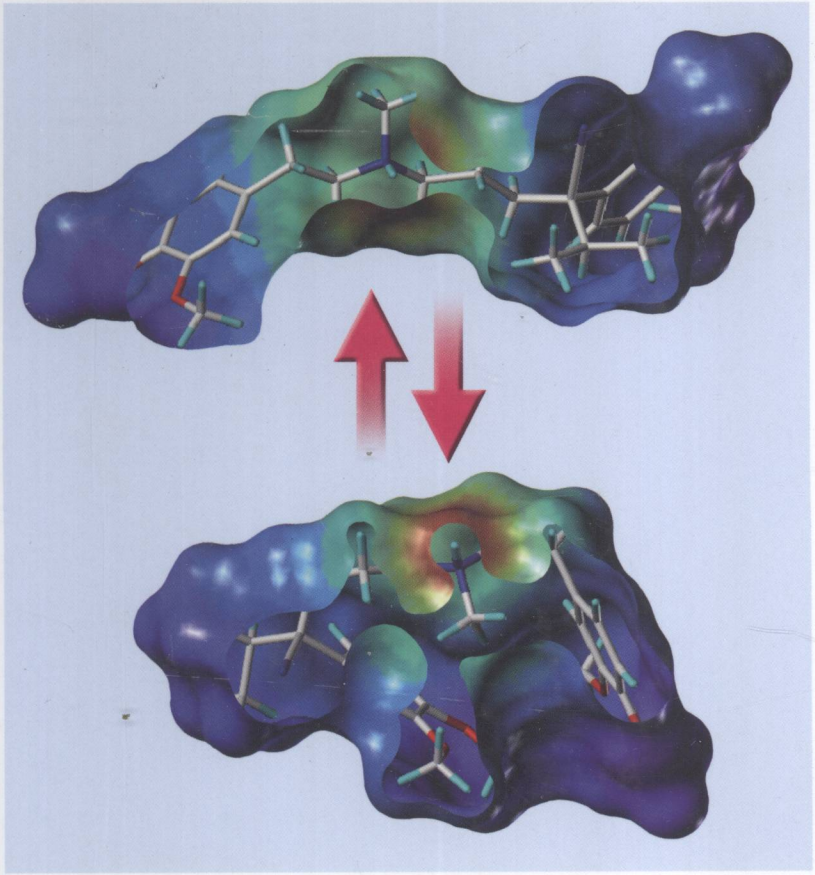
Methods and Principles in Medicinal Chemistry

Edited by
Raimund Mannhold

WILEY-VCH

Molecular Drug Properties

Measurement and Prediction



Volume 37

Series Editors:
R. Mannhold,
H. Kubinyi,
G. Folkers

Molecular Drug Properties

Measurement and Prediction

Edited by
Raimund Mannhold

WILEY-VCH Verlag GmbH & Co. KGaA

Series Editors

Prof. Dr. Raimund Mannhold
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstrasse 1
40225 Düsseldorf
Germany
mannhold@uni-duesseldorf.de

Prof. Dr. Hugo Kubinyi
Donnersbergstrasse 9
67256 Weisenheim am Sand
Germany
kubinyi@t-online.de

Prof. Dr. Gerd Folkers
Collegium Helveticum
STW/ETH Zurich
8092 Zurich
Switzerland
folkers@collegium.ethz.ch

Volume Editor

Prof. Dr. Raimund Mannhold
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstrasse 1
40225 Düsseldorf
Germany
mannhold@uni-duesseldorf.de

Cover Illustration

Molecular lipophilicity potentials for an extended, more lipophilic and a folded, less lipophilic conformer of verapamil are shown ($\Delta\log P_{MLP} = 0.6$). Violet regions: higher lipophilicity; blue regions: medium lipophilicity; yellow regions: weakly polar; red regions: strongly polar (Preparation of this graph by Pierre-Alain Carrupt is gratefully acknowledged.)

