Journal of Medicinal Chemistry

VOLUME 15, NUMBER 6 JUNE 1972

Interrelationship of the Regression Models Used for Structure-Activity Analyses

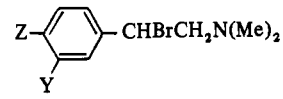
Arthur Cammarata

Temple University, School of Pharmacy, Laboratory of Physical Medicinal Chemistry, Philadelphia, Pennsylvania 19140.
Received January 10, 1972

The two most frequently used regression models for structure-activity analyses, the additive and the linear multiple regression models, are shown to be fundamentally interrelated. Each approximation used in reducing the additive model to its linear multiple regression counterpart leads to useful insight into how these regression methods are best applied, at least when dealing with organic compounds. Interaction models have little evidence in their favor when applied to biological data. An illustration of the systematic method of analysis that is provided by the theoretical development is presented.

Based on literature applications, there would appear to be two separate regression models that may be applied to the correlation of biological data. These are the additive model of Free and Wilson[1] and the linear multiple regression model of Hansch and his coworkers.[2-4] Cammarata has indicated, however, how these regression models can be interrelated in approaching the analysis of biological data using molecular orbital indexes,[5,6] and he has also specified the conditions under which the regression models might be taken as equivalent when using free energy related substituent indexes.[6,7] A number of statistically based structure-activity studies would seem to support the view that the additive and linear multiple regression models can be made equivalent, but each of these studies is presently tenuous. In most instances it is not clear what level of approximation is involved.[7-10] In one case there is a definite conflict. Based on a theoretical development, Singer and Purcell[11] concluded that an additive model would tend to break down when biological activities are related parabolically to a substituent constant. Yet Clayton and Purcell[8] have shown estimates of the pI_{50} values for a series of butyrylcholinesterase inhibitors to compare favorably when calculated on the basis of an additive model and of a linear regression model in which the pI_{50}'s were related in a parabolic manner to π.

In this report the interrelationship between the regression models used in structure-activity studies is developed in detail. Graham and Kamar's data[12] (Table I) on the adrenergic blocking potencies of *N,N*-dimethyl-2-bromophenethylamines (I) in the rat is used in support of the development. Hansch and Lien[13] have taken a linear multiple regression approach in analyzing this data. This same compilation is presently used for the contrasts that are afforded. While confirming the regression equation reported by Hansch and Lien, the more detailed method of approach which is the subject of this paper provides correlations that may be interpreted following four different rationales rather than only one as earlier reported.[13] These new correlations can be applied in suggesting additional substances whose test results potentially would allow distinctions to be made between the differing interpretations.



I

Basis for the Methods. To avoid any possible misunderstanding of the manner in which the regression methods are interrelated, at least as applied to structure-activity studies,

Table I. Adrenergic Blocking Potencies of Some *N,N*-Dimethyl-2-bromophenethylamines

Substituent variations										$\log(1/ED_{50})^a$	
Meta					Para						
F	Cl	Br	I	Me	F	Cl	Br	I	Me	Obsd	Estd
1										8.16	7.86
	1									8.68	8.28
		1								8.89	8.54
			1							9.25	9.25
				1						9.30	8.78
					1					7.52	7.52
						1				8.16	7.98
							1			8.30	8.47
								1		8.40	8.40
									1	8.46	8.22
1						1				8.19	8.38
1							1			8.57	8.87
1									1	8.82	8.62
	1					1				8.89	8.80
	1						1			8.92	9.29
	1								1	8.96	9.04
		1				1				9.00	9.06
		1					1			9.35	9.55
		1							1	9.22	9.30
				1					1	9.30	9.54
				1			1			9.52	9.79
									Unsubstituted	7.46	7.46[b]

[a] From ref 12. [b] By definition.

each step of the development will be presented as succinctly as possible. A statement of the assumptions involved precedes its mathematical counterpart and the implications that are associated with each assumption follow.

