Topics in Current Chemistry

114

Steric Effects in Drug Design

Springer-Verlag Berlin Heidelberg NewYork Tokyo

Steric Effects in Drug Design

Editors: M. Charton and I. Motoc

With Contributions by
V. Austel, A. T. Balaban, D. Bonchev,
M. Charton, T. Fujita, H. Iwamura,
O. Mekenyan, I. Motoc

With 40 Figures and 19 Tables

Springer-Verlag
Berlin Heidelberg New York Tokyo
1983

This series presents critical reviews of the present position and future trends in modern chemical research. It is addressed to all research and industrial chemists who wish to keep abreast of advances in their subject.

As a rule, contributions are specially commissioned. The editors and publishers will, however, always be pleased to receive suggestions and supplementary information. Papers are accepted for "Topics in Current Chemistry" in English.

ISBN 3-540-12398-9 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-12398-9 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data. Main entry under title: Steric effects in drug design.
(Topics in current chemistry; v. 114) Bibliography: p. Includes index.
1. Chemistry, Pharmaceutical — Addresses, essays, lectures. 2. Structure-activity relationship (Pharmacology) — Addresses, essays, lectures. I. Austel, V. (Volkhard), 1939–. II. Series. [DNLM: 1. Chemistry, Pharmaceutical. 2. Drugs. Wl T0539LK v.114/QV 744 S838]
QDl.F58 vol. 114 [RS410] 540s [615'.191] 83-4674

Printed in GDR

2152/3020-543210

Table of Contents

Introduction

Marvin Charton[1] and Ioan Motoc[2]

1 Department of Chemistry, School of Liberal Arts and Sciences, Pratt Institute, The Clinton Hill Campus, Brooklyn, NY 11205, USA
2 Max-Planck-Institut für Strahlenchemie, Stiftstraße 34–36, 4330 Mülheim a. d. Ruhr, FRG

Table of Contents

1 Introduction

Modern approaches to the design of bioactive molecules such as drugs, insecticides, herbicides, and fungicides is based on the quantification of bioactivity as a function of molecular structure. The seeds of this concept lie in the work of Meyer [1] and of Overton [2] who successfully demonstrated a dependence of bioactivity on a physicochemical parameter, the partition coefficient, which is a function of molecular structure. Vital to the development of the field was the concept of the receptor site, heralded by the work of Longley [3], stated and developed in depth by Ehrlich [4]. Biological activity according to this model depends on the recognition of a bioactive substrate (bas) by a receptor site, followed by binding of the bas to the receptor site. Realization of the dependence of bioactivity on configurations [5] led to the recognition that steric effects of one kind or another were a major factor in the potency of a bas. The work of Meyer and Overton was developed further particularly by Ferguson [6], whose work is fully described by Albert [7].

In a parallel development, structural effects on the chemical reactivity and physical properties of organic compounds were modelled quantitatively by the Hammett equation [8]. The topic is well reviewed by Shorter [9]. Hansen [10] attempted to apply the Hammett equation to biological activities, while Zahradnik [11] suggested an analogous equation applicable to biological activities. The major step forward is due to the work of Hansch and Fujita [12], who showed that a correlation equation which accounted for both electrical and hydrophobic effects could successfully model bioactivities. In later work, steric parameters were included [13].

Our objective in this work is to present surveys of the methods now available for the quantitative treatment of steric effects in the design of bioactive molecules. Commonly, this consists in the modification of a lead compound by structural changes which result in a set of bioactive substances. The bioactivity is determined and then related to structure. This is generally carried out by means of multiple linear regression analysis using a correlation equation of the type

$$Q_{ba,X} = T_1\tau_X + T_2\tau_X^2 + L\sigma_{IX} + D\sigma_{DX} + S\zeta_X + B_0 \qquad (1)$$

τ is a transport parameter,
σ_I and σ_D are the localized (field and/or inductive) and delocalized (resonance) electrical effect parameters, and
ζ is a steric parameter.

The correlation of bioactivity data with Eq. 1 or some relationship derived from it results, if successful, in a correlation equation called a quantitative structure activity relationship (QSAR).