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Composition SNP Best-set Typesetter Ltd., Hong Kong

Printing Betz-Druck GmbH, Darmstadt

Bookbinding Litges & Dopf GmbH, Heppenheim

Cover Design Grafik-Design Schulz, Fußgönheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN 978-3-527-31755-4

Molecular Drug Properties

Edited by
Raimund Mannhold

Methods and Principles in Medicinal Chemistry

Previous Volumes of this Series:

G. Cruciani (ed.)

Molecular Interaction Fields

Vol. 27

2006, ISBN 978-3-527-31087-6

M. Hamacher, K. Marcus, K. Stühler, A. van Hall, B. Warscheid, H. E. Meyer (eds.)

Proteomics in Drug Research

Vol. 28

2006, ISBN 978-3-527-31226-9

D. J. Triggle, M. Gopalakrishnan, D. Rampe, W. Zheng (eds.)

Voltage-Gated Ion Channels as Drug Targets

Vol. 29

2006, ISBN 978-3-527-31258-0

D. Rognan (ed.)

Ligand Design for G Protein-coupled Receptors

Vol. 30

2006, ISBN 978-3-527-31284-9

D. A. Smith, H. van de Waterbeemd, D. K. Walker

Pharmacokinetics and Metabolism in Drug Design, 2nd Ed.

Vol. 31

2006, ISBN 978-3-527-31368-6

T. Langer, R. D. Hofmann (eds.)

Pharmacophores and Pharmacophore Searches

Vol. 32

2006, ISBN 978-3-527-31250-4

E. Francotte, W. Lindner (eds.)

Chirality in Drug Research

Vol. 33

2006, ISBN 978-3-527-31076-0

W. Jahnke, D. A. Erlanson (eds.)

Fragment-based Approaches in Drug Discovery

Vol. 34

2006, ISBN 978-3-527-31291-7

J. Hüser (ed.)

High-Throughput Screening in Drug Discovery

Vol. 35

2006, ISBN 978-3-527-31283-2

K. Wanner, G. Höfner (eds.)

Mass Spectrometry in Medicinal Chemistry

Vol. 36

2007, ISBN 978-3-527-31456-0

Dedicated with love
to my wife Barbara
and my daughter Marion

List of Contributors

Alex Avdeef
*p*ION INC
5 Constitution Way
Woburn, MA 01801
USA

Jonas Boström
AstraZeneca R&D
Department of Lead Generation
43183 Mölndal
Sweden

Giulia Caron
CASMedChem laboratory
Dipartimento di Scienza
Tecnologia del Farmaco
Università di Torino
Via P. Giuria 9
10125 Torino
Italy

Pierre-Alain Carrupt
Unit of Pharmacochemistry
School of Pharmaceutical
Sciences
University of Geneva,
University of Lausanne
Quai Ernest-Ansermet 30
1211 Geneva 4
Switzerland

Giuseppe Ermondi
CASMedChem laboratory
Dipartimento di Scienza
Tecnologia del Farmaco
Università di Torino
Via P. Giuria 9
10125 Torino
Italy

Peter Ertl
Novartis Institutes
for Biouedical Research
4002 Basel
Switzerland

Bernard Faller
Novartis Pharma AG
Lichtstrasse 35
4056 Basel
Switzerland

Andreas Frank
Institute for Organic Chemistry and
Biochemistry
Technical University Munich
Lichtenbergstrasse 4
85747 Garching
Germany

Molecular Drug Properties. Measurement and Prediction. R. Mannhold (Ed.)

