replaced by normal equations with fewer unknowns.

Kubinyi and Kehrhahn give a number of more typical examples, and in particular their Table VI (ref 10, p 1045) shows results parallel to our Table III.

References and Notes

(1) These include at least one of our own, and we are grateful to Drs. H. Kubinyi and O. T. Kehrhahn of Knoll, AG, Ludwigshafen, for pointing out that the Free–Wilson conditions are eq 41 and 42 of the present paper rather than eq 4 of ref 2.
(2) L. J. Schaad, R. H. Werner, L. Dillon, L. Field, and C. E. Tate, *J. Med. Chem.*, 18, 344 (1975).
(3) F. B. Hildebrand, "Methods of Applied Mathematics", Prentice-Hall, New York, N.Y., 1952, Chapter 1.
(4) G. E. Shilov, "Linear Algebra", Prentice-Hall, Englewood Cliffs, N.J., 1971.
(5) C. Hansch, *Acc. Chem. Res.*, 2, 232 (1969), and earlier references cited in this review.
(6) R. F. Gould, Ed., "Biological Correlations–The Hansch Approach", American Chemical Society, Washington D.C., 1972.
(7) S. M. Free, Jr., and J. W. Wilson, *J. Med. Chem.*, 7, 395 (1964).
(8) W. P. Purcell, G. E. Bass, and J. M. Clayton, "Strategy of Drug Design", Wiley, New York, N.Y., 1973.
(9) P. N. Craig, Chapter 8 in ref 6.
(10) H. Kubinyi and O. H. Kehrhahn, *J. Med. Chem.*, 19, 578, 1040 (1976).
(11) H. Kubinyi, *J. Med. Chem.*, 19, 587 (1976).
(12) R. Courant, "Differential and Integral Calculus", Vol. II, Interscience, New York, N.Y., 1936, p 27.
(13) T. Fujita and T. Ban, *J. Med. Chem.*, 14, 148 (1971).
(14) A. Cammarata and S. J. Yau, *J. Med. Chem.*, 13, 93 (1970).

Quantitative Structure–Activity Relationships. 7.[1] The Bilinear Model, a New Model for Nonlinear Dependence of Biological Activity on Hydrophobic Character

Hugo Kubinyi

Research Institute of Knoll AG, D 6700 Ludwigshafen/Rhein, Federal Republic of Germany.

Received October 5, 1976

The bilinear model, $\log 1/C = a \log P - b \log (\beta P + 1) + c$, a new model for nonlinear dependence of biological activity on hydrophobic character, is applied to 57 data sets of biological activity values in homologous series. From a comparison of the statistical parameters and the residuals obtained with the bilinear model and the parabolic model, the superiority of the bilinear model for a precise quantitative description of both linear and nonlinear parts of structure–activity relationships can be derived; the bilinear model explains the particular effect that in homologous series the relationship between biological activity and hydrophobic character is strictly linear for the lower members, while for higher members this relationship is nonlinear.

The biological response elicited by a drug is determined by its intrinsic activity and by its ability to reach a definite receptor site. While the intrinsic activity of a drug molecule depends on various physicochemical properties and the geometry of the molecule, drug permeation is considered to be—in most cases—a passive transport process which is influenced only by the lipophilicity of the molecule. Thus, in homologous series, where the intrinsic activity of all members can be assumed to be identical biological activity should be a simple function of lipophilicity.

Indeed linear relationships between biological activity and hydrophobic character (eq 1)[2-8] are obtained for a large number of homologous series. However, this linear relationship cannot go on infinitely, otherwise compounds with infinite activity would exist. A cut-off point[9] is reached in each homologous series: biological activity increases with increasing lipophilicity, reaches a maximum, and then decreases with further increase of hydrophobic character.

$$\log 1/C = a \log P + b \quad (1)$$

In 1964 Hansch[5-7] proposed a parabolic model (eq 2) for the dependence of biological activity on hydrophobic character on the basis of a "random walk" process; on the way from the outer phase, where the drug is applicated, to their receptor sites the drug molecules have to penetrate a number of lipophilic and hydrophilic barriers. While hydrophilic molecules tend to remain in the aqueous phases and lipophilic molecules tend to go into the lipid (membrane) phases, molecules with an optimal hydrophilic–lipophilic balance will have the best chance to penetrate all barriers and to reach the receptor sites.

$$\log 1/C = a(\log P)^2 + b \log P + c \quad (2)$$

Although the parabolic model has been supported by consideration of a kinetic model[7,10,11] and its suitability for practical purposes has been proven with some hundred examples, there remains a discrepancy between the linear model and the parabolic model; from the linear model at least the left side of the "parabola" should be strictly linear, while from the parabolic model both sides of the structure–activity relationship should be more or less curved.