Assumption 1. An additive model applies to each and every compound in a structure-activity compilation. Thus, for a set of N compounds, the biological responses, A, for each compound n in the series, under equivalent conditions of assay, can be written

$$A_n = \sum_p \sum_s a_{n,ps} + \mu \; (n = 1,2, ..., N) \tag{1}$$

in which μ represents the biological effect of a parent structure and there is a value for a_{ps} corresponding to the biological effect due to each substituent s at a position p of this structure.

Implication i. The biological effect imparted by a substituent attached at different points on a parent structure differs

$$a_{n,ps} \neq a_{n,qs} \tag{1i}$$

Hence, a *m*-Cl differs from a *p*-Cl in its biological effect; and an α-Me differs from a β-Me in its biological effect.

Implication ii. The biological effect imparted by a substituent in one compound n differs from the biological effect imparted by the same substituent in a second compound m in the series:

$$a_{n,ps} \neq a_{m,ps} \tag{1ii}$$

In other words, the biological effect imparted by a substituent situated at the same position of a parent structure differs as other substituents are placed on the parent structure. This behavior in the biological effect associated with a *particular* substituent in passing from one derivative to another is indicative of an interaction mechanism. With an interaction mechanism operative, it can be said that the biological effect of NO_2 in, say, *m*-nitroanisole differs from the biological effect of NO_2 in *m*-nitrotoluene when using the same test system. Such interaction effects cannot usually be very great, since it is common practice to define physical[14] and biological[15] substituent values by subtracting a free energy measure for a monosubstituted (or more highly substituted parent) compound from corresponding measures for the more highly substituted derivatives. This leads to a simplifying assumption.

Assumption 2. Following Free and Wilson,[1] each substituent may be considered, at least initially, as making essentially the same biological contribution in each derivative possessing the substituent

$$a_{n,ps} = a_{m,ps} \tag{2}$$

Implication i. In the event of an interaction effect, as may occur with NO_2 and OH groups that are situated para to one another on an aromatic nucleus, the interacting substituents can be identified as "new" groups which are unrelated to corresponding noninteracting substituents. The additivity of eq 1 can thus be maintained. According to this implication, the interaction model used by Boček, Kopecký, and their coworkers[16,17] in analyzing the toxicities of substituted benzenes toward mice is not generally applicable to eq 1. As a result, the attempt of Singer and Purcell[11] to interrelate the regression methods in terms of the Boček-Kopecký model is of no consequence to the present development.

Assumption 3. Each substituent contribution a_{ps} can be interpreted as a weighted average biological effect due to differing physical properties of each substituent. A linear combination of physically meaningful substituent parameters X can then be written

$$a_{n,ps} = \sum_X (b_{ps,X} X_{ps})_n \; (n = 1,2, ..., N) \tag{3}$$

where the weighting factors b_X give the fractional contribution of each substituent property toward the biological effect associated with the substituent. It follows necessarily that $\Sigma_X b_X = 1$, which is a normalizing condition.

Implication i. The fractional contribution made by any one substituent property X is not necessarily the same for each of the substituents at a given position of substitution (eq 3i). As a consequence, the substitution of eq 3 into eq 1, even when made in accord with assumption 2, leads to a set of relationships which contains so many independent coeffi-

$$b_{ps,X} \neq b_{pt,X} \tag{3i}$$

cients $b_{ps,X}$ that, for practical purposes, it is often impossible to gain an estimate for their values based on a knowledge of the values for A and for the various X. An additional simplifying assumption is thus indicated.

Assumption 4. The weighting factor associated with a given substituent property can be taken as the same for all substituents situated at a specified position p.

$$b_{ps,X} = b_{pt,X} \tag{4}$$

Implication i. For all substituents at a given position of substitution, a linear multiple regression model serves in relating the biological substituent effects a to the various physically based substituent constants X. Equation 3 can thus be written

$$a_p = \sum_X b_{p,X} X_p \tag{4i}$$

in which the subscripts n and s have been deleted from eq 3 since eq 4i applies to all compounds having a number of differing substituents at position p.