ISBN: 978-3-527-31755-4

Alexandra Galland
Unit of Pharmacochemistry
School of Pharmaceutical Sciences
University of Geneva,
University of Lausanne
Quai Ernest-Ansermet 30
1211 Geneva 4
Switzerland

J Andrew Grant
AstraZeneca Pharmaceuticals
Mereside
DECS Global Compound Sciences
Alderley Park,
Cheshire SK10 4TG
UK

Davy Guillarme
Laboratory of Analytical Pharmaceutical Chemistry
School of Pharmaceutical Sciences
University of Geneva,
University of Lausanne
Boulevard d'Ivoy 20
1211 Geneva 4
Switzerland

Yveline Henchoz
Unit of Pharmacochemistry
School of Pharmaceutical Sciences
University of Geneva,
University of Lausanne
Quai Ernest-Ansermet 30
1211 Geneva 4
Switzerland

Ovidiu Ivanciuc
Sealy Center for Structural Biology and Molecular Biophysics
Departments of Biochemistry and Molecular Biology
University of Texas Medical Branch
301 University Boulevard
Galveston, TX 77555-0857
USA

Horst Kessler
Institute for Organic Chemistry and Biochemistry
Technical University Munich
Lichtenbergstrasse 4
85747 Garching
Germany

Andreas Klamt
COSMO*logic* GmbH & Co. KG
Burscheider Str. 515
51381 Leverkusen
Germany

Institute of Physical and Theoretical Chemistry
University of Regensburg
93040 Regensburg
Germany

Christopher A. Lipinski
Scientific Advisor
Melior Discovery
10 Connshire Drive
Waterford, CT 06385-4122
USA

Franco Lombardo
Novartis Institute for Biomedical Research
250 Massachusetts Avenue
Cambridge, MA 02139
USA

Burkhard Luy
Institute for Organic Chemistry
and Biochemistry
Technical University Munich
Lichtenbergstraße 4
85747 Garching
Germany

Raimund Mannhold
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstraße 1
40225 Düsselsorf
Germany

Sophie Martel
Unit of Pharmacochemistry
School of Pharmaceutical
Sciences
University of Geneva,
University of Lausanne
Quai Ernest-Ansermet 30
1211 Geneva 4
Switzerland

Sorel Muresan
AstraZeneca R&D
Computational Chemistry
431 83 Mölndal
Sweden

Claude Ostermann
Nycomed GmbH
Byk-Gulden-Str. 2
78467 Konstanz
Germany

Alessandro Pedretti
Istituto di Chimica Farmaceutica
Facoltà di Farmacia
Università di Milano
Via Mangiagalli 25
20131 Milano
Italy

Gennadiy I. Poda
Pfizer Global R & D
700 Chesterfield Parkway West
Mail Zone BB2C
Chesterfield, MO 63017
USA

Oleg Raevsky
Department of Computer-Aided
Molecular Design
Institute of Physiologically
Active Compounds
Russian Academy of Sciences
Severnii proezd, 1
142432, Chernogolovka,
Moscow region
Russia

Serge Rudaz
Laboratory of Analytical
Pharmaceutical Chemistry
School of Pharmaceutical Sciences
University of Geneva,
University of Lausanne
Boulevard d'Ivoy 20
1211 Geneva 4
Switzerland

Jens Sadowski
AstraZeneca
Lead Generation KJ257
43183 Mölndal
Sweden

Marina Shalaeva
Pfizer Global Research
and Development
Groton Laboratories
Groton, CT 06340
USA

Brian J Smith
The Walter and Eliza Hall
Institute of Medical Research
Department of Structural Biology
1G Royal Parade, Parkville,
Victoria 3050
Australia

Bernard Testa
Pharmacy Department
University Hospital Centre
CHUV-BH 04
46 Rue du Bugnon
1011 Lausanne
Switzerland

Igor Tetko
GSF – National Research Centre
for Environment and Health
Institute for Bioinformatics
(MIPS)
Ingolstädter Landstraße 1
85764 Neuherberg
Germany

Suzanne Tilton
Novartis Institute
for Biomedical Research
250 Massachusetts Avenue
Cambridge, MA 02139
USA

Jean-Luc Veuthey
Laboratory of Analytical
Pharmaceutical Chemistry
School of Pharmaceutical Sciences
University of Geneva,
University of Lausanne
Boulevard d'Ivoy 20
1211 Geneva 4
Switzerland

