Journal of Medicinal Chemistry

VOLUME 16, NUMBER 7 JULY 1973

On Model Building in Structure-Activity Relationships. A Reexamination of Adrenergic Blocking Activity of β-Halo-β-arylalkylamines†

Stefan H. Unger and Corwin Hansch*

Department of Chemistry, Pomona College, Claremont, California 91711. *Received December 8, 1972*

Five criteria are presented for selecting a "best equation" to correlate biological activity: selection of independent variables, statistical justification of the choice of independent variables in the "best equation," principle of parsimony, ratio of data to variables, and, most importantly, agreement between qualitative and quantitative models; that is, the equation must be in accord with the known physical-organic and biomedicinal chemistry of the process under study. These criteria are applied to a reexamination of the blocking action of β-halo-β-phenylalkylamines at various levels of factorization of the variables. An equation involving σ^+, π, and $r_{v,p}$ (van der Waals' radii from the para position) is selected. The electronic nature of the process is confirmed by a new set of substituent constants (Swain-Unger) **S** and **P** which are both optimized and orthonormalized. Agreement with the qualitative model of Chapman, *et al.*, is briefly discussed.

The formulation of structure-activity relationships (SAR) in biomedicinal chemistry involves both a qualitative and quantitative aspect. Qualitatively, one attempts to define the interaction of a drug with a receptor and to delineate the basic physical or chemical processes which produce the given biological response. Various models are employed in rationalizing the action of archetypal drugs and one usually worries relatively little about explaining variations in the activities of derivatives. Quantitatively, however, it is just this variation in the activity of congeneric series which provides the subject matter for investigation. No matter from which point of view one approaches the problem, however, one must attempt to put together a picture which is both qualitatively and quantitatively self-consistent and in accord with all of the facts.

One must rely heavily on statistics in formulating a quantitative model but, at each critical step in constructing the model, one must set aside statistics and ask questions. Does the model agree with what is known about the biochemistry and/or molecular biology which may be involved? Do the weights given the various physical and chemical parameters employed seem reasonable in light of our knowledge of physical-organic chemistry and other SAR in related systems? What limitations do the amount, quality, and range of data place on our results? Without such a qualitative perspective, one is apt to generate statistical unicorns, beasts that exist on paper but not in reality. For example, it has recently become[1] all too clear that one can correlate a set of dependent variables using random numbers as independent variables. Such correlations meet the usual criteria of high significance according to the F statistic and correlation coefficient. Topliss and Costello's results[1] add substance to the remark that if one fails to obtain a SAR for a set of data, it is really a reflection on his library (*i.e.*, he hasn't tried enough variables). What is one to do to avoid cluttering the literature with poor or meaningless correlations?[2]

At this stage in the development of model building there are five criteria which must be considered before one settles on a "best equation" to correlate a set of congeners.

1. Selection of independent variables. The widest possible number of independent variables must be considered. Despite the remark referred to above, enough experience is still not in hand to say that any given set of parameters such as σ, π, and E_s is sufficient for the task. While the above three have been most widely used, one must not overlook, for example, polarizability, the various models of steric effects (*e.g.*, molar volume), dummy parameters, and MO parameters. Of course the parameters selected in the "best equation" should be essentially independent of each other and parameters which are thought to be "purer" measures of effects should take precedence over hybrids as an aid in interpretation.

2. Justification of the choice of independent variables. When one is satisfied that all "reasonable" parameters have been considered, each term in the highest order equation must be validated by an appropriate statistical procedure[2] such as all possible regressions, backward elimination, forward selection, or stepwise regression, etc. The authors have found it advantageous to examine all possible regressions and then to use a forward selection procedure with sequential F tests to obtain the "best equation," generally that with the lowest standard deviation and all terms significant.

3. Principle of parsimony (Occam's Razor). All things being equal, one accepts the simplest model.

4. Number of terms. Topliss and Costello's analysis[1] suggests that one should have at least five to six data points per variable in order to avoid chance correlations.

