Translated from *Uspekhi Khimii*, 47, 739-772 (1978) U.D.C. 615.015.11

Quantitative "Chemical Structure-Biological Activity" Relations

S.V.Nizhnii and N.A.Epshtein

The principal methods used to obtain quantitative relations between chemical structure and biological activity are reviewed and the postulates, possibilities, and prospects of these methods are examined.
The bibliography includes 299 references.

CONTENTS

I.	Introduction	383
II.	The method of substructural analysis. The pattern recognition method	385
III.	The mathematical additive model	386
IV.	The multiparameter regression method	387
V.	Analysis of the structure-activity relation by Hansch's method	390
VI.	The quantum-statistical approach	391
VII.	The use of factorial analysis in the derivation of structure-activity relations	392
VIII.	The use of semiempirical methods in the study of the mechanisms of reactions of biologically active substances	393
IX.	The scope of practical applications. Advances, difficulties, and prospects	394
X.	Conclusion	395

I. INTRODUCTION

The determination of the relations between the chemical structures of substances and their pharmacological activities is the most important task in the synthesis of highly effective and selective medicinal drugs. Biochemical pharmacology is as yet incapable of formulating rigorously the problems of specific synthesis. Nevertheless, an enormous scientific effort and capital investment are being devoted to the development of new drugs[1]. In this situation one is justified in attempting to compile empirical and semiempirical rules linking individual characteristics of the structures of substances and their effect on biological systems.

Extensive experimental data have now accumulated showing that there are definite relations between the physicochemical properties of a substance and its biological activity. The elucidation of the origin of the empirical relations between the properties of compounds *in vitro* and their biological action on the organism should make it possible to arrive at ideas concerning the mechanisms of biologically important chemical reactions and to establish on this basis well-founded recommendations for chemists working in the field of synthesis and, on the other hand, to predict the possible specificity and effectiveness of the biological action of newly synthesised substances, shortening thereby the laborious and expensive stages involving biological tests.

The possibility of determining structure-activity relations arose approximately one hundred years ago when it was observed[2] that the intensity of the narcotic action of certain substances decreases with increase of their solubility in water. A little later a relation was noted between narcotic activity and the distribution of substances in a water-lipid system[3,4]. A whole series of physicochemical parameters of substances, whose relation with narcotic activity and toxicity has been elegantly described by Lazarev[5,6], was established subsequently.

Despite the limited and qualitative nature of the majority of structure-activity relations observed in the first half of the XXth century, the justification of the hypothesis that the biological activity of a substance can be accounted for in individual instances by *in vitro* studies and described on the basis of the laws of physical chemistry and in terms of its concepts must be regarded as the general outcome of such relations.

Modern approaches to the structure-activity problem (based largely on the principles of physical and theoretical chemistry) are examined below. With their aid one may now hope to restrict the number of total syntheses and the biological screening of all substances which seem to be promising as medicinal drugs on the basis of general considerations.

The assertion that the structure of a substance and the biological response caused by it are involved in an unambiguous cause and effect relation, which is usually readily accepted on trust, must in fact be regarded as only one of the postulates underlying the approach to the structure-activity problem[7,8]. The most general postulates of this type may be reduced to the following.

1. The structure S of a chemical compound Y, designated by S_y, can be expressed in terms of a set of structural parameters $S_y = \{S_y^i\}$, where $i = 1, 2, \ldots, n$.

2. There is a relation between the structure of a substance and its biological activity, i.e. there is a function $F_y(S_y, A_y)$ relating the activity A of the substance Y to its individual characteristics—$F_y(S_y^i, A_y)$.

3. If the form of the function $F(S, A)$ has been established, the function may be extrapolated to compounds of the series Y_1, Y_2, Y_3,... with similar structures, differing to the substituents X_1, X_2,..., or by their positions in the molecule.

4. The sequence of the activities of a series of compounds, for example, $A_{X_1} > A_{X_2} > A_{X_3}$ may be deduced by analysing the set of functions $F_{Y_1}(F_{Y_1}, A_{Y_1})$, $F_{Y_2}(S_{Y_2}, A_{Y_2}), \ldots$.

The prediction of biological activity is restricted to a series of compounds related by common structural features, it being implied that a "leading" structure has already been established and that the application of a particular approach should reduce the number of experimental

tests on the biological activities of the members of the series.

In order to obtain a working model [i.e. $F(S, A)$], the reliability of which is to be subsequently estimated, it is suggested that the molecules of a biologically active substance undergo a series of physicochemical and chemical transformations in the organism: dissolution, sorption, partition, binding, chemical reactions, segregation, etc. Evidently, any of these stages or their combination may determine in each specific instance the specificity of the pharmacological action of the substance or may constitute the rate-limiting stage in the overall effect, determined as biological activity.

If the first condition holds, i.e. the physicochemical characteristics of the substance reflect its structure by a set of some of its parameters, then, in order to determine the form of $F_Y(S_Y, A_Y)$, it is necessary to obtain A in a quantitative form, namely $[A]$†. The quantity $[A]$, usually determined by means of biological tests, is, firstly, a statistical mean quantity and, secondly, it can reflect only the final results of the effect of the substance on the organism or the system simulating it. In fact the final result may be a consequence of one or several stages of the action of the substance. Hence it is clear that the most complex task in the determination of $F_Y(S_Y, A_Y)$ is the parametrisation of $[A]$.

If the observed biological response BR to the organism or the system simulating it to a certain substance is a consequence of a sequence of physicochemical processes involving the substance, then the time t elapsed between the onset of the action of the substance and its final result may be related to a stage controlling the rate of the process. If the substance at an initial concentration C_0, giving rise to a standard biological response, is introduced into the biological system, there is a probability N that an amount of the substance C_0N will reach sites with which it reacts and this leads to [BR]† after a period Δt. In this case the rate of the process can be represented in the following form (for a detailed treatment, see Section IV)[8,11,12]:

$$\frac{d[BR]}{dt} = KC_0N\bar{k}_i, \qquad (1)$$

where K is a proportionality coefficient relating the rate of the biological response[11,12] to the rate of the controlling stage, while $\bar{k}_i$ is the rate constant or equilibrium constant for this stage.

The above considerations and Eqn. (1) have not been rigorously proved. The point is that the standard biological response (ED_{50}, LD_{50}, etc.) does not contain in an explicit form information about the rate $d[BR]/dt$. Furthermore, in a real case the rate constant for the process $\bar{k}_i$ may be determined by any individual stage, a multiple repetition of such stages, or their combination. Function (1) should therefore be regarded merely as formally reflecting the relation between [BR]and a hypothetical rate-determining steady-state physical or chemical process. However, by introducing the additional postulate

$$\frac{d[BR]}{dt} = \text{const}, \quad \text{i.e.} \quad KC_0N\bar{k}_i = \text{const}, \qquad (2)$$

†Here and henceforth $[A]$ and [BR] are a quantitative measure of the activity, expressed in terms of the concentration, and a quantitative measure of the biological effect, expressed in terms of arbitrary units, respectively.

It is possible to obtain a set of functions $F_Y(S_y^i, A_Y)$ analogous to the experimentally observed set[9-11] (for further details, see Section IV), taking into account the fact that $1/C_0$ is usually employed as $[A]$.

The choice of $1/C_0$ as a measure of the effectiveness of the biological action of the substance $[A]$ cannot at present be rigorously justified. Nevertheless, the legitimacy of employing $1/C_0$ as a measure of A follows to some extent from experimental data[13-16] and a model of the equilibrium distribution[17]. Indeed if the chemical potentials of the substance at the site of its administration and at the site of its action are the same[17], its initial concentration (C_0) must determine its concentration in a specific region of the biophase (C_B), binding to which (i.e. reaction with this region) gives rise to the biological effect, since in this case the law of the partition of the substance between the phases $C_B/C_0 = \text{const}$ holds, whence the validity of $1/C_0$ as a criterion of the effectiveness of the biological action is evident. On the other hand, if the biological action of the substance is determined by the inhibition of a specific enzyme of the biophase, its inhibition rate constant may be used to estimate $[A]$ (O'Brien[13]); experimental evidence for the relation between $1/C_0$ and the inhibition rate constant has been obtained[15,16,18].

The theoretical case where the time of attainment of a definite state of inhibition of the enzyme is inversely proportional to C_0 is obtained for second-order reactions. Indeed, in this case $[A] = K^i_{0.5} = k_1(1/C_0)$ to within the constant k_1. In fact postulate (2) or its logarithmic form

$$\lg[A] = \lg(1/C_0) = \text{const} + \lg N + \lg K + \lg \bar{k}_i \qquad (3)$$

imply merely that, for the same rate of the rate-limiting stage, all the substances in the series, i.e. related compounds, should give rise to an identical biological response over a fixed time interval regardless of their structure and composition. Because of this, the quantity $-dC_{cr}/dt$, in the form $-dC_{cr}/dt = \bar{k}_i NC_0$, should be introduced directly into Eqn. (1) and not the formal quantity $d[BR]/dt$. In this form the problem is stated correctly and reduces to the determination of the dependence of $\lg(1/C_0)$ on various physicochemical parameters.

In order to facilitate the mathematical interpretation of Eqn. (3), it is assumed that the reaction limiting the rate of the process at a given stage is of first (or pseudofirst) order. This is possible if (1) only a small proportion of the reacting active centres in the organism interact with the test substance and (2) the substance is present in the organism in a considerable excess relative to the number of reactive sites. The second instance is frequently obtained under *in vitro* conditions, i.e. in a model physicochemical system, while the first probably obtains under the real conditions occurring when an organism is acted upon by the given substances.

The origin of the structure–activity relations is deduced by comparing the results of independent physicochemical and biological studies on the relevant substances. Table 1 presents the forms of biological action for which a correlation between the type of biological action of the substance and its physicochemical characteristics has been established empirically.

If it is assumed that Eqn. (3) is correct, then, for a quantitative expression of the influence of substituents on $\lg \bar{k}_i$ and hence on $\lg[A]$, it is necessary to take into account, to a first approximation, the hydrophobic, electronic, and steric components of the interaction, which may be represented in the form of their linear combination:

$$\lg \bar{k}_i = f_h(S_h) + f_e(S_e) + f_s(S_s). \qquad (4)$$

On combining Eqns. (3) and (4), we obtain

$$\lg [A] = \lg (1/C_0) = f_h(S_h) + f_e(S_e) + f_s(S_s) + \lg N + \text{const.} \quad (5)$$

The specific form and significance of each term of Eqn.(5) can be found in different ways as a function of the model adopted or the available experimental observations (see below).

Table 1. The forms of biological action for which a quantitative correlation has been established with the physicochemical parameters of substances and the parameters of their model interactions.

Physicochemical characteristics of substance	Type of biological action	References	Sources of tabulated values of physicochemical parameters
Logarithm of partition coefficient lg P	Bacteriostatic activity	77,84,87, 147-149	
	Narcotic activity	12,135,150, 151	41,176-178
	Inhibition of enzymes	12,18,77, 152,153	
	Protein binding	12,53,135, 154-156	
	Anticonvulsant activity	77,157	
	Local anaesthetic activity	77,151	
	Permeability of tissues or artificial membranes	12,158,159	
	Spasmolytic activity	77,135	
	Carcinogenic activity	160	
	Inhibition of absorption of oxygen	135	
	Localisation of boron in brain	12,161	
	Toxicity to multicellular organisms	162	
	Inhibition of cell division	135	
	Androgenic activity	163	
	Anabolic activity	163	
	Antitumour activity	164	
	Inhibition of luminescence	135	
	Action on central nervous system	165-167	
	Antagonism to histamine, serotonin, etc.	135	
	Fibrinolytic activity	156,168	
	Inhibition of Hill reaction	169-171	
	Analgesic activity	150	
	Anti-inflammatory activity	172	
	Growth-activating activity	173	
	Anti-electric shock activity	174	
	Antifungal activity	175	
Lipophilic parameter π	Bacteriostatic activity	77,81,147 179-182	41,52,72, 177,178
	Narcotic activity	12,60	210,211
	Inhibition of enzymes	18,86,179, 183-191	
	Protein binding	53,156,192	
	Antimalarial activity	193	
	Insecticidal activity	194	
	Local anaesthetic activity	195	
	Antihelminthic activity	88	
	Sweetness (relative)	196	
	Hypotensive activity	197	
	Localisation of boron in brain	12	
	Synergistic activity	11,76	
	Inhibition of Hill reaction	198	
	Natrivretic activity	91	
	Haemolysis of erythrocytes	128,148	
	Adrenergic blocking activity	199,200	
	Antiarrhythmic activity	94	
	Action on potential of axon	128	
	Inhibition of complement	201	
	Progesterone etc. activity	144,202	
	Antifungal activity	203	
	Phytotoxic activity	81,179	
	Inhibition of cell division	204	
	Antiphagocytic activity	205	
	Toxicity to multicellular organisms	179	
	Acaricidal activity	206	
	Antogonism to histamine, serotonine, etc.	207	
	Inhibition of saliva secretion	179	
	Fungicidal activity	98	
	Herbicidal activity	68,208	
	Anti-inflammatory activity	179	
	Depolarising action on electrically excited membranes	179	
	Inhibition of mitochondrial electron transport	209	
Thin-layer chromatographic (t.l.c.) parameters R_M and ΔR_M	Bacteriostatic activity	57,182,212	57,178,212,
	Toxicity to multicellular organisms	55,212	216,217
	Protein binding	96,178,212	
	Androgenic activity	163	
	Synergestic activity	163	
	Haemolysis of erythrocytes	148,178,212	
	Hypoglycaemic activity	96	
	Anabolic activity	163,212	
	Inhibition of mitochondrial electron transport	209,212,213	
	Progesterone etc. activity	214	
	Inhibition of Hill reaction	179	
	Analgesic activity	212	
	Fungicidal activity	212	
	Herbicidal activity	212	
δ t.c.	Inhibition of Hill reaction	179,215	215
Parachor Pr, $g^{1/4}$ cm^3 s$^{1/2}$ mole^{-1}	Bacteriostatic activity	42,60,62	43,60,61 220-222
	Narcotic activity	42,60,61, 218	
	Protein binding	60	
	Inhibition of embryonic development	60	
	Inhibition of cell division	42	
	Thyromimetic activity	62	
	Local anaesthetic activity	62	
	Fibrinolytic activity	62	
	Haemolysis of erythrocytes	62	
	Action on potential of axon	62	
	Sympatheticomimitic activity	62	
	Parasympatheticolytic activity	62	
	Anti-inflammatory activity	219	
Molar volume V_M, cm^3 mole^{-1}	Bacteriostatic activity	42,70	44
	Narcotic activity	42,61,70	
Molecular connectivity ^xM	Narcotic activity	67	67
	Enzyme inhibition	67	
	Antifungal activity	67	
Molecular weight M, g mole^{-1}	Narcotic activity	60,61	223
	Protein binding	60	
	Inhibition of embryonic development	60	
Logarithm of solubility in water lg C_S (mole litre^{-1})	Bacteriostatic activity	42	224,225,300
	Narcotic activity	42	
Partial pressure (in model system) p, mmHg	Anaesthetic activity	179,26	
Hildebrand's solubility parameter δ, cal$^{1/2}$ cm$^{3/2}$	Haemolysis of erythrocytes	227	65,228,229
Molar attraction constant Fa, cal$^{1/2}$ cm$^{3/2}$ mole^{-1}	Narcotic activity	60,61,66	60,66
	Inhibition of embryonic development	60	
Van der Waals' constant a, litre^2 atm mole^{-2}	Narcotic activity	142	230
Hammett's constant σ (σ_M, σ_π)	Bacteriostatic activity	12,147,180	69,72,210
		181	233-236
	Enzyme inhibition	12,18,77 184-188, 231,232	
	Haemolysis of erythrocytes	12	
	Local anaesthetic activity	195,237	
	Antagonism to histamine, serotonin, etc.	238	
	Sweetness (relative)	196	
	Inhibition of mitochondrial electron transport	209	
	Toxicity to multicellular organisms	162,179	
	Andrenergic blocking activity	199,200	
	Inhibition of hormonal activity	239	
	Fibrinolytic activity	156	
	Hypotensive activity	197	
	Antimalarial activity	193	
	Antiphagocytic activity	205	
	Acaricidal activity	206	
	Growth activating action	173	
	Antielectric shock activity	174	
	Antifungal activity	175	
	Herbicidal activity	68,208	
Inductive and resonance components σ_I and σ_R of Hammett's constant	Enzyme inhibition	18,129,185, 240	235,241
	Antagonism to histamine, serotonin, etc.	238	
	Antiviral activity	242	
Electron donor and electron acceptor substituent constants σ^+, σ^-, $\Delta\sigma_R^+$	Enzyme inhibition	12,18,77 129,179	69,90,235 241,243-24
	Haemolysis of erythrocytes	12	
	Inhibition of complement	201	
	Toxicity to multicellular organisms	12	

