Journal of Mathematical Chemistry 13(1993)115–132

Application of the linear free energy principle for derivation of nonlinear quantitative structure–activity relationships. A steady-state kinetic model

Artūras K. Dubickas
Department of Mathematics, University of Witwatersrand, PO WITS 2050, Johannesburg, South Africa

and

A.A. Petrauskas*
A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russia

Received 18 November 1991

All the quantitative structure–activity relationships (QSAR) which so far have been practically observed can be simple interpreted within the framework of kinetic formalism. Based on the assumption that the drug action can be modeled by the steady-state kinetic scheme which includes only reactions of the first order, and the rate constants of all the elementary reactions can be formally described by means of linear free energy relationships (LFER), an algorithm for the derivation of nonlinear QSA relationships is proposed. It is shown that all the possible rate equations for the kinetic schemes of a given complexity can be subdivided into a limited number of classes according to the type of concentration polyhedron. Namely, if one assumes that the kinetic scheme includes no more than n species, then there exist $p(n) - 1$ general rate equations. Based on such classification, the possibility of finding the necessary QSA relationship proceeding only from the experimental data on the final biological responses is discussed.

1. Introduction

Among the variety of approaches used to analyze the physico-chemical and biological properties of compounds, the application of linear free energy relationships (LFER) principle still remains the most attractive since potentially it can provide clear information about the physico-chemical nature of the processes analyzed. So far, the most excellent results have been achieved in the application of the LFER principle for correlation of rate or equilibrium constants of simple chemical reactions. In particular, the extrathermodynamical approach became attractive in recent years when Taft, Kamlet and Abraham derived the complete set of solvatochromic parameters,

*Present address: Marion Merrell Dow Inc., Research Institute, 2110 East Galbraith Road, P.O. Box 156300, Cincinnati, OH 45215-6300, USA.

which in their linear combination were able to describe quantitatively almost all types of the single-stage physico-chemical processes occurring in condensed phase (see, for example, refs. [1–3]).

The major problem of the application of the LFER principle for biological data correlation is conditioned by the complexity of biological systems. It was long recognized that the final biological response of drug action results from various branched or consecutive elementary reaction steps; each of these steps may be rate-determining for any new biologically active compound. Therefore, the linear relationships between the final biological response and various compound predictor variables may be obtained only in very particular cases, e.g. when the same rate-determining step defines the biological action of all the analyzed compounds. On the other hand, the precise mechanism of drug action at the molecular level is never known; therefore, it is impossible to derive theoretically the required structure–activity relationship which will certainly describe the biological action of the compounds under investigation.

Because of these difficulties, so far the purely empirical methodology has been used by pharmaceutical chemists to derive the necessary nonlinear QSA relationships. It implied a simple visualization of the structure–activity dependence by plotting activity against the certain predictor variables in the two-dimensional space, and subsequent fitting of the obtained curve into some empirical function. Thus, it was long recognized that most of the structure–activity curves are composed of two portions, ascending and descending, both usually being linear. In 1964, based on the hypothetical "random walk" model of drug action, Hansch proposed to fit such "two-portion" curves into the parabolic function [4–6]

$$A = C_0 + C_1 x + C_2 x^2, \tag{1}$$

where A represents the rate of biological response or the pharmacokinetic constant, C_0, C_1, C_2 are the regression coefficients, and x is the hydrophobic parameter of the variable substituent or the whole drug molecule. In 1976, Kubinyi generalized Hyde's equilibrium model [7] and McFarland's probability model [8], and proposed the use of the bilinear equation [9,10]

$$A = C_0 + C_1 x + C_2 \log\left(1 + 10^{C_3 + x}\right). \tag{2}$$

Although parabolic and bilinear functions were initially derived based on the more or less realistic models of drug action, their application to the correlation of the biological data in most of the practical cases must be considered as purely empirical.

First, in order to derive parabolic and bilinear equations, a number of over-simplifications have been made. For example, the final biological response was assumed to result simply from a number of consecutive partitions (compartments) between various organic/water phases of the drug molecule, and subsequent single-stage interaction of the drug molecule with the biological receptor. Obviously, such

drug action modeling within the consecutive multicompartment framework represents only a very narrow case among all the possible interpretations of the biphasic structure–activity dependencies. Further, it was assumed that the rate or equilibrium constants between various water/organic systems are strictly interrelated, while the recent results of Taft, Kamlet and Abraham [1–3] indicate that such relationships may be observed only for the substrate homologous series, or in the case of very close organic/water systems.

Second, although initially these equations were derived only for the activity–lipophilicity correlations, they were extensively used to correlate the connectivity indexes [11,12], steric [13], electronic and semiempirical quantum chemical parameters [14].

Due to these uncertainties, only very conditional physico-chemical meaning could be attributed to the optimized regression coefficients in eqs. (1) and (2). Consequently, a number of other empirical equations were proposed to describe the above-mentioned two-portion curves, e.g. the hyperbolic [11,12], moduli [13], and even sinus [12] functions were used,

$$A = C_0 + C_1 x + C_2/x, \tag{3}$$

$$A = C_0 + C_1 |x - C_2|, \tag{4}$$

$$A = C_0 + C_1 \sin x. \tag{5}$$

Among other proposed functions, the Kubinyi function (2) proved to give the most reliable fits of experimental data, since it could correctly describe the non-symmetric and bilinear nature of the visualized structure–activity curves. (The curvilinear nature of certain types of empirical functions has been extensively criticized in refs. [9, 10].)

In 1978, Franke summarized certain cases where the QSA relationships were stated to be described by curves composed of three and four linear or curvilinear parts, and proposed to fit them either into the fourth-order power series (6) or into the complex five-part function (7) [15],

$$A = C_0 + C_1 x + C_2 x^2 + C_3 x^3 + C_4 x^4, \tag{6}$$

$$A = \begin{cases} C_0 + C_1 x, & \text{if } x \le x_{\mathrm{I}}, \\ C_2 + C_3 x + C_4 x^2, & \text{if } x_{\mathrm{I}} \le x \le x_{\mathrm{II}}, \\ C_5 + C_6 x + C_7 x^2, & \text{if } x_{\mathrm{II}} \le x \le x_{\mathrm{III}}, \\ C_8 + C_9 x, & \text{if } x_{\mathrm{III}} x \le x_{\mathrm{IV}}, \\ C_{10} + C_{11} x + C_{12} x^2, & \text{if } x \ge x_{\mathrm{IV}}. \end{cases} \tag{7}$$

Earlier, the third-order power series were used to establish QSA relationships by Hansch and Clayton [16].