Besides the parabolic model several other models for nonlinear dependence of biological activity on hydrophobic character have been presented in the last years.[12-16] Among these models the most interesting seems to be the Mc Farland model:[12] Mc Farland used a simple hypothetical system, made up of alternating aqueous phases and lipid (membrane) phases of equal volumes and considered the probability of a drug molecule to cross a definite number of barriers. The Mc Farland model can be represented by eq 3[17] (corresponds to eq 15 of ref 12). Symmetrical curves with linear ascending and descending sides and a parabolic part within the range of $\log P = 0$ are obtained from eq 3. Although Mc Farland recognized systematic deviations between his model and the parabolic model (the same deviations can be recognized between the computer plot from the kinetic model[7,10,11] and the parabola fitted to this plot), he considered the parabolic model to be sufficient for all practical purposes.

$$\log 1/C = a \log P - 2a \log (P + 1) + c \quad (3)$$

To adapt the Mc Farland model to biological reality, the model was reconsidered using a slightly modified system (Figure 1).[17] Only four relevant phases are regarded, e.g., as a model of a simple bacterial cell or an isolated tissue

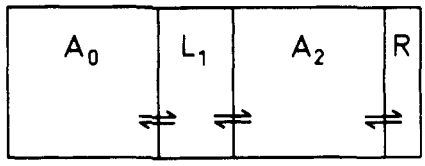

Figure 1. Hypothetical biological system used for the derivation of the bilinear model:[17] A_0 = outer aqueous phase, L_1 = lipid (membrane) phase, A_2 = inner aqueous phase, R = receptor "phase".

cell: an outer aqueous phase A_0, a lipid (membrane) phase L_1, an inner aqueous phase A_2, and a receptor "phase" R; the different volumes of the aqueous phases, the lipid phase, and the receptor "phase" are indicated in Figure 1.

Following the probability approach of Mc Farland, eq 4 was derived[17] from this modified system (the assumptions and theoretical considerations leading to the model are discussed in ref 17); in eq 4 β is the volume ratio of the

$$\log 1/C = a \log P - b \log (\beta P + 1) + c \quad (4)$$

lipid phase and the aqueous phases, V_l/V_a. If the equilibrium approach of Higuchi and Davis[13] is followed, a very similar equation (eq 5)[17] can be derived from the system

$$\log 1/C = a \log P - \log (\beta P + 1) + c \quad (5)$$

presented in Figure 1 (the Higuchi–Davis model was extended recently to a generalized model for ionizable substances by Martin and Hackbarth;[18] however, no applications of the model-derived equations to practical examples were given). It should be noted that eq 3 ($b = 2a$; $\beta = 1$) and eq 5 ($b = 1$) are special forms of eq 4.

Unsymmetrical curves with linear ascending and descending sides and a parabolic part within the range of optimal lipophilicity are obtained from eq 4; unlike the Mc Farland model, from which only symmetrical curves with $\log P_0 = 0$ result, eq 4 is appropriate for a quantitative description of structure–activity relationships in all practical cases. Because of the characteristic curves resulting from eq 4 the name "bilinear model" was proposed.

Equation 4 is a four-parameter equation, the term β being a nonlinear term, which must be determined by stepwise iteration or by a Taylor series iteration method; the application of both calculation procedures to the bilinear model has been described in detail.[17b]

For a practical comparison of the bilinear model and the parabolic model, 57 data sets of biological activity values in homologous series were selected (Table I); only sets having six or more data points are included in Table I. With a few exceptions, the variance of log P values for each set is five or more log P units. Structurally heterogeneous sets and homologous series with a log P variance of less than 3.5 units were excluded. All equations were calculated with a Taylor series iteration program.

Results

The equations resulting from the bilinear model for the data sets of Table I are given in Table II together with the 95% confidence intervals of the linear terms a, b, and c and log P_0 values calculated from the bilinear model and the parabolic model. β is given in its logarithmic form because the values of β vary over several powers of ten. Not too much importance should be attributed to the confidence intervals of the linear terms a, b, and c; in several cases the confidence intervals are extremely wide (e.g., eq 14, Table II), although the overall regression is statistically significant; however, such wide confidence intervals are well known in nonlinear regression analysis[25]

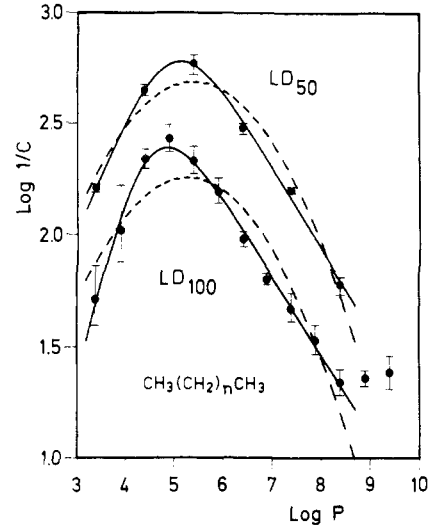

Figure 2. LD_{50} and LD_{100} values of alkanes in mice (Table I, eq 6 and 7) fitted with the bilinear model (—) and the parabolic model (– – –);[17] confidence limits are indicated for each value.