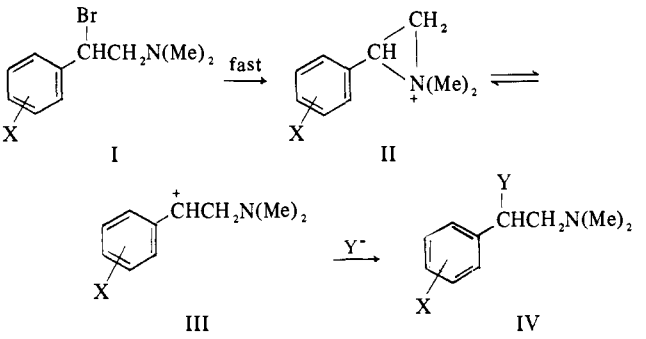
5. Qualitative model. It is most important that one have a

†This work was supported by Grant CA 11110 from the National Institute of Health.

qualitative model which is consistent with the known physical-organic and biomedicinal chemistry of the process under consideration.

A recent report by Cammarata[3,‡] has prompted the reexamination of our[4] earlier analysis of the adrenergic blocking activity of β-halo-β-arylalkylamines (I). These compounds are unstable at physiological pH and are thought to undergo solvolysis as shown in Scheme I.[5] Chapman, *et al.*,[5b] continu-

Scheme I



ing the work of Chapman and Triggle,[5a] suggest that III is rapidly formed from the ethyleniminium ion II and then reacts with a nucleophilic center at the α-receptor site. Garner, *et al.*,[5d] have extended studies of I to the dog, rabbit, and rat, substantiating the concept of an ethyleniminium ion and α-receptor activity. Belleau[§] has suggested that the initial attachment of III is important in determining the activity, and Chapman, *et al.*,[5b] develop this theory to include a more detailed model of the receptor, emphasizing the importance of bonding to the β-C atom of III. We present results below which substantiate the general outline presented by Chapman, *et al.*[4,5]

Our original analysis[4] of the data of Graham and Karrar[6] gave eq 1. Equation 2 is the "best" result from Cammarata's recent analysis,[3] that is, the equation one might select simply on the basis of "best fit," a dangerous procedure *not suggested by Cammarata.* In eq 1, σ and π are simply summed for the disubstituted derivatives. Both in our work

	n	r	s	
$pC = 1.22\pi - 1.59\sigma + 7.89$	22	0.918	0.238	(1)
$pC = 0.75\pi_m - 0.91\sigma_m + 1.67r_{v,p} + 5.77$	22	0.961	0.168	(2)

and in Cammarata's, one data point (4-phenyl) has been omitted because of very bad fit. The n represents the number of data points, r the correlation coefficient, and s the standard deviation. Cammarata's eq 2 is interesting in that it does not assign hydrophobic effects to para substituents. As will be shown later, the van der Waals radii for this limited set of para substituents are accidently correlated by the hydrophobic and electronic effects (*cf.* eq 15). The most disconcerting aspect of eq 2, however, is that an electronic effect for just the meta substituents is indicated. To our knowledge, this is without precedent in the literature of physical-organic chemistry. While it is conceivable that some unusual balance of forces could result in a canceling of the electronic effect from the para substituents, this also seems unlikely in view of the proposed mechanism; namely, the carbonium ion intermediate III would be expected to exhibit a very strong electronic substituent effect, and one would expect that σ^+ would perform better than σ in this process. This view has been formulated as shown in eq 3 which is an improvement

	n	r	s	
$pC = 1.15\pi - 1.47\sigma^+ + 7.82$	22	0.944	0.197	(3)

over eq 1, but not quite as good as eq 2. By the principle of parsimony one might prefer eq 3 but, more importantly, it agrees with the physical-organic chemistry.

While σ^+ constants[7] are derived from a system similar to III ($X\text{-}C_6H_4C^+Me_2$), the system is, in fact, not identical with III. Since any σ must rationalize both the resonance and inductive effects of substituents and since these effects may not be varying in parallel fashion in the two systems, one cannot expect a single parameter to give the best possible answer (in general). Taft and Lewis[8] attempted to deal with this problem by factoring σ into σ_I and σ_R. In this approach the following assumptions were made: (1) the total electronic substituent effect is the simple sum of inductive and resonance contributions; (2) inductive contributions to σ_m and σ_p are equivalent; (3) a unique proportionality constant $\propto$ exists which relates resonance effects between meta and para positions; (4) σ_I can be obtained from an independent system (substituted acetic acids) in which resonance effects are absent. Recently, Swain and Lupton[9] presented a slightly different approach using a variation of assumptions 1 and 4. The inductive-field effect was calculated from a regression of σ' (bicyclooctanecarboxylic acid ionization; set 5[9]) *vs.* σ_m and σ_p. These $\mathcal{F}$ values were then used with σ_p to calculate $\mathcal{R}$ *via* $\sigma_p = \alpha\mathcal{F} + \mathcal{R}$ where α was evaluated by assuming that $N^+(CH_3)_3$ contained no resonance effect. [Dr. A. Leo has pointed out to us that log K/K_H was used instead of σ values (the factor $\rho = 1.65$ missing) which causes $\mathcal{F}$ to be out of line with other σ.]