It is clear that all these QSA relationships cannot provide the necessary information about the biological processes analyzed because one never can be sure that the optimized function correctly represents the biological mechanism of drug action. On the other hand, all the above-mentioned visualized structure–activity dependencies could be simply interpreted within the framework of kinetic formalism. Thus, the bilinear nature of the two-portion type SA curves can be intuitively considered to be conditioned by the presence of only two rate-determining elementary reactions, the LFER principle being applicable to both of them. Analogously, the complex curves composed of three and four linear or curvilinear parts could be simply explained by the drug action modeling by the four-stage kinetic scheme. Then the following questions reasonably arise: how complex can QSA relationships be if we assume that the biological action of the compound is described by a certain kinetic scheme? Are there limited or unlimited kinds of possible QSA relationships? Is it possible to reproduce the kinetic scheme of drug action and, hence, to choose the necessary QSA relationship proceeding only from the experimental data on the final biological responses? The purpose of the present paper is to obtain at least partial answers to these questions. Here, we will consider only the steady-state kinetic schemes assuming that all the reactions included are of the first order, since only in this case the required structure–activity relationships can be obtained in the form of explicit algebraic equations.

2. Mathematical description of the kinetic model and analysis of its properties

Suppose that the biological action of a set of bioactive compounds can be described by the same kinetic scheme. Let this kinetic scheme include n internal species $X_1, X_2, \ldots, X_n$ which may represent either kinetically important components of the biological system (e.g. various biological target receptors, enzymes produced by the biological system and catalyzing various side reactions of drug molecules, etc.) or various forms of drug molecules (e.g. drug molecules in the administration, membrane or receptor phases, drug molecules bound to the target receptors by the covalent or noncovalent bonds, etc.). The concentrations of internal species $X_i(t)$ have a significant time variation and therefore are dynamical variables. Suppose that the rate of biological response is proportional to the rate of X_{n+1} species production, i.e. to $\omega(t) = \partial X_{n+1}(t)/\partial t$. Our aim is to find the dependence of $\omega(t)$ on the drug molecule predictor variables. Generally, these dependencies can be derived in the form of explicit algebraic expressions only if all the reactions included in the kinetic scheme are of the first order, and all the rate constants are related to the drug molecule predictor variables by the certain functions.

Thus, let all the elementary reactions in our scheme be of the first or the pseudo-first order. For the sake of simplicity, assume that all the reactions included in the kinetic scheme are of the following forms:

$$X_i \xrightarrow{k(i,j,s)} X_j + X_s, \quad X_i \xrightarrow{k(i,0,s)} X_s,$$

where $1 \le i \le n$, $1 \le j \le s \le n+1$, $j, s \ne i$, and $k(i, j, s)$ is a rate constant ($k(i, j, s) \ge 0$). (It will be readily seen that all the considerations below, including theorems, are true for any kinetic scheme of the first order.) Below, we will consider that the rate constants of all the elementary reactions can be formally related to a certain number of drug molecule predictor variables by means of the linear free energy relationships

$$\log k(i, j, s) = a_0(i, j, s) + \sum_{i=1}^{k} a_i(i, j, s) x_i. \tag{8}$$

Here, $a_0, a_1, \ldots, a_k$ are the linear regression coefficients, and $x_1, \ldots, x_k$ represent the complete set of drug molecule parameters such that their linear combination can describe any single-stage reaction occurring in the biological system. There may be rate constants which do not depend on drug molecule properties; then formally all $a_i = 0$, $i = 1, \ldots, k$.

The rate law of drug action will be described by the following system of differential equations:

$$\partial \overline{X}(t) / \partial t = \mathbb{K} \overline{X}(t), \quad \overline{X}(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{pmatrix}, \tag{9}$$

$$\mathbb{K} = \| k_{ij} \|_{i,j=1,2,\ldots,n},$$

where

$$k_{ij} = \begin{cases} \sum\limits_{u=0, u \ne j}^{i-1} k(j, u, i) + 2k(j, i, i) + \sum\limits_{u=i+1, u \ne j}^{n+1} k(j, i, u), & \text{if } i \ne j, \\ - \sum\limits_{0 \le u, v \ne j \le n+1} k(j, u, v), & \text{if } i = j. \end{cases} \tag{10}$$

The rate of biological response is proportional to

$$\omega(t) = \sum_{u=1}^{n} \left(2k(u, n+1, n+1) + \sum_{v=0, v \ne u}^{n} k(u, v, n+1) \right) X_u(t).$$

Obviously, $\omega(t)$ is a function of substrate parameters $x_1, \ldots, x_k$ and the time of its action t. However, the biological activity of a given compound is usually characterized by its initial concentration causing the standard biological response (e.g. ED_{50}, LD_{50}, MIC, MBC, etc.) at a fixed time moment. These experimental data are informative enough if we suppose that they are time independent, i.e. $|\omega(t_1) - \omega(t_2)| < \varepsilon$ for any $t_1, t_2 \in [0; +\infty)$ (here, ε is an experimental error). At least, this inequality must be satisfied for $t_1, t_2 \in [t_0; +\infty)$. In the chemical sense, this

means that the experimental data have been obtained under the conditions near the steady state. Mathematically, this inequality is satisfied if there exists $\lim_{t\to\infty}\omega(t)=\omega^*$. Therefore, we suppose that there exist $\lim_{t\to\infty}X_i(t)=X_i^*$, $i=1,2,\ldots,n$. Then

$$\mathbb{K}\overline{X}^*=0,\quad \overline{X}^*=\begin{pmatrix}X_1^*\\X_2^*\\\vdots\\X_n^*\end{pmatrix}, \tag{11}$$

$$\omega^*=\sum_{u=1}^{n}\left(2k(u,n+1,n+1)+\sum_{v=0,v\neq u}^{n}k(u,v,n+1)\right)X_u^*. \tag{12}$$

ω^* is the experimentally obtained value, and $\omega^*>0$. For convenience, and without loss of generality, we shall assume that all $X_i^*>0$. Indeed, $U=\{u\,|\,1\le u\le n, X_u^*=0\}$. Written in the form

$$\sum_{j=1}^{n}k_{ij}X_j^*=0,\quad i=1,2,\ldots,n, \tag{13}$$

where k_{ij} are defined in eq. (10), the system (11) does not contain the symbols $k(u,v,w)$, where $u\in U$. If $U=\sigma_n$ (here and below, $\sigma_n=\{1,2,\ldots,n\}$), then $\omega^*=0$. Conversely, if $U\neq\sigma_n$, then for $v\in\sigma_n\backslash U$, $X_v^*>0$, and $k(u,v,w)=0$, if u or $w\in U$ (see eqs. (10) and (13)). Thus, all the symbols X_u^*, $u\in U$, $k(u,v,w)$, where u, v, or $w\in U$, can be excluded from eqs. (11) and (12), and the kinetic scheme with $n-\operatorname{card}U$ species can be considered ($\operatorname{card}U$ is the number of elements in the set U).