Table 1. (cont'd).

Physicochemical characteristics of substance	Type of biological action	References	Sources of tabulated values of physicochemical parameters
Taft's inductive constant σ^*	Bacteriostatic activity	87,149	235,245,246
	Enzyme inhibition	18,195,240	
	Anticonvulsant activity	157	
	Adrenergic blocking activity	127	
	Action on central nervous system	167	
	Inhibition of saliva secretion	179	
	Antielectric shock activity	174	
	Depolarising action on excited membranes	179	
	Antifungal activity	175	
Radical substituent constants $\sigma^{\cdot}$ and E_R	Bacteriostatic activity	77,148	74,76
	Enzyme inhibition	18,77	
	Local anaesthetic activity	77	
	Synergistic activity	12,76	
Swain and Lupton's parameters $\mathcal{F}$ and $\mathcal{R}$	Antagonism to histamine, serotonin, etc.	238	69,72,210
	Enzyme inhibition	18,185	
	Progesterone activity	202	
Charge transfer constant C_T	Enzyme inhibition	18,78	78
	Carcinogenic activity	79,80	
Change in ionisation constant ΔpK_a	Bacteriostatic activity	81,79,181, 182	247,248,284
	Enzyme inhibition	12,13,18,77, 179-185,191	
	Inhibition of oxygen adsorption	135	
	Inhibition of cell division	135	
	Inhibition of mitochondrial electron electron transport	209	
	Anti-inflammatory activity	179	
	Natriuretic activity	91	
	Phytotoxic activity	81	
Molar electronic polarisation P_E, cm^3 $mole^{-1}$	Bacteriostatic activity	85,87	
	Enzyme inhibition	18,86,185	
	Anaesthetic activity	249	
Polarisability α, cm^3	Local anaesthetic activity	83	60,83
Molecular refraction MR, cm^3 $mole^{-1}$	Enzyme inhibition	18,153,188 232	72,210,222, 251,252
	Inhibition of complement	201	
	Narcotic activity	61	
	Progesterone activity	202,250	
Dipole moment μ, D	Antihelminthic activity	88	253
	Enzyme inhibition	18,152	
Ionisation potential I, eV $mole^{-1}$	Narcotic activity	254,255	83,257,258
	Local anaesthetic activity	83	
	Carcinogenic activity	79,256	
Polarographic half-wave potential $\varphi_{1/2}$, V	Bacteriostatic activity	84	230
Taft's steric parameter E_S	Bacteriostatic activity	37,149	128,235
	Enzyme inhibition	18,129,185, 186,188, 190,240	
	Anticonvulsant activity	157	
	Antagonism to histamine, serotonin, etc.	102	
	Antiviral activity	8	
	Synergistic activity	11,76	
	Motor activity	259	
	Fibrinolytic activity	156,168	
	Antiphagocytic activity	205	
	Antielectric shock activity	174	
	Antiallergic activity	97	
	Antifungal activity	175	
Corrected Taft's steric parameter E_s^c	Enzyme inhibition	195	235
	Adrenergic blocking activity	127	
	Depolarising action on excited membranes	179	
Van der Waals' radius r_V, Å	Bacteriostatic activity	37	260
	Adrenergic blocking activity	200	
Surface tension γ, dyn cm^{-1}	Narcotic activity	261,262	221,265,266
	Action on central nervous system	263	
	Local anaesthetic activity	264	
Surface pressure π_s, dyn cm^{-1}	Action on central nervous system	226,263,257	265
	Local anaesthetic activity	237	
Hydrophilic-lipophilic balance HLB	Modification of herbicidal activity	68	268
Molar heat of adsorption Q_{ads}, cal $mole^{-1}$	Permeability of tissues and artificial membranes	269	269
	Inhibition of oxygen absorption	269	
	Inhibition of luminescence	269	
Logarithm of specific rate of dissolution $\lg K_{diss}$, μmole cm^{-2} s^{-1}	Local anaesthetic activity	59	
Constant for binding by a model receptor (protein, etc.) K_b	Analgesic activity	145	41
	Antitumour activity	270	
	Carcinogenic activity	271,272	
Molar diamagnetic susceptibility χ, cm^3 $mole^{-1}$	Narcotic activity	61	202

Table 1. (cont'd).

Physicochemical characteristics of substance	Type of biological action	References	Source of tabulated values of physicochemical parameters
Chemical shift of atoms X in NMR spectrum δ^X, p.p.m.	Hypoglycaemic activity	90,96	274-279
	Natriuretic activity	91	
	Fungicidal activity	98	
Spectroscopic data: vibration frequency ν, s^{-1} wavenumber ω, cm^{-1}	Bacteriostatic activity	95	
	Antiarrhythmic activity	94	
	Antiallergic activity	97	
Peripheral parameter characteristing the properties of substituents in steroids c	Progesterone, etc. activity	144	250
Principal quantum-chemical parameters			
overall charge of π-electrons at r atom q_r or ϵ,C	Bacteriostatic activity	120	
	Enzyme inhibition	18,109,232	
overall $(\sigma + \pi)$ charge of electrons at r atom Q_r^1, C		114,120,	
	Bacteriostatic activity	280	
	Anticholinergic activity	113	
	Enzyme inhibition	18,110,152	
Energies of highest occupied and lowest unoccupied molecular orbitals E_{HOMO} and E_{LUMO}, eV	Bacteriostatic activity	280	
	Local anaesthetic activity	83	
	Analgesic activity	281	
	Hypotensive activity	115	
	Action on central nervous system	282,283	
	Inhibition of cell division	204	
Electrophilic (E) and nucleophilic (N) capacities for ultradelocalisation S_r^E and S_r^N	Bacteriostatic activity	114,120	
	Anticholinergic activity	113	
	Hypotensive activity	115	
	Nicotine-like activity	12	
Free valence index F_r	Bacteriostatic activity	114	
Atomic-orbital coefficient of r atom C_r	Bacteriostatic activity	120	

We shall adopt the convention that all the approaches to the structure–activity problem, including the examination [in terms of Eqn. (3)] of specific physicochemical characteristics of the given substance, are semiempirical.

The contribution of the second group of methods (empirical methods) consists in a systematic analysis of studies of biological activity in order to discover the structural elements and their combinations characteristic of a particular type of activity. We have included in the same group methods for the estimation of the mathematical contribution of definite structural elements to the biological activity of the compound.

II. THE METHOD OF SUBSTRUCTURAL ANALYSIS. THE PATTERN RECOGNITION METHOD

Both these methods are based on the hypothesis that the biological activity of a substance is due to the presence in its molecule of definite structural elements (substructures).

The frequency of the activity of the substructure L characterises the contribution which the substructure makes to the probability that the compound is biologically active [19]. It is expressed numerically as the ratio of the number of biologically active substances containing a definite structural unit (B) to the total number of compounds analysed where this substructure is also encountered (T), i.e. $L = B/T$.

Using a descriptive system of fragment coding [20,21], capable of distinguishing 1200 structural fragments—functional groups, chains, rings, etc.—and 110 types of

their combinations, a formalism[22] and a computer program were developed; the latter was capable of determining L for compounds differing appreciably in their chemical structure and exhibiting a known type of biological activity, which were compared with the same values for compounds analysed for potential biological activity.

The attempt[19] to estimate the antiarthritic-immunoregulatory activity of 850 preparations of different structures and with different mechanisms of their pharmacological activity in terms of the criterion L, regardless of tests based on physiological methods[23], led to the following findings:

(1) simple identification of the structural fragments characteristic of the most active substances of this group does not permit the choice, as was to be expected, of the optimum compounds and hence the determination of the sequence based on their increasing or decreasing activities;

(2) the experimental frequency tables for the activities of substructures, including the values of L for 492 structural fragments encountered in the most active compounds with a given type of activity, made it possible to establish a series of compounds based on decreasing activities and consistent with the experimental series‡.

Next the same program was used to search for substances with antiarthritic-immunoregulatory activity among 703 substances which had not yet been tested. It was shown[19] that the criterion L is suitable not only for the determination of the potential activities of a wide variety of types of compounds but also for the recommendation of the required combination of substructures in the synthesis of compounds previously unknown as antiarthritic-immunoregulatory agents.

Naturally in the form in which it has been proposed the criterion L cannot be referred directly to structure-active relations. When the pattern recognition method is applied[24,25], apart from the substructural principle of characterising substances their physicochemical parameters are introduced as well. In this case, a series of preselected chemical compounds with a definite specificity of their biological action (learning series including more than 100 substances) is represented by a surface in n-dimensional space, where the coordinates are formal equivalents of the structural fragment and the physicochemical parameters of the substance; the points on the surface correspond to individual compounds. The test substance or group of substances is treated formally in terms of the same features as the learning series after which one finds how close they are located in the given n-dimensional space to the points corresponding to the active compounds in the learning series. A special method[25] makes it possible to formulate individual groups of features in terms of space vectors, which makes it possible to reduce the number of variables and to discover fragments characteristic of active compounds. According to Chu et al.[24], the combined use of fragment and physicochemical criteria to characterise substances permitted the prediction with 80–90% probability of antitumour activity in a series of 17 previously uninvestigated compounds.

‡In order to avoid "discoveries" compounds constituting the basis of the compilation of the Tables, these compounds were excluded from the total number of substances tested.

III. THE MATHEMATICAL ADDITIVE MODEL

The mathematical additive model (MAD)[26] is based on the hypothesis of the independent additive contribution of substituents to $[A]$ for a number of compounds Y_1, Y_2, Y_3,..., differing by the substituents X_1, X_2, X_3,..., the mathematical contribution of the molecular skeleton being assumed to be constant. Accordingly one can write

$$[A] = \sum_i a_i x_j + \lambda, \tag{6}$$

where a_i is the mathematical expression for the contribution of the ith substituent to $[A]$, $x_j = 0$ or 1 depending on the absence or presence of the ith substituent, and λ is the constant mathematical contribution of the skeleton to $[A]$.

The MAD imposes an additional limitation on the overall contribution of substituents to biological activity for the same position in the entire series of compounds. For example, if three substituents A, B, and C may exist in the same structural element, then

$$\sum_i^k a_i(\mathrm{A}) + \sum_i^l a_i(\mathrm{B}) + \sum_i^q a_i(\mathrm{C}) = k a_i(\mathrm{A}) + l a_i(\mathrm{B}) + q a_i(\mathrm{C}) = 0, \tag{7}$$

where k, l, and q are the numbers of compounds containing the substituents A, B, and C respectively in the given position. The contribution of any substituent can therefore be expressed in terms of a linear combination of the contributions of the remaining substituents. Evidently, if the series of compounds tested contains h structural elements, in which there may be m substituents, then the solution of the system of equations thus formulated requires n equations with $m - h$ independent variables. The individual contributions of the substituents can be used to calculate $[A]$ for compounds of the series which have not yet been investigated.