A more direct approach to the factorization of electronic effects has now been made, based on two independent mathematical procedures:[10] (a) optimization[#] with respect to a large number of chemically diverse data sets of the two electronic parameters **S** (field-inductive-σ bond perturbation) and **P** (resonance-π bond perturbation; perpendicular to **S**), and (b) orthonormalization[10] of the optimized **S** and **P**, with **S** referred to the bicyclooctanecarboxylic acid ionization (set 5[9]) and **P** forced to be orthogonal to **S**. The chemical constraint is made in such a manner that the individual data points are not given infinite weight. (Although the log K/K_H values were used, since **S** and **P** are normalized, the above-mentioned error is eliminated.) Both **S** and **P** are normalized to N, the total number of substituent constants studied; this maintains average values when more substituents are added to the current 42. The entire procedure may be profitably couched in vector terminology;[10,11] thus, the electrical substituent space (SS_e) is spanned by the two orthonormal basis vectors **S** and **P** (which describe an optimum plane with respect to a large amount of data). Any other set of data **Z** is a vector whose electrical contribution may be characterized by its projection on SS_e.

$$\mathbf{Z} = a\mathbf{S} + b\mathbf{P} (+c)^{\#} \quad (4)$$

The total SS is defined analogously, including mutually

‡Note that in this report, the meta and para labels in Table I are reversed. The parenthetical numbers associated with the coefficients in eq 6-10 appear to be standard deviations.

§*Cf.* discussion in ref 5b with that of Belleau.[5e]

#C. G. Swain, S. H. Unger P. Strong, and N. Rosenquist, unpublished results. Optimization is here with respect to $Z = aS + bP + c$ where $c \simeq 0$ since **Z** is referenced to hydrogen. The substituent constants are essentially identical; r for correlation between the two sets is >0.99 (those obtained with and without c in the optimization). % P, a, and b are statistically identical; however, these newer values should be used in subsequent work.

Table I. Data[a]

No.	S_m	P_m	S_p	P_p	σ^+ σ_m^+	σ^+ σ_p^+	r_p	r_m	π π_m^+	π π_p^+	Log $1/ED_{50}$[b] obsd	
1	0.0	0.0	0.0	0.0	0.0	0.0	1.200	1.200	0.0	0.0	7.460	H
2	0.0	0.0	0.911	−0.406	0.0	−0.070	1.470	1.200	0.0	0.150	8.160	4-F
3	0.0	0.0	0.916	−0.125	0.0	0.110	1.750	1.200	0.0	0.700	8.680	4-Cl
4	0.0	0.0	0.905	−0.127	0.0	0.150	1.850	1.200	0.0	1.020	8.890	4-Br
5	0.0	0.0	0.826	−0.098	0.0	0.140	1.980	1.200	0.0	1.260	9.250	4-I
6	0.0	0.0	0.017	−0.376	0.0	−0.310	1.970	1.200	0.0	0.520	9.300	4-Me
7	0.911	−0.406	0.0	0.0	0.350	0.0	1.200	1.470	0.130	0.0	7.520	3-F
8	0.916	−0.125	0.0	0.0	0.400	0.0	1.200	1.750	0.760	0.0	8.160	3-Cl
9	0.905	−0.127	0.0	0.0	0.410	0.0	1.200	1.850	0.940	0.0	8.300	3-Br
10	0.826	−0.068	0.0	0.0	0.360	0.0	1.200	1.980	1.150	0.0	8.400	3-I
11	0.017	−0.376	0.0	0.0	−0.070	0.0	1.200	1.970	0.510	0.0	8.460	3-Me
12	0.916	−0.125	0.911	−0.406	0.400	−0.070	1.470	1.750	0.760	0.150	8.190	4-F, 3-Cl
13	0.905	−0.127	0.911	−0.406	0.410	−0.070	1.470	1.850	0.940	0.150	8.570	4-F, 3-Br
14	0.017	−0.376	0.911	−0.406	−0.070	−0.070	1.470	1.970	0.510	0.150	8.820	4-F, 3-Me
15	0.916	−0.125	0.916	−0.125	0.400	0.110	1.750	1.750	0.760	0.700	8.890	3,4-di-Cl
16	0.905	−0.127	0.916	−0.125	0.410	0.110	1.750	1.850	0.940	0.700	8.920	4-Cl, 3-Br
17	0.017	−0.376	0.916	−0.125	−0.070	0.110	1.750	1.970	0.510	0.700	8.960	4-Cl, 3-Me
18	0.916	−0.125	0.905	−0.127	0.400	0.150	1.850	1.750	0.760	1.020	9.000	4-Br, 3-Cl
19	0.905	−0.127	0.905	−0.127	0.410	0.150	1.850	1.850	0.940	1.020	9.350	3,4-di-Br
20	0.017	−0.376	0.905	−0.127	−0.070	0.150	1.850	1.970	0.510	1.020	9.220	4-Br, 3-Me
21	0.017	−0.376	0.017	−0.376	−0.070	−0.310	1.970	1.970	0.510	0.520	9.300	3,4-di-Me
22	0.905	−0.127	0.017	−0.376	0.410	−0.310	1.970	1.850	0.940	0.520	9.520	4-Me, 3-Br