Implication ii. In applying an additive model to biological data, each substituent has its biological effect determined in an independent fashion. As a consequence the biological substituent effects associated with each of the substituents at a specified position may first increase and subsequently decrease in value. To relate these biological substituent effects to physically based substituent constants using multiple regression techniques higher powers of each physical constant may have to be included in eq 4i to take this effect into account. Thus, the additive model of eq 1 can indeed apply when biological substituent effects are parabolically related to a given substituent property. This implication is in agreement with the findings of Clayton and Purcell[8] but contrasts sharply with the conclusion of Singer and Purcell's theoretical analysis.[11]

Implication iii. The multiple regression model relating the activities of the compounds in a structure-activity compilation to physically based substituent constants is obtained by substituting eq 4i into eq 1. If substituent variations are made at only two positions for the compounds in a set, the regression model can be written

$$A = \left(\sum_X b_{u,X} X_u\right) + \left(\sum_X b_{v,X} X_v\right) + \mu \tag{4ii}$$

where each grouping of terms is appropriate to one of the positions of substitution. In general, there will be one such term for each position of substitution on a parent drug structure. Equation 4ii does not contain the subscripts n and s as does eq 1, since it applies to all compounds and all possible substituent variations found in a set of data. The significance of the substitution of eq 4i into eq 1 in arriving at eq 4ii in relation to the statistics of eq 4ii is discussed in the section dealing with the application of these relationships.

Assumption 5. In principle, no additional assumptions are necessary beyond those used in arriving at eq 4ii to have a satisfactory regression model. This is because in applying eq 4ii to a set of data it can be shown whether the weighting coefficients $b_{p,X}$ appearing before a physical parameter X is essentially the same at each position of substitution, *i. e.*,

$$b_{u,X} = b_{v,X} \tag{5}$$

The assumption represented by eq 5 is frequently unknowingly made, however, by the simple expedient of adding a physical parameter which refers to differing positions of substitution prior to conducting a regression analysis.

Implication i. Accepting the assumption made by eq 5, eq 4i can be written

$$a = \Sigma b_X X \tag{5i}$$

where the subscript p is deleted from eq 4i since eq 5i is applicable to all positions of substitution. Both Cammarata and Yau[7] and Fujita and Ban[10] tacitly worked at this level of approximation when, ostensibly to increase the number of statistical degrees of freedom, they included ortho-para and meta-para biological activity contributions, respectively, into a single multiple regression model corresponding to eq 5i. Viewed in terms of the number of assumptions involved in arriving at eq 5i, neither Cammarata and Yau's nor Fujita and Ban's correlations adequately demonstrate the equivalence of additive and linear multiple regression models.

Implication ii. The substitution of eq 5 into eq 4ii leads to the regression model in which additivity of substituent parameters define the independent variables of the equation. Considering only two physical properties of substituents to contribute to the biological effect eq 4ii can thus be written

$$A = b_1\Sigma\sigma + b_2\Sigma\pi + \mu \tag{5ii}$$

where for this particular case electronic and lipophilic substituent properties are designated as important. The problem with accepting a regression model such as eq 5ii *a priori* obviously is that a nonequivalence in the biological behavior of two or more differing positions of substitution may be masked. Less obvious is the fact that certain physical influences may be discarded as unimportant on statistical grounds when using such a model, when in actuality there are two separate physical interpretations possible for the data. This point is discussed for the example which follows.