Let Γ be a stoichiometric matrix $M\times(n+1)$ of the kinetic scheme, where M is the number of elementary reactions. For the row $(\gamma_1,\gamma_2,\ldots,\gamma_{n+1})$ corresponding to the reaction $X_i\to aX_j+bX_s$, $i,j=1,2,\ldots,n$, $s=1,2,\ldots,n+1$, $i\neq j$, $i\neq s$, $j\neq s$, we have

$$\gamma_u=\begin{cases}0, & \text{if } u\notin\{i,j,s\},\\ -1, & \text{if } u=i,\\ a, & \text{if } u=j,\\ b, & \text{if } u=s.\end{cases}$$

Hence, every row of matrix Γ contains only three non-zero elements $(-1,1,1)$ or two non-zero elements $(-1,1,\text{or }-1,2)$. Let Γ_1 be a submatrix of Γ, obtained by striking out the last column from Γ. Note that according to (10),

$$\det\mathbb{K}=\sum_{u_i,v_i\neq i}\prod_{i=1}^{n}k(i,u_i,v_i)\Gamma(u_1,v_1,u_2,v_2,\ldots,u_n,v_n), \tag{14}$$

where $\Gamma(u_1, v_1, u_2, v_2, \ldots, u_n, v_n)$ is the determinant $n \times n$ consisting of the rows of Γ_1 which correspond to the reactions $k(1, u_1, v_1)$, $k(2, u_2, v_2)$, $\ldots$, $k(n, u_n, v_n)$. It is clear that for every i, $1 \le i \le n$, at least one $k(i, u_i, v_i) > 0$, i.e. the kinetic scheme includes at least one reaction of the type $X_i \to X_s$ or $X_i \to X_j + X_s$. Indeed, according to eqs. (10) and (13), the condition $k(i, u, v) = 0$ for all $u, v \neq i$ leads to $k(u, v, w) = 0$ for $i \in \{v, w\}$. Hence, one particle X_i^* can be omitted in the kinetic scheme, and the case with $n-1$ species can be considered.

Obviously, the system (11) has non-zero solution $\overline{X}^*$ if and only if $\det \mathbb{K} = 0$. Suppose that $\det \mathbb{K} \not\equiv 0$. Then $\det \mathbb{K}$ is a polynomial in the variables $k(u, v, w)$. For any $\delta > 0$, there exist δ_{uvw}, $0 \le \delta_{uvw} \le \delta$, such that $\det \mathbb{K} \neq 0$ after replacing all $k(u, v, w)$ by $k(u, v, w) + \delta_{uvw}$. For these new $k(u, v, w)$, we shall obtain $X_1^* = X_2^* = \ldots = X_n^* = 0$ and $\omega^* = 0$. Such a replacement is correct, since all the predictor variables $x_1, \ldots, x_k$ and, hence, the rate constants can be experimentally determined only with the certain nonzero precision. Therefore, $\det \mathbb{K} \equiv 0$.

Using the same argument, we have $\Gamma(u_1, v_1, u_2, v_2, \ldots, u_n, v_n) = 0$. Hence, all the rows of the matrix Γ_1 are linearly dependent, and rank $\Gamma_1 \le n-1$. Consequently, there exists $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{R}$ such that $\mu_1^2 + \mu_2^2 + \ldots + \mu_n^2 \neq 0$ and

$$\Gamma_1 \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = 0. \tag{15}$$

Since (9), (10) and (15) yield

$$\mu_1 \partial X_1(t)/\partial t + \mu_2 \partial X_2(t)/\partial t + \ldots + \mu_n \partial X_n(t)/\partial t \equiv 0,$$

the following equality holds:

$$\mu_1 X_1(t) + \mu_2 X_2(t) + \ldots + \mu_n X_n(t) \equiv \text{const}. \tag{16}$$

This means that under our assumptions for the kinetic scheme, there exists at least one equation of conservation constraint. Thus, there exists a set $\Pi \subset \mathbb{R}^n$, $\Pi \neq 0$, such that

$$\Pi = \{(\mu_1, \mu_2, \ldots, \mu_n) \in \mathbb{R}^n \mid \mu_1 X_1(t) + \mu_2 X_2(t) + \ldots + \mu_n X_n(t) \equiv \text{const}\}.$$

Π is called the concentration polyhedron, and it depends only on the topological structure of the kinetic scheme (graph). Π is a linear space and it has a basis. By e_H, $H \subseteq \sigma_n$, we denote the vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, where

$$\alpha_i = \begin{cases} 1, & \text{if } i \in H, \\ 0, & \text{if } i \notin H. \end{cases}$$

THEOREM 1

For any concentration polyhedron Π corresponding to the kinetic scheme of the first order with n species, there exists such basis $e_{H_1} e_{H_2}, \ldots, e_{H_l}$ that $H_i \cap H_j = \varnothing$, card $H_i \geq 2$, $i, j = 1, 2, \ldots, l$, $i \neq j$, and $H_1 \cup H_2 \cup \ldots \cup H_l \subseteq \sigma_n$.

Proof

Let $f = (\mu_1, \mu_2, \ldots, \mu_n) \in \Pi$, $f \neq 0$ (see eqs. (15) and (16)). Denote $A_0 = \{i \mid \mu_i = 0\}$. Without loss of generality, assume that $A_0 = \{s+1, s+2, \ldots, n\}$ (if $A_0 = \varnothing$, then $s = n$). Let Γ_2 be a matrix obtained from Γ_1 by striking out the last $n - s$ columns and, after that, by striking out all zero rows from the obtained matrix. Then, according to (15),

$$\Gamma_2 \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_s \end{pmatrix} = 0. \tag{17}$$

So, every row of Γ_2 has at least two non-zero elements. Let us consider the sum

$$\partial X_1(t)/\partial t + \partial X_2(t)/\partial t + \ldots + \partial X_s(t)/\partial t$$

$$= \sum_{i=1}^{n} X_i(t) \sum_{i \in \{u,v,w\}}^{n} k(u, v, w) S(u, v, w). \tag{18}$$

Here, $S(u, v, w)$ is the sum of the elements of the row which corresponds to $k(u, v, w)$ in matrix Γ_2. Hence, $S(u, v, w) \geq 0$. Since $\lim_{t \to \infty} \partial X_i(t)/\partial t = 0$ and $\lim_{t \to \infty} X_i(t) = X_i^* > 0$, every $S(u, v, w)$ in eq. (18) is equal to zero. Consequently, every row in Γ_2 has only two non-zero elements -1 and 1.