The scope of the MAD has been frequently analysed[27–32] and attempts have been made to improve the model[33,34]. Using the same hypothesis of additivity, Eqn. (6) can be written in the form

$$\lg \frac{[A_1]}{[A_0]} = \sum_i G_i x_j, \tag{8}$$

where the subscripts 0 and 1 refer to $[A]$ for the unsubstituted and substituted compounds, G_i is the logarithm of the change in $[A]$ when the ith substituent is introduced instead of $X_1 = H$, and x_j has the same significance as in Eqn. (6). For a homologous series, $\lg [A_0] = \text{const.}$ and Eqn. (8) becomes

$$\lg [A_1] = \sum_i G_i x_j + \text{const.} \tag{9}$$

This stage already simplifies the MAD greatly: firstly, the constant term proves to be the activity of the initial unsubstituted compound and not a derivative of the molecular skeleton; secondly, since $[A]$ in the modified model is considered in relation to the activity of a hydrogen-containing compound, one may disregard the restrictive relation (7), which reduces the number of equations.

Despite the artificiality of the hypotheses underlying the MAD (the additivity and independence of the contributions of the substituents), the agreement between the experimental and calculated values of $[A]$ may be satisfactory, provided that (a) successive structural changes in series of compounds are investigated and (b) the results of biological tests used for the determination of the relative contributions of substituents are sufficiently accurate[27].

The fundamental effectiveness of the MAD may be demonstrated. We shall assume that it is necessary to test a series of compounds where the substituents may be located in four positions ($h = 4$) and the number of substituents in the compound is varied: $m = 2$, 3, 4, or 5. The theoretically possible number of variants of molecules is then 120. The number of structural elements which must be introduced into the MAD is $\Pi = \sum_{i}^{h} m_i + 1 = 15$, where unity represents the contribution of the molecular skeleton. Consequently, in order to determine theoretically the values of $[A]$ for 120 compounds, it is sufficient to carry out only 11 tests of their biological activity (i.e. $\Pi - h$ tests). In view of the low accuracy of biological experiments, in order to obtain a reliable result in practice, it is desirable to have the largest possible number of experimental data [28].

Allowance for the mutual influences of the substituents in the MAD constitutes a more realistic approach [31,32]. For this purpose, the terms G_i^*, taking into account the change in $[A]$ owing to the interactions of the substituents, are introduced into Eqn. (9):

$$\lg [A] = \sum_i G_i x_j + \sum_i G_i^* x_i x_j \ldots + \text{const.} \qquad (10)$$

For example, if the compound contains the substituents X_1, X_2, and X_3, then, when account is taken of their mutual influence, we have $\lg [A] = G_1 x_1 + G_2 x_2 + G_3 x_3 + G_1^* x_1 x_2 x_3 + G_2^* x_1 x_2 + G_3^* x_1 x_3 + G_4^* x_2 x_3 + \text{const}$. It has been shown [34,35] that the use of this procedure makes it possible to take into account the influence of the relative positions of the substituents on biological activity. Naturally, in this case there is a sharp increase of the number of terms in the equation and hence of the number of necessary biological tests, which reduces the practical scope of the application of the mathematical model.

Attempts have been described to relate the mathematical additive model to the mechanism of the action of substances [34,36-40].

The results accumulated hitherto concerning the application of empirical methods for the estimation of $[A]$ show that fairly similar values of $[A]$ may be obtained by independent computational and experimental methods. The practical use of the above methods requires a generalisation, systematic arrangement and representation in a form suitable for computer calculations of the extensive experimental data from studies on the biological activity of chemical compounds, and the development of a single system for the mathematical representation of the structures of substances.

IV. THE MULTIPARAMETER REGRESSION METHOD

The qualitative relations established between the biological activity and physicochemical properties of substances [2-4] stimulated studies designed to discover the essential features of these relations. The first attempt of this kind involves the equilibrium distribution model (EDM). [17] It is based on the hypothesis that, after the introduction of the substance into a biological system, chemical equilibrium, determined by the equality of the chemical potentials (μ) of the substance in different phases, is established between its concentration at the site of administration and the concentration in the vicinity of the site of response. The constancy of the isonarcotic concentrations in the lipid phase C_{lip} followed from the results of tests on the narcotic activity of a wide variety of compounds [17]. At low concentrations $\mu = \mu_0 + RT \ln C$; hence, setting the values of μ for the lipid and aqueous phases equal and assuming that C_{lip} is constant, it is possible to obtain a relation between $\lg(1/C_0)$ and the partition coefficient of the substance in the biological system (P_b). The latter is as a rule unknown, but it has been established [17,41,42] that P for a model system (for example olive oil–water and octanol–water systems) may be used instead:

$$\lg (1/C) = a \lg P + b. \qquad (11)$$

In Eqn. (11) $\lg P$ may be replaced by other physicochemical parameters of the substance, such as, for example, the parachlor (Pr), the molar volume (V_M), and the solubility ($\lg C_S$). The most complete information about structure–activity correlations based on the EDM has been published by McGowan [42]. On the basis of the theory of solubility, it is shown [42] that $\lg P$ can be represented by $\lg P = k_1 Pr + E_A$. On the other hand, the parachor may be related to $\lg C_S$ and V_M with the aid of the approximate relations $\lg C_S \simeq \lg C_l + k_2 Pr$ and $Pr \simeq k_3 V_M$, where E_A is a term taking into account the association of substances or their capacity for the formation of a hydrogen bond, C_l is the molar concentration of the substance recalculated for the pure liquid phase, and k are the corresponding proportionality coefficients. The relation between $\lg P$, on the one hand, and Pr, V_M, and $\lg C_S$, on the other hand, in the given model can be accounted for in terms of thermodynamic considerations, since the main contribution to $\lg P$ comes from the change in the entropy of the system owing to the rearrangement of the solvation shells of the molecules on passing from the aqueous to the organic phase. The degree of dehydration depends on a number of factors, in the first place on the size and polarity of the molecules of the substance and their packing density. $\lg C_S$, Pr, and V_M may be used as characteristics of these factors.

The possibility of using the additivity rules to predict the biological activity of substances on the basis of the additivity of the values of Pr and V_M was demonstrated for the first time with the aid of the equilibrium distribution model [42-44]. The limitations of the EDM are evident, since it takes no account of either the kinetics of the attainment of equilibrium in the system or of the mechanism of the chemical reactions occurring in it.

At the beginning of the 1960s Hansch and coworkers proposed an approach fundamentally different from the EDM. They suggested that the overall physicochemical process causing the biological response be regarded as a steady-state and not an equilibrium process [9]. They began with the assumption that the time interval corresponding to the biological test is insufficient for the attainment of chemical equilibrium. Hansch and Fujita [9] used the familiar scheme [45] for the biological activity of the substance:

$$\text{Compound in extra cellular phase} \xrightarrow[\text{stage I}]{} \text{Site of action in cellular phase} \xrightarrow[\text{stage II (rate-determining)}]{k_{II}} \cdots \xrightarrow[\text{stage III}]{} \cdots \xrightarrow[\text{stage } n]{} \text{Biological response}. \qquad (12)$$

It was found that the principal stages at I and II—the migration of molecules from the very dilute solution outside the cell to a definite site within the cell and the chemical or physical process determining the overall rate of the biological response respectively. It was postulated that other stages (III...n) do not affect the observed biological response.

In order to derive a formula relating $\lg(1/C_0)$ to physicochemical parameters, Hansch and Fujita[9] used a phenomenological approach (see Section I). It follows from Eqn. (2), which can be rewritten in the form

$$\lg(1/C_0) = \lg N + \lg \bar{k}_{II} + \text{const} \quad (13)$$

for scheme (12), that the problem reduces to expressing $\lg(1/C_0)$ in terms of the physicochemical parameters determining the values of $\lg N$ and $\lg \bar{k}_{II}$. On the basis of the relation observed by Collander[46,47] between $\lg P$ and the rate of migration of organic compounds through tissues and the nearly parabolic relation between $\lg(1/C_0)$ and $\lg P$ observed in a number of instances[9,12,48], it was assumed that N obeys the normal distribution law with respect to $\lg P$:

$$N = N_0 \cdot \exp\left[-\frac{(\lg P - \lg P_{op})^2}{b}\right], \quad (14)$$

where P_{op} is the optimum value of P corresponding to the maximum value of $N = N_0$ and b is a constant. Hammett's equation

$$\lg \bar{k}_x = \rho\sigma + \lg \bar{k}_0, \quad (15)$$

in which $\bar{k}_X$ and $\bar{k}_0$ represent the rate or equilibrium constants for *meta*- or *para*-substituted and unsubstituted benzene derivatives respectively (for the same reaction and under identical conditions), ρ is a constant for the given reaction, and σ the Hammett constant for the substituent, was used for the first time for the term $\lg \bar{k}_X$. After substituting Eqns. (14) and (15) in Eqn. (13), we obtain

$$\lg(1/C_0) = a_1(\lg P)^2 + a_2 \lg P + a_3\sigma + a_4, \quad (16)$$

where a_1–a_4 are constants.

Eqn. (15) is a consequence of the application of the principle of linear free energy relations (PLE).[49-51] It can be used to express the rate or equilibrium constants ($\lg \bar{k}_X$) for a series of related compounds involved in similar reactions. One then obtains for different types of interactions equations which are analogous in form to Eqn. (15). Expressing $\lg \bar{k}_X$ in Eqn. (13) with their aid, it is possible to obtain a wide variety of equations taking into account the resonance, conformational, steric, and other factors influencing $\bar{k}_X$; in particular, the following expression has been obtained[11]:

$$\lg(1/C_0) = -a_1(\lg P)^2 + a_2 \lg P + a_3\sigma + a_4E_s + a_5, \quad (17)$$

where E_S is the steric Taft parameter.

In order to take into account the influence of individual substituents on biological activity, i.e. in order to examine a series of compounds and also to calculate the values of $\lg P$ for substances which have not as yet been investigated, the substituent constant $\pi_X = \lg(P_X/P_0)$ was introduced by analogy with σ.[52] Here P_0 is the partition coefficient of the unsubstituted initial compound in a model system and P_X is the partition coefficient of the compound with the substituent X. It can be shown that π can also be obtained as a result of the application of the PLE to hydrophobic interactions. For a series of substances involved in a reaction j, variation of the substituent X yields

$$\lg(\bar{k}_x/\bar{k}_0) = \varkappa_j \lg(K_{xi}/K_{oi}), \quad (18)$$

where K_i are the equilibrium constants for the standard reference reaction i, $\bar{k}_j$ are the rate constants for reaction j (K_{Xi} and $\bar{k}_{Xj}$ when the molecule contains the substituent X and K_{0i} and $\bar{k}_{0j}$ in its absence), and κ_j is the constant of reaction j. The type of interaction between the substituent and the reaction centre for reactions i and j is assumed to be the same. If the interaction is assumed to be hydrophobic, then evidently the equilibrium constant in a partition system of the oil–water type is none other than the partition coefficient of the substance P. The partition of the substance (i.e. its interaction with the solvent of the "solvation ⇌ desolvation" type with respect to one of the solvents) in the model system, for example octanol–water, may be used as the standard reference reaction. Then substitution of P_X and P_0 for the equilibrium constants K_{Xi} and K_{0i} in Eqn. (18) yields an equation similar to Hammett's equation in which the constant $\pi = \lg(P_X/P_0)$ is analogous to the Hammett constant σ:

$$\lg \bar{k}_x = \varkappa\pi + \lg \bar{k}_0. \quad (19)$$

Like other reaction constants of the substituents, the constant π depends on the system in which it is determined[41,52], owing to the interaction of the substituents with other parts of the molecule. In practice this is taken into account by selecting for the calculation of π the corresponding model systems (molecules) similar to the molecules analysed.

By definition there is a linear relation between $\lg P_X$ and π_X, which made it possible to employ the corresponding values of π instead of $\lg P$ in equations of type (17):

$$\lg(1/C_0) = a_1\pi^2 + a_2\pi + a_3\sigma + a_4E_s + a_5 \quad . \quad (20)$$

These can be calculated[41,52].

In contrast to the kinetic approach constituting the basis of Hansch's method, Franke and Oehme[8] described an attempt to obtain similar structure–activity relations on the basis of an equilibrium model. It is assumed in this case that the entire chain of chemical reactions leading to the final biological effect operates only in the establishment of the equilibrium at each stage. Thus the entire system may be characterised by the overall equilibrium constant

$$K_c = \prod_i K_i, \quad (21)$$

which describes the equilibrium between the initial state (active substance D + receptor R) and the final state, identified with a "biological stimulus" St, corresponding to the Law of Mass Action. The concentration of the reaction product C_X, which is the main one for the observed biological action, was introduced[8] as a measure of St:

$$St = C_x = K_cC_DC_R. \quad (22)$$

Subsequently a general equation of type (3) was obtained[8] by expressing the biological activity $[A]$ in the form

$$[A] = f(St) = f(K_cC_DC_R), \quad (23)$$

taking into account the linear sections on the dose–activity curves, where

$$[A] = kSt, \quad (24)$$

and introducing the transfer functions W (analogous to the function N in Hansch's method) in order to find the concentration of the active substance near the receptors. Eqns. (16), (17), and (20) are special cases of the type (3) equation. It is readily seen that in this case the use of Eqns. (21)–(24) constitutes a formal procedure which makes it possible to bring the physicochemical parameters determining K_C and W into correspondence with $\lg[A]$ or $\lg(1/C_0)$.

It is clear from the foregoing that there is at present no unique physicochemical theory corresponding to both the kinetic and equilibrium approaches. However, an attempt may be made to group the physicochemical features of compounds in accordance with their physical significance in order to take into account to the maximum possible

extent the structural characteristics responsible for the biological activity shown by the compound. In many instances this facilitates the interpretation of structure–activity correlations and gives them a concrete physicochemical content.