Giulio Vistoli
Istituto di Chimica Farmaceutica
Facoltà di Farmacia
Università di Milano
Via Mangiagalli 25
20131 Milano
Italy

Han van de Waterbeemd
AstraZeneca
DECS – Gobal Compound Sciences
Mereside 50S39
Macclesfield
Cheshire SK10 4TG
UK

Preface

Despite enormous investments in pharmaceutical research and development, the number of approved drugs has declined in recent years. The attrition of compounds under development is dramatically high. Safety, insufficient efficacy and, to some extent, absorption, distribution, metabolism, excretion and toxicity (ADMET) problems are the responsible factors. Formerly, drugs were discovered by testing compounds synthesized in time-consuming multistep processes against a battery of *in vivo* biological screens. Promising compounds were then further tested in development, where their pharmacokinetic (PK) properties, metabolism and potential toxicity were investigated. Adverse findings were often made at this stage and projects were re-started to find another clinical candidate. Drug discovery has undergone a dramatic change over the last two decades due to a methodological revolution including combinatorial chemistry, high-throughput screening and *in silico* methods, which greatly increased the speed of the process of drug finding and development.

More recently, the bottleneck of drug research has shifted from hit-and-lead discovery to lead optimization, and more specifically to PK lead optimization. Some major reasons are (i) the imperative to reduce as much as feasible the extremely costly rate of attrition prevailing in preclinical and clinical phases, and (ii) more stringent concerns for safety. The testing of ADME properties is now done much earlier, i.e. before a decision is taken to evaluate a compound in the clinic.

As the capacity for biological screening and chemical synthesis has dramatically increased, so have the demands for large quantities of early information on ADME data. The physicochemical properties of a drug have an important impact on its PK and metabolic fate in the body, and so a good understanding of these properties, coupled with their measurement and prediction, are crucial for a successful drug discovery programme.

The present volume is dedicated to the measurement and the prediction of key physicochemical drug properties with relevance for their biological behavior including ionization and H-bonding, solubility, lipophilicity as well as three-dimensional structure and conformation. Potentials and limitations of the relevant techniques for measuring and calculating physicochemical properties of drugs are critically discussed and comprehensively exemplified in 17 chapters from 35 distinguished authors, from both academia and the pharmaceutical industry.

Molecular Drug Properties. Measurement and Prediction. R. Mannhold (Ed.)

ISBN: 978-3-527-31755-4

We are indebted to all authors for their well-elaborated chapters, and we want to express our gratitude to Dr Andreas Sendtko and Dr Frank Weinreich from Wiley-VCH for their valuable contributions to this volume and the ongoing support of our series *Methods and Principles in Medicinal Chemistry*.

Raimund Mannhold, *Düsseldorf*
Hugo Kubinyi, *Weisenheim am Sand*
Gerd Folkers, *Zürich*

August 2007

A Personal Foreword

Several editors of previous volumes in this series lised the platform of the Personal Foreword to reflect routes and contents of their scientific lives and in particular to appreciate the invaluable support by rewarded colleagues. It is a pleasure for me to continue this tradition.

After the study of pharmaceutical sciences in Frankfurt/Main I joined the Department of Clinical Physiology at the Heinrich-Heine-Universität Düsseldorf to start my PhD work dedicated to pharmacological studies of the calcium channel blocker verapamil under the supervision of Raimund Kaufmann. He was a very liberal scientific teacher and he allowed me to fine-tune the contents of my PhD work according to my personal preferences.

Frequent contacts with the manufacturer of verapamil, the Knoll company in Ludwigshafen, enabled an intense communication with Hugo Kubinyi, working at that time as a medicinal chemist for Knoll. As a consequence of frequent fruitful discussions with Hugo I included quantitative structure–activity relationship (QSAR) studies on verapamil congeners in my PhD work and continued working in the QSAR field till the present.