[a] See text for sources. [b] Log $1/ED_{50}$ (mol/kg): antagonism of *N,N*-dimethyl-2-bromophenethylamines to adrenaline in the rat.[4]

orthogonal basis vectors for lipophilicity, stericity, etc.** The extent to which **Z** fails to fall in the optimum SS is reflected by the angle θ between **Z** and SS. The cosine of this angle is equal to the (uncorrected) correlation coefficient r *via* the scalar product[11]

$$\mathbf{Z}'_{obsd} \cdot \mathbf{Z}'_{calcd} = \Sigma \mathbf{Z}'_{obsd}\mathbf{Z}'_{calcd} = \sqrt{\Sigma \mathbf{Z}'^2_{obsd}}\sqrt{\Sigma \mathbf{Z}'^2_{calcd}} \cos\theta \quad (5)$$

where the $\mathbf{Z}'$ are expressed relative to their respective means. The per cent contribution of **P** to the total observed electrical effect is then

$$\%\,\mathbf{P} = 100\%b/(|a| + |b|) \quad (6)$$

where %**P** is signed according to the quadrant in which **Z** (or its projection) falls in SS_e (quadrant I $[+a,+b]$, II $[+a,-b]$, III $[-a,-b]$, IV $[-a,+b]$). Approximate %**P** for some common Hammett-type σ's are[10] 30% for σ_m, 53% for σ_p, 56% for σ_p^-, 29% for σ_m^+, and 79% for σ_p^+, but only 18% for σ_I, 73% for σ_R, and 9% for $\sigma^*_{XCH_2COOR}$. It is possible to calculate ideal blended σ values by the reverse of eq 4. In addition, by a judicious choice†† of weighting factors a and b, the $\sigma_{x\%\,P}$ will remain normalized.[10,††] The advantages of $\sigma_{x\%\,P}$ are for those without benefit of digital computers, for graphical representation of the results, as well as to benefit from the optimization of S and P. The full papers should be consulted for complete details.[10,#]

Using the data in Table I, higher levels of factorization, both from the point of view of dependent as well as independent variables, can now be considered. Values of r_v are those used by Cammarata (use of E_s in place of this variable gave identical results); other values are from our earlier paper,[4] Brown and Okamoto,[7] or Unger.[10] Since it is meaningless to sum radii, fully factored r_v have been used throughout. Note that π values from the phenoxyacetic acid system have been used and hence $\pi_m \neq \pi_p$, while in Cammarata's analysis it is assumed that $\pi_m = \pi_p$.

If the data are first factored into two sets of monosubstituted meta and para derivatives, it is found that the best single variable is $r_{v,m}$ or $r_{v,p}$, respectively. Unfortunately, with only six points one cannot profitably check out higher order equations. Completely parallel results are obtained by combining the two sets of monosubstituted derivatives, eq 7

$$\begin{array}{lccc} & n & r & s \\ pC = 1.33\,(\pm 0.32)\,r_{v,m} + 2.41 & 11 & 0.988 & 0.104 \\ \quad (\pm 0.32)\,r_{v,p} + 2.90\ \ (\pm 0.86) & & & \end{array} \quad (7)$$

(figures in parentheses are 95% confidence limits). No other terms are selected at any Level of factorization although all linear combinations of the indicated (*cf.* eq 8-13) variables were tested.

