MOLECULAR PHARMACOLOGY, 6, 195–205

Electronic Molecular Pharmacology: The Benzothiadiazine Antihypertensive Agents

II. Multiple Regression Analyses Relating Biological Potency and Electronic Structure

ARNOLD J. WOHL

Department of Pharmacology, Schering Corporation, Bloomfield, New Jersey 07003

(Received November 12, 1969)

SUMMARY

Extended Hückel Theory calculations have been carried out on a number of antihypertensive benzothiadiazine-1,1-dioxides in their preferred tautomeric and conformational structures [A. J. Wohl, *Mol. Pharmacol.* **6**, 189 (1970)] in an attempt to define by multiple regression analyses the electronic and steric factors which quantitatively control biological potency. The results indicate that position 5 in the benzenoid portion of the molecule is a center for receptor-drug electron transfer, determined by the energy of the most energetic occupied molecular orbital. Also of importance is the degree of positive charge localized on atom 4 and the substituent group on position 3, as well as the charge on the sulfone oxygen atoms. An intimate sequence of molecular events comprising the mechanism of action is proposed, and a model of receptor structure is offered which accounts for the competitive behavior of these compounds with Ba^{++} and Ca^{++} and is used to predict the potencies of a wide variety of nonselected compounds with high accuracy.

INTRODUCTION

It is clear that both the type and degree of biological activity of chemical compounds must be related to their chemical structure. Attempts to deduce such relationships in the past, however, have had only limited success qualitatively, and virtually no success at all in a quantitative sense. Recently, however, multiple regression techniques, as used by Hansch (1), employing a substituent-parameter approach, and by Cammarata and Stein (2) and Neely *et al.* (3), applying the indices of molecular electronic structure provided by molecular orbital calculations, have yielded useful results.

A recent theoretical work by Cammarata (4) builds up multiple regression equations for biological activity based principally on Klopman's concept of charge- and frontier-controlled reactions (5, 6). These analyses led to theoretical regression equations for frontier-controlled reactions, based upon frontier orbital charge density, and for charge-controlled reactions, based upon atomic charge and superdelocalizabilities. The equations to be reported support this conceptualization for the benzothiadiazine antihypertensive agents.

Still lacking, however, is an example of the application of multiple regression analyses to a large series of chemical agents in which the type and degree of biological action are

well characterized, accurately measurable, and directly referable to a disease process against which the drugs are useful. That this lack exists is understandable, in that a number of conditions must be fulfilled before such an undertaking is practicable.

1. It is necessary that all chemical agents included in both the retrospective and predictive regression equations be shown to exert a particular biological effect via the identical molecular mechanism of action, insofar as is possible.

2. It is also necessary to know the nature of the chemical species responsible for biological activity. This restriction is meant to include conformational as well as ionization structure, as discussed in the preceding paper (7).

3. Finally, it is vital that the biological activity being used as the dependent variable in the regression analysis be measured in a simple medium, at least for preliminary study. This restriction ensures that the drug-receptor interaction will be uncomplicated quantitatively by membrane transport, metabolism, distribution of drug, and so on. It should also be clear that the actual measurement of activity must be highly accurate. These conditions virtually presuppose procedures *in vitro*, which, while perhaps not always physiologically relevant, may yield valuable additional information regarding mechanism of action which may permit the construction of verifiable hypotheses regarding receptor structure.

The benzothiadiazine-1,1-dioxides (8, 9) meet all of the above ideal requirements. The antihypertensive mechanism of action is consistent with recent experiments showing the compounds to be competitive antagonists of Ca^{++}-induced vasoconstriction *in vitro* (10–12); the chemical species is the neutral molecule;[1] and the compounds have been studied with regard to their conformational and tautomeric preferences (7). As a final advantage, the potency of each compound in the series as a competitive antagonist of Ca^{++}-induced rat aortic vasoconstriction *in vitro* [Ba^{++} was used for ease of solubility (10)] is easily measured with high accuracy by standard techniques.

[1] J. G. Topliss, personal communication.

METHODS

Since no information is available regarding the electronic nature of the benzothiadiazine–"Ca^{++} receptor" interaction, a variety of quantum chemical indices obtained from Extended Hückel Theory calculations, performed as in the preceding paper (7), were considered as independent variables.

