QSAR: Hansch Analysis and Related Approaches

by Hugo Kubinyi

Weinheim · New York
Basel · Cambridge · Tokyo

Hugo Kubinyi

QSAR: Hansch Analysis and Related Approaches

Methods and Principles in Medicinal Chemistry

Edited by
R. Mannhold
P. Krogsgaard-Larsen
H. Timmerman

QSAR: Hansch Analysis and Related Approaches

by Hugo Kubinyi

Weinheim · New York
Basel · Cambridge · Tokyo

Author:
Prof. Dr. Hugo Kubinyi
ZHV/W, A30
BASF AG
D-67056 Ludwigshafen

Prof. Povl Krogsgaard Larsen
Dept. of Organic Chemistry
Royal Danish School of Pharmacy
DK-2100 Copenhagen
Danmark

Editors:
Prof. Raimund Mannhold
Biomedical Research Center
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstraße 1
D-40225 Düsseldorf

Prof. Hendrik Timmerman
Faculty of Chemistry
Dept. of Pharmacochemistry
Free University of Amsterdam
De Boelelaan 1083
NL-1081 HV Amsterdam
The Netherlands

Published jointly by
VCH Verlagsgesellschaft mbH, Weinheim (Federal Republic of Germany)
VCH Publishers, New York, NY (USA)

Editorial Director: Dr. Hans F. Ebel
Production Manager: Dipl. Wirt.-Ing. (FH) Hans-Jochen Schmitt

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

Die Deutsche Bibliothek – CIP Einheitsaufnahme
Kubinyi, Hugo:
QSAR: Hansch analysis and related approaches / by Hugo Kubinyi. – Weinheim ; New York ; Basel ; Cambridge ; Tokyo : VCH, 1993
(Methods and principles in medicinal chemistry ; Vol. 1)
ISBN 3-527-30035-X (Weinheim ...)
ISBN 1-56081-768-2 (New York)
NE: GT

Printing and Bookbinding: Druckhaus "Thomas Müntzer" GmbH, D-99947 Bad Langensalza
Printed in the Federal Republic of Germany

Distribution:

VCH, P. O. Box 101161, D-69451 Weinheim (Federal Republic of Germany)

Switzerland: VCH, P. O. Box, CH-4020 Basel (Switzerland)

United Kingdom and Ireland: VCH (UK) Ltd., 8 Wellington Court, Cambridge CB1 1HZ (England)

USA and Canada: VCH, 220 East 23rd Street, New York, NY 10010–4606 (USA)

Japan: VCH, Eikow Building, 10-9 Hongo 1-chome, Bunkyo-ku, Tokyo 113 (Japan)

ISBN 3-527-30035-X (VCH, Weinheim) ISBN 1-56081-768-2 (VCH, New York)

Dedicated to Corwin Hansch

Preface

The present monograph is the first volume in a new series of handbooks entitled "Methods and Principles in Medicinal Chemistry". The prime focus of this series is an educational introduction into the current knowledge of methodological aspects and basic principles in the rapidly developing field of Medicinal Chemistry.

Potentials and limitations of techniques will be critically and comparatively discussed and comprehensively exemplified. It is intended to provide the reader with the appropriate information for applying the adequate techniques to a given problem and to avoid misleading interpretations due to the improper use of methodology. Main topics under the scope of this new publication are:

- The determination of chemical properties of biologically relevant molecules.
- Innovative approaches in the characterization of biological activity.
- Methodological aspects in deriving SAR and QSAR analyses.
- Current developments in the physiological and biochemical understanding of diseases.
- Future perspectives in the development of Medicinal Chemistry.

The first volume in the series deals with Hansch analysis and related approaches. Publication of the Hansch model in the early sixties represents the starting point of modern QSAR methodology and correspondingly the present monograph focuses on these aspects of Medicinal Chemistry. But not the historical reasons have primarily led the editors to start the series with this topic. The "classical" QSAR methods also nowadays play an important role in Medicinal Chemistry. Despite the advances in protein crystallography, molecular modeling, and structure-derived molecular design, Hansch analysis and related approaches are continuously useful tools to quantitatively derive and prove hypotheses on structure-activity relationships. In addition, the quantitative treatise of kinetic aspects of drug action remains an exclusive domain of these methods.

According to the aim of this new series Hugo Kubinyi gives a practice-oriented introduction into Hansch analysis and related approaches which familiarizes the reader with the proper application of these methodologies. The comprehensive list of references gives an excellent access to current literature and comfortably introduces the reader to fields of his special interest.