Application. An additive model (eq 1 with assumption 2) was applied to the data found in Table I to derive the biological substituent effects a_m and a_p, shown in Table II, which apply to the meta and para positions of substitution, respectively. The biological response measure (ED_{50}) was converted to logarithmic form (log $1/ED_{50}$) and the activity of the unsubstituted compound was set equal to μ in order to define a_m and a_p in the same relative manner as are physically based substituent constants.[6,7,10] Inspection of the derived values (Table II) shows that corresponding substituents lead to a differing biological effect depending on whether they are substituted meta or para. These biological substituent constants may be considered additive based on the agreement obtained between the observed and the estimated log ($1/ED_{50}$) for the compounds ($n = 22$; $R = 0.911$; $s = 0.214$).

Table II. Group Contributions

	Meta					Para			
Y	a_m	σ_m	π_m	r_v	Z	a_p	σ_p	π_p	r_v
I	0.84[a]	0.35	1.26	1.98	I	1.79[a]	0.28	1.26	1.98
Me	0.76	−0.07	0.52	1.97	Me	1.32	−0.17	0.52	1.97
Br	1.01	0.39	1.02	1.85	Br	1.08	0.23	1.02	1.85
Cl	0.52	0.37	0.70	1.75	Cl	0.82	0.23	0.70	1.75
F	0.06[a]	0.33	0.15	1.47	F	0.40[a]	0.06	0.15	1.47
H	0.00[b]	0.00[b]	0.00[b]	1.20	H	0.00[b]	0.00[b]	0.00[b]	1.20

[a] Single-point determination. [b] By definition.

Assumptions 3 and 4 lead to a simplification of the additive model. When these assumptions are valid, biological substituent effects that are appropriate to a given position of substitution should correlate with physically based substituent parameters by a linear multiple regression model (eq 4i). In seeking agreement with this prediction, two different relationships appropriate to each position of substitution are found. These relationships, which are based on the values given in Table II, are specified by eq 6a, 6b, 7a, and 7b, in which the standard errors associated with the determination of the coefficients are included.

A comparison of eq 6 and 7 shows that the coefficients of eq 7 are about double those of eq 6. The intercepts to eq 6a and 7a are nearly zero in each case, as they should be according to eq 4i when the biological and physical substituent constants are defined in the same relative manner. The intercepts of eq 6b and 7b differ substantially from zero since the biological substituent constant a is defined in a relative manner whereas the physical substituent constant r_v (the van der Waals radius) may be considered as an absolute-type measure. The about two times greater value for the intercept of eq 7b relative to the intercept of eq 6b most probably reflects the corresponding ratio of the slopes to these equations, since equivalent values for r_v are used in deriving the two equations.

Based on the theoretical development, eq 4i and 4ii are but two alternative ways of expressing the same type of correlation. There are 4 possible combinations of eq 6 and 7, and accordingly 4 differing equations are found for the correlation of log ($1/ED_{50}$) (eq 8a–8d). The coefficients of eq 8a–8c are essentially identical with the coefficients of eq 6 and 7 as must be the case if, by the theoretical development, eq 4i and 4ii are equivalent. Because of this equivalence, there are a total of $n^* = 12$ independent points leading to eq 8 and *not* $n = 22$ points. The latter gives the total number of compounds on which the 12 independent points are based whereas the former gives the number of independent a values. Less satisfying agreement is found between the coefficients of eq 6b, 7b, and 8d, most probably because of the intercorrelation between $r_{v,m}$ and $r_{v,p}$.

The analyses to this point show the coefficients of eq 6a, 7a, and of eq 6b, 7b to differ. In principle, because the coefficients differ, the quantities on which eq 6a, 7a and eq 6b, 7b are based should not be "mixed together" to derive,

Meta

	n	n^*	R	s	
$a_m = -0.645\ (\pm 0.509)\sigma_m + 0.919\ (\pm 0.215)\pi_m + 0.119$	6		0.931	0.196	(6a)
$a_m = 1.238\ (\pm 0.276)r_v - 1.578$	6		0.913	0.191	(6b)