The molecular orbital calculations performed by EHT[2] yield net atomic charges for each atom in the molecule; these are further partitioned into σ- and π-electronic charges [$q_{(N)}{}^{\sigma}$, $q_{(N)}{}^{\pi}$].

Summed regional π- and σ-charges were calculated for 3-substituents, as Streitweiser (13) has recently done for a reaction problem in organic chemistry. This is symbolized $q(3R)^{\sigma \text{ or } \pi}$.

The energy levels of the molecular orbitals are also part of the computer printout. The highest occupied (most energetic) molecular orbital, which contains a full electron complement (of 2 electrons), frequently participates in chemical reactions (14, 15) and has been found to have biological significance (16, 17) along with the least energetic (lowest unoccupied) orbital. These numbers are negative quantities. As multiplied by -1, these terms are symbolized E_{HOMO} and E_{LEMO}, respectively.

Additional quantities are the approximate (frontier) nucleophilic and electrophilic superdelocalizabilities ($S'_{(N)}E$, $S'_{(N)}N$), where the subscript refers to the atom in question. These indices measure the tendency (18, 19) for incoming (interacting) electron-rich or electron-poor centers to be localized at an atom, and have found chemical as well as biological use (2, 20, 21).

[2] The abbreviations used are: EHT, extended Hückel theory; HOMO, highest occupied molecular orbital; LEMO, least energetic molecular orbital.

These superdelocalizabilities refer to π-electron orbitals, and thus probably relate to π-π drug-receptor interaction at specific atoms in the drug molecule.

An additional set of molecular orbital quantities are the atomic and bond polarizabilities (20). These quantities roughly measure the electric moment induced at an atom by itself—self-atom polarizability; and at two adjacent atoms—atom-atom polarizability. These are also reaction indices and have found chemical use (14). The frontier approximation employed in calculating superdelocalizability, which restricts attention to the HOMO and LEMO as the reaction-determining orbitals, was also used to calculate polarizabilities ($\pi'_{N,N}$, $\pi'_{N,N+1}$).

Although misordering of orbitals has sometimes been noted with EHT, a σ-orbital rather than a π-orbital occasionally appearing as the HOMO, calculations on the benzothiadiazines have invariably shown there to be a π-orbital at a very similar energy whenever this misordering of energy levels occurred. For the polarizabilities, as well as the superdelocalizabilities discussed above, the computations were done on the frontier π-orbital.

Once preliminary results utilizing the above information were obtained, which selected the particular atoms in the molecule that could be regarded, based on the results of the regression analyses, as being involved in the interaction with the receptor, attempts were made to refine the work by introducing several additional quantities.

Coulomb's law gives the direct electrostatic (charge-charge) interaction between two points (centers) of charge as

$$F = \frac{q_d\, q_r}{\epsilon R_{dr}} \tag{1}$$

where q_d is a charge on a drug atom d and q_r is the corresponding charge on receptor atom r, ϵ is the dielectric constant of the medium, and R_{dr} is the separation between atoms d and r. For all atoms in drug and all atoms in the receptor,

$$F = \sum_d \sum_r \frac{q_d\, q_r}{R_{dr}} \tag{2}$$

But we may consider the receptor as constant, and the dielectric constant as unchanged by drug. The preliminary results to be discussed implicated certain atoms in the drugs as participants in the interaction, and so unit positive charges symbolizing q_r were located at points in space topographically corresponding to the apparently involved drug atoms. One of seven such points is denoted by [A] (Fig. 1), and Coulombic interactions between all drug atoms and points [A] through [G] were calculated by

$$F^{[A]} = \sum_d \frac{q_d}{R_{d[A]}} \tag{3}$$

A second type of drug atom-receptor atom intermolecular energy may be sim-

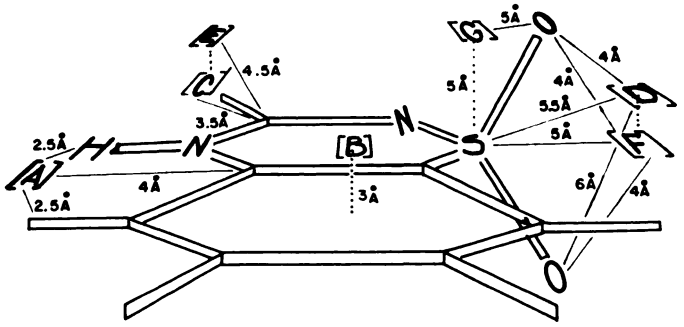


FIG. 1. *Perspective drawing of benzothiadiazine molecule, showing hypothetical "receptor sites" selected for intermolecular force calculations, with distances to nearest atomic centers of interest*