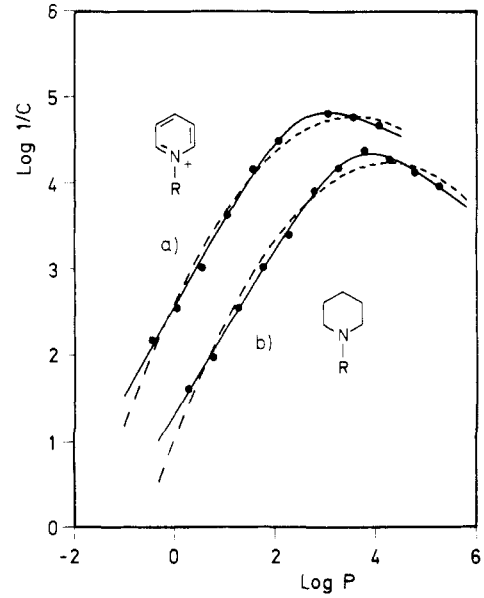

Figure 3. Hemolytic activity of (a) *N*-alkylpyridinium compounds (Table I, eq 57) and (b) *N*-alkylpiperidines (Table I, eq 35) fitted with the bilinear model (—) and the parabolic model (– – –).

in those cases where the data do not accurately conform to the used model. In the case of eq 9 and 10 the wide confidence intervals of the b term are caused by the fact that the value of b is determined by only one data point in each equation. No meaningful confidence intervals could be derived for the nonlinear term β from the Taylor series iteration method[17b].

The statistical parameters obtained for each data set with the bilinear model and the parabolic model are given in Table I together with a partial F test (see below). Figures 2–5 illustrate the differences between the bilinear model and the parabolic model and demonstrate the goodness of fit obtained with each model.

As compared with the parabolic model the bilinear model contains an additional parameter; therefore, it is not surprising that higher correlation coefficients r are obtained with the bilinear model in most cases. However, the standard deviations s account for this extra parameter; if the s values from both models are compared, lower standard deviations are obtained in 52 out of 57 examples with the bilinear model. To check the statistical signif-

Table I. Comparison of the Bilinear Model and the Parabolic Model: Data Sets and Statistical Parameters