Para

$a_p = -2.014\ (\pm 0.731)\sigma_p + 1.646\ (\pm 0.258)\pi_p + 0.108$	6		0.968	0.206	(7a)
$a_p = 1.984\ (\pm 0.307)r_v - 2.477$	6		0.955	0.213	(7b)
$\log (1/ED_{50}) = -1.004(\pm 0.302)\sigma_m + 0.791(\pm 0.150)\pi_m - 1.993(\pm 0.402)\sigma_p + 1.479(\pm 0.139)\pi_p + 7.914$	22	12	0.946	0.203	(8a)
$\log (1/ED_{50}) = -0.911(\pm 0.249)\sigma_m + 0.747(\pm 0.123)\pi_m + 1.666(\pm 0.124)r_{v,p} + 5.769$	22	12	0.961	0.168	(8b)
$\log (1/ED_{50}) = -2.127(\pm 0.459)\sigma_p + 1.539(\pm 0.159)\pi_p + 0.651(\pm 0.164)r_{v,m} + 7.066$	22	12	0.924	0.235	(8c)
$\log (1/ED_{50}) = 0.618(\pm 0.139)r_{v,m} + 1.722(\pm 0.156)r_{v,p} + 4.893$	22	12	0.942	0.200	(8d)
$a_{m/p} = -1.445\ (\pm 0.494)\sigma + 1.255\ (\pm 0.216)\pi + 0.212$	11		0.899	0.259	(9a)
$a_{m/p} = 1.687(\pm 0.374)r_v - 2.169$	11		0.832	0.308	(9b)
$\log (1/ED_{50}) = -1.543(\pm 0.269)\Sigma\sigma + 1.173(\pm 0.124)\Sigma\pi + 7.905$	22		0.907	0.251	(10a)
$\log (1/ED_{50}) = 1.143(\pm 0.173)\Sigma r_v + 4.949$	22		0.828	0.326	(10b)

respectively, a single regression equation. This "mixing" of a_m and a_p to derive a single regression equation (corresponding to eq 5i) intrinsically corresponds to a recognition that assumption 5 is valid. Without the previous analyses as a guide, one could have followed Cammarata and Yau[7] and Fujita and Ban[10] in deriving the relations 9a and 9b, which, while having an acceptable number of statistical degrees of freedom, tend to mask the differing biological behavior of the meta and para positions. The poorer fits provided by these equations tend to suggest the inadequacy of assumption 5 rather than the need for any additional variables.

Only one set of values for H rather than the two shown in Table II was used in deriving eq 9a and 9b to avoid assigning an undue weight to reference points. This is the reason $n = 11$ and not $n = 12$ appears in designating the number of points on which eq 9a and 9b are based.

Assumption 5 is also tacitly recognized as valid whenever, for multisubstituted compounds, additivity in a physical parameter is used as a basis for describing a physical characteristic due to the substituents prior to a linear multiple regression analysis. The regression model in this case (eq 5ii) is equivalent, in principle, to the regression model (eq 5i) used to derive eq 9. Hence, the counterparts to eq 9 are given by eq 10a and 10b. These also tend to mask the differing biological behavior of the meta and para positions.

Equation 10a is identical with the correlation previously reported by Hansch and Lien.[13] In fairness to these workers it should be noted that including $\Sigma\sigma$, $\Sigma\pi$, and Σr_v as the independent variables to a linear multiple regression model leads to a regression equation in which Σr_v is statistically insignificant. This points out one of the advantages to conducting statistical analyses of biological data along lines consistent with a theoretical model; one is less apt to discard as insignificant a physical property which itself may account for the data.