plified, for a series of molecules, as the field created by a series of charges upon the midpoint of a bond located a given distance away (23, 24). The midpoints of receptor intramolecular bonds were also considered to be located at points [A]–[G], and the electric field was calculated by

$$E^{[A]} = \sum_d \frac{q_d}{R^3_{d[A]}} \cdot \mathbf{r}_{d[A]} \qquad (4)$$

where $\mathbf{r}_{d[A]}$ is the vector distance. These two calculations were performed for all molecules in the series.

The following variables were employed in the preliminary and final regression analyses.

q_N^σ = net σ-charge on atom N
q_N^π = net π-charge on atom N
Q_N = net total ($\sigma + \pi$) charge on atom N
$q(3R)^{\sigma,\pi}$ = summed net regional charge over all atoms in $3R$ group (σ or π)
E_{HOMO} = energy$\cdot(-1)$ of HOMO, in electron volts
E_{LEMO} = energy$\cdot(-1)$ of LEMO, in electron volts
$S'_{(N)}N$ = approximate (frontier) nucleophilic superdelocalizability on atom N
$S'_{(N)}E$ = approximate (frontier) electrophilic superdelocalizability on atom N
$\pi'_{N,N}$ = frontier self-atom polarizability of atom N
$\pi'_{N,N+1}$ = frontier atom-atom polarizability between atom N and atom $N + 1$
$F^{[A]}$ = intermolecular Coulombic interaction energy for a molecule and point [A]
$E^{[A]}$ = electric field created at point [A] by a set of charges on a molecule

Except for the energy levels (E_{HOMO} and E_{LEMO}), which ranged between 7.1 and 12.6 eV, all other variables had values between -2 and $+2$.

All compounds studied were treated by EHT computation at the tautomer and conformer combination found to have the lowest energy in the preceding paper (7).

An appropriate measure of pharmacological potency in a series of competitive blocking agents is a quantity termed pA_2. This measurement, as used by Arunlakshana and Schild (25), and later by Van Rossum (26), is the negative logarithm of the molar concentration of antagonist which necessitates a doubling of agonist concentration for an equivalent effect when the antagonist is present. The pA_2 is mathematically equal to the drug-receptor dissociation constant, although partitioning uncertainties between tissue bath drug concentration and that in the biophase make the latter term unrealistic.

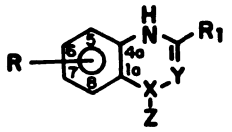
All compounds studied were evaluated and pA_2 determinations made as indicated by Van Rossum (26) and as was done for diazoxide (10). These values, as well as all compounds used in the study, are shown in Table 1.

The compounds studied were treated as individual "observations" in which the dependent variable, pA_2, was assumed to correlate with one or more of the numerous EHT–molecular orbital quantities treated as independent variables. The assumption implicit in this treatment is that receptor structure and mechanism of action hold constant while pA_2 varies throughout the series, presumably with one or more of the independent variables.

The standard stepwise multiple regression techniques used to treat these data can be evaluated by all three criteria: the multiple correlation coefficient (r), the standard error of the estimates ($S_{y\cdot x}$), and the F-statistic. Of the three, r is the least reliable as an estimate of "goodness of fit" of the derived equation to the Y-values, for r may increase to very high values ($\rightarrow$1.0), while $S_{y\cdot x}$ increases. This circumstance indicates multicolinearity—that is, high correlations among the molecular orbital variables treated—which mathematically results in a