Let $\beta_1, \beta_2, \ldots, \beta_r$ represent the distinct real numbers among $\mu_1, \mu_2, \ldots, \mu_s$. Denote $C_1 = \{i \mid \mu_i = \beta_1\}, C_2 = \{i \mid \mu_i = \beta_2\}, \ldots, C_r = \{i \mid \mu_i = \beta_r\}$. Let $\gamma_{uv} = -1, \gamma_{uw} = 1$ for the uth row of Γ_2. Then (17) implies $\mu_v = \mu_w$ and, hence, there exists such j, $1 \leq j \leq r$, that $v, w \in C_j$. Let Γ^j, $1 \leq j \leq r$, be a matrix obtained from Γ_2 by retaining only the columns with indexes C_j. It is readily seen that every non-zero row in Γ^j has only two non-zero elements -1 and 1. Similarly to (18), it leads to

$$\sum_{i \in C_j} \partial X_i(t)/\partial t \equiv 0.$$

Hence, $e_{C_1}, e_{C_2}, \ldots, e_{C_r} \in \Pi$, card $C_j \geq 2$, $C_i \cap C_j = \varnothing$, $i \neq j$, and

$$f = \beta_1 e_{C_1} + \beta_2 e_{C_2} + \ldots + \beta_r e_{C_r}. \tag{19}$$

Note that

$$e_{C_1} + e_{C_2} + \ldots + e_{C_r} = e_{C_1 \cup C_2 \cup \ldots \cup C_r} \in \Pi. \tag{20}$$

Let rank $\Pi = 1$. For every basis B of Π

$$f_1 = (\mu_1^1, \mu_2^1, \ldots, \mu_n^1), \ldots, \quad f_l = (\mu_1^l, \mu_2^l, \ldots, \mu_n^l), \tag{21}$$

we define a natural number

$$N(B) = \sum_{\substack{1 \le i \le n, 1 \le j \le l \\ \mu_i^j \ne 0}} 1.$$

Let (21) be a basis of Π with the minimal number $N(B)$. For every j, $1 \le j \le l$, the vector f_j can be expressed in the form (19). Let us assume that there exists such j, for example $j = 1$, for which $r > 1$. Then

$$f_1 = \beta_1 e_{D_1} + \beta_2 e_{D_2} + \ldots + \beta_r e_{D_r}, \quad e_{D_i} \in \Pi.$$

The vectors $f_1, f_2, \ldots, f_l$ are linearly independent; therefore, for at least one i, $1 \le i \le r$, the vectors $e_{D_i}, f_2, f_3, \ldots, f_l$ are linearly independent. Hence, $e_{D_i}, f_2, f_3, \ldots, f_l$ is a new basis B_1 of Π, and $N(B_1) < N(B)$. Consequently, $f_1 = \theta_1 e_{H_1}, f_2 = \theta_2 e_{H_2}, \ldots, f_l = \theta_l e_{H_l}$, where $\theta_i \in \mathbb{R}$, $\theta_i \ne 0$, $H_i \subseteq \sigma_n$, $\operatorname{card} H_i \ge 2$. Therefore, $e_{H_1}, e_{H_2}, \ldots, e_{H_l}$ is a new basis B_2 of Π, and $N(B_2) = N(B) = \sum_{i=1}^{l} \operatorname{card} H_i$. To prove the theorem, it suffices to show that $H_i \cap H_j = \varnothing$, $i, j = 1, 2, \ldots, n$, $i \ne j$. For convenience, we shall prove that $H_1 \cap H_2 = \varnothing$. Suppose that $H_1 \cap H_2 = H \ne \varnothing$. Obviously, $H_1 \ne H_2$, since $e_{H_1} \ne e_{H_2}$. If $H_1 = H$, then $e_{H_2 \setminus H_1} = e_{H_2} - e_{H_1} \in \Pi$. So $e_{H_1}, e_{H_2 \setminus H_1} = e_{H_3}, \ldots, e_{H_l}$ is a new basis of B_3 of P, and $N(B_3) < N(B_2)$, i.e. a contradiction. Analogously, $H_2 \ne H$. Note that $e_{H_1 \setminus H} - e_{H_2 \setminus H} = e_{H_1} - e_{H_2} \in \Pi$. $e_{H_1 \setminus H} - e_{H_2 \setminus H}$ can be expressed in the form (19); therefore, from (20) we find that

$$e_{H_1 \setminus H} + e_{H_2 \setminus H} = e_{(H_1 \setminus H) \cap (H_2 \setminus H)} \in \Pi.$$

Hence, $e_{H_1 \setminus H}, e_{H_2 \setminus H} \in \Pi$ and $e_H = e_{H_1} - e_{H_1 \setminus H} \in \Pi$. Thus, $e_H, e_{H_2}, e_{H_3}, \ldots, e_{H_l}$ or $e_{H_1 \setminus H}, e_{H_2}, e_{H_3}, \ldots, e_{H_l}$ is a new basis B_4 of Π, and $N(B_4) < N(B_2)$, i.e. contradiction to the assumption $H_1 \cap H_2 \ne \varnothing$ is achieved. □

When $t \to \infty$, eq. (16) comes to

$$\mu_1 X_1^* + \mu_2 X_2^* + \ldots . \mu_n X_n^* = \mu_1 X_1(0) + \mu_2 X_2(0) + \ldots + \mu_n X_n(0).$$

Let rank $\Pi = 1$. Applying theorem 1, we have

$$\sum_{j \in H_i} X_j^* = \sum_{j \in H_i} X_j(0), \quad i = 1, 2, \ldots, l. \tag{22}$$

It is clear that rank $\Gamma_1 = n - 1$. Let $\tau_1 \in H_1$, $\tau_2 \in H_2$, . . . , $\tau_l \in H_l$. After striking out the τ_1th, τ_2th, . . . , and τ_lth equations in system (11), we shall obtain (together with l equations (22)) $n - l + l = n$ linearly independent equations. The obtained system of linear equations can be easily solved. After substituting the obtained X_u^* into equation (12), we will obtain

$$\omega^* = \sum_{i=1}^{l} \sum_{j \in H_i} X_j(0) f_i(\ . \ . \ . \ k(u, v, w) \ . \ . \ . \ ), \tag{23}$$

where f_i depends only on the rate constants $k(u, v, w)$. After substituting all the rate constants according to eq. (8), one will obtain the necessary QSA relationships.