Two Dutch colleagues and friends have strongly influenced me since the early 1980s. I first met Roelof Rekker, one of the fathers of log *P* calculation approaches, on the occasion of one of the famous Noordwijkerhout meetings. Roelof fascinated me with his elegant lipophilicity studies. After years of fruitful cooperation I had the privilege to coauthor with him our booklet "Calculation of Drug Lipophilicity" updating the Σf system, the first fragmental approach for lipophilicity calculation.

My first personal contact to Henk Timmerman happened on the wonderful island of Capri during a symposium on pharmaceutical sciences. Henk Timmerman headed one of the largest and most important departments of Medicinal Chemistry in European academia. It was very impressive to face his views on our research field, and his integrated and straightforward way to guide research projects. For several years I collaborated with his group and, as an added bonus, became a great fan of Amsterdam.

In the early 1990s, I founded the book series *Methods and Principles in Medicinal Chemistry* with Verlag Chemie; Henk Timmerman and Povl Krogsgaard Larsen joined me on the initial board of series editors. Hugo Kubinyi followed Povl Krogsgaard Larsen after the first three volumes were released. Henk contributed

Molecular Drug Properties. Measurement and Prediction. R. Mannhold (Ed.)

ISBN: 978-3-527-31755-4

to the series very intensely and successfully for many years, and I want to thank him for the times of coediting this book series. When retiring from the chair of Medicinal Chemistry at the Vrije Universiteit of Amsterdam, he forwarded his work in the series to Gerd Folkers from ETH, Zurich.

In the late 1990s another fruitful and pleasant cooperation arose in Perugia, Italy, with the chemometric group of Sergio Clementi and Gabriele Cruciani, two guys with excellent skills and scientific enthusiasm. Since 1997 I have spent weeks up to months each year in Perugia for joint projects on three-dimensional (3D) QSAR and virtual screening studies. Fortunately, these stays also enable a further specialization in Italian food and wine.

The present volume is dedicated to the measurement and the prediction of key physicochemical drug properties with relevance for their biological behavior, including ionization and H-bonding, solubility, lipophilicity as well as 3D structure and conformation.

In the *Introductory section*, Bernard Testa, Giulio Vistoli and Alessandro Pedretti give us "A Fresh Look at Molecular Structure and Properties", which are key concepts in drug design, but may not mean the same to all medicinal chemists. This chapter serves as a general opening, and invites readers to stand back and reflect on the information contained in chemical compounds and on its description. The authors base their approach on a discrimination between the "core features" and the physicochemical properties of a compound.

Han van de Waterbemd focuses on "Physicochemical Properties in Drug Profiling". These properties play a key role in drug metabolism and pharmacokinetics (DMPK). Their measurement and prediction is relatively easy compared to DMPK and safety properties, where biological factors come into play. However, the latter depend to some extent on physicochemical properties as they dictate the degree of access to biological systems. The change in work practice towards high-throughput screening (HTS) in biology using combinatorial libraries has also increased the demands on more physicochemical and absorption, distribution, metabolism and excretion (ADME) data. Han's chapter reviews the key physicochemical properties, both how they can be measured as well as how they can be calculated in some cases.

Alex Avdeef opens the section on *Electronic Properties* considering "Drug Ionization and Physicochemical Profiling". The ionization constant tells the pharmaceutical scientist to what degree the molecule is charged in solution at a particular pH. This is important to know, since the charge state of the molecule strongly influences its other physicochemical properties. After an in-depth discussion of the accurate determination of ionization constants, Alex focuses on three physicochemical properties where the ionization constant relates to a critical distribution or transport function: (i) octanol–water and liposome–water partitioning, (ii) solubility, and (iii) permeability.

Ovidiu Ivanciuc describes the computation of "Electrotopological State (E-state) Indices" from the molecular graph and their application in drug design. The E-state encodes at the atomic level information regarding electronic state and topo-

logical accessibility. Computing of E-state indices is based exclusively on the molecular topology and it can be done efficiently for large chemical libraries. Comparative QSAR models from a large variety of descriptors show that the E-state indices are often selected in the best QSAR models.