Considering the 12 disubstituted derivatives in the same systematic fashion, one finds a different SAR‡‡ which, although not identical with, is quite similar to that found for all 22 derivatives.§§ Equations 8-13 correlate all 22 data points and are arranged by the level of factorization. The "best equation" is that with both the lowest s and all terms justified by a sequential F test. In level I a sequential $F_{\alpha=0.05}$ test indicates the following significance of terms: $\sigma^+ \rightarrow \pi \rightarrow r_{v,p}$; in II, $\sigma_m^+ \rightarrow \sigma_p^+ \rightarrow \pi \rightarrow r_{v,p}$; in III, $\mathbf{P}_p \rightarrow S_p \rightarrow S_m \rightarrow \pi \rightarrow r_{v,p}$; in IV, $S_m \rightarrow \pi_m \rightarrow r_{v,p}$; in V, $\sigma_m^+ \rightarrow \pi_m \rightarrow \sigma_p^+ \rightarrow \pi_p$; in

**Preliminary results indicate that $\pi_{benzene}$ and E_s are orthogonal to S and P.

†† If $x = \%\,P/100$, then let $a = \pm(1 - b^2)^{1/2}$, where $b = \pm(x^2/1 + 2[x^2 - x])^{1/2}$ and use $+a,+b$ for quadrants I and III and $+a,-b$ for quadrants II and IV.

‡‡ This perhaps suggests, *but by no means proves because of the limited nature of the data set*, a special dichotomy between mono- and disubstituted derivatives. If bonding to the receptor is the rate-determining step, then in the presence of disubstitution the transition state changes slightly from one based on size (allowing rotation of the phenyl, *e.g.*) to one based on hydrophobicity and electronic effects as well. Note that the monosubstituted data are not much different in response from the disubstituted so this is probably not an artifact of the data (*i.e.*, mono- ≠ disubstitution). Note also the very bad fit for the 4-phenyl substituent (which was not included in the analysis). This substituent may be so large that it acts to effectively prevent rotation or fit of the phenyl (*cf.* ref 4). One must proceed cautiously, however, because of covariance as expressed in eq 15. The next best equation ($pC = 1.06\pi - 1.30\sigma^+ + 0.57r_{v,p} + 7.07$; $n = 11$, $r = 0.975$, $s = 0.159$) or the one obtained by substituting for $r_{v,p}$ ($pC = 1.36\pi - 1.77\sigma^+ - 0.23r_{v,m} + 8.10$; $n = 11$, $r = 0.960$, $s = 0.202$) correlate less well, but are not completely poor choices.

§§ It is interesting that none of the eq 8-13 would have been selected by a stepwise regression which would have stopped after addition of $r_{v,m}$ to $r_{v,p}$ at $s = 0.201$. Thus, this entire very profitable "standard deviation sink" would have been overlooked if one had relied entirely on stepwise regression.

VI, $\pi_p \rightarrow \sigma^+ \rightarrow \pi_m \rightarrow r_{v,p}$. In eq 9, one obtains essentially the same result using either $r_{v,p}$ or $r_{v,m}$.

It is apparent that factorization of the independent variables *in this system* is not accompanied by an improvement in fit; that is, the independent variables are additive by position. [Of course, if a higher order term (*e.g.*, π^2) had been important, one would not expect "additivity" since combining effects could push the response over the maximum. Thus, additivity has two meanings: lack of interaction between substituents yielding a simple sum of effects or a lack of interaction coupled with a linearity in response.]

Compared to eq 7, electronic and hydrophobic terms now dominate the picture. Although r_v explains a large amount of the variance, the coefficients show it to be of average importance. It does not pay to factor π if one includes a term in $r_{v,p}$; note that $r_{v,m}$ does not enter eq 8–13. This