The case where $l = 1$ and $H_1 = \sigma_n$ is identical to the kinetic schemes of the cell-free enzyme catalyzed reactions which have been extensively investigated in the literature. A number of different algebraic and graphical methods for deriving ω^* for these simplified kinetic schemes have been proposed [17–25]. All these methods can be applied for deriving ω^* in the general case also.

3. Classification of all possible algebraic expressions for w^*

Let σ be a permutation of the set σ_n. If in expression (23) all $X_j(0)$ are replaced by $X_{\sigma(j)}(0)$ and all $k(u, v, w)$ are replaced by $k(\sigma(u), \min(\sigma(v), \sigma(w)), \max(\sigma(v), \sigma(w)))$, a number of new rate functions ω_σ^* will be obtained (here, $\sigma(0) = 0$, $\sigma(n + 1) = n + 1$). For the given kinetic scheme, this procedure will mean a simple renumeration of the species $X_1, X_2, \ldots, X_n$; therefore, we will say that the functions ω^* and ω_σ^* are equivalent. Obviously, any given class of such equivalent functions includes a function in the form of e. (23), where $H_1 = \{1, 2, \ldots, r_1\}$, $H_2 = \{r_1 + 1, r_1 + 2, \ldots, r_2\}, \ldots, H_i = \{r_{l-1} + 1, r_{l-1} + 2, \ldots, r_l\}$, $r_l \leq n$, $\operatorname{card} H_1 \geq \operatorname{card} H_2 \geq \ldots \geq \operatorname{card} H_l \geq 2$.

For example, for the kinetic scheme defined by the reactions

$$X_1 \rightarrow X_4, \quad X_1 \rightarrow 2X_4, \quad X_2 \rightarrow X_3, \quad X_3 \rightarrow X_2, \quad X_3 \rightarrow X_1 + X_2 \quad (n = 3),$$

the following rate equation is obtained:

$$\omega^* = [X_2(0) + X_3(0)] \frac{k(2,0,3)k(3,1,2)[k(1,0,4) + 2k(1,4,4)]}{[k(2,0,3) + k(3,1,2) + k(3,0,2)][k(1,0,4) + k(1,4,4)]}.$$

If we define $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 2$, then

$$\omega_\sigma^* = [X_1(0) + X_2(0)] \frac{k(1,0,2)k(2,1,3)[k(3,0,4) + 2k(3,4,4)]}{[k(1,0,2) + k(2,1,3) + k(2,0,1)][k(3,0,4) + k(3,4,4)]}, \tag{24}$$

and the function ω_σ^* corresponds to the kinetic scheme

$$X_3 \to X_4, \quad X_3 \to 2X_4, \quad X_1 \to X_2, \quad X_2 \to X_1, \quad X_2 \to X_1 + X_3.$$

If the algebraic expression for ω_0^* ($\omega_0^* \neq 0$) is obtained from ω^* by replacing some of the symbols $k(u, v, w)$ with zeros, we will say that ω_0^* is simpler than ω^*. For example, in expression (24) let $k(3, 4, 4) = 0$. Then

$$\omega_0^* = [X_1(0) + X_2(0)] \frac{k(1,0,2)k(2,1,3)}{k(1,0,2) + k(2,1,3) + k(2,0,1)}, \tag{25}$$

and expression (25) is simpler than (24).

Denote by $\mathrm{Cl}\,\omega^*$ the class of algebraic expressions which contain all the expressions equivalent to ω^* and all the expressions which are simpler than any one of them. For example, let $n = 3$ and

$$\omega^* = [X_1(0) + X_2(0)] \frac{k(1,0,2)k(2,1,3)[k(3,0,4) + 2k(3,4,4)]}{[k(1,0,2) + k(2,1,3)][k(3,0,4) + k(3,4,4)]} \tag{26}$$

corresponds to the scheme $X_1 \to X_2, X_2 \to X_1 + X_3, X \to X_4, X_3 \to 2X_4$. Then there are $6 = 3!$ expressions ω_σ^* equivalent to ω^*, where σ is a permutation of the set $\{1, 2, 3\}$. In accordance to each of these expressions, there are two simpler expressions.

Let Ω_n be a set of all possible algebraic expressions of the rate equation ω^* for the kinetic schemes with no more than n species.

THEOREM 2

There are $p(n) - 1$ algebraic expressions $\omega_1^*, \omega_2^*, \ldots, \omega_{p(n)-1}^*$ corresponding to the kinetic schemes with n species, such that

$$\Omega_n = \bigcup_{i=1}^{p(n)-1} \mathrm{Cl}\,\omega_i^*.$$

Here, $p(n)$ is the well-known function of "partitions" in number theory. A partition of n is a representation of n as a sum of any positive integral parts where the order of parts is irrelevant. Thus, $2 = 1 + 1$ has two partitions (i.e. $p(2) = 2$), $3 = 2 + 1 = 1 + 1 + 1$ has three partitions ($p(3) = 3$), $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$ has five partitions ($p(4) = 5$), $5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1$ has seven partitions ($p(5) = 7$), and so on. For example, $p(200) = 3\,972\,999\,029\,388$. For $n \to \infty$, the following asymptotic formula holds [26]:

$$p(n) \sim \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{2n/3}}.$$

Proof

Let ω^* be the algebraic expression corresponding to a given kinetic scheme with $\leq n$ species, and Π be the concentration polyhedron of this scheme. Assume

that $\Pi = \langle e_{H_1}, e_{H_2}, \ldots, e_{H_l} \rangle$ i.e. $e_{H_1}, e_{H_2}, \ldots, e_{H_l}$ is a basis of Π. Also, let $H_0 = \sigma_n \backslash \bigcup_{i=1}^{l} H_i$, where $H_1, H_2, \ldots, H_l$ are defined as above. Consider the "complete" kinetic scheme with the polyhedron Π which includes all possible reactions of the following forms:

$$X_i \to X_s, \quad \text{where } i, s \in H_1;\ i, s \in H_2;\ \ldots;\ i, s \in H_1;\ i \in H_0, s \in H_0 \cup \{n+1\}.$$

$$X_i \to 2X_s, \text{ where } i \in H_0, s \in H_0 \cup \{n+1\},$$

$$X_i \to X_j + X_s, \quad j \neq s,$$

$$\text{where } i, j \in H_1, s \in H_0 \cup \{n+1\};\ i, j \in H_2, s \in H_0 \cup \{n+1\};\ \ldots;$$
$$i, j \in H_l, s \in H_0 \cup \{n+1\};\ i \in H_0, j, s \in H_0 \cup \{n+1\}.$$