We shall consider certain possibilities of a group of parameters (based on their physical significance) for the derivation of correlations. According to Eqn. (5), when N is expressed in terms of hydrophobic properties of the substance, then

$$\lg [A] = \lg (1/C_0) = f'_h(S_h) + f_e(S_e) + f_s(S_s) + \text{const}, \qquad (25)$$

where f'_h is a function combining the partition phenomena and the hydrophobic interaction with the receptor.

Not only physicochemical parameters of the entire molecule but also certain substituent constants may be chosen as the S_i in Eqn. (25). The following parameters can be used under these conditions.

(1) as S_h: (a) $\lg P$ [11,12,41,48,53] or quantities associated with it—the Hansch constant π (the lipophilicity parameter)[11,12,41,52]; $R_M = (1/R_f - 1)$ or ΔR_M (analogue of π) from thin-layer partition chromatography[54-56]; the protein binding constants (determined for a model system)[12,41,57,58]; $\lg C_S$§, where C_S is the solubility in water[42,58,59]; (b) quantities related to molecular size and packing: the molar volume (V_M), the parachor (Pr), molar refraction (MR), the solubility parameter (δ), the diamagnetic susceptibility (χ_m), the molecular connectivity (χ),[42,60-67], and the molecular weight M [60]; (c) the molar attraction constant (F_a) [60,63,64]; (d) the hydrophilic–lipophilic balance (HLB) [68];

(2) as S_e: (a) different electronic substituent constants σ characterising the electron density at the reaction centre[8,11,12] if they can be expressed[69] as a linear combination of the constants $\mathcal{F}$ and $\mathcal{R}$, where $\mathcal{F}$ is the inductive constant, which takes into account only the inductive effect of the substituent on the functional group and $\mathcal{R}$ is the resonance constant, taking into account all the resonance interactions (the use of $\mathcal{F}$ and $\mathcal{R}$ [70-72] makes it possible to avoid the consideration of different types of σ in the correlation equations)¶. (b) The charge-transfer constant (C_T), which is a measure of the stability of the complex formed via charge transfer[78-80]; (c) various parameters characterising the electron density, bond orders, and bond lengths; parameters of the first kind include the changes in the ionisation constant ΔpK,[10,12,81,82] the ionisation potential I,[83] and the polarographic half-wave potential $\varphi_{1/2}$ [84]; parameters of the second kind include the molar electronic polarisation (P_E) [85-87] or simply the polarisability (α),[60,83] the dipole moment (μ),[12,88,89] and spectroscopic data (infrared, ultraviolet, NMR, etc.)[90-98]; (d) various quantum-mechanical parameters (see below).

(3) As S_S: (a) steric Taft constants (E_S),[11,12,99] determined analogously to σ, and the corrected steric interaction parameters $E_S^S = E_S + 0.306(n_H - 3)$ [12,100]; (b) the van der Waals radii (r_V).[37,101-103].

The literature data for structure–activity correlations have shown that the degree of the functions f_h, f_e, and f_s in Eqn. (25) is as a rule not greater than the second, i.e. in the general case they correspond to a parabolic relation and are therefore characterised by certain extremum values corresponding to S_i^0. The extremum is determined from the condition

$$\left(\frac{\partial \lg (1/C)}{\partial S_{i=k}}\right)_{i \neq k} = 0$$

and corresponds to the optimum value of the corresponding physicochemical parameter for biological activity[10-12,104]. The possibility of an optimum value of S_h^0 can be explained in different ways[104]. It appears that the following most important factors for the occurrence of S_h^0, S_e^0, and S_s^0 may be distinguished: (a) metabolism or competing metabolic reactions in the biological system; (b) the obligatory complementarity of the molecule (or its active component) and the reacting region of the receptor; (c) a low concentration of the active molecules in the vicinity of receptors. Furthermore, in the study of enzymes *in vitro* the occurrence of S_i^0 can be due to micelle formation with entrainment of active molecules.

The number of parameters quoted above may be reduced if they are expressed by a quantum-mechanical procedure. It has been shown[105] that π as well as $\lg P$ can be expressed as a linear function of quantum-chemical parameters:

$$\pi = k_1 \sum_r |Q_r^T| + k_2 \sum_r S_r^E + k_3, \qquad (26)$$

where $|Q_r^T|$ is the absolute value of the overall charge, S_r^E is the electrophilicity caused by the superdelocalisation of the electrons at the atom r, and k_1–k_3 are constants.

However, since the use of experimental values of π and $\lg P$ to determine structure–activity correlations yields satisfactory results in individual instances, it appears much more useful to employ quantum-chemical methods in order to express the term $f_e(S_e)$,[12,106-116], using various computational methods[116].

The following quantum-chemical quantities are used as parameters characterising reactivity†: the overall π-electronic charge $q_r(\epsilon)$‡ on the atom r; the overall $(\sigma + \pi)$ charge on the atom r; the highest occupied (E_{HOMO}) and lowest unoccupied (E_{LUMO}) molecular orbitals; the electrophilic (S_r^E) and nucleophilic (S_r^N) characteristics of the influence of superdelocalisation[118,119]; the free valence index F_r (mainly for homolytic reactions)[114]; the atomic-orbital coefficient C_r.[120]

In the general case it is not clear which quantum-chemical parameters should be preferred for use in the structure–activity correlations. This depends on the proposed mechanism of the biological action of the compound and is usually related to the charge on the heteroatom (for example q_N on the nitrogen atom), to the energies E_{HOMO} and

---

§It has been shown[59] that, in the study of the biological activities of substances introduced into a biological system in a solid form, the rate of dissolution of this form must be taken into account in structure–activity correlations. This is because the correlation between the rate and C_S is poor.

¶In radical reactions it is still necessary to employ $\sigma^\bullet$ and E_R, because they are poorly correlated with $\mathcal{F}$ and $\mathcal{R}$.[73-77]

†They are calculated using parameters obtained from spectroscopic measurements, the ionisation potentials, etc. The corresponding values for the extensive Hückel method have been tabulated[117].

‡Here and henceforth the subscript r means that the corresponding quantity refers to the atom r in the molecule.

E_{LUMO} or to the difference $E_{HOMO} - E_{LUMO}$, and to S^N.[8,10,12,114,116] Hypotheses of this kind are not often found to be true in practice. This is probably because only the potential reactivity of the compound is considered by this procedure without taking into account the possible steric, conformational, and other parameters influencing the biological activity. Attempts have therefore been made to use various linear combinations of quantum-chemical parameters[114]. However, this procedure has not as yet been adequately investigated and its scope is not entirely clear.

V. ANALYSIS OF THE STRUCTURE–ACTIVITY RELATION BY HANSCH'S METHOD

As stated above, Hansch's method is used to express the biological activities of members of series of related compounds in the form of functions of physicochemical parameters characterising the structure of these compounds. As a rule, the relative importance of taking into account these parameters in the description of biological activity is not known beforehand, which makes it necessary to test a series of model equations containing various combinations of parameters. These equations are obtained by regression analysis, after which the best one is selected on the basis of the significance of the correlation coefficients. It is recommended[8,103] that certain rules be observed.

1. It is desirable to consider the greatest possible number of independent parameters.

2. The parameters S_i considered above are as a rule integral quantities characterising molecular properties in a complex manner; for this reason, there is a definite relation between them. The literature data show that the correlation between the S_i referring to the same interactions (for example to hydrophobic interactions) is much better than between parameters characterising different types of interactions (for example between $\lg F$ and σ).[70,121] This observation can be used as one of the empirical rules for the selection of various combinations of independent parameters S_i§.

3. After the preliminary examination of the possible versions of the set of parameters and comparison with the aid of regression analysis of a series of model equations, it is necessary to determine by means of Student's t distribution the levels of significance of individual coefficients. Parameters with a low level of significance are disregarded.

4. The multiple regression coefficient[122,123] R, the standard deviation, and the F test are used for the statistical assessment of the equations obtained. The square of the correlation coefficient (R^2) indicates the proportion of variations in data obtained from the corresponding equations. The closer it is to unity, the closer the functional relation obtained to the one sought. For similar values of R, it is necessary to take into account the standard deviations and to use equations with the smallest standard deviations. The F test is used to estimate the probability that the suitability of a particular type of equation is not fortuitous.

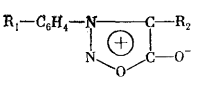
5. In order to exclude random correlations, it is recommended that at least five–six results of biological measurements are obtained for each physicochemical parameter tested[124].

6. Other conditions being equal, preference should be given to simpler equations.

7. In order to reduce the amount of calculations, to simplify them, and to avoid fundamental errors, it is necessary to take into account the reality of the mechanism of the phenomenon under consideration (see Section VIII), since it is possible to obtain on the basis of formal characteristics equations with satisfactory values of R^2 and a high significance according to the F test but having little in common with reality[125].

If possible, one should divide the substances compared into a series of definite substructures with common features, for example, substances with *cis*- and *trans*-positions of a certain group, series obtained by changing one substituent while those in other positions are kept constant, etc.[8] Such subdivision naturally simplifies the problem (a similar empirical approach was discussed in Section III). In order to derive equations for such a series of compounds, it is necessary to introduce an auxiliary parameter D, taking into account the difference between the substructures. For example, it has been used to take into account the steric difference between the *meta*- and *para*-derivatives of diethyl phenyl phosphate in the inhibition of cholinesterase[126]—it was set equal to unity for the former and to zero for the latter. The value of n_H, the number of hydrogen atoms linked to a definite atom in the molecule, for example, the nitrogen atom of an amino-group, may be chosen as D.[127] The auxiliary parameter may be used to combine equations taking into account the differences between the subgroups in order to determine the biological activity of particular stereoisomers only. In those cases where there are several subgroups, it is possible to introduce two or more auxiliary parameters[128].

Examination of polysubstituted compounds presents the greatest difficulties. In such cases it is not always possible to obtain an equation describing the correlation between the biological activity and the sum of identical substituent constants i.e. ($\sigma_i + \sigma_j$, $\pi_i + \pi_j$, etc.), because the influence of different types of substitution may have different effects on the biological activity. This problem is solved by substituting in the equation parameters characteristic of the particular type of substitution. For example, in describing the inhibition of oxidative dealkylation by compounds having the general formula



an equation with a high correlation coefficient is obtained only when the constants π and σ are taken into account separately for R_1 and R_2.[129]

The most complex case obtains when the biological activity of substances is influenced appreciably by their metabolism.

VI. THE QUANTUM-STATISTICAL APPROACH

Attempts to relate the biological activity of a substance to microscopic characteristics of the interactions in the substance–receptor system by a combination of quantum-statistical and statistical thermodynamic methods have been undertaken[94,130,131]. In order to obtain a working

---

§In order to characterise the absence of a relation between physicochemical parameters, the so called individual correlation coefficients r, determined as dimensionless quantities, varying in the range $-1 \leqslant r \leqslant 1$, are used. The closer the value of r to 0, the more independent are the corresponding physicochemical parameters.

model, the following assumptions were made: (1) in a multistage process occurring from the instant of the introduction of the substance into a biological system until its interaction with the receptor resulting in biological activity, the probability of each stage is independent of that of other stages; (2) substances reach the sites of reception by virtue of Brownian-like motion; (3) the reaction sites consist of localised atoms; (4) the binding of the biologically active substance at a single receptor or group of receptor sites is independent of its binding at other sites in the system.

According to these hypotheses, the initial equation can be expressed in the form

$$\frac{d[BR]}{dt} = KC_0p_1p_2, \tag{27}$$

where K is a proportionality coefficient, C_0 the initial concentration of the substance, p_1 the probability of the completion of the stage involving the penetration of the substance to the receptor, and p_2 the probability of the binding of the substance to the receptor. Next is is postulated that p_1 varies in parallel with the parameter π, i.e. with the lipophilic properties, which are responsible for the passage of the substance through biological membranes,, and that p_2 varies in parallel with the adsorption of the substance at the receptor site or group of sites. If it is assumed that the binding of the substance to the receptor occurs at energetically indistinguishable sites, then p_2 should be wholly determined by the electronic structure of the heteroatoms of the substance adsorbed. For the specific examples which have been examined[94,131], it was postulated that the binding takes place via a carbonyl group, the energetic characteristics of the probability of the binding of which to receptors were determined from spectroscopic data.

The mathematical expression[131] relating the activity of the substance to the permeability of membranes to it and the probability of its interaction with the receptor includes the following parameters: the lipophilicity, the spectroscopically estimated distribution of the potential reactive centres (groups), and the chemical potential of the test substance in solution. It follows from the analysis of this expression that the most active compound may be found or synthesised when the substituent or its position in the molecule is chosen in such a way that it influences simultaneously the lipophilic properties of the compound and the energy state of the reactive group.

Insufficient reliable data are as yet available for the assessment of the predictive power of the method[130]. At the present time one can estimate only its fundamental advantage: the necessary parameters—the lipophilicity and the wavenumbers of the suggested reactive groups—may be determined experimentally fairly readily.

VII. THE USE OF FACTORIAL ANALYSIS IN THE DERIVATION OF STRUCTURE–ACTIVITY RELATIONS

Evidently the limitation of the above methods to any one type of biological activity is artificial. In order to solve rigorously the problem of the relation between structure and activity, it is necessary to take into account all possible physiological processes and to relate them to the entire set of physicochemical parameters characterising or determining each stage in the action of the given substance. The difficulties in this procedure are the unusual complexity of biological systems, the wide variety of reactions occurring in them, and the imperfection of our knowledge of their mechanisms and interrelations. Virtually the first attempt to take into account several possible versions of the interaction is a study[132] where the method of factorial analysis was used.