TABLE 1

Structure and numbering of benzothiadiazines and similar compounds



COMPOUND NUMBER	R	R_1	X	Y	Z	pA_2 ± S.E.M.
1	7-AZA	H	S	N	O_2	3.76 ± 0.17
2	H	CH_3	S	N	O_2	3.76 ± 0.14
3	6-Cl	H	S	N	O_2	4.92 ± 0.15
4	H	H	S	N	O_2	3.62 ± 0.04
5	7-Cl	CH_3	S	N	O_2	5.11 ± 0.01
6	6-Cl	CH_3	S	N	O_2	5.40 ± 0.13
7	8-Cl	CH_3	S	N	O_2	4.39 ± 0.10
8	5-Cl	CH_3	S	N	O_2	3.94 ± 0.17
9	6,8-Cl_2	CH_3	S	N	O_2	4.63 ± 0.19
10	5,7-Cl_2	CH_3	S	N	O_2	4.68 ± 0.09
11	6,7-Cl_2	CH_3	S	N	O_2	6.12 ± 0.21
12	6-CH_3	CH_3	S	N	O_2	4.34 ± 0.16
13	7-CH_3	CH_3	S	N	O_2	4.26 ± 0.21
14	6-Cl-7-CH_3	CH_3	S	N	O_2	5.82 ± 0.13
15	5-CH_3-7-Cl	CH_3	S	N	O_2	4.63 ± 0.09
16	6,7,8-Cl_3	CH_3	S	N	O_2	5.85 ± 0.18
17	7-F	CH_3	S	N	O_2	4.68 ± 0.13
18	6-NO_2-7-Cl	CH_3	S	N	O_2	5.54 ± 0.13
19	7-Cl	CH_3	S	N	O	4.44 ± 0.17
20	6-Cl	⊙	S	N	O_2	5.41 ± 0.21
21	6-Cl	⇩	S	N	O_2	4.16 ± 0.20
22[a]	5-THIA-6-Cl	CH_3	S	N	O_2	3.60 ± 0.23
23[b]	6,7-Cl_2	C_5H_9	S	N	O_2	7.73 ± 0.13
24	6,7-Cl_2	Δ^1-C_5H_7	S	N	O_2	4.59 ± 0.11
25	6,7-Cl_2	Δ^2-C_5H_7	S	N	O_2	7.00 ± 0.17
26	6,7-Cl_2	Δ^3-C_5H_7	S	N	O_2	7.84 ± 0.21
27[c]	6,7-Cl_2	C_7H_7	S	N	O_2	6.64 ± 0.27
28	6-CH_3-7-Cl	H	S	N	O_2	5.46 ± 0.12
29	7-Cl	H	S	N	O_2	4.12 ± 0.25
30	7,8-Cl2	CH_3	S	N	O_2	5.34 ± 0.08
31	6-CH_3-7-Cl	CH_3	S	N	O_2	5.70 ± 0.15
32	5-CH_3	CH_3	S	N	O_2	3.86 ± 0.14
33	6-Cl-7-SH	Δ^3-C_5H_7	S	N	O_2	7.27 ± 0.18
34	6,7-Cl_2	CH_3	C	N	O	5.19 ± 0.07
35	6-Cl	CH_3	S	CH	O_2	4.85 ± 0.23
36	6-OCH_3	CH_3	S	N	O_2	4.78 ± 0.16

[a] Thieno Ring Fused
[b] This and Following Three Compounds are Cyclopentyl and pentenyl
[c] 1-Cycloheptatrienyl

high r. There are often significant intercorrelations between the molecular orbital variables; only when computer inclusion of intercorrelated variables, an unusual occurrence, decreases the $S_{y \cdot x}$ are such variables considered to be of significance (27).

The F-statistic measures the significance of the coefficient of a variable as it contributes to the over-all regression, and only highly significant contributions are considered.

The assumption is made that a high r, coupled with a low $S_{y \cdot x}$ for a regression equation, the variables whose coefficients are highly significant, indicates something of biological import. It is, strictly speaking, invalid to imply causality to a correlation, however high it may be. From a pragmatic point of view, however, if such a regression equation, applied to compounds other than those treated, yields a prediction of the pA_2 within the statistical limits of experimental data, it is felt that the equation in question is at least a working description of the elec-

tronic factors which determine the receptor attack of a molecule and its consequent biological effects.