When more than one substituent is present, the substituent effects may be assumed to be additive, or alternatively, each substituent may be parameterized separately.

Generally, the *transport parameter* used is the logarithm of the partition coefficient of the bas (bioactive substance) or some quantity derived from it. The partition coefficient is almost always determined between water and 1-octanol. Parameters obtained by chromatographic methods are being used with increasing frequency however. The term in τ^2 is introduced to account for the frequently observed parabolic dependence of a data set of bas on the transport parameter. Models other than

parabolic have also been used, in particular, the bilinear model of Kubinyi [14]. In general, this behavior is accounted for by the fact that when a bas crosses a biomembrane it must first transfer from the aqueous phase to the biomembrane, and must then transfer from the other side of the biomembrane back into an aqueous phase.

Frequently, the *electrical effect parameters* are combined into a composite substituent constant. The well-known Hammett σ_m and σ_p constants [8,9] are examples of such composite electrical effect parameters. It is the steric parameter with which we are particularly concerned here.

In order to understand the *nature of steric effects* in molecular bioactivity it is necessary to have a model of the path by which a bas exerts its effect. Such a model has been proposed by MacFarland [15]. It involves the following steps:

1. The bas enters the organism and moves to a receptor site. In the course of its trip the bas will cross one or more membranes. This may be termed the transport step.

2. The bas is recognized by the receptor (rep) and a bas-rep complex is formed. Generally, this complex is held together by intermolecular forces and its formation is reversible. This may be called the complex formation step.

3. The bas-rep complex may undergo chemical reaction resulting in the formation and/or cleavage of covalent bonds. This step may be termed the chemical reaction step.

The structure of a bas may be written in the form XGY where X is the variable substituent (or substituents), Y is the active site at which bond formation or cleavage takes place, and G is the skeletal group to which they are attached. If no reaction occurs, then the bas has the form XG.

In the *transport step* the bas must be transferred from an aqueous phase to the membrane which may be represented as a lipid phase. The transport step can be considered to be a function of the difference in intermolecular forces between water and bas and those between lipid and bas. For a neutral bas these intermolecular forces include some combination of the following:

hydrogen bonding	(hb)
dipole—dipole	(dd)
induced dipole—induced dipole	(ii)
dipole—induced dipole	(di)
charge transfer	(ct)

For an ionic bas we may add to the list above:

ion—dipole	(Id)
ion—induced dipole	(Ii)

Of this list, hb interactions can be sensitive to steric effects. This has been established for amino acid transport parameters such as the hydrophobicities and partition coefficients. There is some evidence that charge transfer interactions are also subject to steric effects.

Of the remaining interactions, those which involve the dipole moment of the bas may be subject to an indirect steric effect as the preferred conformation of the bas may depend on steric effects and the dipole moment will depend on the conformation.

The Id and Ii interactions may be subject to steric effects resulting from steric hindrance to the solvation of the group which ionizes with the nature of the substituent.

Only the ii interaction is independent of steric effects.

In the *second step* the bas is recognized by the receptor site and the bas-rep complex forms. As was noted above, the complex is generally bonded by intermolecular forces. The bas is transferred from an aqueous phase to the receptor site. The receptor site is very much more hydrophobic than is the aqueous phase. It follows, then, that complex formation depends on the difference in intermolecular forces between the bas-aqueous phase and the bas-receptor site. The importance of a good fit between bas and receptor site has been known for many years. The configuration and conformation of the bas can be of enormous importance. Also important is the nature of the receptor. If the receptor is a cleft, as is the case in some enzymes, steric effects may be maximal as it may not be possible for a substituent to relieve steric strain by rotating into a more favorable conformation. In such a system, more than one steric parameter will very likely be required in order to account for steric effects in different directions. Alternatively, the receptor may resemble a bowl, or a shallow, fairly flat-bottomed dish. Conceivably it may also be a mound. In a bowl or dish, steric effects are likely to be very different from those in a cleft. Possible examples are shown in Fig. 1, 2, and 3.