Mathematical factorial analysis may be formally applied to any complex multidimensional problem, provided that it can be represented by a sum of interrelated functions of certain variables, i.e. in the form

$$P(i, k) = \sum_{j=1}^{m} U(i, j)\, V(j, k), \tag{28}$$

where $P(i,k)$ is the measured property of the system in the ith case when k is varied, $U(i,j)$ is the jth factor for the ith case, and $V(j,k)$ is the jth variable factor for the given variable k; the summation is carried out with respect to all the factors j which are important[132]. In the relevant specific instance of the structure–activity relation, we have

$$[A](i, k) = \sum_{j} D(i, j)\, H(j, k), \tag{29}$$

where $[A](i,k)$ is the kth measured biological activity of substance i, $D(i,j)$ the jth property (physical or chemical) of the ith compound, and $H(j,k)$ the jth parameter (physiological) of the kth biological activity; the summation is carried out with respect to all j independent parameters which are necessary to take into account the results of tests.

The form of factorial analysis applied to the case described differs from the standard form by the fact that it includes the method of regression analysis as one of the factors[132,133] in the form of the identity factor. Thus, provided that the accuracy of the experimental data is known, it is possible to determine the number of independent abstract factors ("eigenvectors" according to Weiner and Weiner[132]) necessary for the complete characterisation of the factorial space, i.e. one can determine the number of variables influencing the system tested. Leaving on one side the discussion of the mathematical formalism used by Weiner and Weiner, we shall mention only that in this first stage the number of eigenvectors is equivalent to the dimension of the factorial space. The next step is the identification of each factor in the form of the specific level of its significance for each physiological response. For this purpose, the entire set of variables is compared with all possible physicochemical properties or parameters of the model interactions which are presumed to be important. Although Weiner and Weiner[132] used the characteristics of substances usually employed for semi-empirical models, this does not imply that one must necessarily employ the polarisability, the partition coefficients, etc. (Table 1).

Since the term $[A](i,k)$ in Eqn. (29) may be expressed and measured experimentally as the biological activity manifested in a kth case and the term $H(j, k)$ relates any physiological parameter to any property of the substance and the quantity $[A]$, the method of factorial analysis may be used to predict the uniqueness of the compound tested or to specify the set of properties giving rise to the maximum useful activity for a minimum number of side effects. The method of factorial analysis was tested by comparing the calculated data with the experimental results obtained by measuring the activities of 16 substances in a series of substituted compounds having the general formula

R_1–N–R_3 (with R_2 in the vertical position) in 11 types of biological and physiological action (including action on a specific type of activity, the antagonistic behavioural reflexes of the animal, and objective

parameters of the state of the organism). It was established[132] that, in order that the model should reproduce the results of studies to within ±0.1 on a logarithmic scale, it is necessary to employ eight parameters describing, according to Weiner and Weiner[132], the main structural characteristics of substances. Unfortunately, a list of the properties employed is not given[132]; nevertheless the potential possibilities of this approach should stimulate the expansion and intensification of this field of research.

VIII. THE USE OF SEMIEMPIRICAL METHODS IN THE STUDY OF THE MECHANISMS OF REACTIONS OF BIOLOGICALLY ACTIVE SUBSTANCES

Apart from the selection of the optimum compounds, it is desirable to obtain, as a result of the analysis of the structure–activity relation, an idea about the mechanism of their biological activity. We shall consider one of the possible procedures for this purpose.

1. The conclusion as to which properties of the substances are the most important for their biological activity may be reached on the basis of the statistical significance of the physicochemical parameters in correlation equations.

2. Next one must determine the ways in which these properties influence the biological activity, beginning with the form of the function $f_i(S_i)$ [Eqn. (25)]. Here account must be taken of the fact that a parabolic form of the relation with $\lg P$ (or with other S_h) may be a consequence not only of the partition of the substance between the phases[12,134] but also of its hydrophobic binding to the receptor[58]. Because of this, only in the case of a linear relation with $\lg P$ can one conclude that hydrophobic binding is in fact the determining process[8].

In certain cases where a cross-product of terms of the type S_iS_j is introduced in Eqn. (25), the correlation is improved[10,135]. If S_i and S_j are substituent constants referring to a single type of interaction, such cross-product takes into account the mutual influence of the substituents[50]; otherwise, it takes into account the mutual perturbations of different types of interaction[8].

3. When the influence of hydrophobic binding is decisive, one can estimate approximately the dimensions of the binding region on the receptor from the values of the parameters S_h for which a discontinuity is observed in the functional dependence of $[A]$ on S_h.[58]

4. Certain conclusions about the mechanism of the motion of a given substance may be reached by analysing auxiliary parameters, which must be included in the equation to improve the correlation.

5. If possible one should calculate $\lg P_{op}$ (or the values of other S_h^0) and compare them with the corresponding values for other series of compounds giving rise to a similar biological response. For example, similar values of $\lg P_{op}$ obtain for a non-specific action of the substance. In particular, in the case of narcotic action identical values $\lg P_{op} \simeq 2$ were obtained for a series of substances with fundamentally different structures and compositions: barbiturates, thiamorpholidines, acetylenic alcohols, acetylenecarbamates, etc. This permits the conclusion that narcosis is due to a non-specific action by the substances and that substances with $\lg P_{op} \simeq 2$ are in general the most effective narcotics[12].

An approximate model of the biological activity of the substance may be obtained from the information derived by the above procedure but must be subsequently tested and refined by additional experiments.

Interesting prospects for the understanding of the mechanism of the most complex stage in the interaction of the substance in the organism—the formation of a complex with the receptor—arise from a theoretical treatment of Hansch's method[36,120]. Since the experimental determination of the type of substance–receptor bond is difficult, it has been suggested that the total energy change ΔE occurring on formation of the substance–receptor complex be used as a bond characteristic. If one considers a series of substances with the same type of biological activity, which is due to binding to particular receptors, then the difference between the activities of these substances may be expressed to a first approximation by the change in the free energy of formation of the substance–receptor complex:

$$\Delta [A] = -\left(\frac{1}{RT}\right)\Delta G + \text{const}, \qquad (30)$$

where $\Delta G = \Delta E + \text{const}$.

ΔE represents, according to the authors[36,120], the sum of the independent contributions of the electronic (ΔE^e), solvation (ΔE^d), steric (ΔE^s), and conformational (ΔE^p) components of the energy change. This hypothesis corresponds to a linear relation between the free energies[36]. The next problem is to express the different contributions ΔE^i in terms of experimental or calculated quantities. In order to calculate ΔE^e, for example, using the perturbation theory[136,137], we obtain

$$\Delta E^e = \sum_r \sum_s E_{rs} = \sum_r \sum_s \left(\underbrace{-\frac{Q_r Q_s e^2}{D_{rs} R_{rs}}}_{\text{(I)}} + \sum_m \sum_n \underbrace{\frac{2C_{mr}^2 C_{ns}^2 \beta^2}{E_m^* - E_n^*}}_{\text{(II)}} \right), \qquad (31)$$

where E_{rs} is the electronic energy of the interaction between atoms r and s, Q_r and Q_s are the overall charges, e is the electronic charge, R_{rs} is the distance between atoms r and s, D_{rs} is the dielectric constant of the medium, $E_m^* - E_n^*$ is the energy difference between the highest occupied molecular orbital of the receptor and the lowest unoccupied molecular orbital of the compound, β is the exchange integral taking into account the interaction between the molecular orbitals, and C_{mr} and C_{ns} are the atomic orbital coefficients. Each pair of interacting atoms, in a molecule of the compound (s) and the receptor (r) is assumed to make an independent contribution to ΔE^e.

Since Eqn. (3) includes properties characterising the receptor, it cannot be used directly. For this reason, limiting cases are considered[36,120], corresponding to different relations between E_m^* and E_n^* in Eqn. (31).

1. The Case "Controlled by the Boundary"

If $E_m^* \simeq E_n^*$, the interaction between the corresponding orbitals is decisive for E_{rs}. Then term (I) can be neglected in Eqn. (31) compared with term (II). This case corresponds to the formation of a covalent bond between the molecules of the substance and the receptor. In order to determine the necessary properties corresponding to the receptor, it was suggested that, for a series of related compounds combining with the same receptor site, the coefficients C_{mr} are constant for each compound. Because of this we have

$$\Delta E^e = \sum_s \alpha_s C_{ns},$$

where α_s includes C_{mr}. The following final expression is obtained for ΔE taking into account the solvation energy $\Delta E = \Delta E^e + \Delta E^d$:

$$\Delta E = \sum_s (\alpha_s C_{ns} - \lambda_{ns} Q_s \pm \lambda'_{ns} Q_s^2), \quad (32)$$

where, as mentioned above, the α_s are approximately constant for a series of substances interacting with the same receptor, while λ_{ns} and λ'_{ns} are constants obtained by examining ΔE^d.

An example of the correlation between $\lg(1/C_0)$ and the parameters included in Eqn. (32) is provided by the equation

$$\lg(1/C_0) = 8.8C_{nN} + 237Q_N - 1670Q_N^2 + 6.6, \quad (33)$$

obtained for the bacteriostatic activity of sulphonamides against *E. Coli* [120]. The subscript N in Eqn. (33) refers to the nitrogen atom.

2. The Case "Controlled by the Charge"

This obtains if the difference between E_m^* and E_n^* is large. Then term (I) in Eqn. (31) is of decisive importance, compared with terms (II), and the interaction between atoms r and s is determined mainly by the total charges Q_r and Q_s. In this type of interaction the bond between the compound and the receptor is ionic. When account is taken of this factor, Eqn. (31) can be written as follows:

$$\Delta E = \sum_r \sum_s \left[-\frac{Q_r Q_s e^2}{D_{rs} R_{rs}} + 2\left(\sum_m C_{mr}^2\right)\left(\sum_n C_{ns}^2\right)\gamma \right], \quad (34)$$

where γ is the average value of the ratio $\beta^2/(E_m^* - E_n^*)$. Having assumed the constancy of C_{mr} and Q_r in relation to C_{ns} and Q_s respectively for a series of related compounds interacting with the same receptor and taking into account the fact that ΔE^d in the case controlled by the charge is zero, it is possible to obtain the following equation [36]:

$$\Delta E = \Delta E^e = \sum_s (\nu_{ns} Q_s + \xi_{ns} S_s^E), \quad (35)$$

where ν_{ns} and ξ_{ns} are constants and S_s^E is the parameter of the capacity for ultradelocalisation (see Section IV). The case commonest in practice is probably intermediate between the boundary cases examined above. It has been suggested that the following expression for ΔE be used for it [36¶]:

$$\Delta E = \Delta E^e + \Delta E^d = \sum_s (\nu_{ns} Q_s + \xi_{ns} S_s^E) + h \lg P. \quad (36)$$

The parameter π may be expressed similarly in terms of quantum-chemical characteristics [105] (see Section IV). Without considering the scope of Cammarata's method in the determination of biological activity, which is clearly limited by the complexity of the calculations, one must emphasise yet again that the principles underlying it may be productive in the treatment of the mechanism of the formation of the substance–receptor complex.

¶ In Eqn. (36) $\lg P$ may be replaced by π.

IX. THE SCOPE OF PRACTICAL APPLICATIONS, ADVANCES, DIFFICULTIES AND PROSPECTS

As can be seen from the foregoing discussion, there exist at present a number of approaches to the solution of the problem of the correlation between structure and activity, of different complexity and developed to different extents. They are compared in Table 2—the principles underlying the methods, their applications, and certain disadvantages are indicated. The practical value of correlations usually depends on three factors [138]: accuracy, simplicity, and the type of information necessary for use in an empirical equation. Evidently, as regards the "accuracy and simplicity" criteria, preference cannot be given to any one of the approaches described and only the future will show whether the apparently most progressive method of factorial analysis will justify itself. If the existing approaches are considered from the standpoint of the information employed in them, then in an implicit (additive model, substructure analysis) or explicit (factorial analysis) form such information can be expressed in terms of one of the specific types of the multiparameter regression model. Indeed it has been shown [139] that there is a relation between the contribution of substituents to biological activity and their physicochemical characteristics (the substituent constants π, σ, etc.), i.e. all the models enumerated contain common information, which is most clearly expressed in Hansch's model. This makes it possible to consider different approaches from the same standpoint and to justify to a large extent the attention which has been devoted in the literature to the multiparameter regression model.

More than 2000 equations of types (16), (17), and (20) for more than 20 000 chemical compounds have now been described [140]. Many of these have all the formal features (high correlation coefficients and statistical levels of significance) which at first sight permit the conclusion that a change of the substituents in the structure alters its hydrophobic, electronic, steric, and other properties, and hence the biological response elicited by the compound. According to Hansch's method, the last factor may be taken into account and expressed mathematically by the summation of the physicochemical properties of the substituted compounds. Approximately 50 relations, which have been used to predict the efficiency of the biological action of substances, are quoted in the literature [98,141-143]; the structure–activity correlatons are used much more widely to discover the mechanisms of the action of biologically active substances (see, for example, Refs. 8, 10–12, 18, 36, 95, 144, 145). The following causes responsible for a certain gap between theory and practice may be suggested.

1. The resultant biological activity of substances depends on their influence on many physiological processes occurring in the organism. The experimental limitation of the above method to any one type of biological activity, expressed in terms of the concentration of the substances giving rise to the given effect, is therefore too rough an approximation to reality to be used as a parameter. It appears that, only for a reliably established mechanism, for example, when the latter is determined by a particular enzyme, can the concentration of substances giving rise to a standard change in the concentration of the enzyme be used correctly to calculate $[A]$ in terms of the corresponding properties and characteristics of model interactions.

Table 2. Principal methods for the determination of structure–activity correlations.