RESULTS AND DISCUSSION

Preliminary data. Regression Eq. I is the result obtained from stepwise analysis of compounds 1–7, 9, 11–14, and 16–19 in Table 1. The pA_2 range of this set of 16 compounds is 3.76–6.12, a 230-fold potency spread.

$$pA_2(\hat{\Gamma}) = 28.34 - 1.96E_{\text{HOMO}} - 52.99S'_{(5)}E + 48.38S'_{(5)}N - 56.07S'_{(4)}E + 98.57S'_{(6)}N \quad \text{(I)}$$

$$r^2 = 0.947, \quad S_{y\cdot x} = 0.23, \quad F = 48.31 \quad (p < 0.0005)$$

The statistical significance of the regression equation is high, for, in addition to the high correlation coefficient and F-test value, the standard error of the estimate is such that the potential inaccuracy of the equation goes from a low of 3.8% to a maximum of 6.1%. Predictively, however, the equation is useful only for compounds included within the relatively narrow selection of structural types indicated. Predictions of activity for compounds possessing R groups other than H of CH_3 are uniformly low by an order or two of magnitude, and predictions for 5-substituted compounds such as Nos. 8, 10, 15, and 32 are uniformly too high by approximately an order of magnitude. It is clear, then, that Eq. I, while a regression of high statistical significance, covers too narrow a structural type to be useful.

There are salient features of Eq. I, however, which are of interest. No charge terms are included; the interaction would, however, seem to be charge-controlled (5, 6). The presence of E_{HOMO} would imply, since the coefficient for this term is negative, that the more energetic (the smaller this number, the higher the pA_2) is HOMO, the more facile electron donation from the compound would be, and consequently that the compounds act by electron donation, at least in part.

The negative coefficients for position 4 and 5 electrophilic superdelocalizabilities (SE) and the positive coefficients for the nucleophilic superdelocalizabilities (SN) at positions 5 and 6 confirm the above reasoning, and indicate more accurately the areas of importance for this effect within the structure of the drugs.

Equation II, which considers the 23 compounds 1–7, 9, 11–14, 16–21, and 23–27, and covers the 12,000-fold pA_2 range 3.65–7.84, is a more satisfactory description.

$$pA_2(\hat{\Gamma}) = 64.12 - 5.16E_{\text{HOMO}} + 55.09S'_{(5)}N + 115.78S'_{(6)}N + 5.16q(3R)^{\pi} \quad \text{(II)}$$

$$r^2 = 0.960, \quad S_{y\cdot x} = 0.28, \quad F = 62.31 \quad (p < 0.0005)$$

The statistical significance of Eq. II is roughly equivalent to that of Eq. I, with the standard error going from 3.6 to 7.7% and the high correlation coefficient and F-value corroborating the accuracy of the regression. Equation II, however, permits distinctly adequate predictions over a much wider range of structural prototypes than does Eq. I. The form of this equation is similar to that of Eq. I except for the apparently important inclusion of the regional charge on the 3-substituent, implying, in addition to electron donation from electron-rich areas on the aromatic ring, a charge-charge interaction which may be involved in stabilization of the drug-receptor complex.

The absence of the seemingly important sulfonyl group is, in all probability, an artifact of the basically limited compound selection, for all the compounds studied are quite similar. The involvement of these atoms, however, becomes clear when intermolecular forces are considered in the next section of this paper.

The general usefulness of Eq. II is still limited, however, for it is inadequate to predict (account for) the potency of 5-substituted compounds in that it consistently overestimates. Since it seemed that the breadth of structural types included in Eq. II was quite large, it was difficult to account for the inability of the regression equation to deal successfully with 5-substitution unless the assumption was made that position 5 was so intimately involved in the reaction mechanism that either steric or direct electronic interactions had to be considered in order to establish a generally useful regression equation.

Improved results. Equation III should be compared with Eq. I, for it is the result of performing a multiple regression analysis on the same 16-compound set as in the earlier equation, but also includes the intermolecular energy calculations discussed under METHODS.

$$pA_2(\hat{Y}) = 38.34 - 2.99E_{\mathrm{HOMO}} + 13.78S'_{(5)}N - 0.004\pi'_{6,7} + 0.04\pi'_{1a,8} + 11.03E^{[A]} \quad \text{(III)}$$

$$r^2 = 0.996, \quad S_{y\cdot x} = 0.07, \quad F = 202.57 \quad (p \ll 0.0005)$$

Equation III differs from Eq. I in two salient respects: a far better, almost perfect, fit of the experimental data has been obtained, and it shows the likelihood of a receptor point located at point [*A*] in Fig. 1. The 1a–8 and 6–7 polarizabilities are correlated with the *SN* of position 6 and consequently represent that term in Eq. I. It is evident from the size of the coefficient in $\pi'_{6,7}$ that this term is of limited importance. The electrophilic superdelocalizabilities at atoms 4, 5, and 6 are all highly correlated with $E^{[A]}$ and represent those terms of Eq. I in Eq. III. The interpretation of Eq. III, therefore, is the same as was given for Eq. I, and the greatly increased goodness of fit is likely the result of the inclusion, in the stepwise procedure, of $E^{[A]}$, the electric field upon a presumed receptor site located in the plane of the nucleus and in the area of atoms 4, 4a, and 5.