Then the following complete function corresponds to this complete scheme:

$$\omega^*(\Pi) = \sum_{i=1}^{l} \sum_{j \in H_i} X_j(0) f_i^c(\ \ldots\ k(u, v, w)\ \ldots\ ). \tag{27}$$

By eliminating some of the reactions in the complete scheme, we can obtain every scheme with the concentration polyhedron Π. Since the expression ω^* still has the form (23), it can be obtained from (27) by replacing the proper rate constants $k(u, v, w)$ by zeros. Since card $H_1 \geq$ card $H_2 \geq \ldots \geq$ card $H_l \geq 2$ and $\sum_{i=1}^{l}$ card $H_i \leq n$, the number of different concentration polyhedrons equals $q(2) + q(3) + \ldots + q(n)$. Here, $q(m)$ is the number of representations of m as a sum of any integers from the set $\{2, 3, 4, \ldots\}$. To prove the theorem, it remains to show that

$$q(2) + q(3) + \ldots + q(n) = p(n) - 1. \tag{28}$$

Define $p(0) = 1$, $q(0) = 1$, $q(0) = 0$. It is well known that

$$\prod_{u=1}^{\infty} \frac{1}{1 - x^u} = \sum_{n=0}^{\infty} p(n) x^n$$

(e.g. see ref. [26]). Analogously,

$$\prod_{u=2}^{\infty} \frac{1}{1 - x^u} = \prod_{u=2}^{\infty} \sum_{m_u=0}^{\infty} x^{um_u} = \sum_{v=0}^{\infty} q(v) x^v.$$

Hence,

$$\sum_{n=0}^{\infty} p(n) x^n = \prod_{u=1}^{\infty} \frac{1}{1 - x^u} = \frac{1}{1 - x} \prod_{u=2}^{\infty} \frac{1}{1 - x^u} = \sum_{w=0}^{\infty} x^w \sum_{v=0}^{\infty} q(v) x^v$$

$$= \sum_{n=0}^{\infty} x^n \sum_{\substack{v+w=n \\ v, w \geq 0}} q(v) = \sum_{n=0}^{\infty} [q(0) + q(1) + q(2) + \ldots + q(n)] x^n.$$

Consequently, $p(n) = q(0) + q(1) + q(2) + \ldots + q(n) = 1 + q(2) + \ldots + q(n)$, and the equality (28) is proved. □

4. Examples of the simplest rate equations and nonlinear quantitative structure–activity relationships

According to theorem 2, if one assumes that the drug action is modeled by the kinetic scheme which includes no more than two kinetically important compounds (species), then there exists one general kinetic equation which will describe all the particular mechanisms of drug action. Indeed, according to the complete scheme for $n = 2$,

$$X_1 \xrightarrow{k_1} X_2,\ X_2 \xrightarrow{k_2} X_1,\ X_1 \xrightarrow{k_3} X_2 + X_3,\ X_2 \xrightarrow{k_4} X_1 + X_3,$$

where $k(1, 0, 2) = k_1$, $k(2, 0, 1) = k_2$, $k(1, 2, 3) = k_3$, $k(2, 1, 3) = k_4$, the rate equation is obtained:

$$\omega^* = [X_1(0) + X_2(0)]\,\frac{k_2k_3 + k_1k_4 + 2k_3k_4}{k_1 + k_2 + k_3 + k_4}. \tag{29}$$

Analogously, there exist only two general kinetic equations which will describe all possible kinetic schemes with $n \le 3$.

According to the first complete scheme

$$X_1 \xrightarrow{k_1} X_2,\ X_2 \xrightarrow{k_2} X_1,\ X_1 \xrightarrow{k_3} X_2 + X_3,\ X_1 \xrightarrow{k_4} X_2 + X_4,$$

$$X_2 \xrightarrow{k_5} X_1 + X_3,\ X_2 \xrightarrow{k_6} X_1 + X_4,\ X_3 \xrightarrow{k_7} X_4,\ X_3 \xrightarrow{k_8} 2X_4,$$

where $k(1, 0, 2) = k_1$, $k(2, 0, 1) = k_2$, $k(1, 2, 3) = k_3$, $k(1, 2, 4) = k_4$, $k(2, 1, 3) = k_5$, $k(2, 1, 4) = k_6$, $k(3, 0, 4) = k_7$, $k(3, 4, 4) = k_8$, the rate equation is

$$\omega^* = [X_1(0) + X_2(0)]$$

$$\times \frac{(k_2k_4 + k_4k_5 + 2k_4k_6 + k_1k_6 + k_3k_6)(k_7 + k_8) + (k_2k_3 + 2k_3k_5 + k_3k_6 + k_1k_5 + k_4k_5)(k_7 + 2k_8)}{(k_1 + k_2 + k_3 + k_4 + k_5 + k_6)(k_7 + k_8)}.$$