	r	s	
Level I ($r_{v,m}$, $r_{v,p}$, π, σ^+)			
$pC = 0.82\ (\pm0.27)\pi - 1.02\ (\pm0.45)\sigma^+ + 0.62\ (\pm0.43)r_{v,p} + 7.06\ (\pm0.55)$	0.964	0.164	(8)
Level II ($r_{v,m}$, $r_{v,p}$, π, σ_m^+, σ_p^+)			
$pC = 0.83\ (\pm0.27)\ \pi - 1.21\ (\pm0.60)\ \sigma_p^+ - 0.91\ (\pm0.49)\sigma_m^+ + 0.64\ (\pm0.43)\ r_{v,p} + 7.02\ (\pm0.55)$	0.966	0.163	(9)
Level III ($r_{v,m}$, $r_{v,p}$, π, $\mathbf{S}_m$, $\mathbf{P}_m$, $\mathbf{S}_p$, $\mathbf{P}_p$)			
$pC = 0.86\ (\pm0.30)\ \pi + 0.47(\pm0.26)\mathbf{S}_m - 0.36\ (\pm0.21)\ \mathbf{S}_p - 0.92\ (\pm0.61)\mathbf{P}_p + 0.62(\pm0.49)\ r_{v,p} + 7.08\ (\pm0.62)$	0.967	0.167	(10)
Level IV ($r_{v,m}$, $r_{v,p}$, π_m, π_p, $\mathbf{S}_m$, $\mathbf{P}_m$, $\mathbf{S}_p$, $\mathbf{P}_p$)			
$pC = 0.85\ (\pm0.30)\ \pi_m - 0.47\ (\pm0.27)\ \mathbf{S}_m + 1.64\ (\pm0.26)\ r_{v,p} + 5.83\ (\pm0.46)$	0.961	0.170	(11)
Level V ($r_{v,m}$, $r_{v,p}$, π_m, π_p, σ_m^+, σ_p^+)			
$pC = 0.83\ (\pm0.27)\ \pi_m + 1.33\ (\pm0.20)\ \pi_p - 0.92\ (\pm0.50)\ \sigma_m^+ - 1.89\ (\pm0.57)\ \sigma_p^+ + 7.80\ (\pm0.17)$	0.966	0.164	(12)
Level VI ($r_{v,m}$, $r_{v,p}$, π_m, π_p, σ^+)			
$pC = 0.82\ (\pm0.28)\ \pi_m + 0.65\ (\pm0.51)\ \pi_p - 0.94\ (\pm0.49)\sigma^+ + 0.85\ (\pm0.70)\ r_{v,p} + 6.77\ (\pm0.91)$	0.965	0.166	(13)

supports Cammarata's point (eq 2) that there is a steric effect in the para position, although no $r_{v,p}$ term appears in eq 12. In this equation, the coefficient with π_p is larger than π_m. Since π and r_v (for these six substituents) are not really independent (r^2 for correlation between $r_{v,m}$ and π_m is 0.7 and 0.75 for $r_{v,p}$ and π_p), it is not clear whether there is an increased hydrophobic effect by para substituents or if there is a steric effect in addition to the hydrophobic effect.

The most interesting equation is 10. In this result it is seen that the newly formulated **S** and **P** do as well as σ^+. This is exciting because while **S** is based on σ', **P** is not related to any given model. It is simply the electronic component of the substituent effect not included in the inductive-field effect **S**. The fact that **S** and **P** correlate a wide variety of data very well[10] constitutes direct proof of the long held assumption that σ is made up of two components, inductive and resonance. While in the present instance these parameters do not offer a great advantage over σ^+, it is most likely that in the systems differing from the model σ's there will be distinct advantages. Also, when the mechanism is not known, use of **S** and **P** can be used to estimate % **P** and thus $\sigma_{x\%\,\mathbf{P}}$.

No $\mathbf{P}_m$ term is found in eq 10 (0% $\mathbf{P}_m$,72% $\mathbf{P}_p$). This appears to be an artifact of the choice of and "noise" in the data which becomes apparent when these σ^+ are related to **S** and **P**.

Equation 14 indicates 30% $\mathbf{P}_m$ and 75% $\mathbf{P}_p$, in excellent

	n	r	s	
$\sigma^+ = 0.77\ (\pm0.06)\mathbf{P}_p + 0.26\ (\pm0.02)\mathbf{S}_p + 0.20\ (\pm0.06)\mathbf{P}_m + 0.47\ (\pm0.02)\mathbf{S}_m - 0.01\ (\pm0.02)$	22	0.998	0.020	(14)

agreement with data from physical-organic chemistry (*vide supra*).[10] Note the interrelationships of the coefficients. Therefore, the loss of the $\mathbf{P}_m$ term ($F_{\alpha=0.05}$ fails to justify addition) is due to the small amount of variance contributed by the meta substituents. The $\mathbf{P}_m$ term is also the last to enter eq 14 by a sequential F test. Finally, construction of either $\sigma_{30\%\,\mathbf{P}m,\ 75\%\,\mathbf{P}p}$ or $\sigma_{0\%\ \mathbf{P}m,\ 72\%\,\mathbf{P}p}$ and replacement in eq 10 gives nearly identical results.