In the *third step*, when it occurs, chemical bonds are made and/or broken. Steric effects in this case should resemble those generally observed for chemical reactivity

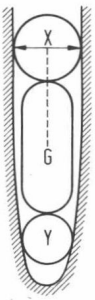


Fig. 1. "Cleft" receptor site (stylized). Steric effect of the substituent X is normal to the group axis, XG, and is shown by the arrows

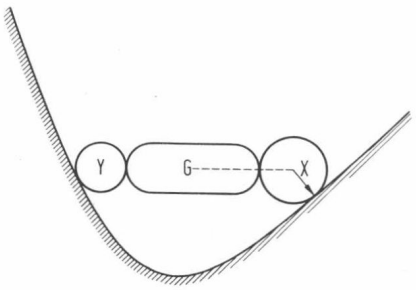


Fig. 2. "Bowl" receptor site

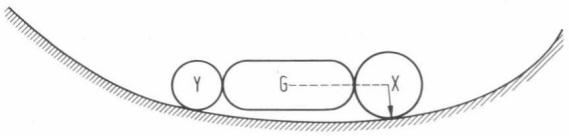


Fig. 3. "Dish" receptor site

if the substituent is in close proximity to the active site. Substituents distant from the active site may still exert a steric effect on the fit with the receptor site.

There are several distinct methods available for the parameterization of steric effects. The first group of parameters is defined from chemical reactivities (Taft E_S parameters and their modifications [16]), from Van der Waals radii and molecular geometries (Verloop [17]), and from a combination of these sources (Charton [18]). An alternative to these parameters is the use of the topological methods such as DARC-PELCO (Dubois and coworkers [19]), the branching equation (Charton [20]), molecular connectivity (Kier and Hall [21]) and minimal steric difference (Simon and Szabaday [22]). Steric parameters may also be obtained from force field calculations.

A second major area of interest in the design of bioactive molecules is that of the process of molecular recognition which is vital to the formation of the bas-rep complex. The functional groups required for recognition and activity, in their appropriate arrangement in space, or conformation, constitute the pharmacophore of a bas. The study of the possible conformations of a bas may be carried out by molecular mechanics or quantum chemical calculations. Another important factor is the nature and geometry of the receptor site.

Finally, it is necessary to consider the significance of "bulk" parameters such as molar volume, parachor and related quantities which have frequently been suggested as a measure of steric effects.

In the reviews which follow this introduction, we shall present a variety of methods for the treatment of steric effects.

2 Acknowledgements

We acknowledge with thanks the cooperation and support we have received from Springer Verlag in the course of this work. We would especially like to thank Dr. F. L. Boschke for his aid and guidance which were essential to our success.

3 References

1. Meyer, H.: Arch. Exptl. Pathol. Pharmakol. *42*, 110 (1899)
2. Overton, E.: Studien über die Narkose, Fischer, Jena 1901
3. Langley, J. N.: J. Physiol (London) *1*, 339 (1978)
4. Ehrlich, P.: Ber. dtsch. chem. Ges. *42*, 17 (1909)
5. Cushny, A. R.: Biological Relations of Optically Isomeric Substances, Balliere, Tindall and Cox, London 1926
6. Ferguson, J.: Proc. Royal Soc. *127B*, 387 (1939)
7. Alpert, A.: Selective Toxicity, Chapter 13, Methuan, London 1965

8. Hammett, L. P.: J. Am. Chem. Soc. *59*, 96 (1937)
9. Shorter, J.: Correlation Analysis in Organic Chemistry, Clarendon Press, London 1973
10. Hansen, O. R.: Acta Chem. Scand. *16*, 1593 (1962)
11. Zahradnik, R.: Arch. Int. Pharmadyn. Ther. *125*, 311 (1962)
12. Hansch, C., Fujita, T.: J. Am. Chem. Soc. *86*, 1616 (1964)
13. Hansch, C., Deutsch, E. W., Smith, R. N.: ibid. *87*, 2738 (1965)
14. Kubinyi, H.: Prog. Drug Res. *23*, 98 (1979)
15. McFarland, J. W.: ibid. *15*, 123 (1971)
16. a) Taft, R. W.: The Separation of Polar, Steric and Resonance Effects in Reactivity in: Steric Effects in Organic Chemistry (Newman, M. S. ed.), p. 556, Wiley, New York 1056
16. b) Shorter, J.: The Separation of Polar, Steric and Resonance Effects by the use of Linear Free Energy Relationships, in: Advances in Linear Free Energy Relationships (Chapman, N. B., Shorter, J. eds.), Plenum, New York 1972
17. Verloop, A., Hoogenstraaten, W., Tipker, J.: Drug Design *7*, 165 (1976)
18. Charton, M.: The Prediction of Chemical Lability Through Substituent Effects, in: Design of Biopharmaceutical Properties Through Prodrugs and Analogs (Roche, E. B. ed.), Am. Pharm. Assoc. Acad. Pharm. Sci., Washington, D.C. 1977
19. Dubois, J. E., Laurent, D., Aranda, A.: J. Chim. Phys. *70*, 1616 (1973)
20. Charton, M.: J. Org. Chem. *43*, 3995 (1978)
21. Kier, L. B., Hall, L. H.: Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York 1976
22. Simon, Z., Szabaday, Z.: Studia Biophys. (Berlin) *39*, 123 (1973)