Eq	Compounds	Biological activity	n	Bilinear model r	Bilinear model s	Parabolic model r	Parabolic model s	Partial F test[a]	Ref (eq)
6	Alkanes	Mice, iv LD_{50}	6	0.998	0.038	0.976	0.101	19.38*	17, 19
7	Alkanes	Mice, iv LD_{100}	11	0.996	0.039	0.930	0.148	111.41*	17, 19
8	Alkanes	Mice, iv AD_{100}	12	0.978	0.101	0.909	0.189	23.79*	19
9	Alcohols	Frog ventricle, narcotic action	10	0.998	0.133	0.978	0.458	76.66*	20 (25, 26)
10	Alcohols	Tadpoles, narcotic action	10	1.000	0.068	0.995	0.215	63.30*	20 (29, 30)
11	Ethers	Mice, iv LD_{100}	7	0.995	0.066	0.925	0.215	39.84*	19
12	Ethers	Mice, iv AD_{100}	7	0.999	0.022	0.955	0.152	181.33*	19
13	Ketones	Mice, iv LD_{100}	6	0.980[c]	0.188	0.962	0.213	1.85	19
14	Ketones	Mice, iv AD_{100}	7	0.969	0.209	0.908	0.305	5.48	19
15	$RCOO^-$	*P. omnivorum*, MIC (100%)	14[b]	0.959	0.191	0.844	0.346	26.13*	21 (48)
16	$RCOO^-$	Red cell dove, MHC	8	0.992	0.147	0.979	0.213	6.49	22 (18)
17	$RCHBrCOO^-$	Red cell sheep, $C_{H_{100}}$	8	0.941	0.362	0.891	0.436	3.26	22 (19)
18	$RCHBrCOO^-$	*V. cholerae*, MKC (pH 7.5)	7	0.985	0.185	0.958	0.264	5.15	11 (23)
19	$RCHBrCOO^-$	*V. cholerae*, MKC (pH 8.5)	6	0.980[c]	0.180	0.959	0.206	1.92	11 (29)
20	$RCHBrCOO^-$	*V. cholerae*, MKC (pH 6.0)	6	0.986	0.187	0.971	0.219	2.10	11 (46)
21	$RCHBrCOO^-$	*B. lepisepticus*, MKC (pH 6.0)	6	0.981[c]	0.235	0.985	0.171		11 (47)
22	$RCHBrCOO^-$	*B. lepisepticus*, MKC (pH 7.5)	6	0.991	0.158	0.959	0.277	7.20	11 (49)
23	$RCHBrCOO^-$	*D. pneumoniae*, MKC (pH 6.5)	8	0.939	0.519	0.898	0.593	2.54	11 (53)
24	$RCHBrCOO^-$	*S. hemolyticus*, MKC (pH 6.0)	8	0.959	0.444	0.898	0.620	5.72	11 (54)
25	$RCHBrCOO^-$	*D. pneumoniae*, MKC (pH 7.5)	8	0.934	0.530	0.893	0.596	2.33	11 (82)
26	$RCHBrCOO^-$	*D. pneumoniae*, MKC (pH 8.5)	8	0.942	0.488	0.909	0.542	2.16	11 (84)
27	$RCHBrCOO^-$	*S. hemolyticus*, MKC (pH 7.5)	8	0.942	0.523	0.872	0.685	4.58	11 (87)
28	$RCHBrCOO^-$	*S. hemolyticus*, MKC (pH 8.5)	8	0.943	0.522	0.872	0.686	4.61	11 (88)
29	$RCHOHCOO^-$	*B. lepisepticus*, MKC (pH 7.5)	7	0.995	0.106	0.963	0.245	18.31*	11 (16)
30	$RCHOHCOO^-$	*B. lepisepticus*, MKC (pH 8.5)	6	0.993	0.191	0.997	0.108		11 (21)
31	$RCHOHCOO^-$	*D. pneumoniae*, MKC (pH 7.5)	7	0.996	0.176	0.987	0.259	5.68	11 (74)
32	$ROSO_3^-Na^+$	Red cell sheep, $C_{H_{50}}$	6	0.999	0.069	0.995	0.145	11.30	22 (27)
33	α-Monoglycerides	Red cell dove, MHC	7	0.999	0.057	0.988	0.150	24.64*	22 (26)
34	α-Monoglycerides	Red cell dove, MHC	8	0.998	0.049	0.989	0.105	18.95*	11 (95)
35	*N*-Alkylpiperidines	Red cell dove, MHC	11	0.999	0.050	0.992	0.142	56.88*	22 (23)
36	Alkylguanidines	Rat liver mitochondria, I_{50}	10[b]	0.997	0.103	0.989	0.180	15.44*	11 (11)
37	Alkylguanidines	*M. fructicola*, ED_{50}	6	0.999	0.017	0.994	0.030	6.61	11 (17)
38	Alkylguanidines	*S. pastorianus*, ED_{50}	6	0.996	0.051	0.990	0.065	2.84	11 (33)
39	Diamidines	Vs. *S. typhosa*	8	0.995	0.119	0.989	0.152	4.20	23 (13)
40	Diamidines	Vs. *S. aureus*	8	0.999	0.041	0.997	0.073	11.75*	23 (37)
41	Diguanidines	Vs. *S. typhosa*	8	0.994	0.202	0.995	0.156		23 (14)
42	Diguanidines	Vs. *S. aureus*	8	0.993	0.193	0.979	0.296	7.82*	23 (38)
43	Diguanidines	Vs. *Strep. viridans*	8	0.990	0.286	0.984	0.314	2.01	23 (49)
44	$RN^+(CH_3)_3$	*S. typhosa*, MKC	7	0.997	0.169	0.986	0.305	10.06	11 (40)
45	$RN^+(CH_3)_3$	*S. aureus*, MKC	6	0.999	0.107	0.996	0.161	4.72	11 (83)
46	$C_6H_5CH_2N^+(R)(CH_3)_2$	*P. aeruginosa*, MIC	9	0.937	0.168	0.880	0.208	4.19	11 (19)
47	$C_6H_5CH_2N^+(R)(CH_3)_2$	*S. typhosa*, MIC	10	0.975	0.140	0.949	0.183	5.91	11 (20)
48	$C_6H_5CH_2N^+(R)(CH_3)_2$	*P. vulgaris*, MIC	10	0.974	0.132	0.974	0.123	0.12	11 (30)
49	$C_6H_5CH_2N^+(R)(CH_3)_2$	*P. vulgaris*, MKC	10	0.973	0.153	0.960	0.172	2.87	11 (36)
50	$C_6H_5CH_2N^+(R)(CH_3)_2$	*S. aureus*, MIC	12	0.993	0.100	0.979	0.158	14.62*	11 (42)
51	$C_6H_5CH_2N^+(R)(CH_3)_2$	*Cl. welchii*, MIC	12	0.983	0.172	0.966	0.230	8.14*	11 (43)
52	$C_6H_5CH_2N^+(R)(CH_3)_2$	*P. aeruginosa*, MKC	12	0.990	0.122	0.962	0.219	21.07*	11 (45)
53	$C_6H_5CH_2N^+(R)(CH_3)_2$	*Cl. welchii*, MKC	12	0.992	0.125	0.980	0.190	12.60*	11 (48)
54	$C_6H_5CH_2N^+(R)(CH_3)_2$	Red cell sheep, $C_{H_{50}}$	6	0.996	0.171	0.992	0.213	2.67	11 (55)
55	$C_6H_5CH_2N^+(R)(CH_3)_2$	*C. albicans*, MKC	11	0.984	0.170	0.961	0.243	9.31*	11 (58)
56	$C_6H_5CH_2N^+(R)(CH_3)_2$	Red cell sheep, $C_{H_{50}}$	6	0.996	0.154	0.998	0.085		11 (76)
57	*N*-Alkylpyridinium bromides	Red cell dove, MHC	9	0.998	0.071	0.993	0.137	17.46*	22 (21)
58	*N*-Alkylnicethamide chlorides	Vs. *S. aureus*	20	0.970	0.265	0.961	0.292	4.60*	23 (43)
59	5-Alkyl-8-hydroxyquinolines	*A. niger*, Inh (RBR)	11	0.977	0.082	0.955	0.106	6.55*	11 (80)
60	5-Alkyl-8-hydroxyquinolines	*A. niger*, Inh (RBR)	9	0.989	0.073	0.986	0.076	1.44	11 (103)
61	4-R-Lincomycins	*S. lutea* (RBR)	8	0.998	0.063	0.992	0.105	9.81*	11 (50)
62	Phorbol 12,13-diesters	Irritant activity, ED_{50}	6	1.000	0.041	0.978	0.320	182.67*	17 (44, 45)