Equations 8 are statistically acceptable at each level of approximation (assumptions 1–4) involved in formulating the alternative regression models. Because of the difference in the regression coefficients of eq 6 and 7, the application of assumption 5 is inappropriate. Thus, eq 8 should serve as a basis for making inferences regarding the physical properties that influence the adrenergic blocking potencies of compounds having the structure I. It is suggested that the alternative interpretive models might be at least narrowed to 2 possibilities by designing compounds following eq 8 prior to presenting a physical rationale. Table III shows some possible substituent variations that may be used and the predicted $\log (1/ED_{50})$ for compounds with these substituents based on each of eq 8. The first 3 compounds may allow a choice between eq 8a and 8c and eq 8b and 8d. In the first instance these compounds are expected to be fairly potent blocking agents, while in the latter instance they are expected to be relatively much less potent. Compounds 3 and 4 may confirm whatever choice is made. The change from *m*-NH_2 to *p*-NH_2 should lead to a decrease in potency if in accord with eq 8a and 8c but an increase in potency if in accord with eq 8b and 8d. A distinction is possible between eq 8a and 8c using *m*- and *p-tert*-Bu derivatives. The change from *m-tert*-Bu to *p-tert*-Bu should lead to a pronounced decrease in potency according to eq 8a but a pronounced increase in potency according to eq 8c. No clear distinction is possible between eq 8b and 8d, at least for the substituent variations in Table III, but this may not be a problem if the above predictions are born out.

Conclusions

Most investigations of structure-activity relationships are based either explicitly or implicitly on assumption 1, *i. e.*, an additive model applies to each compound in a structure-activity compilation. In fact, once the condition is made that a statistical model must agree with accepted physical principles or physical interpretations in an extrathermodynamic sense, it becomes difficult to provide evidence in favor of any other model over an additive one. The one

Table III. Possible Compounds Enabling Interpretive Distinction

No.	Substituents Y	Z	Predicted $\log (1/ED_{50})$ Eq 8a	Eq 8b	Eq 8c	Eq 8d
1	CF_3	H	8.34	6.19	8.64	6.39
2	OMe	H	7.89	5.75	8.05	5.83
3	NH_2	H	7.16	5.05	8.07	5.85
4	H	NH_2	4.89	8.35	6.70	7.56
5	*tert*-Bu	H	9.36	7.13	8.88	6.61
6	H	*tert*-Bu	6.47	10.41	13.84	9.69

example of the use of an interaction model for the analysis of biological data[16,17] does not discount the generality of the additive model, since the data for this analysis when grouped according to congeneric series provide good linear regression fits with physically based substituent parameters.[18,19] Certain reported[1] examples of the inadequacy of an additive model to correlate biological activities seem most appropriately interpreted as reflecting the inadequacy of assumption 2. In Free and Wilsons's paper,[1] it was pointed out that the bacteriostatic activities for multisubstituted tetracyclines were not well correlated by the use of an additive model. But in this statistical analysis certain groups, *e. g.*, NO_2, were not recognized as physically "different," when intramolecularly H bonded or when strongly conjugated through resonance with an OH group, from corresponding groups not acted upon by these effects.

A break-down of assumption 2 for one or more members of a series can be taken as a basis for the majority of studies directed toward mapping the hydrophobic regions of enzymes.[20,21] Homologous series of compounds are usually involved in these investigations, and inferences regarding the nature of the enzyme binding site are drawn from the behavior of the binding constants as a molecular side chain, usually linear aliphatic, has its length increased. An increment of change in the binding constant as the side chain is increased most frequently is least when the last-added CH_2 group is in a hydrophilic environment. This corresponds to a breakdown in assumption 2, since by this assumption each added CH_2 group should make an equivalent contribution to the binding constant if the binding behavior for the members of an homologous series is strictly additive.

From the large number of linear multiple regression equations that have been reported to correlate various sets of biological data[2-4,18,19] assumptions 2-4 seem generally applicable. Most probably, however, assumption 4 is not strictly followed. Rather, the fractional biological effect due to a substituent property may vary for each of the members of the series, but the variation may often be sufficiently small so that this fraction is well approximated by an average value—the coefficient determined by the regression analysis. In practice, the coefficients of a linear multiple regression equation are not identified as a fractional biological effect due to a substituent property, *i. e.*, the coefficients to the regression equation are not normalized. Including a normalization condition into the least-squares solution of a linear multiple regression model, while desirable from a theoretical standpoint, most probably will not influence appreciably earlier conclusions drawn off of the relative magnitudes of the coefficients in a regression equation.