Features and Problems of Practical Drug Design

Volkhard Austel

Dr. Karl Thomae GmbH, Biberach an der Riss

Table of Contents

1 Introduction

The search for new drugs is one of the main objectives of medicinal chemistry. The biological properties of these drugs must be significantly superior to those of known drugs so as to improve therapeutical treatment of diseases. The desirable improvements are specified by the medicinal objective and comprise pharmacodynamic, e.g. type of effect, potency, selectivity, as well as pharmacokinetic properties, such as absorption, distribution and metabolic behaviour. These properties constitute the activity profile of a drug. Since there are no generally valid rules which relate chemical structure to certain activity profiles, new drugs can only be searched for in an empirical manner. Considering the complexity of the matter it is not surprising, therefore, that nowadys on an average between 6000 and 10000 compounds have to be synthesized and tested before one of them can be introduced into the market as a new drug. The demands on the biological properties of new drugs are increasing as are those on the corresponding experimental and clinical investigations. This does not only reduce the rate of success in improving medicamental therapy but also raises the costs of new developments considerably. Since the funds for drug research are limited one has to find means of reducing the rise in these costs. The medicinal chemist, like any other scientist involved in drug research, can contribute to this aim by using experimental capacity as economically as possible.

2 Drug Design in Practice

2.1 Economization of Empirical Drug Design

How can one proceed economically in fields of research which can only be explored empirically? This is possible by designing experiments so that a maximum of new information is obtained, i.e. by optimizing the information-expense ratio. For the medicinal chemist, the term "experiment" signifies in this context test compounds and the information refers to structure-activity information. At this point two main problems arise, i.e.

— how can structure-activity information be quantified?
— how can the amount of structure-activity information which a certain compound will give be determined prior to synthesis and biological testing?

Strictly speaking, there is no objective way of quantifying structure-activity information. At best one can subjectively estimate the amount of structure-activity information a group of compounds will give relative to another one. What criteria can such an estimate be based on? It is a generally accepted rule that similar compounds will show similar biological properties and hence also give similar structure-activity information. Assuming that in an idealized case every one out of a group of potential test compounds requires the same experimental expense for synthesis and testing and that two of these compounds are to be selected, we can envisage two extreme cases, i.e. one in which the two compounds are structurally very similar and alternatively, one in which two compounds do not very much resemble each other.

In the first case both compounds would give roughly the same structure-activity information, whereas in the second case largely independent information will be obtained. Therefore, the first case exemplifies a comparatively unfortunate choice since the information-expense ratio is only half as good as that of the second case. This very simplified example shows one of the main ways in which the medicinal chemist can influence the costs of drug research, i.e. through proper design and selection of test compounds. It is an essential even though not sufficient requirement of economical drug research that test compounds are chose so as to be mutually dissimilar.

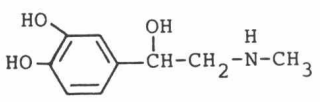
2.2 Selection of Test Compounds

2.2.1 Similarity of Chemical Compounds

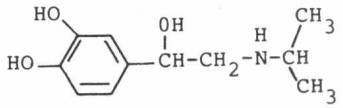
When are chemical compounds similar and how can similarity be objectively determined?