Düsseldorf
Kopenhagen
Amsterdam
Summer 1993

Raimund Mannhold
Povl Krogsgaard-Larsen
Hendrik Timmerman

A Personal Foreword

The first lipophilicity-activity relationship was published by Charles Richet in 1893, exactly 100 years ago. From his quantitative investigations of the toxicities of ethanol, diethyl ether, urethane, paraldehyde, amyl alcohol, acetophenone, and essence of absinthe (!) he concluded *"plus ils sont solubles, moins ils sont toxiques"* (the more they are soluble, the less toxic they are). One year later Emil Fischer derived the lock and key model of ligand-enzyme interactions from his results on the stereospecificity of the enzymatic cleavage of anomeric glycosides.

In the following decades the receptor concept evolved from investigations of Paul Ehrlich; a continuous development of medicinal chemistry began, leading to better and better drugs against many diseases. However, despite important contributions by Meyer, Overton, Traube, Moore, Warburg, Fühner, and Ferguson to the dependence of nonspecific biological activities of drugs on their lipophilicity (most often expressed by oil/water partitioning), the field of quantitative relationships between chemical structures and their biological activities lay dormant for about 70 years.

The discipline of quantitative structure-activity relationships (QSAR), as we define it nowadays, was initiated by the pioneering work of Corwin Hansch on growth-regulating phenoxyacetic acids. In 1962–1964 he laid the foundations of QSAR by three important contributions: the combination of several physicochemical parameters in one regression equation, the definition of the lipophilicity parameter π, and the formulation of the parabolic model for nonlinear lipophilicity-activity relationships.

This was the time when I started my Ph. D. thesis on irritant and tumor-promoting phorbol esters, their isolation, partial synthesis, and structure-activity relationships at the Max Planck Institute of Biochemistry in Munich. Indeed, one diagram in this book (Figure 43, chapter 7.4) refers to these compounds. Although I recognized a nonlinear relationship between the biological activities and the chain length of the ester groups (I even measured partition coefficients and found a nice linear dependence on the lipophilicity of the compounds), the small step from drawing a diagram to formulating a mathematical model, *i.e.* deriving a parabolic equation, was too large for me at that time. Shortly afterwards, then doing research in pharmaceutical industry, I became aware of the work of Corwin Hansch, Toshio Fujita, William Purcell, and others on quantitative structure-activity relationships. Like some of my colleagues in pharmaceutical industry I noticed this new approach but did not consider to apply it to practical drug design. For years I lived with the prejudice that QSAR is a tool to describe only more or less nonspecific biological effects, like antibacterial, antifungal, hemolytic, narcotic, and toxic activities.

My conversion from Saulus to Paulus happened after a discussion with Rudolf Gompper in Munich in 1974. In his seminar on theoretical chemistry he also mentioned the pioneering contributions of Corwin Hansch to medicinal chemistry. I presented my scepticism but, at the same time, felt ashamed of my ongoing

ignorance and decided to read some more papers. Three fortunate circumstances worked hand in hand: William Purcell's book *"Strategy of Drug Design. A Molecular Guide to Biological Activity"* had just arrived in our library and I read it in one day, fascinated by its content and style. An experienced technician helped me with his statistics programs (some months later I had discussions with a professional statistician who insisted that everything we QSAR people do is forbidden for this or that reason. I never would have started QSAR work if I had spoken to him first; now it was too late, I already was infected). A colleague provided a data set on antihistaminic compounds for which, another day later, a beautiful $\pi-\sigma$ relationship could be derived. A compound of this series came to preclinical and clinical development, but unfortunately it turned out to be only a drug for guinea pigs; it had almost no activity in humans.

After this big start I tried to understand the underlying theories and recalculated many published equations. My knowledge and experience increased, but I found a lot of numerical and also logical errors in the early QSAR literature. The consequence was to refine old models, to develop new ones, and to write scientific papers. My attempts to publish them were a difficult task. The comments of the reviewers ranged from *"much ado about nothing"* to *"wrong"* and it took a lot of patience, insistence, and several rebuttal letters to place them in the Journal of Medicinal Chemistry.

The publications of Corwin Hansch helped me to proceed. A two-month sabbatical in his group at the Pomona College followed in 1978. This visit led to a deeper understanding of quantitative structure-activity relationships and their physicochemical and biological foundations on my side. On the other hand, it stimulated Corwin Hansch to apply the bilinear model to the QSAR of enzyme inhibitors; the most interesting applications of this new model resulted from his work, from 1980 onwards.

Nowadays drug development is much too expensive to be guided by trial and error. QSAR, molecular modeling, and protein crystallography are important and valuable tools in computer-assisted drug design. The aim of this book is to give an introduction to QSAR methodology for beginners and practitioners and to present selected examples of typical applications. Comments are derived from about 20 years of practical applications, from thousands of calculated and recalculated QSAR equations. It still is my attitude to check other people's equations, especially when reviewing manuscripts. Some warnings are given and the limitations of QSAR methods will be discussed. As the commonly used methods are Hansch analysis, the Free Wilson model, and, recently coming up, comparative molecular field analysis (CoMFA), the focus is on these approaches.