Method	Principle (hypothesis) underlying the method	Field of application	Disadvantages	Remarks
Mathematical additive model	The effectiveness of the biological activity of the chemical compound is determined by its structure and may be calculated as the sum of the contributions of the substituents in its molecule, the contribution of each substituent being independent of the presence of others (hypothesis of the additivity and independence of the contributions of substituents)	Prediction of the effectiveness of the biological activity in series of related compounds and hence reduction of the number of biological tests	Is inconsistent with the ideas concerning the influence of substituents on one another and hence on the physicochemical properties of the molecules of the given substance. The method is applicable only to series of compounds with the same type of biological activity	In order to improve it, the model may be (a) modified to take into account the mutual influence of substituents and (b) may be related to the multiparameter regression model. In both cases the model becomes considerably more complex
Substructural analysis	Arbitrary structural elements (fragments) are characterised by the frequency of their repetition in known biologically active compounds. The compound tested is analysed from the standpoint of the presence of such fragments. (Hypothesis—the biological activity of the substance is determined by the presence of particular structural elements.)	Non-experimental screening—prediction of the direction and effectiveness of the biological activity of substances. Determination of original structures	Has no physicochemical basis. The quantitative predictions of the effectiveness of biological activity are unreliable	In order to improve quantitative estimates, physicochemical characteristics of substances are introduced—this version of the procedure is referred to as the pattern recognition method
Multiparameter regression method	The set of parameters characterising model physicochemical processes occurring with participation of the given substance is correlated with its biological activity. (Hypothesis—the biological activity is determined by the structure and reactivity of the molecules of the given compound)	Prediction of the effectiveness of biological activity in series of related compounds without biological tests. The study of the mechanisms of biologically important chemical reactions	Is applicable to series of compounds. Does not allow the prediction of the direction of biological action. Conclusions concerning the effectiveness of substances are frequently erroneous	In order to reduce the number of parameters, one can use calculated or experimental quantum-chemical parameters—the quantum-statistical approach, the Cammarata method.
Factorial analysis	Relates the characteristics of the structure of the compound to the set of possible physiological results of its action. By sorting the possible variants, it established the ways in which the biological activity of substances arises and determines its effectiveness	Non-experimental screening—prediction of the direction and effectiveness of biological activity. Prediction of the uniqueness of the structure and form of activity and of side effects	Complexity. The scope of the method has not been adequately tested	Includes the multiparameter regression model as one of the factors

2. Apart from its evident advantages, the extensive application of mathematical statistical methods, which is characteristic of semiempirical approaches, to the structure–activity problem made the true pathways followed in the biological action of substances and reliable results statistically insignificant against the background of the enormous number of the results of biological tests obtained by different investigators. The low accuracy and reproducibility of biological experiments naturally have an effect under these conditions.

3. The use of regression methods to obtain correlation equations requires much greater caution than is frequently exercised in the literature. It is known that a number of independent variables can be correlated successfully using random terms if the latter are also independent. In reality, this situation means that our information about the mechanisms and the ways in which a particular type of biological activity is manifested is still insufficient to take into account the maximum number of parameters important for the manifestation of biological activity.

The above difficulties are fundamental and an attempt has been made only recently to eliminate some of them by means of factorial analysis [132]. Nevertheless there is at present no alternative to Hansch's method and its different versions in the formulation of quantitative structure–activity relations. This is mainly because the foundations of Hansch's method rest on the most general physicochemical principles, which probably serve as the basis of the mechanism of the action of biologically active compounds.

X. CONCLUSION

Summarising the foregoing review, it appears useful to put forward certain considerations applicable to the problem as a whole. The above methods for the determination of the types of functions relating quantitatively the characteristics of the structure of a substance and its biological activity are a natural consequence of the theoretical foundations of these methods and the types of interactions (in the set of all possible interactions) in the substance–biological object system which are taken into account.

It has been established reliably that the best correlation between biological activity in a series of substances and the partition coefficient ($\lg P$), the lipophilic constants of the substituent (π), and other characteristics associated with them is shown by substances exhibiting a non-specific type of biological action†. This implies that the chemical reaction which determines the outcome of the action of the substance occurs on the surface or within the cell membrane and that the site of the effective action of the compound is not localised. In practice such simple behaviour is probably rarely encountered.

---

†For example, it has been suggested [60] that a satisfactory correlation between $\lg P$ and π be used as a criterion for the estimation of the non-specificity of the biological action of the substance.

It is quite evident that the value of any structure–activity relation, not including the specificity characteristics as a factor of the manifestation of biological activity, is low. However, in the case of the specific binding of substances it is at present impossible to specify the number and types of parameters whose combination would be sufficient for an exhaustive and general solution of the problem of the determination of the structure–activity relation. This factor probably in fact explains the number of different parameters and their combinations (Table 1) which have been used to establish correlations and which are in essence difficult to extend further. It would appear that a satisfactory description of a theoretical relation in terms of experimental data should indicate that the model adopted reflects correctly the essential features of real processes and that the leading role among them is played precisely by those interactions which are taken into account in the corresponding correlation. Nevertheless, "astonishing freaks" in Hansch's words [146] (i.e. equations which are unrealistic in terms of their physical significance or ineffective in their predictive power but have high correlation coefficients), are discovered all too often when the results of empirical and semiempirical approaches are tested, which indicates most clearly certain significant miscalculations underlying the derivation of the corresponding equations. Whether this situation is a consequence of the impossibility of allowing completely for the entire set of different types of interactions between the substance and the organism and the lack of clear-cut principles whereby each specific situation can be reduced to a single equation of all the necessary parameters, or whether it is due to other less evident causes, is difficult to establish. Apart from the objective complexity of the problem, we believe that this is aggravated by the excessive employment of different versions of the approaches described above, frequently without adequate critical analysis of the results of biological tests used and the selected physicochemical parameters.

Future progress in the application of the procedures, outlined above, for the determination of quantitative structure–activity relations will probably be achieved by setting up a few fundamental experiments in combination with their profound theoretical elaboration and not by the excessive accumulation of new results. Furthermore, the main task—the parametrisation of biological activity, which cannot be regarded as finally completed in the form which exists at the present time—remains of primary importance.

In this connection it appears that real advances in the semiempirical approaches may be expected as a result of the narrowing of their fields of application, for example by restricting them to those cases where a definite type of biological activity may be characterised by the behaviour of molecular-biological, biochemical, and other models, and not by their limited extension to all types of biological activity *in vivo*.

o0o

The above review covers the literature up to 1975 inclusive. Among the papers published recently, mention should be made of the reviews on the mathematical additive model [286-287], the relations between this model and the multiparameter regression method (MRM), [287] the possibility of using the parachor in correlations between the chemical structures and biological activities of substances [288], a new "bilinear" model for a non-linear relation between the biological activities and hydrophobic properties of molecules [290], model equations for MRM for substances capable of ionising [291], a classification of therapeutic agents in conformity with pharmacophors for the pattern recognition method [292], the elucidation of the relative importance of steric parameters in relation to steroid hormones [293], and a compilation of tables of lipophilic parameters [294].

It has been suggested that new parameters be used in MRM: partition coefficients taking into account both the neutral and ionic forms of the given substance [295]; surface activity and the work of adsorption from the aqueous phase at the oil–water interface [296]; "hydrogen bonding" parameters [297]. Increasing attention is being attracted by different modifications of the capacity for molecular binding, which is readily susceptible to calculation whatever the complexity of the given compounds [298,299].

REFERENCES

1. C. Hansch, S. H. Unger, and A. B. Forsythe, J. Medicin. Chem., 16, 1217 (1973).
2. C. Richet, Compt. rend. Soc. biol., 45, 775 (1893).
3. H. H. Meyer, Arch. Expt. Pathol. Pharmacol., 42, 109 (1899).
4. E. Overton, "Studien über die Narkose", Fischer, Jena, 1901.
5. N. V. Lazarev, "Neelektrolity" (Non-Electrolytes), Izd. Voen.-Mor. Med. Akad., Leningrad, 1944.
6. N. V. Lazarev, "Narkotiki" (Narcotics), Izd. Inst. Gigieny Truda, Leningrad, 1940.
7. S. A. Giller, A. V. Glaz, V. E. Golender, L. A. Rastrigin, and A. B. Rozenblit, Khim.-Farm. Zhur., No. 12, 18 (1972).
8. R. Franke and P. Oehme, Pharmazie (Berlin), 28, 489 (1973).
9. C. Hansch and T. Fujita, J. Amer. Chem. Soc., 86, 1616 (1964).
10. I. Schwartz, Stud. Cercet Chim. (Bucuresti), 21, 721 (1973).
11. C. Hansch, Chem. Res., 2, 232 (1969).
12. C. Hansch, in "Drug Design" (Edited by E. Ariens), Acad. Press, New York, 1971, Vol. 1, p. 271.
13. R. O'Brien, "Toxic Phosphate Esters" (Translated into Russian), Izd. Mir, Moscow, 1964.
14. K. H. Meyer and H. Hemmi, Biochem. Z., 277, 39 (1935).
15. I. V. Berezin, A. V. Levashov, and K. Martinek, Vth Federated European Biological Society Meeting, Prague, 1968;
16. I. V. Berezin, A. M. Klibanov, and K. Martinek, Uspekhi Khim., 44, 17 (1975) [Russ. Chem. Rev., No. 1 (1975)].
17. J. Ferguson, Proc. Roy. Soc. B, 127, 387 (1939).
18. C. Hansch, Adv. Pharmacol. Chemother., 13, 45 (1975).
19. R. D. Cramer, G. Redl, and C. E. Berkoff, J. Medicin. Chem., 17, 533 (1974).
20. M. F. Lynch, J. M. Harrison, W. G. Towne, and J. E. Ash, in "Computer Handling of Chemical Structural Information", Elsevier, New York, 1972, p. 67.
21. P. N. Craig and U. M. Ebert, J. Chem. Docum., 9, 141 (1965).
22. P. J. Harrison, J. Appl. Statist., 17, 226 (1967).
23. O. T. Walz, M. J. Di Martino, and A. Misher, J. Pharm. Exp. Ther., 178, 223 (1971).