[a] An asterisk indicates a statistically significant value ($p < 0.05$). [b] Isopropyl compound not included. [c] $p > 0.05$; all other regressions are statistically significant ($p < 0.05$).

icance of these differences between both models a partial F test[17b] was applied (Table I): in 25 out of 53 cases, where the bilinear model gives a better fit, this difference is statistically significant ($p < 0.05$). It is no argument against the bilinear model that the difference between both models is statistically significant in only 25 cases; in most examples the number of data points is too small to achieve statistical significance. If only examples with ten or more data points are considered, the bilinear model gives a statistically significant better fit in 14 out of 17 cases (more than 80%), while the four examples, where the parabolic model gives a better fit (eq 21, 30, 41, 54; Table I) consist only of six to eight data points. In cases where only such a few data points are present, fortuitous deviations may lead either to a better fit of the bilinear model or to a better fit of the parabolic model, due to the similarity of both models.

If the residuals resulting from the bilinear model and from the parabolic model are compared,[17b,26,27] the differences between both models can be seen at a glance (Figure 6). While the residuals from the bilinear model are without any detectable regularity, the residuals resulting from the parabolic model are not only significantly greater but also systematically distributed. The curves are

Table II. Bilinear Model. Log $1/C = a \log P - b \log (\beta P + 1) + c$. Results and Log P_0 Values