According to the second complete scheme

$$X_1 \xrightarrow{k_1} X_2,\ X_1 \xrightarrow{k_2} X_3,\ X_2 \xrightarrow{k_3} X_1,\ X_2,\ X_1 \xrightarrow{k_4} X_3,\ X_3 \xrightarrow{k_5} X_1,$$

$$X_3 \xrightarrow{k_6} X_2,\ X_1 \xrightarrow{k_7} X_2 + X_4,\ X_1 \xrightarrow{k_8} X_3 + X_4,\ X_2 \xrightarrow{k_9} X_1 + X_4,$$

$$X_2 \xrightarrow{k_{10}} X_3 + X_4,\ X_3 \xrightarrow{k_{11}} X_1 + X_4,\ X_3 \xrightarrow{k_{12}} X_2 + X_4,$$

where $k(1, 0, 2) = k_1$, $k(1, 0, 3) = k_2$, $k(2, 0, 1) = k_3$, $k(2, 0, 3) = k_4$, $k(3, 0, 1) = k_5$, $k(3, 0, 2) = k_6$, $k(1, 2, 4) = k_7$, $k(1, 3, 4) = k_8$, $k(2, 1, 4) = k_9$, $k(2, 3, 4) = k_{10}$, $k(3, 1, 4) = k_{11}$, $k(3, 2, 4) = k_{12}$, the rate equation is

$$\omega^* = [X_1(0) + X_2(0) + X_3(0)] \times \frac{\begin{vmatrix} -(k_1 + k_2 + k_7 + k_8) & k_3 + k_9 & k_5 + k_{11} \\ k_1 + k_7 & -(k_3 + k_4 + k_9 + k_{10}) & k_6 + k_{12} \\ k_7 + k_8 & k_9 + k_{10} & k_{11} + k_{12} \end{vmatrix}}{\begin{vmatrix} -(k_1 + k_2 + k_7 + k_8) & k_3 + k_9 & k_5 + k_{11} \\ k_1 + k_7 & -(k_3 + k_4 + k_9 + k_{10}) & k_6 + k_{12} \\ 1 & 1 & 1 \end{vmatrix}}.$$

After substituting all the rate constants to $\exp\{(a_0(u, v, w) + \sum_{i=1}^{k} a_i(u, v, w)x_i)\ln 10\}$ into the rate equations corresponding to the complete kinetic schemes, we will obtain the most general forms of the possible QSA relationships:

$$\omega^* = \sum_{i=1}^{l} \sum_{j \in H_i} X_j(0) f_i^c \left(\ldots \exp\{(a_0(u, v, w) + \sum_{i=1}^{k} a_i(u, v, w)x_i) \ln 10\} \ldots \right). \quad (30)$$

Here, $\sum_{j \in H_i} X_j(0)$ and all a_i, $i = 0, \ldots, k$, represent a set of nonlinear regression coefficients, and all x_i, $i = 1, \ldots, k$, represent a set of variables parameters.

Note that the drug biological potency should normally be measured by the net rate of biological responses ω^* caused by the constant drug initial concentration in the administration phase $X_D(0)$, rather than by the variable drug concentration initiating the constant rate of biological response, i.e. ED_{50}, LD_{50}, MIC, MBC, etc. The point is that if one wishes to vary the initial substrate concentration $X_D(0)$, then one must take into account that it may be included in any of the first-order rate constants. Hence, the necessary QSA relationships in the form of explicit algebraic expressions for the dose–response correlations will be obtained after the following transformations of eq. (30):

(i) substitute the variable ω^* for the constant representing standard biological response;

(ii) substitute the proper coefficient(s) $a_0(u, v, w)$ for the $X_D(0)a_0'(u, v, w)$, where $a_0'(u, v, w)$ is the drug concentration-independent coefficient;

(iii) solve the obtained algebraic equation with respect to the variable $X_D(0)$.

After accounting for all possible changes in (ii), we will obtain many more types of possible QSARs than follows from the classification according to theorem 2. In addition, many of the obtained equations will be unsolvable in the explicit form with respect to $X_D(0)$.

It is easy to show that the most widely used bilinear equation represents the particular case of the general QSA relationship which is in accordance with the kinetic scheme with $n = 2$. Indeed, assuming that in expression (29) $k_2 = k_3 = 0$, $k_1 = 10^{a_1+a_2x}$, and $k_4 = 10^{a_3+a_4x}$, we obtain

$$\log(1/\omega^*) = \log\left[\frac{1}{X_1(0) + X_2(0)}\left(\frac{1}{k_1} + \frac{1}{k_4}\right)\right]$$

$$= -\log[X_1(0) + X_2(0)] + \log\left(10^{-a_1-a_2x} + 10^{-a_3-a_4x}\right)$$

$$= -\log[X_1(0) + X_2(0)] - a_1 - a_2x + \log\left(1 + 10^{a_1-a_3+(a_2-a_4)x}\right)$$

$$= C_0 + C_1x + \log\left(1 + 10^{C_2+C_3x}\right), \tag{31}$$

where $C_0 = -\log[X_1(0) + X_2(0)] - a_1$, $C_1 = -a_2$, $C_2 = a_1 - a_3$, $C_3 = a_2 - a_4$. Expression (31) was shown to be identical to the bilinear equation (2) from the viewpoint of practical use [10].

Assuming that in expression (29) $k_4 = 0$, $k_1 = 10^{a_1+a_2x}$, $k_2 = 10^{a_5+a_6x}$, and $k_3 = 10^{a_7+a_8x}$, we obtain

$$\log(1/\omega^*) = \log\left[\frac{1}{X_1(0) + X_2(0)}\left(\frac{k_1}{k_2k_3} + \frac{1}{k_2} + \frac{1}{k_3}\right)\right]$$

$$= -\log[X_1(0) + X_2(0)] + \log\left(10^{a_1-a_5-a_7+(a_2-a_6-a_8)x} + 10^{-a_5-a_6x} + 10^{-a_7-a_8x}\right)$$

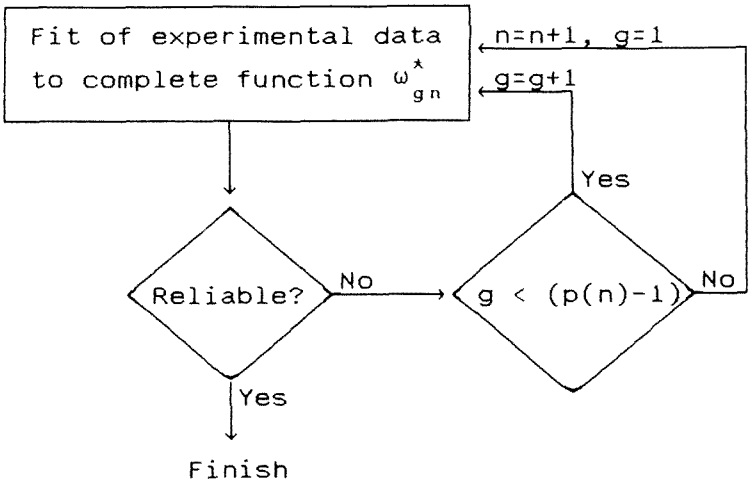
$$= C_0 + C_1x + \log\left(1 + 10^{C_2+C_3x} + 10^{C_4+C_5x}\right), \tag{32}$$

where $C_0 = -\log[X_1(0) + X_2(0)] + a_1 - a_5 - a_7$, $C_1 = a_2 - a_6 - a_8$, $C_2 = -a_1 + a_7$, $C_3 = -a_2 + a_8$, $C_4 = -a_1 + a_5$, $C_5 = -a_2 + a_6$. Under the proper values of $a_1, \ldots, a_8$, expression (32) will describe the structure–activity plots with three linear or curvilinear ascending and descending parts. Analogously, it can be readily shown that the proper kinetic schemes with n elementary reactions will give all the other practically observed n-portion structure–activity dependencies described in the introduction.