Despite the adequacy of Eq. III for predicting the activity of compounds falling within its narrow structural breadth, it fails quite as completely as the much less accurate Eq. I with regard to 3-substitution, although Eq. III is the first equation which satisfactorily accounts for the potency of 5-substituted compounds.

Equation IV includes the 24 compounds 1–9 and 11–25 in Table 1. While the requirements of compound variety have forced the addition of several additional terms, Eq. IV can be seen to be an entirely satisfactory regression equation spanning the 17,000-fold potency range of 3.60–7.84.

$$pA_2(\hat{Y}) = 1.95 + 8.63S'_{(1)}N - 135.08S'_{(4)}E + 78.14E^{[D]} + 54.52S'_{(3)}N - 0.03\pi'_{5,5} - 0.03\pi'_{1a,1a} + 61.60F^{[B]} + 1.49q(3R)^r \quad \text{(IV)}$$

$$r^2 = 0.994, \quad S_{y\cdot x} = 0.12, \quad F = 265.34 \quad (p \ll 0.0005)$$

Several interesting differences are clear between Eqs. III and IV. The E_{HOMO} requirement has disappeared, evidently because, in this much wider structural range, it is now of less determinative importance for activity than the limit imposed by the statistical significance threshold of the regression analysis. It has, in effect, been diluted in importance by the presence of large numbers of compounds that have roughly equal E_{HOMO} values but whose potency is primarily determined by other quantities.

The $E^{[A]}$ quantity found in Eq. III has been replaced by the correlated self-atom polarizability at position 5 (which allows correct predictions for the 5-thia compound (No. 22 in Table 1) and $S'_{(4)}E$. The sulfur and oxygen atoms are now shown to be im-

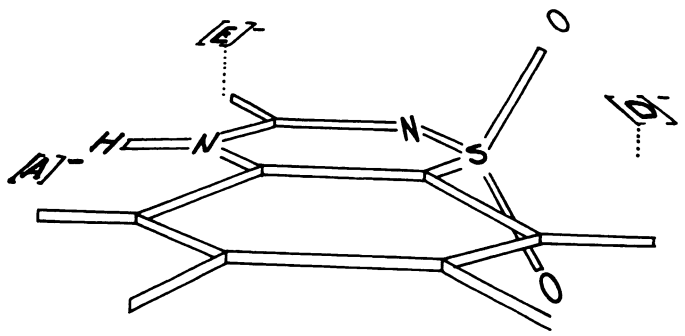


FIG. 2. *Perspective drawing of benzothiadiazine molecule, showing hypothetical "receptor sites" chosen from those in Fig. 1 by multiple regression analysis as having potency-correlated biological relevance*

TABLE 2

Correlation matrix for Eq. IV

	Y	q_σ^3	$S'N^{(1)}$	$S'E^{(4)}$	$E^{[D]}$	$S'N^{(3)}$	$\pi'_{5,5}$	$\pi'_{1a,1a}$	$F^{[E]}$
Y	1.0	.795	.375	-.484	.251	.088	-.283	-.306	.600
q_σ^3		1.0	.047	-.336	.084	-.210	-.276	-.126	.702
$S'N^{(1)}$			1.0	.158	.493	.059	.415	-.363	.166
$S'E^{(4)}$				1.0	.462	.065	.493	-.125	-.109
$E^{[D]}$					1.0	.067	.280	-.369	.368
$S'N^{(3)}$						1.0	.123	.003	-.448
$\pi'_{5,5}$							1.0	-.382	-.216
$\pi'_{1a,1a}$								1.0	.013
$F^{[E]}$									1.0

portant by the inclusion of $E^{[D]}$ and $S'_{(1)}N$, and thus permits tentative placement of another receptor site at $[D]$, making three altogether, $[A]$, $[D]$, and $[E]$, out of the seven originally chosen as likely candidates (Fig. 2).