The similarity measure intuitively used by organic chemists is the number of structural features and their mutual arrangement which two compounds have in common. From this point of view adrenalin (*1*) and isoproterenol (*2*) would look rather similar and in fact both have similar biological properties as they are both agonists of β-adrenergic receptors. Both compounds are comparatively dissimilar to propranolol (*3*) even though a certain resemblance is still maintained. This situation is again reflected in the biological properties. Thus, propranolol can still interact with β-receptors but can no longer stimulate them, i.e. contrary to the former two compounds propranolol is an antagonist.

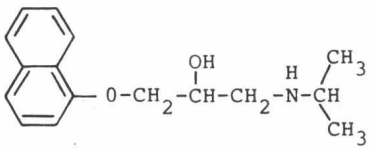
On the other hand toluene (*4*) and 5-methoxy-indole (*5*) would not be considered very similar from a conventional point of view, yet their hemolytic activity is the same ($\log 1/C = 1.93$, where C is the minimal molar concentration which causes 100% hemolysis in rabbit erythrocytes [1]). Obviously, topological similarity or dissimilarily is not always a factor which is relevant for biological activity. In the present example the relevant feature is a physicochemical property of the compounds, i.e. lipophilicity as expressed by the octanol/water partition coefficient



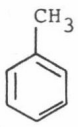
1



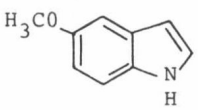
2



3



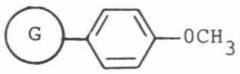
4



5

(log $P_{h/w}$ = 2.11 and 2.10 for *4* and *5* respectively) [2]. The idea that the biological properties of chemical compounds could be a function of their physico-chemical properties was proposed in the last century by Crum-Brown and Frazer [3,4]. Decisive progress in the understanding of the relationships between chemical structures of compounds and their biological activity came from the work of Hansch. In 1963, Hansch [5] showed that biological properties can be quantitatively related to physico-chemical properties of compounds with the aid of an extrathermodynamic model of drug action. Since then thousands of quantitative structure-activity-relationships have been found. Lists of several hundred examples have been published [6–8]. Considering this large background of evidence, there can be little doubt nowadays that the biological properties of chemical compounds are determined exclusively by their physico-chemical properties. For quantitative description one uses parameters which represent the corresponding property either explicitly or implicitly.

Thus, the commonly applied parameters π, σ and E_s are explicit representations of lipophilic, electronic and steric properties respectively. Indicator variables, on the contrary, frequently refer to fixed combinations of physico-chemical properties. For example, an indicator variable which denotes the presence of a 4-methoxy-group in general structure *6* refers simultaneously to all physico-chemical properties of this substituent, i.e. to its contribution to the lipophilic (π), electronic (σ), and steric properties (E_s) of the system.



6

G is some basic pharmacophoric system

The Free-Wilson method of deriving quantitative structure-activity-relationships [10] uses implicit representations of physico-chemical properties and there are also numerous examples where indicator variables have been successfully included in the Hansch approach.

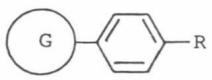
The possibility of describing chemical structures numerically with the aid of physico-chemical parameters and indicator variables puts us in the position to determine similarity or dissimilarity of chemical compounds more objectively. Chemical compounds can be represented as points in an n-dimensional space whose coordinates are formed by the parameters which are used to characterize the compounds. This space is therefore called parameter space. The distance of two

points in parameter space can be taken as a measure of the similarity of the corresponding compounds [11]. The longer the distance the less similar the two compounds are.

As has been mentioned previously, mutual dissimilarily of the test compounds is a necessary but not sufficient criterion for a properly designed test series. In addition, the parameters which have been chosen to characterize the compounds must vary independently of oneanother. In order to fulfil both criteria an appropriate test series has to be designed so that the corresponding points are evenly distributed over the respective parameter space.