24. K. C. Chu, R. J. Feldmann, M. B. Shapiro, G. F. Hazard, and R. I. Geran. J. Medicin. Chem., 18, 539 (1975).
25. K. C. Chu, Analyt. Chem., 46, 1181 (1974).
26. S. M. Free and J. W. Wilson, J. Medicin. Chem., 7, 395 (1964).
27. W. P. Purcell, Biochim. Biophys. Acta, 105, 201 (1965).
28. W. P. Purcell and J. M. Clayton, J. Medicin. Chem., 11, 199 (1965).
29. J. A. Singer and W. P. Purcell, J. Medicin. Chem., 10, 1000 (1967).
30. D. R. Hudson, G. E. Bass, and W. P. Purcell, J. Medicin. Chem., 13, 1184 (1970).
31. J. Kopecký, Experientia (Bucuresti), 20, 667 (1967).
32. J. Kopecký and K. Bocek, Experientia (Bucuresti), 23, 125 (1967).
33. T. Ban and T. Fujita, J. Medicin. Chem., 12, 353 (1969).
34. T. Fujita and T. Ban, J. Medicin. Chem., 14, 148 (1971).
35. P. Pratesi and E. Grana, Adv. Drug. Res., 2, 134 (1965).
36. A. Cammarata, in "Molecular Orbital Studies in Chemical Pharmacology" (Edited by L. B. Kier), Springer-Verlag, Berlin–New York, 1970, p. 156.
37. A. Cammarata and S. J. Yau, J. Medicin. Chem., 13, 93 (1970).
38. A. Cammarata, J. Medicin. Chem., 8, 150 (1968).
39. J. M. Clayton and W. P. Purcell, J. Medicin. Chem., 12, 1087 (1969).
40. P. N. Craig, J. Medicin. Chem., 15, 144 (1972).
41. A. Leo, C. Hansch, and D. Elkins, Chem. Rev., 71, 525 (1971).
42. J. C. McGowan, J. Appl. Chem., 1S, 120 (1951); 2, 323 (1952); 2, 651 (1952); 4, 41 (1954).
43. O. Exner, Coll. Czech. Chem. Comm., 32, 24 (1967).
44. O. Exner, Coll. Czech. Chem. Comm., 32, 1 (1967).
45. E. I. Ariens, Arch. Int. Pharmacodyn., 127, 459 (1960).
46. R. Collander, Acta Chem. Scand., 4, 85 (1950); 5, 774 (1951).
47. R. Collander, Physiol. Plantarum., 7, 420 (1954).
48. C. Hansch and W. J. Dunn, J. Pharm. Sci., 61, 1 (1972).
49. L. P. Hammett, "Physical Organic Chemistry" (Translated into Russian), Izd. Mir, Moscow, 1972.
50. V. Palm, "Osnovy Kolichestvennoi Teorii Organicheskikh Reaktsii" (Fundamentals of a Quantitative Theory of Organic Reactions), Izd. Khimiya, Leningrad, 1967.
51. J. E. Leffler and E. Grunwald, "Rates and Equilibria of Organic Reactions", John Wiley, New York, 1963.
52. T. Fujita, J. Iwasa, and C. Hansch, J. Amer. Chem. Soc., 86, 5175 (1964).
53. W. J. Dunn and C. Hansch, Chem.-biol. Interactions, 9, 75 (1974).
54. E. C. Bate-Smith and R. G. Westall, Biochim. Biophys. Acta, 4, 427 (1950).
55. C. B. Boyce and B. V. Milborrow, Nature, 208, 537 (1965).
56. G. L. Biagi, A. M. Barbaro, M. C. Guerra, G. C. Forti, and M. E. Fracasso, J. Medicin. Chem., 17, 28 (1974).
57. F. Helmer, K. Kehs, and C. Hansch, Biochemistry, 7, 2858 (1968).
58. R. Franke, Acta Biol. Medicin. Ger., 25, 789 (1970).
59. J. H. Collett, J. C. Dearden, and E. Tomlinson, J. Pharm. Pharmacol., 25, Suppl., 136P (1973).
60. A. Leo, C. Hansch, and C. Church, J. Medicin. Chem., 12, 766 (1969).
61. I. Moriguchi, Chem. Pharm. Bull. (Tokyo), 23, 247 (1975).
62. P. Ahmad, C. A. Fyfe, and A. Mellors, Biochem. Pharmacol., 24, 1103 (1975).
63. J. A. Ostrenga, J. Medicin. Chem., 12, 766 (1969).
64. A. Cammarata and S. J. Yau, J. Medicin. Chem., 14, 1211 (1971).
65. J. H. Hildebrand, J. M. Prausnitz, and R. L. Scott, "Regular and Related Solutions", Van Nostrand Reinhold, New York, 1970.
66. J. A. Ostrenga, J. Medicin. Chem., 12, 349 (1969).
67. L. B. Kier and W. J. Murray, J. Medicin. Chem., 18, 1272 (1975).
68. W. Van Valkenberg and A. F. Yapel, Adv. Chem. Ser., No. 114, 252 (1972).
69. C. G. Swain and E. C. Lupton, J. Amer. Chem. Soc., 90, 4328 (1968).
70. P. N. Craig, J. Medicin. Chem., 14, 680 (1971).
71. K. Bowden and R. Young, J. Medicin. Chem., 14, 1211 (1971).
72. C. Hansch, A. Leo, S. H. Unger, K. H. Kim, D. Nikaitani, and E. J. Lien, J. Medicin. Chem., 16, 1207 (1973).
73. D. Farcasiu, Stud. Cercet. Chim. (Bucuresti), 14, 37 (1966).
74. T. Yamamoto and T. Otsu, Chem. and Ind., 4, 787 (1967).
75. T. Otsu, T. Ito, Y. Fujii, and M. Imoto, Bull. Chem. Soc. Japan, 41, 204 (1968).
76. C. Hansch, J. Medicin. Chem., 11, 920 (1968).
77. C. Hansch and R. Kerley, J. Medicin. Chem., 13, 957 (1970).
78. B. Hetnarski and R. D. O'Brien, J. Medicin. Chem., 18, 29 (1975).
79. B. Pullman and A. Pullman, Nature, 199, 467 (1963).
80. J. Izenberg, A. Szent-György, Proc. Nat. Acad. Sci. U. S. A., 44, 857 (1958); 45, 519 (1959).
81. T. Fujita, J. Medicin. Chem., 9, 797 (1966).
82. T. Fujita and C. Hansch, J. Medicin. Chem., 10, 991 (1967).
83. D. Agin, L. Hersh, and D. Holtzman, Proc. Nat. Acad. Sci. U. S. A., 53, 952 (1965).
84. E. Kutter, H. Machleidt, W. Reuter, R. Sauter, and A. Wildfeuer, Adv. Chem. Ser., No. 114, 98 (1972).
85. A. Cammarata, J. Medicin. Chem., 10, 525 (1967).
86. E. Coats, W. R. Glave, and C. Hansch., J. Medicin. Chem., 13, 913 (1970).
87. C. Hansch, K. Nakamoto, M. Gorin, P. Denisevi, E. R. Garrett, S. M. Hemanack, and C. H. Won, J. Medicin. Chem., 16, 917 (1973).
88. J. W. McFarland, L. H. Conover, A. L. Howes, J. E. Lynch, O. R. Courtney, and D. H. Morgan, 155th Nat. Meet. Amer. Chem. Soc. (1968); quoted after Ref. 12.
89. J. W. McFarland, Progr. Drug Res., 15, 124 (1971).
90. C. D. Ritchie and W. F. Sager, Progr. phys. org. Chem., 2, 342 (1964).
91. N. Kakeya, N. Yata, A. Kamada, and M. Aoki, Chem. Pharm. Bull. (Tokyo), 18, 191 (1970).
92. W. A. Seth-Paul and A. Van Duyse, Spectrochim. Acta, (A), 28, 211 (1972).

93. J. P. Tollenaere, J. Medicin. Chem., **16**, 971 (1973).
94. T. K. Lin, Y. W. Chien, R. R. Dean, J. E. Dutt, H. W. Sause, C. H. Yen, and P. K. Yonan, J. Medicin. Chem., 17, 751 (1974).
95. A. Rastelli, P. G. Debenede, G. G. Battistu, and A. Albasini, J. Medicin. Chem., 18, 963 (1975).
96. J. K. Seydel, H. Ahrens, and W. Losert, J. Medicin. Chem., 18, 234 (1975).
97. B. J. Broughton, P. Chaplen, P. Knowles, E. Lunt, S. M. Marshall, D. L. Pain, and K. R. Wooldrid, J. Medicin. Chem., 18, 1117 (1975).
98. W. H. Dekker, H. A. Selling, and J. C. Overeem, J. Agr. Food Chem., 23, 785 (1975).
99. R. W. Taft, J. Amer. Chem. Soc., 74, 3120 (1952).
100. C. K. Hancock and C. P. Falls, J. Amer. Chem. Soc., 83, 4214 (1961).
101. M. Charton, J. Amer. Chem. Soc., 91, 615 (1969).
102. E. Kutter and C. Hansch, J. Medicin. Chem., 12, 647 (1969).
103. S. H. Unger and C. Hansch, J. Medicin. Chem., 16, 745 (1973).
104. C. Hansch and J. M. Clayton, J. Pharm. Sci., 62, 1 (1973).
105. A. Cammarata and K. S. Rogers, J. Medicin. Chem., 14, 269 (1971).
106. O. Chalvet, J. Theoret. Biol., 8, 51 (1962).
107. B. Pullman and A. Pullman, "Quantum Biochemistry" (Translated into Russian), Izd. Mir, Moscow, 1965.
108. B. Pullman, Electronics Aspects of Biochemistry, Acad. Press, New York, 1964.
109. C. Hansch, E. W. Deutsch, and R. N. Smith, J. Amer. Chem. Soc., 87, 2738 (1965).
110. W. P. Purcell, J. Medicin. Chem., 9, 294 (1966).
111. W. B. Neely, Mol. Pharmacol., 3, 108 (1967).
112. W. B. Neely, H. C. White, and A. Rudzik, J. Pharm. Sci., 57, 1176 (1968).
113. A. Cammarata and R. L. Stein, J. Medicin. Chem., 11, 829 (1968).
114. O. E. Schultz, H. Hollmann, and E. Weber, Arzneim.-Forsch., 20, 1529 (1970).
115. A. J. Wohl, Mol. Pharmacol., 6, 195 (1970).
116. L. B. Kier, "Molecular Orbital Theory in Drug Research", Acad. Press, New York, 1971.
117. W. P. Purcell and J. A. Singer, J. Chem. Eng. Data, 12, 235 (1967).
118. K. Fukui, C. Nagata, and A. Imamura, Science, 132, 87 (1960).
119. K. Fukui, T. Yonezawa, and C. Nagata, J. Chem. Phys., 27, 1247 (1957).
120. A. Cammarata, J. Medicin. Chem., 11, 1111 (1968).
121. C. Hansch, R. M. Muir, T. Fujita, P. P. Maloney, F. Geiger, and M. Streich, J. Amer. Chem. Soc., 85, 2817 (1963).
122. R. A. Fischer, "Statistical Methods for Research Workers", Oliver and Boyd, Edinburgh, 1970.
123. G. W. Snedecor and C. W. Gemmel, "Statistical Methods", Iowa State Univ. Press, Ames, Iowa, 1967.
124. J. G. Topliss and R. J. Costello, J. Medicin. Chem., 15, 1066 (1972).
125. P. J. Goodford, Adv. Pharm. Chemotherapy, 11, 51 (1973).
126. C. Hansch, J. Org. Chem., **35**, 620 (1970).
127. C. Hansch and E. J. Lien, Biochem. Pharmacol., 17, 709 (1968).
128. Y. C. Martin and K. R. Lynn, J. Medicin. Chem., 14, 1162 (1971).
129. R. Franke, E. Gäbler, and P. Oehme, Acta. biol. Medicin. Ger., 32, 545 (1974).
130. T. K. Lin, J. Medicin. Chem., 17, 151 (1974).
131. T. K. Lin, J. Medicin. Chem., 17, 745 (1974).
132. P. H. Weiner and M. L. Weiner, J. Medicin. Chem., 16, 655 (1973).
133. P. H. Weiner, E. R. Malinows, and A. R. Levinsto, J. Phys. Chem., 74, 4537 (1970).
134. J. W. McFarland, J. Medicin. Chem., 13, 1192 (1970).
135. C. Hansch and S. M. Anderson, J. Medicin. Chem., 10, 745 (1967).
136. G. Clopman and R. F. Hudson, Theoret. Chim. Acta, 8, 165 (1967).
137. G. Clopman, J. Amer. Chem. Soc., 90, 223 (1968).
138. R. C. Reid and T. K. Sherwood, "The Properties of Gases and Liquids: Their Estimation and Correlation" (Russian translation edited by V. B. Kogan), Izd. Khimiya, Leningrad, 1971, p. 22.
139. P. N. Craig, Adv. Chem. Ser., No. 114, 115 (1972).
140. C. Hansch, J. Medicin. Chem., 19, 1 (1976).
141. W. P. Purcell, European J. Medicin. Chem., 10, 335 (1975).
142. J. J. Kaufman and W. S. Koski, Int. J. Quantum Chem. Symp., 1975, No. 2, p. 35.
143. P. G. Goodford and F. E. Norrington, in "Quantitative Structure–Activity Relationships", (Edited by M. Tichy), Akadémiai Kiadó, Budapest, 1976, p. 79.
144. J. G. Topliss and E. L. Shapiro, J. Medicin. Chem., 18, 621 (1975).
145. M. E. Rogers, H. H. Ong, E. L. May, and W. A. Klee, J. Medicin. Chem., 18, 1036 (1975).
146. J. A. Montgomery, J. G. Mayo, and C. Hansch, J. Medicin. Chem., 17, 477 (1974).
147. E. J. Lien, C. Hansch, and S. M. Anderson, J. Medicin. Chem., 11, 430 (1968).
148. G. L. Biagi, M. C. Guerra, and A. M. Barbaro, J. Medicin. Chem., 13, 944 (1970).
149. F. R. Quinn, J. S. Driscoll, and C. Hansch., J. Medicin. Chem., 18, 332 (1975).
150. E. Kutter, A. Herz, H. J. Teschema, and R. Hess, J. Medicin. Chem., 13, 801 (1970).
151. C. Hansch, A. Vittoria, C. Silipo, and P. Y. C. Jow, J. Medicin. Chem., 18, 546 (1975).
152. W. P. Purcell, J. Medicin. Chem., 9, 297 (1966).
153. C. Hansch and S. Silipo, J. Medicin. Chem., 17, 661 (1974).
154. F. Helmer, K. Kiehs, and C. Hansch, Biochemistry, 7, 2858 (1968).
155. C. Hansch and F. Helmer, J. Polymer Sci., **A-1**, 6, 3295 (1968).
156. C. Hansch and K. N. Kaula, Biochem. Pharmacol., 19, 2193 (1970).
157. G. H. Hamor and E. J. Lien, Farmaco (Pavia), 24, 704 (1969).
158. G. L. Flynn and S. H. Yalkowsky, J. Pharm. Sci., 61, 838 (1972).
159. A. A. Sinkula and S. H. Yalkowsky, J. Pharm. Sci., 64, 181 (1975).
160. J. A. Montgomery and J. G. Mayo, J. Medicin. Chem., 17, 477 (1974).
161. C. Hansch, A. R. Steward, and J. Iwasa, Mol. Pharmacol., 1, 87 (1965).
162. O. R. Hansen, Acta Chem. Scand., 16, 1593 (1962).
163. M. A. Q. Chaudry and K. C. James, J. Medicin. Chem., 17, 157 (1974).
164. E. M. Hodnett and W. I. Dum, J. Medicin. Chem., 13, 768 (1970).