Eq	a[a]	b[a]	c[a]	Log β	Log P_0, calcd from: Bilinear model	Log P_0, calcd from: Parabolic model
6	0.524 (±0.36)	0.886 (±0.33)	0.443 (±1.30)	−4.956[b]	5.12[b]	5.38[b]
7	0.956 (±0.35)	1.306 (±0.34)	−1.498 (±1.21)	−4.412[b]	4.85[b]	5.26[b]
8	0.799 (±0.69)	1.152 (±0.66)	−0.794 (±2.41)	−4.577[b]	4.93[b]	5.23[b]
9	1.114 (±0.09)	4.344 (±4.38)	−0.883 (±0.24)	−6.433[c]	5.97[c]	6.73[c]
10	1.192 (±0.05)	6.131 (±19.14)	−0.797 (±0.16)	−6.229[c]	5.61[c]	8.82[c]
11	1.095 (±0.71)	1.384 (±0.68)	0.859 (±0.60)	−2.037	2.62	3.59
12	0.647 (±0.11)	0.951 (±0.11)	1.469 (±0.14)	−2.623	2.95	3.56
13	0.733 (±1.34)	1.382 (±1.22)	1.321 (±1.66)	−2.668	2.72	2.85
14	0.682 (±1.47)	1.121 (±1.39)	1.809 (±1.58)	−2.325	2.52	2.71
15	0.398 (±0.10)	2.596 (±3.71)	3.826 (±0.30)	−0.728	−0.01	−0.30
16	0.960 (±0.29)	1.413 (±0.53)	2.511 (±0.21)	−1.743	2.07	2.40
17	0.598 (±0.42)	1.378 (±0.84)	2.498 (±0.62)	−3.621	3.51	3.21
18	0.707 (±0.33)	1.471 (±0.60)	2.426 (±0.35)	−1.918	1.88	1.76
19	0.840 (±0.95)	1.389 (±0.92)	2.117 (±0.55)	−1.560	1.74	1.97
20	0.710 (±0.59)	1.869 (±1.22)	2.970 (±0.50)	−2.455	2.24	1.93
21	0.902 (±1.00)	1.895 (±1.15)	2.912 (±0.65)	−1.883	1.84	1.79
22	0.827 (±0.57)	1.939 (±0.86)	2.388 (±0.43)	−2.208	2.08	1.89
23	0.809 (±0.65)	1.996 (±1.12)	3.369 (±0.89)	−3.444	3.28	2.91
24	0.734 (±0.48)	2.291 (±1.16)	2.996 (±0.76)	−3.828	3.50	2.82
25	0.794 (±0.62)	1.957 (±1.22)	2.765 (±0.91)	−3.599	3.43	3.05
26	0.857 (±0.62)	1.836 (±1.03)	2.476 (±0.84)	−3.394	3.34	3.15
27	0.784 (±0.57)	2.194 (±1.36)	2.382 (±0.90)	−3.829	3.57	3.02
28	0.785 (±0.57)	2.195 (±1.36)	2.079 (±0.89)	−3.829	3.57	3.02
29	1.100 (±0.44)	1.267 (±0.42)	3.784 (±0.80)	0.133	0.69	2.03
30	0.885 (±0.59)	1.147 (±1.29)	3.078 (±0.89)	−0.953	1.48	1.85
31	0.881 (±0.25)	1.097 (±0.45)	3.164 (±0.33)	−1.977	2.59	3.38
32	1.236 (±0.39)	1.191 (±0.37)	2.276 (±0.21)	−1.750		3.75
33	0.993 (±0.35)	2.212 (±0.30)	1.149 (±0.76)	−3.454	3.36	3.36
34	0.831 (±0.14)	1.650 (±0.15)	1.516 (±0.35)	−4.100	4.11	4.16
35	0.962 (±0.06)	1.408 (±0.17)	1.305 (±0.10)	−3.555	3.89	4.25
36	0.825 (±0.11)	0.925 (±0.43)	3.498 (±0.34)	0.594	0.32	1.44
37	0.146 (±0.10)	0.679 (±0.20)	5.060 (±0.12)	−2.806	2.24	1.93
38	0.279 (±0.29)	1.206 (±0.58)	5.119 (±0.35)	−2.814	2.29	1.99
39	0.846 (±0.24)	1.031 (±1.74)	−1.282 (±1.18)	−7.059[c]	7.72[c]	8.73[c]
40	0.912 (±0.11)	1.022 (±0.19)	−0.896 (±0.53)	−6.409[c]	7.33[c]	7.60[c]
41	0.792 (±0.23)	0.734 (±0.60)	−0.236 (±0.94)	−6.710[c]		9.20[c]
42	0.806 (±0.22)	1.150 (±0.59)	0.382 (±0.88)	−6.756[c]	7.13[c]	7.75[c]
43	0.803 (±0.28)	0.835 (±1.27)	0.275 (±1.20)	−7.195[c]	8.59[c]	10.21[c]
44	1.002 (±0.23)	1.787 (±1.14)	2.512 (±0.33)	−2.309	2.42	2.62
45	1.191 (±0.33)	1.394 (±0.81)	2.147 (±0.32)	−2.127	2.89	3.35
46	0.766 (±0.73)	1.033 (±0.66)	2.866 (±0.31)	−1.099	1.56	2.13
47	0.552 (±0.20)	0.937 (±0.30)	3.998 (±0.15)	−2.030	2.19	2.38
48	0.675 (±0.26)	1.186 (±0.31)	2.878 (±0.17)	−1.980	2.10	2.21
49	0.683 (±0.29)	1.439 (±0.35)	2.879 (±0.19)	−2.065	2.02	1.97
50	1.047 (±0.19)	1.507 (±0.19)	4.757 (±0.12)	−1.438	1.79	2.18
51	0.942 (±0.26)	1.274 (±0.31)	3.774 (±0.18)	−1.800	2.25	2.68
52	0.983 (±0.19)	1.715 (±0.22)	2.643 (±0.13)	−1.735	1.86	1.99
53	1.061 (±0.20)	1.376 (±0.23)	3.723 (±0.14)	−1.656	2.18	2.65
54	2.139 (±1.85)	2.118 (±1.71)	3.871 (±1.82)	−0.358		2.59
55	1.150 (±0.31)	2.041 (±0.37)	3.159 (±0.21)	−2.047	2.16	2.25
56	1.006 (±0.50)	1.191 (±0.95)	2.919 (±0.40)	−1.986	2.72	3.15
57	1.028 (±0.12)	1.316 (±0.36)	2.559 (±0.10)	−2.499	3.05	3.56
58	0.561 (±0.12)	0.702 (±0.20)	3.272 (±0.38)	−5.974[c]	6.57[c]	7.62[c]
59	0.381 (±0.11)	0.929 (±0.24)	−0.604 (±0.36)	−5.128	4.97	4.74
60	0.547 (±0.14)	0.936 (±0.19)	−1.230 (±0.44)	−4.814	4.96	5.03
61	1.104 (±0.18)	1.772 (±0.30)	−0.545 (±0.11)	−1.451	1.67	1.78
62	0.193 (±0.06)	1.054 (±0.10)	9.373 (±0.37)	−9.983[c]	9.33[c]	7.69[c]