"Polar Surface Area" (PSA) is the topic of Peter Ertl's chapter. PSA has been shown to provide very good correlations with intestinal absorption, blood–brain barrier penetration and several other drug characteristics. It has also been effectively used to characterize drug-likeness during virtual screening and combinatorial library design. The descriptor seems to encode an optimal combination of H-bonding features, molecular polarity and solubility properties. PSA can be easily and rapidly calculated as a sum of fragment contributions using only the molecular connectivity of a structure.

Lastly, Oleg Raevsky discusses "H-bonding Parameterization in QSAR and Drug Design". Studies based on direct thermodynamic parameters of H-bonding and exact 3D structures of H-bonding complexes have essentially improved our understanding of solvation and specific intermolecular interactions. These studies consider the structure of liquid water, new X-ray data for specific H-bonding complexes, partitioning in water–solvent–air systems, a refinement in the PSA approach, improvement of GRID potentials, and calculation schemes of optimum H-bonding potential values for any concrete H-bonding atoms. Oleg exemplifies the successful application of direct H-bonding descriptors in QSAR and drug design.

Conformational Aspects are covered in the next section. First, Jens Sadowski discusses automatic "Three-dimensional Structure Generation" as a fundamental operation in computational chemistry. It has become a standard procedure in molecular modeling and appropriate software has been available for many years. Several of the most common concepts as well as their strengths and limitations are shown in detail. An evaluation study of the two most commonly used programs, CONCORD and CORINA, indicates their general applicability for robust, fast and automatic 3D structure generation. Within the limitation of single conformation generation, reasonable rates of reproducing experimental geometries and other quality criteria are reached. For many applications, the obtained 3D structures are good enough to be used without any further optimization.

Then, Jonas Boström and Andrew Grant review "Exploiting Ligand Conformations in Drug Design". Section 1 gives a theoretical outline of the problems and presents details of various implementations of computer codes to perform conformational analysis. Section 2 describes calculations illustrative of the current accuracy in generating the conformation of a ligand when bound to proteins (the bioactive conformer) by comparisons to crystallographically observed data. The final section concludes by presenting some practical applications of using knowledge of molecular conformation in actual drug discovery projects.

Finally, Burkhard Luy, Andreas Frank and Horst Kessler discuss "Conformational Analysis of Drugs by Nuclear Magnetic Resonance Spectroscopy". The determination and refinement of molecular conformations comprehends three main methods: distance geometry (DG), molecular dynamics (MD) and simulated annealing (SA). In principle, it is possible to exclusively make use of DG, MD or

SA, but normally it is strongly suggested to combine these methods in order to obtain robust and reliable structural models. Only when the results of different methods match should a 3D structure be presented. There are various ways of combining the described techniques and the procedural methods may differ depending on what kind of molecules are investigated. In this chapter, the authors give instructions on how to obtain reliable structural models.

Solubility is a fundamental characteristic of drug candidates. In synthetic chemistry, low solubility can be problematic for homogeneous reactions, and in preclinical experimental studies, low solubility may produce experimental errors or precipitation.

First, Chris Lipinski debates "Experimental Approaches to Aqueous and Dimethylsulfoxide Solubility". The emphasis is on the discovery stage as opposed to the development stage. The reader will find numerous generalizations and rules-of-thumb relating to solubility in a drug discovery setting. The solubility of drugs in water is important for oral drug absorption. Drug solubility in dimethylsulfoxide (DMSO) is important in the biological testing of a compound formatted as a DMSO stock solution. Solubility in aqueous media and DMSO is discussed in the context of both similarities and differences.

Then, Andreas Klamt and Brian Smith discuss the "Challenge of Drug Solubility Prediction". While standard models have emerged for log P, no such convergence can be observed for log S, probably due to its inherent nonlinear character. Thus, nonlinear models are required, but it is questionable whether neural network techniques will ever yield reliable models, because the number of good quality data required will be of the order of hundreds of thousands. In the authors' view, the best way is to make use of the fundamental laws of physical chemistry and thermodynamics as much as possible. Using the supercooled state of the drug as intermediate state, and splitting log S into one smaller contribution arising from the free energy of fusion and a large contribution from the solubility of the supercooled drug, appear to be the only sensible way for reasonable calculation.