Corwin Hansch initiated QSAR and he contributed the most to its development. Correspondingly, this book is dedicated to him on the occasion of his 75th anniversary in October 1993. He taught us how to apply QSAR in a proper manner to gain more insight into structure-activity relationships and biological mechanisms. The one and only way to thank him is to feel responsible to use and to develop the QSAR discipline in his sense. Thus, the book shall also be understood as a stimulus to further research on the real relationships between chemical structures and biological activities.

Heidelberg and Ludwigshafen — Hugo Kubinyi

Content

1. Introduction 1

1.1. History and Development of QSAR 4
1.2. Drug-Receptor Interactions 7

2. Biological Data. The Additivity of Group Contributions 15

3. Parameters 21

3.1. Lipophilicity Parameters 27
3.2. The Measurement of Partition Coefficients and Related Lipophilicity Parameters 32
3.3. Lipophilicity Contributions and the Calculation of Partition Coefficients 34
3.4. Polarizability Parameters 40
3.5. Electronic Parameters 42
3.6. Steric Parameters 47
3.7. Other Parameters 49
3.8. Indicator Variables 54

4. Quantitative Models 57

4.1. The Extrathermodynamic Approach (Hansch Analysis) 57
4.2. The Additivity Model (Free Wilson Analysis) 62
4.3. The Relationships between Hansch and Free Wilson Analysis (The Mixed Approach) 65
4.4. Nonlinear Relationships 68
4.5. Dissociation and Ionization of Acids and Bases 77
4.6. Other QSAR Approaches 85

5. Statistical Methods 91

5.1. Regression Analysis 91
5.2. The Significance and Validity of QSAR Regression Equations 95
5.3. Partial Least Squares (PLS) Analysis and Other Multivariate Statistical Methods 100

6. Design of Test Series in QSAR 109

7. Applications of Hansch Analysis 115

7.1. Enzyme Inhibition 116
7.2. Other *in vitro* Data 121
7.3. Pharmacokinetic Data 126
7.4. Other Biological Data 129
7.5. Activity-Activity Relationships 133

8. Applications of Free Wilson Analysis and Related Models 139

9. 3D QSAR Approaches 149

9.1. Stereochemistry and Drug Action 149
9.2. Active Site Interaction Models 151
9.3. Comparative Molecular Field Analysis (CoMFA) 159
9.4. Molecular Similarity QSAR Analyses 172

10. Summary and Conclusions 179

References 183

Index 211

1. Introduction

The interactions of drugs with their biological counterparts are determined by intermolecular forces, *i.e.* by hydrophobic, polar, electrostatic, and steric interactions. Quantitative structure-activity relationships (QSAR) derive models which describe the structural dependence of biological activities either by physicochemical parameters (Hansch analysis), by indicator variables encoding different structural features (Free Wilson analysis), or by three-dimensional molecular property profiles of the compounds (comparative molecular field analysis, CoMFA).

Drugs, which exert their biological effects by interaction with a specific target, be it an enzyme, a receptor, an ion channel, a nucleic acid, or any other biological macromolecule, must have a three-dimensional structure, which in the arrangement of its functional groups and in its surface properties is more or less complementary to a binding site. As a first approximation the following can be concluded: the better the steric fit and the complementarity of the surface properties of a drug to its binding site are, the higher its affinity will be and the higher may be its biological activity.

A complication arises from the functionalities of the biological macromolecules typically involved in ligand-protein interactions: certain structural features of the ligand determine whether a compound is

- a substrate (having a functional group which is hydrolyzed, acylated, oxidized, *etc.*, by an enzyme),
- an inhibitor (exhibiting affinity to the binding site of an enzyme, but containing no such group),
- a competitive receptor antagonist (having affinity to an agonist binding site, but mediating no receptor response),
- an allosteric receptor antagonist (binding to a different site, see below),
- a functional receptor antagonist (having no affinity to the receptor molecule, but inhibiting the receptor response *via* a different mechanism of action),
- a receptor agonist (displaying intrinsic activity in addition to affinity, *i.e.* containing certain structural features which cause the receptor to respond in a certain manner), or
- an allosteric effector molecule (binding at a different site of a protein and changing its 3D structure in such a way that a certain property of the protein, *e.g.* conformational flexibility or affinity to a substrate, an agonist, a cofactor, or any other small or large ligand is significantly changed).

The fit of the three-dimensional structure and the complementarity of the surface properties of a drug to its binding site are conditions for its biological activity. Another one, at least equally important, is that the drug has to reach this binding site. Even in simple *in vitro* systems, *e.g.* in enzyme inhibition, the surrounding water