165. E. J. Lien, M. Hussain, and M. P. Golden, J. Medicin. Chem., 13, 623 (1970).
166. C. F. Barfknecht and D. E. Nichols, J. Medicin. Chem., 18, 208 (1975).
167. R. I. Mrongovius, European J. Medicin. Chem., 10, 474 (1975).
168. M. Yoshimito, K. N. Kaulla, and C. Hansch, J. Medicin. Chem., 18, 950 (1975).
169. K. H. Büchel, W. Draber, A. Trebst, and E. Pistorius, Z. Naturforsch., 21, 243 (1966).
170. W. Draber, K. Dickore, K. H. Büchel, A. Trebst, and E. Pistoriu, Naturwissenschaften, 55, 446 (1968).
171. W. Draber, K. H. Büchel, K. Dickore, A. Trebst, and E. Pistoriu, Progr. Photosynthesis Res., 3, 1789 (1969).
172. P. Moser, K. Jakel, P. Krupp, R. Menasse, and A. Sallmann, European J. Medicin. Chem., 10, 613 (1975).
173. T. Fujita, Chem. Regulation Plants (Tokyo), 5, 124 (1970).
174. G. H. Hamor and E. J. Lien, Farmaco (Pavia), 24, 704 (1969).
175. E. J. Lien and C. Hansch, Adv. Chem. Ser., 114, 155 (1972).
176. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1965, Vol. 4, p. 118.
177. A. Leo, P. Low, C. Silipo, and C. Hansch, J. Medicin. Chem., 18, 865 (1975).
178. G. L. Biagi, A. M. Barbaro, O. Gandolfi, M. C. Guerra, and G. Cantelli, J. Medicin. Chem., 18, 873 (1975).
179. T. Fujita, Adv. Chem. Ser., No. 114, 1 (1972).
180. K. Bowden and M. P. Henry, Adv. Chem. Ser., No. 114, 130 (1972).
181. G. H. Miller, P. H. Doukas, and J. K. Seydel, J. Medicin. Chem., 15, 700 (1972).
182. T. Miyagishima, Chem. Pharm. Bull. (Tokyo), 22, 2288 (1974).
183. Y. C. Martin, J. Medicin. Chem., 13, 145 (1970).
184. R. C. Allen, G. L. Carson, and C. J. Cavallito, J. Medicin. Chem., 13, 909 (1970).
185. E. Coats, W. R. Glave, and C. Hansch, J. Medicin. Chem., 13, 913 (1970).
186. T. Fujita, J. Medicin. Chem., 16, 923 (1973).
187. E. A. Coats, J. Medicin. Chem., 16, 1102 (1973).
188. C. Silipo and C. Hansch, Mol. Pharmacol., 10, 955 (1974).
189. C. F. Wilkinson, K. Hetnarski, G. P. Cantwell, and F. J. Dicarlo, Biochem. Pharmacol., 23, 2377 (1974).
190. R. Labbe-Bois, C. Laruelle, and J.-J. Godfroid, J. Medicin. Chem., 18, 85 (1975).
191. Y. C. Martin, W. B. Martin, and J. D. Taylor, J. Medicin. Chem., 18, 883 (1975).
192. T. Fujita, Adv. Chem. Ser., No. 114, 80 (1972).
193. P. Craig and C. Hansch, J. Medicin. Chem., 16, 661 (1973).
194. W. B. Neely and W. K. Whitney, J. Agr. Food. Chem., 16, 571 (1968).
195. C. Hansch, Farmaco (Pavia), 23, 293 (1968).
196. E. V. Deutsch and C. Hansch, Nature, 211, 75 (1966).
197. J. G. Topliss and M. D. Yudis, J. Medicin. Chem., 15, 394 (1972).
198. C. Hansch and E. W. Deutsch, Biochim. Biophys. Acta, 112, 381 (1966).
199. S. H. Unger and C. Hansch, J. Medicin. Chem., 16, 745 (1973).
200. A. Cammarata, J. Medicin. Chem., 15, 573 (1972).
201. M. Yoshimoto, C. Hansch, and P. Y. Jow, Chem. Pharm. Bull. (Tokyo), 23, 437 (1975).
202. M. E. Wolff and C. Hansch, J. Medicin. Chem., 17, 898 (1974).
203. W. Dittmar, E. Druckrey, and H. Urbach, J. Medicin. Chem., 17, 753 (1974).
204. I. Simiti, I. Schwartz, M. Farkas, and E. Chindris, Rev. roum. biochim., 11, 145 (1974)
205. R. Franke, W. Zschiesche, K. Augsten, J. Guettner, and G. Hesse, RES, J. Reticuloendothel. Soc., 16, 87 (1974).
206. B. Tinlano, Farmaco, Ed. Sci., 30, 423 (1975).
207. R. F. Rekker, W. T. Nauta, T. Bultsma, and C. G. Waringa, European J. Medicin. Chem., 10, 557 (1975).
208. A. F. Yapel, Adv. Chem. Ser., No. 114, 183 (1972).
209. K. H. Büchel and W. Draber, Adv. Chem. Ser., No. 114, 141 (1972).
210. F. E. Norrington, R. M. Hyde, S. G. Williams, and R. Wootton, J. Medicin. Chem., 18, 604 (1975).
211. A. Leo, P. Y. C. Jow, C. Silipo, and C. Hansch, J. Medicin. Chem., 18, 865 (1975).
212. E. Tomlinson, J. Chromatog., 113, 1 (1975).
213. J. D. Turnbull, G. L. Biagi, A. J. Merola, and D. Cornwell, Biochem. Pharmacol., 20, 1383 (1971).
214. G. L. Biagi, A. M. Barbaro, and M. C. Guerra, Experientia (Basel), 27, 918 (1971).
215. W. Draber, K. H. Büchel, and K. Dickore, "Proc. II Int. Congr. Pestic Chem.," (Edited by A. S. Tahori), Gordon and Breach, New York, 1972, Vol. 5, p. 153.
216. "Spravochnik Khimika" (The Chemist's Handbook), Moscow–Leningrad, 1965, Vol. 4, p. 136.
217. G. L. Biagi, O. Gandolfi, M. C. Guerra, A. M. Barbaro, and G. Cantelli, J. Medicin. Chem., 18, 868 (1975).
218. R. H. Adamson, Life Sci., 3, 1131 (1964).
219. P. Ahmad, C. A. Fyfe, and A. Mellors, Canad. J. Biochem., 53, 1047 (1975).
220. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Leningrad, 1971, Vol. 1, p. 390.
221. O. R. Quayle, Chem. Rev., 53, 484 (1953).
222. A. I. Vogel, W. T. Cresswell, G. H. Jeffrey, and J. Leicester, J. Chem. Soc., 514 (1952).
223. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1971, Vol. 2, p. 398.
224. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1965, Vol. 3, p. 306.
225. "Spravochnik po Rastvorimosti" (Handbook of Solubility), Izd. Khimiya, Moscow–Leningrad, 1961, Vol. 1; 1963, Vol. 2; 1970, Vol. 3.
226. J. A. Clements and K. M. Wilson, Proc. Nat. Acad. Sci. U. S. A., 48, 1008 (1962).
227. A. Cammarata, S. J. Yau, and K. S. Rogers, Pure Appl. Chem., 35, 495 (1973).
228. G. M. Bristow and W. F. Watson, Trans. Faraday Soc., 54, 1731 (1958).
229. A. F. M. Barton, Chem. Rev., 75, 731 (1975).
230. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1965, Vol. 4, p. 515.
231. P. B. Hulbert, Mol. Pharmacol., 10, 315 (1974).
232. A. Chiriac, D. Ciubotar, Z. Szabadai, R. Vilceaanu, and Z. Simon, Rev. roum. Biochim., 12, No. 3, 143 (1975).
233. H. H. Jaffe, Chem. Rev., 53, 191 (1953).
234. D. H. McDaniel and H. C. Brown, J. Org. Chem., 23, 420 (1958).

235. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1965, Vol. 3, p. 935.
236. M. S. Tute, Adv. Drug. Res., 6, 1 (1971).
237. M. Pesak, F. Kopecky, J. Celechovsky, A. Borovansky, and A. Lukas, Cesk. Farm., 23, 207 (1974).
238. K. Bowden and R. C. Young, J. Medicin. Chem., 13, 225 (1970).
239. G. Berg, T. Bultsma, and W. T. Nauta, Biochem. Pharmacol., 24, 115 (1975).
240. J. J. Zimmerman and J. E. Goyan, J. Medicin. Chem., 13, 492 (1970).
241. P. H. Wells, Chem. Rev., 63, 171 (1963).
242. R. Franke, Acta Biol. Medicin. Ger., 30, 467 (1973).
243. H. C. Brown and Y. Okamoto, J. Amer. Chem. Soc., 80, 4979 (1958).
244. T. K. Chih, Acta Chim. Sinica, 32, 107 (1966).
245. L. A. Cohen and W. M. Jones, J. Amer. Chem. Soc., 85, 3398 (1963).
246. R. W. Taft, J. Chem. Phys., 26, 93 (1957).
247. A. Albert and E. P. Serjeant, "Ionisation Constants of Acids and Bases: A Laboratory Manual" (Translated into Russian), Izd. Khimiya, Moscow–Leningrad, 1964.
248. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Moscow–Leningrad, 1965, Vol. 3, p. 77.
249. N. V. Dmitrieva, Biofizika, 16, 958 (1971).
250. G. Teutsch, L. Weber, G. Page, E. L. Shapiro, H. L. Herzog, R. Neri, and E. J. Collins, J. Medicin. Chem., 16, 1370 (1973).
251. C. E. Meloan and R. W. Kiser, "Problems and Experiments in Instrumental Analysis", Ch. E. Merill Books, Columbus, 1963, p. 295.
252. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Leningrad, 1971, Vol. 1, p. 395.
253. O. A. Osipov, V. I. Minkin, and A. D. Garnovskii, "Spravochnik po Dipol'nym Momentam" (Handbook of Dipole Moments), Izd. Vysshaya Shkola, Moscow, 1971.
254. G. Karremann, J. Izenberg, and A. Szent-Guörgy, Science, 130, 1191 (1959).
255. L. E. Lions and J. Mackie, Nature, 197, 589 (1963).
256. A. C. Allison and T. Nach, Nature, 197, 758 (1963).
257. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Leningrad, 1971, Vol. 1, p. 329.
258. L. V. Gurvich, G. V. Karachevtsev, Yu. A. Kondrat'ev, V. A. Lebedev, V. A. Medvedev, V. K. Potapov, and Yu. S. Khodeev, Handbook, "Energii Razryva Khimicheskikh Svyazei. Potentsialy Ionizatsii i Srodstvo k Elektronu" (Chemical Bond Dissociation Energies, Ionisation Potentials and Electron Affinities), Izd. Nauka, Moscow, 1974.
259. R. E. Tessel, J. H. Woods, R. E. Counsell, and M. Lu, J. Pharm. Expt. Therap., 192, 310 (1975).
260. A. Bondi, J. Phys. Chem., 68, 441 (1964).
261. G. Feurstein, Arzneimittel Forsch., 8, 694 (1958).
262. J. Traube, Pflügers Arch., 153, No. 5, 276 (1913).
263. Symposium, "Deistvie Fiziologicheski Aktivnykh Soedinenii na Biologicheskie Membrany" (The Effect of Physiologically Active Compounds on Biological Membranes), Izd. Nauka, Moscow, 1974, pp. 280, 292.
264. N. T. Pryanishnikova and V. A. Pchelin, Dokl. Akad. Nauk SSSR, 126, 1358 (1959).
265. "Spravochnik Khimika" (The Chemist's Handbook), Izd. Khimiya, Leningrad, 1971, Vol. 1, p. 1011; 1965, Vol. 3, p. 726.
266. J. J. Jasper, J. Phys. Chem. Ref. Data, 1, 841 (1972).
267. P. M. Seeman and H. S. Bialy, Biochem. Pharmacol., 12, 1181 (1963).
268. W. C. Griffin, Amer. Perfumer, 65, 26 (1955).
269. R. Vochten and G. Petre, Bull. Soc. chim. belges, 81, 583 (1972).
270. J. A. Hickman, Biochem. Pharmacol., 23, 2833 (1974).
271. R. Franke, Arch. Geschwulstforsch., 36, 30 (1970).
272. R. Franke, Arch. Geschwulstforsch., 37, 45 (1971).
273. L. N. Ferguson, "Modern Structural Theory of Organic Chemistry", Prentice-Hall, Englewoods Cliffs, 1952, p. 54.
274. K. Nukada, O. Yamamoto, T. Suzuki, M. Takeuchi, and M. Ohnishi, Analyt. Chem., 35, 1892 (1963).
275. S. Martin and B. P. Dailey, J. Chem. Phys., 39, 1723 (1963).
276. F. C. Stehling and K. W. Bartz, Analyt. Chem., 38, 1467 (1966).
277. G. Slomp and J. G. Lindberg, Analyt. Chem., 39, 60 (1967).
278. D. Yamamoto, T. Suzuki, M. Yanagisa, K. Hayamizu, and M. Ohnishi, Analyt. Chem., 40, 568 (1968).
279. J. J. Reed, Analyt. Chem., 39, 1586 (1967).
280. F. Peradejordi, A. N. Martin, and A. Cammarata, J. Pharm. Sci., 60, 576 (1971).
281. W. B. Neely, H. C. White, and A. Rudzik, J. Pharm. Sci., 57, 1176 (1968).
282. S. Kang and J. P. Green, Proc. Nat. Acad. Sci. U. S. A., 67, 62 (1970).
283. S. Kang and J. P. Green, Nature, 226, 645 (1970).
284. D. D. Perrin, "Dissociation Constants of Organic Bases in Aqueous Solutions", Butterworths, London, 1973.
285. R. K. Freier, "Aqueous Solutions, Vol. 1: Data for Inorganic and Organic Compounds", de Gryter, Berlin–New York, 1976.
286. E. Gaebler, R. Franke, and P. Oehme, Pharmazie (Berlin), 31, 1 (1976).
287. H. Kubinyi and O. H. Kehrhahn, J. Medicin. Chem., 19, 578, 1040 (1976).
288. S. V. Nizhnic and N. A. Epshtein, Khim.-Farm. Zhur., No. 3, 81 (1977).
289. H. Kubinyi, J. Medicin. Chem., 19, 587 (1976).
290. H. Kubinyi, J. Medicin. Chem., 20, 625 (1977).
291. Y. C. Martin and J. J. Hackbarth, J. Medicin. Chem., 19, 1033 (1976).
292. A. Cammarata and G. K. Menon, J. Medicin. Chem., 19, 739 (1976).
293. R. A. Coburn and A. J. Solo, J. Medicin. Chem., 19, 748 (1976).
294. C. Hansch, S. D. Rockwell, P. Y. C. Jow, A. Leo, and E. E. Steller, J. Medicin. Chem., 20, 304 (1977).
295. R. A. Scherrer and S. M. Howard, J. Medicin. Chem., 20, 53 (1977).
296. N. A. Epshtein and S. V. Nizhnii, Khim.-Farm. Zhur., No. 4, 52 (1977).
297. T. Fujita, T. Nishioka, and M. Nakajima, J. Medicin. Chem., 20, 1071 (1977).
298. L. H. Hall and L. B. Kier, J. Pharm. Sci., 66, 642 (1977).
299. L. B. Kier and L. H. Hall, "Molecular Connectivity in Chemistry and Drug Research", Acad. Press, New York, 1976.

Institute for the Biological Tests
of Chemical Compounds, Moscow