[a] 95% confidence intervals are given in parentheses. [b] Log P = 3.39[24] was used for n-pentane instead of log P = 2.50.[17,19] [c] π was used as hydrophobic parameter instead of log P.

strikingly similar to those presented by Mc Farland;[12] in all cases the calculated log $1/C$ values within the range of log P_0 and far away from log P_0 are underestimated by the parabolic model (positive residuals), while the values between are overestimated (negative residuals).

Discussion

The bilinear model constitutes a further progress in the quantitative description of nonlinear structure–activity relationships. It is mathematically derived from a simple model system, it is generally applicable, and it represents a smooth synthesis of both linear and nonlinear parts of structure–activity relationships; starting with a strictly linear relationship between biological activity and hydrophobic character, a point is reached, where the structure–activity relationship changes within approximately two log P units to a linear decrease of biological activity. Although this cut-off point is no sharp break in the structure–activity correlation, the change is more pronounced in the bilinear model than in the parabolic model.

Identical equations are obtained if a probability approach or an equilibrium approach is used for the derivation of the bilinear model; thus the question, if there

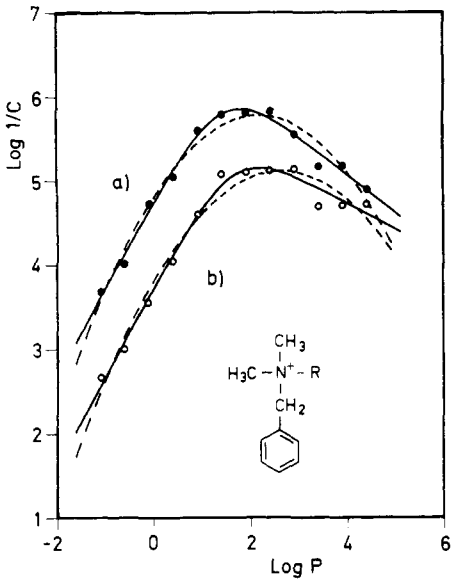

Figure 4. Bactericidal activity of benzyldimethylalkylammonium compounds (a) vs. *S. aureus* (Table I, eq 50) and (b) vs. *Cl. welchii* (Table I, eq 53) fitted with the bilinear model (—) and the parabolic model (- - -).

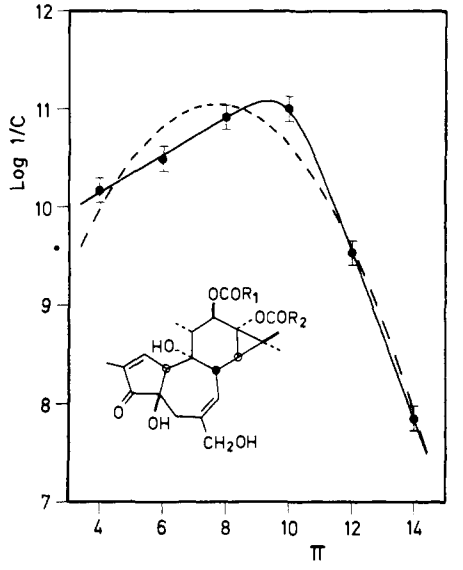

Figure 5. Irritant activity of phorbol 12,13-diesters (Table I, eq 62) fitted with the bilinear model (—) and the parabolic model (- - -);[17] approximate standard deviations are indicated for each value.

are nonequilibrium or quasiequilibrium (steady state) conditions in biological systems, can be ignored if the bilinear model is used.

It is recognized that the bilinear model is—like other models—an attempt to simulate a complex process in a rather simplistic way; there may be other factors which could cause a departure from linearity. However, the model "works"; in a large number of cases a nearly perfect fit results from the bilinear model.