In cases involving the analyses of the biological activities for series containing multisubstituted compounds, and sometimes only monosubstituted compounds, a reinvestigation of much of the earlier literature is indicated. The most approximate form of the additive model, *i. e.*, a linear multiple regression model with assumption 5, has frequently been unwittingly used by (a) fitting the biological activities, for, say, meta- and para-substituted compounds to a single regression equation prior to demonstrating that the coefficients to the regression equations are equal for each position of substitution; by (b) making use of physical substituent constant additivity to obtain the independent variables of a regression model; and by (c) a combination of a and b applied to the same set of data. It can be expected that by following the approach given in this paper many of the earlier interpretations of correlation equations may have to be modified, sometimes appreciably.

It is hoped that the content of this paper is not misconstrued. In pointing out the interrelationship between the present regression models that are applied to biological problems it is unavoidable to also have to point out certain faults in earlier works. The method of approach which has been developed in this paper should help to eliminate many of the factors that could give rise to a misleading correlation. Additional development of this approach no doubt will aid all who are involved in or make use of mathematical methods for the analysis of biological data.

References

(1) S. H. Free, Jr., and J. W. Wilson, *J. Med. Chem.*, 7, 395 (1964).
(2) C. Hansch and T. Fujita, *J. Amer. Chem. Soc.*, 86, 1616 (1964).
(3) T. Fujita, J. Iwasa, and C. Hansch, *ibid.*, 86, 5175 (1964).
(4) J. Iwasa, T. Fujita, and C. Hansch, *J. Med. Chem.*, 8, 150 (1965).
(5) A. Cammarata, *ibid.*, 11, 1111 (1968).
(6) A. Cammarata, "Molecular Orbital Studies in Chemical Pharmacology," L. B. Kier, Ed., Springer-Verlag, New York, N. Y., 1970, p 156.
(7) A. Cammarata and S. J. Yau, *J. Med. Chem.*, 13, 93 (1970).
(8) J. M. Clayton and W. P. Purcell, *ibid.*, 12, 1087 (1969).
(9) P. N. Craig, *ibid.*, 15, 144 (1972).
(10) T. Fujita and T. Ban, *ibid.*, 14, 148 (1971).
(11) J. A. Singer and W. P. Purcell, *ibid.*, 10, 1000 (1967).
(12) J. D. P. Graham and M. A. Kamar, *ibid.*, 6, 103 (1963).
(13) C. Hansch and E. J. Lien, *Biochem. Pharmacol.*, 17, 709 (1968).
(14) J. E. Leffler and E. Grunwald, "Rates and Equilibria of Organic Reactions," Wiley, New York, N. Y., 1963.
(15) R. Zahradnik, *Arch. Int. Pharmacodyn.*, 135, 311 (1962).
(16) K. Boček, J. Kopecký, M. Krivucova, and D. Vlachova, *Experientia*, 20, 667 (1964).
(17) J. Kopecký, K. Boček, and D. Vlachova, *Nature (London)*, 207, 981 (1965).
(18) A. Cammarata and K. S. Rogers, "Advances in Linear Free Energy Relationships," N. B. Chapman and J. Shorter, Ed., Pergamon Press, London, 1972.
(19) A. Cammarata and J. J. Zimmerman, "Pharmacology and Medicinal Chemistry," L. B. Kier, Ed., Marcel-Dekker, New York, N. Y., in press.
(20) B. Belleau, *Ann. N. Y. Acad. Sci.*, 144, 705 (1967).
(21) B. R. Baker, "Design of Active-Site-Directed Irreversible Enzyme Inhibitors," Wiley, New York, N. Y., 1967.