A quite comprehensive section concerns *Lipophilicity*, one of the most informative physicochemical properties in medicinal chemistry and since long successfully used in QSAR studies.

"Chemical Nature and Biological Relevance of Lipophilicity" are the topics of the starting chapter by Giulia Caron and Giuseppe Ermondi. Sections on chemical concepts to understand the significance of lipophilicity, lipophilicity systems, the determination of log P and a general factorization of lipophilicity are dedicated to reflect the chemical nature of lipophilicity. In the second part, the biological relevance of lipophilicity is exemplified for membrane permeation, receptor affinity and the control of undesired human ether-a-go-go-related gene activity.

Pierre-Alain Carrupt and colleagues review "Chromatographic Approaches for Measuring Log P". They present a brief overview of the main features of reversed-phase liquid chromatography (isocratic condition and gradient elution) and capillary electrophoresis (microemulsion electrokinetic chromatography, microemulsion electrokinetic chromatography and liposome/vesicular electrokinetic chromatography) methods used for lipophilicity determination of neutral compounds or the

neutral form of ionizable compounds. Relationships between lipophilicity and retention parameters obtained by reversed-phase liquid chromatography methods using isocratic or gradient condition are reviewed. Advantages and limitations of the two approaches are also pointed out and general guidelines to determine partition coefficients in 1-octanol–water are proposed. Finally, recent data on lipophilicity determination by capillary electrophoresis of neutral compounds and neutral form of ionizable compounds are reviewed.

Raimund Mannhold and Claude Ostermann describe the "Prediction of Log *P* with Substructure-based Methods". Substructure-based methods are either fragmental (use fragments and apply correction factors) or atom based (use atom types and do not apply correction rules). Significant electronic interactions are comprised within one fragment; this is a prime advantage of using fragments. On the other hand, fragmentation can be arbitrary and missing fragments may prevent calculation. An advantage of atom-based methods is that ambiguities are avoided; a shortcoming is the failure to deal with long-range interactions. The predictive power of six substructure-based methods is compared via a benchmarking set of 284 drugs.

Igor Tetko and Gennadyi Poda focus on the "Prediction of Log *P* with Property-based Methods", which are either based on 3D structure representation including empirical approaches, quantum chemical semiempirical calculations, continuum solvation models, molecular dynamics calculations, molecular lipophilicity potential calculations, and lattice energy calculations, or on topological descriptors using graph molecular connectivity or E-state descriptors. Tetko and Poda used the same dataset of 284 drugs, and showed best predictivity for A_S+logP and ALOGPS methods, based on topological descriptors.

Finally, Franco Lombardo and colleagues consider "The Good, the Bad and the Ugly of Distribution Coefficients". The question of "how" and "what" log *D* values we use in our daily work is an important one. Sections on log *D* versus log *P*, issues and automation in the determination of log *D*, pH-partition theory and ion-pairing, and on computational approaches for log *D* are dedicated to answer this question in detail. Computational approaches for log *D* might tempt medicinal chemists to use routinely a computed value as a surrogate of measured values. However, "good" practice should be to determine at least a few values for representative compounds and continue monitoring the performance of computation with additional determinations alongside the medicinal chemistry work.

Physicochemical properties guide *Drug- and lead-likeness* in a dedicated manner. In the concluding chapter, Sorel Muresan and Jens Sadowski discuss simple calculated compound properties and related aspects in this context. The presence or absence of specific chemical features as well as their correlation with each other and with biological potency are of high importance for success in selecting starting points for lead generation and in guiding chemical optimization. A number of important concepts such as property ranges, chemical substructure filters, ligand efficiency and drug-likeness as a classification problem are discussed, and some of them are finally demonstrated in an example of how to select compounds for acquisition.