If only a few data points are available or if the log P values vary within a small range, the parabolic model is a good approximation of the bilinear model and seems to be the superior model. However, if enough data points are present, if the biological data are measured with high accuracy, and if the log P values vary over a wide range, the bilinear model should be preferred. Due to the lack of systematic deviations it will be possible in the future to apply this model also to nonlinear structure–activity

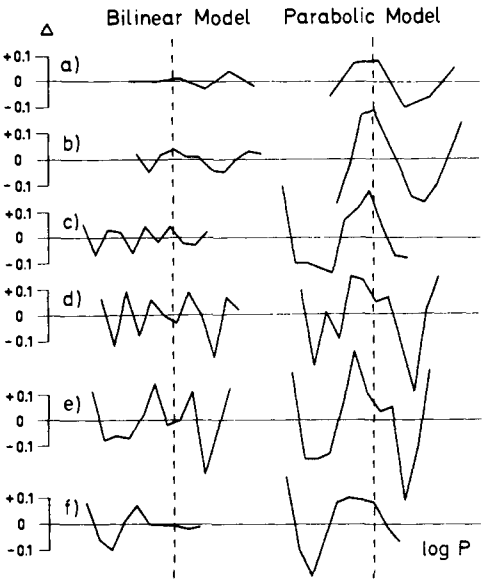

Figure 6. Analysis of residuals obtained with the bilinear model and the parabolic model for (a) eq 6, (b) eq 7, (c) eq 35, (d) eq 50, (e) eq 53, and (f) eq 57 (Table I; Figures 2–4): $\Delta = \log 1/C_{obsd} - \log 1/C_{calcd}$. The dotted lines indicate log P_0 (from the bilinear model).

relationships in heterogeneous sets and to dissociate hydrophobic, electronic, and steric effects more precisely.

Acknowledgment. I am indebted to O. H. Kehrhahn for many stimulating discussions and for his help in running the calculations.

References and Notes

(1) For paper 6, see ref 17b.
(2) H. Meyer, *Arch. Exp. Pathol. Pharmacol.*, **42**, 109 (1899).
(3) E. Overton, *Z. Physiol. Chem.*, **22**, 189 (1897).
(4) E. Overton, "Studien ueber die Narkose", Fischer, Jena, 1901, p 45.
(5) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).
(6) C. Hansch, *Acc. Chem. Res.*, **2**, 232 (1969).
(7) C. Hansch in "Drug Design", Vol. I, E. J. Ariens, Ed., Academic Press, New York, N.Y., 1971, p 271.
(8) C. Hansch and W. J. Dunn, III, *J. Pharm. Sci.*, **61**, 1 (1972).
(9) J. Ferguson, *Proc. R. Soc. London, Ser. B*, **127**, 387 (1939).
(10) J. T. Penniston, L. Beckett, D. L. Bentley, and C. Hansch, *Mol. Pharmacol.*, **5**, 333 (1969).
(11) C. Hansch and J. M. Clayton, *J. Pharm. Sci.*, **62**, 1 (1973).
(12) J. W. Mc Farland, *J. Med. Chem.*, 13, 1192 (1970).
(13) T. Higuchi and S. S. Davis, *J. Pharm. Sci.*, **59**, 1376 (1970).
(14) R. Franke and W. Schmidt, *Acta Biol. Med. Ger.*, 31, 273 (1973).
(15) S. H. Yalkowsky and G. L. Flynn, *J. Pharm. Sci.*, **62**, 210 (1973).
(16) R. M. Hyde, *J. Med. Chem.*, 18, 231 (1975).
(17) (a) H. Kubinyi, *Arzneim.-Forsch.*, **26**, 1991 (1976); (b) H. Kubinyi and O. H. Kehrhahn, *ibid.*, in press.
(18) Y. C. Martin and J. J. Hackbarth, *J. Med. Chem.*, **19**, 1033 (1976).
(19) R. Jeppsson, *Acta Pharmacol. Toxicol.*, **37**, 56 (1975).
(20) C. Hansch, A. R. Steward, J. Isawa, and E. W. Deutsch, *Mol. Pharmacol.*, 1, 205 (1965).
(21) C. Hansch and E. J. Lien, *J. Med. Chem.*, **14**, 653 (1971).
(22) C. Hansch and W. R. Glave, *Mol. Pharmacol.*, **7**, 337 (1971).
(23) E. J. Lien, C. Hansch, and S. M. Anderson, *J. Med. Chem.*, 11, 430 (1968).
(24) A. Leo, P. Y. C. Jow, C. Silipo, and C. Hansch, *J. Med. Chem.*, **18**, 865 (1975).
(25) W. J. Westlake, *J. Pharm. Sci.*, **60**, 882 (1971).
(26) N. R. Draper and H. Smith, "Applied Regression Analysis", Wiley, New York, N.Y., 1966.
(27) C. Daniel and F. S. Wood, "Fitting Equations to Data", Wiley-Interscience, New York, N.Y., 1971.