The very constant coefficient with the π term in eq 8–13, as long as an $r_{v,p}$ term is also included, speaks strongly for the validity of considering a hydrophobic effect from both positions. Note that in eq 13, where π has been factored and $r_{v,p}$ included, the coefficients with the two π terms are close in value.

Cammarata[3] has given a thoughtful set of guides to be used in the formulation of mathematical SAR models, and using these has arrived at a set of equations from which one might select eq 2 as being the "best" model for the SAR of the β-haloamine adrenergic blocking agents. We can now apply our more general criteria to eq 2 and 8–13. Equation 2 fails to meet criterion 5; that is, there are no examples, to our knowledge, in the field of physical-organic chemistry to support a model in which substituents contribute electronically from the meta but not para position. Equation 2 also fails to model electronic effects for the carbonium ion intermediate proposed by Chapman and Triggle.[5] If there is no term in eq 2 for the kind of effect modeled by σ_p^+ or $\mathbf{P}_p$, why does eq 2 give such a high correlation? A possible explanation is afforded by the observation of the following "accidental" correlations.

	n	r	s	
$r_{v,p} = 0.53\ (\pm0.48)\pi_p + 1.38\ (\pm0.36)$	6	0.840	0.188	(15a)
$r_{v,p} = 0.73\ (\pm0.25)\pi_p - 1.05\ (\pm0.69)\sigma_p^+ + 1.26\ (\pm0.18)$	6	0.983	0.073	(15b)

Addition of σ^+ is significant at $F_{\alpha=0.01}$ ($F_{1,3} = 86.83$; $F_{1,3;\,\alpha=0.01} = 34.12$) which again points to the danger of working with limited data possessing a narrow range of response. Bromine and chlorine are essentially identical in electronic effects, relatively close in size, and differ appreciably only in hydrophobicity.

Finally, the results of eq 8–13 show the desirability of examining more than one regression equation.[3] It seems fairly certain that $r_{v,m}$ is not an important factor, that there is a strong resonance factor from both positions of substitution, and that the hydrophobic effect need not be factored, being equivalent from both positions. In accord with criterion 4, therefore, we would select eq 8 as the "best equation," even though there is a slightly lower s with eq 9. The in-

clusion of a σ^+ term excludes II as an important species in the rate-determining step unless the transition state is "late" and essentially resembles IV.

References

(1) J. G. Topliss and R. J. Costello, *J. Med. Chem.*, **15**, 1068 (1972).
(2) N. R. Draper and H. Smith, "Applied Regression Analysis," Wiley, New York, N. Y., 1966, Chapter 6.
(3) A. Cammarata, *J. Med. Chem.*, **15**, 573 (1972).
(4) C. Hansch and E. J. Lien, *Biochem. Pharmacol.*, 17, 709 (1968).
(5) (a) N. B. Chapman and D. J. Triggle, *J. Chem. Soc.*, 1385, 4835 (1963); (b) N. B. Chapman, K. Clarke, and R. D. Strickland, *Proc. Roy. Soc., Ser. B*, **163**, 116 (1965); (c) C. F. Hammer, S. R. Heller, and J. H. Craig, *Tetrahedron*, **28**, 239 (1972); (d) G. E. Garner, R. A. Woodbury, and E. E. Elko, *J. Pharmacol. Exp. Ther.*, **162**, 38 (1968); (e) B. Belleau, *Can. J. Biochem. Physiol.*, 36, 731 (1958).
(6) J. D. P. Graham and M. A. Karrar, *J. Med. Chem.*, 6, 103 (1963).
(7) H. C. Brown and Y. Okamoto, *J. Amer. Chem. Soc.*, **80**, 4979 (1958).
(8) R. W. Taft, Jr., and I. C. Lewis, *ibid.*, **80**, 2436 (1958); **81**, 5343 (1959).
(9) C. G. Swain and E. C. Lupton, Jr., *ibid.*, **90**, 4328 (1968).
(10) S. H. Unger, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Mass., Sept 15, 1970.
(11) H. H. Harman, "Modern Factor Analysis," 2nd ed, University of Chicago Press, Chicago, Ill., 1967, Chapter 4.

Synthesis and Structure–Activity Relationships of Fibrinolytic 1,ω-Diphenyl-1,ω-alkanediamines

Leonard J. Fliedner, Jr.,* Melvyn J. Myers, Joseph M. Schor, and Irwin J. Pachter

Endo Laboratories, Inc., Garden City, New York 11530.