2.2.2 Determination of the Structural Area to be Investigated

Normally it is hardly feasible nor necessary to explore a certain parameter space to the full theoretically possible extent. Rather certain areas of parameter space are examined. These areas can be determined by intuition, synthetic accessibility or by structure-activity-relationships and structure-activity-hypotheses. Consider for example structure *6* to be a lead compound which needs to be optimized with respect to an activity profile consisting of two components, i.e. potency and oral effectivity (gastro-intestinal absorption). Let us for simplicity assume that we want to confine ourselves to variations of the substituent in 4-position (*7*) and that we



7

know size, lipophilicity and electronic properties to be the decisive factors for the biological properties under consideration. If we know for instance that a substituent whose MR-value (taken as a measure for size) exceeds 20 renders the respective compounds more or less inactive and that gastro-intestinal absorption is only significant if the lipophilicity contribution of R is larger than — 0.5 (in π-terms), we can impose appropriate limitations on the area in parameter space which we want to investigate. With respect to electronic properties and to the maximum lipophilicity no limitations would exist a priori. The medicinal chemist will now by intuition and under accessibility considerations conceive a basic set of structures (in this simple case, substituents) which fall within the permissible area of parameter space, e.g. the ones shown in Table 1. In this Table lipophilicity, size and electronic effects are represented by π, MR and σ_p respectively. As can be seen from the values in Table 1, intuition and synthetic accessibility have brought about additional limitations to the investigated area in parameter space, i.e. — 0.84 and 0.78 in σ_p-direction, and 1.98 as the upper limit in π-direction. In MR-direction the lowest possible value (0.92 for fluorine) is reached.

2.2.3 The Number of Test Compounds

Having defined a relevant parameter space or an area within it respectively and a basic set of possible structures we can now return to our original problem, i.e.

Table 1. Set of substituents which fall within a permissible area in a three dimensional parameter space formed by lipophilic (π), steric (MR) and electronic (σ_p) coordinates. The limitations are: MR < 20, $\pi > -0.50$ (values from Ref. [9])

No.	R	π	MR	σ_p
1	H	0.00	1.03	0.00
2	Br	0.86	8.88	0.23
3	Cl	0.71	6.03	0.23
4	F	0.14	0.92	0.06
5	NO_2	−0.28	7.36	0.78
6	CF_3	0.88	5.02	0.54
7	CH_3	0.56	5.65	−0.17
8	OCH_3	−0.02	7.87	−0.27
9	SCH_3	0.61	13.82	0.00
10	$NHCH_3$	−0.47	10.33	−0.84
11	$COOCH_3$	−0.01	12.87	0.45
12	C_2H_5	1.02	10.30	−0.15
13	OC_2H_5	0.38	12.47	−0.24
14	SC_2H_5	1.07	18.42	0.03
15	NHC_2H_5	0.08	14.98	−0.61
16	$N(CH_3)_2$	0.18	15.55	−0.83
17	$COOC_2H_5$	0.51	17.47	0.45
18	OC_3H_7	1.05	17.06	−0.25
19	$CH_2(NCH_3)_2$	−0.15	18.74	0.01
20	t-C_4H_9	1.98	19.62	−0.20

to select test compounds which are evenly distributed over a certain area in parameter space. The first decision which needs to be made refers to the number of test compounds. This number is mainly dependent on the dimensionality of the investigated parameter space and on the stage of the particular drug development project. In order to keep the variables which form the parameter space independent the number of test compounds must exceed the number of variables (dimensions of the parameter space) by at least one. This number will suffice in very early stages of lead optimization when a large number of potentially relevant parameters has to be considered. In more advanced stages when one is more confident about the relevant parameter space and when one tries to obtain a more detailed picture of the distribution of strongly and weakly active compounds in parameter space, two to four test compounds per parameter are a minimum requirement. Such numbers allow the location of the more interesting compounds in parameter space to be estimated. In very advanced stages when one tries to operate with quantitative structure-activity-relationships (QSAR) a minimum of six to ten compounds per dimension of parameter space are required in order to obtain meaningful results.

2.2.4 Methods for Selection of Test Compounds

After the number of test compounds which are to be selected from the basic set has been decided upon one can apply one of the series design techniques reported in the literature. Some of these methods apply to a restricted number of parameters (e.g. two in the Craig graphical method [12,13]) and structural variations (e.g. a limited number of aromatic and aliphatic substitution patterns as in the Topliss manual