5. Discussion and conclusions

How can one establish the minimal kinetic scheme and, hence, the simplest QSA relationship which will certainly describe the experimental data on biological activities? This problem cannot be solved by fitting the experimental data into a single structure–activity equation, since a priori drug action mechanism is never known. An ideal algorithm for establishing the necessary QSA relationships should

imply sequential fitting of experimental data to all possible SA equations with subsequent comparison of the obtained statistical results. Since most of the so far correlated biological data were satisfactorily fitted to the empirical "n-portion"-type functions (see section 1), one may hope that in most cases the drug action can be modeled by the rather simple kinetic schemes. Therefore, according to theorem 2, there can be only a limited number of probable structure–activity functions. Assuming that the kinetic regularities of drug action accomplish our assumptions, one could apply the logical algorithm described in scheme 1. Here, ω^*_{ng} denotes the function (30) which is in accordance with the gth complete kinetic scheme with n



Scheme 1. The principal scheme for establishing QSA relationships.

internal species $X_1, X_2, \ldots, X_n$. Since our aim is to find the "minimal" function describing the biological data within the experimental precision limits, it should be reasonable to start analysis from the simplest complete schemes, i.e. with $n = 2$. For any fixed number n, one should have to analyze $p(n) - 1$ different complete functions. The currently fitted function should be considered to be statistically reliable if for any ith data point the following inequality is satisfied:

$$\frac{|\omega_i^* - \omega_{ngi}^*|}{\omega_i^*} \leq \varepsilon.$$

Here, ω_i^* and ω_{ngi}^* represent, respectively, the experimentally obtained and predicted by the ngth function activities of the ith compound, and ε is the experimental error.

We believe that such a logical algorithm could serve as a basis for the systematic analysis of biological data, providing a new level for the understanding

of drug action. However, its real application in practice will be constrained by a number of fundamental problems which, up to the present, remain unsolved.

First, the problem of deriving the complete set of substrate predictor variables still needs to be solved. All the presently used substrate parameter sets proved to be valid only in the isotropic homogeneous phases, while in the active sites of biological receptors, various local interactions may dominate. This should mean that the correct parameters $x_1, \ldots, x_k$ should be dependent on 3D space coordinates. Consequently, the complete set of substrate predictor variables may appear to be very large. Therefore, only compounds with slightly varied structures could be included in the correlation set.

Second, the problem of collinearity of predictor variables, i.e. the existence of relationships

$$x_i = F_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k),$$

will necessarily arise.

In order to solve these problems and to obtain statistically reasonable results, one should operate with a sufficiently large body of experimental data. In addition, the required QSAR model which will describe the biological data within the experimental precision could be achieved only after consecutive fitting of experimental data to a number of complex nonlinear functions. All this will take a large amount of computer calculation time.

Of course, at present the above-mentioned difficulties can hardly be fully overcome; however, if the pharmaceutical chemists dream of solving the QSAR problem based on analyses of drug action mechanisms at the molecular level, they will have to do this.

References

[1] M.J. Kamlet, R.M. Doherty, P.W. Carr, D. Mackay, M.H. Abraham and R.W. Taft, Environ. Sci. Tech. 22(1988)503.
[2] M.J. Kamlet, R.M. Doherty, G.R. Famini and R.W. Taft, Acta Chem. Scand. B41(1987)589.
[3] M.H. Abraham, P.P. Duce, D.V. Prior, D.G. Barratt, J.J. Morris and P.J. Taylor, J. Chem. Soc. Perkin Trans. II (1989)1355.
[4] C. Hansch and T. Fujita, J. Am. Chem. Soc. 86(1964)1615.
[5] C. Hansch, Accounts Chem. Res. 2(1969)232.
[6] C. Hansch, in: *Drug Design*, Vol. 1, ed. E.J. Ariëns (Academic Press, New York, 1971) p. 271.
[7] R.M. Hyde, Chem. Ind. (1977) 859.
[8] J.W. McFarland, J. Med. Chem. 13(1970)1192.
[9] H. Kubinyi, Arzneim. Forsch. 26(1976)1991.
[10] H. Kubinyi, *Progress in Drug Research*, Vol. 23, ed. E. Jucker (Birkhäuser, Basel, 1979) p. 97.
[11] L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).
[12] L.H. Hall and L.B. Kier, Eur. J. Med. Chem. 13(1978)89.
[13] A.T. Balaban, A. Chiriac, I. Motoc and Z. Simon, *Steric Fit in Quantitative Structure–Activity Relationships* (Springer, Berlin, 1980).

[14] D.B. Boyd, in: *Chemistry and Biology of β-Lactam Antibiotics*, Vol. 1, ed. R.B. Morin and M. Gorman (Academic Press, New York, 1982) p. 437.
[15] R. Franke and R. Kühne, Eur. J. Med. Chem. 13(1978)399.
[16] C. Hansch and J.M. Clayton, J. Pharm. Sci. 62(1973)1.
[17] E.L. King and C. Altman, J. Phys. Chem. 60(1956)1375.
[18] K.-C. Chou and S. Forsen, Biochem. J. 187(1980)829.
[19] K.-C. Chou, J. Theor. Biol. 89(1981)581.
[20] K.-C. Chou, Biophys. Chem. 35(1990)1.
[21] E.P. Whitehead, J. Theor. Biol. 80(1979)355.
[22] G.P. Zhou and M.-H. Deng, Biochem. J. 222(1984)169.
[23] K.J. Indge and R.E. Childs, Biochem. J. 155(1976)567.
[24] A. Cornish-Bowden, Biochem. J. 165(1977)55.
[25] B.N. Goldstein, *Kinetic Graphs in Enzymology* (Nauka, Moscow, 1989).
[26] G.E. Andrews, *The Theory of Partitions* (Addison–Wesley, Reading, MA, 1976).