Received January 31, 1973

The promising results obtained in animal clot lysis experiments with the fibrinolytic bis(tetrahydroisoquinolines) VI prompted the preparation of a related series of 1,ω-diphenyl-1,ω-alkanediamines V. As measured in the standard rat screen (ip) the compounds were found to possess similar fibrinolytic activity. Structural variations discussed include N-substitution, aromatic substitution pattern, and chain length. The three-step synthesis generally involved Friedel–Crafts acylation of an aromatic compound I with a dibasic acid chloride II to yield a bis(ketone) III. The bis(oxime) or bis(methoxime) derivatives IV of the latter were then reduced to the corresponding V.

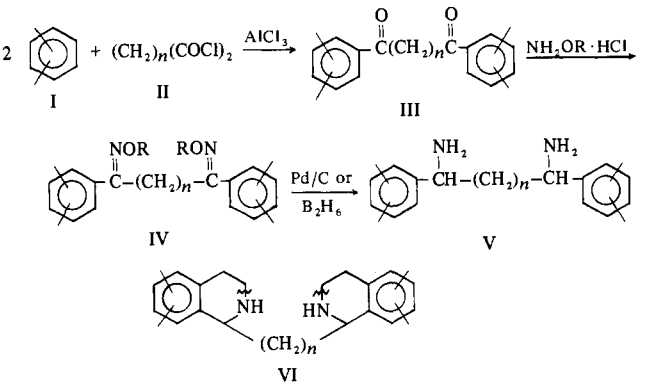
One of the goals of antithrombotic therapy is the development of effective, readily available, fibrinolytic–thrombolytic agents for both acute and prophylactic use. The encouraging results obtained in animal clot lysis experiments with the synthetic bis(tetrahydroisoquinolines) VI† have been described previously along with their structure-activity relationships.[1] Continued molecular modification in this area has uncovered a series of 1,ω-diphenyl-1,ω-alkanediamines V with similar fibrinolytic potential. Their relationship to the bis(tetrahydroisoquinolines) VI may be envisaged by formally cleaving the 2–3 bonds in the heterocyclic rings of the latter.

In general, the potencies of the two classes of compounds were comparable in the standard rat screen, although the most potent VI was approximately six times more active than the most potent V. A limited number of V have been reported previously;[2,3] however, our literature searches have revealed no detailed pharmacological investigations on this class of compounds.

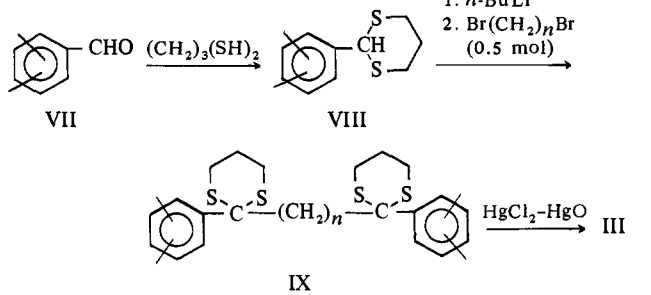
Chemistry. The general synthetic route (Scheme I) to bis(amines) V has been described previously;[2,3] and with suitable modifications depending on the nature of the aromatic substituents, this approach has been consistently employed.

The bis(ketones) III were generally available *via* Friedel–Crafts acylation of 2 mol of the aromatic compound I with 1 mol of a dibasic acid chloride II (Table I, methods A, B, and C). Tetrachloroethane was found to be an excellent solvent for most of these reactions, frequently giving better yields and purer products than literature procedures using either no solvent or other standard solvents (*cf.* Table I). Since acylation of aromatic compounds bearing two different activating substituents would yield a mixture of

Scheme I



Scheme II



products, the ketones **11–13** were prepared by the Corey dithiane procedure.[4] As shown in Scheme II, the substituted aromatic aldehydes VII were converted to their dithiane derivatives VIII by reaction with 1,3-propanedithiol. Alkylation of 2 mol of VIII with 1 mol of the appropriate 1,ω-dibromoalkane in the presence of 2 mol of *n*-butyllithium yielded the bis(dithianes) IX which were in turn hydrolyzed to the desired bis(ketones) III.[5] The dithiane procedure was also used for the synthesis of **1**, since attempted Friedel–

†Roman numerals refer to Schemes I and II and arabic numerals refer to compounds in the tables.