An important qualification for multiple regression analyses, even when performed stepwise, is the possible complications induced by multicolinearity, that is, significant intercorrelations between supposedly independent variables. Table 2 shows the one such event which appears in Eq. IV, a high correlation between $F^{[E]}$ and $q(3R)^\tau$. The statistical significance (F-test) of the equation is, however, markedly reduced by excluding either term, and so it must be assumed that, despite intercorrelation, each term says overlapping but different things about the drug-receptor interaction.

Correlations, however, even of high statistical significance, are of only passing interest unless one is able to derive verifiable hypotheses from them and/or utilize them predictively. Table 3 shows the agreement between predicted and experimentally found pA_2 values using regression Eq. IV. In 6 of the 19 cases the deviation between predicted and experimental values slightly exceeds one standard deviation of the experimental values themselves. For the other 13 cases tabulated, the estimates of potency agree quite closely. The pA_2 range covered by these 19 compounds is some 12,000-fold; the

TABLE 3

Prediction of potency by Eq. IV

COMPOUND NUMBER	OBSERVED pA_2 ± S.D.	PREDICTED pA_2	DEVIATION
10[a]	4.68 ± 0.16	4.52	0.16
15	4.63 ± 0.14	4.68	-0.05
20[a]	5.41 ± 0.27	5.43	-0.02
21	4.16 ± 0.32	4.27	-0.11
22[a]	3.60 ± 0.27	3.59	0.01
23[a]	7.73 ± 0.31	7.83	-0.10
24	4.59 ± 0.11	4.57	0.02
25	7.00 ± 0.34	7.09	-0.09
26	7.84 ± 0.27	7.64	0.20
27	6.57 ± 0.47	6.83	-0.26
28	5.46 ± 0.12	5.13	0.33
29	4.12 ± 0.37	4.59	0.47
30	5.34 ± 0.13	5.20	0.14
31	5.70 ± 0.19	5.56	0.14
32	3.86 ± 0.13	3.68	0.18
33	7.27 ± 0.31	7.60	-0.33
34	5.19 ± 0.12	4.84	0.35
35	4.85 ± 0.24	4.73	0.12
36	4.78 ± 0.27	4.65	0.13

[a]Included in regression.

TABLE 4

Predicted and experimental values for Eq. IV

COMPOUND NUMBER	EXPERIMENTAL RANGE	MEAN	PREDICTED RANGE	MEAN	$\bar{X}_E - \bar{X}_P$
5[a]	1.59 - 2.01 μg/ml	1.79 μg/ml	1.59 - 2.77 μg/ml	2.19 μg/ml	-0.39 μg/ml
22[a]	18.4 - 192.2	59.4	45.1 - 78.3	61.7	-2.3
26[a]	0.001 - 0.011	0.004_2	0.003 - 0.006	0.004_5	-0.0003
27	0.034 - 0.12	0.089	0.037 - 0.064	0.049	0.04
28	0.43 - 1.49	0.80	1.30 - 2.25	1.71	-0.91
30	0.73 - 2.06	1.22	1.27 - 2.21	1.68	-0.46
32	21.0 - 63.4	39.0	52.7 - 91.7	69.5	-30.5
34[bb]	0.87 - 2.51	1.48	2.51 - 4.36	3.31	-1.83
35[c]	1.91 - 5.51	3.24	3.24 - 5.63	4.27	-1.03

[a]Included in Regression
[b]Quinazoline-4-one
[c]1,4-Benzthiazine

range of structural types exceeds the breadth included in the derivation of Eq. IV, compound 35 being a 1,4-benzothiazine rather than a 1,2,4-benzothiadiazine, and compound 34 being a 4-quinazolinone. It should be clear that Eq. IV has excellent predictive properties.

Table 4 shows agreement between experiment and prediction for three compounds included in Eq. IV and six compounds representing structural types which were not included in the derivation. The largest difference (last column) between experiment and prediction is only some 30 μg/ml of drug.

Several compounds not included in Table 1 are inactive, and several more are active but in a different way than would be apparent by comparisons with existing compounds. It is believed that an understanding of the activities of these compounds may better be gained by extrapolation of the results thus far reported. It is of interest to speculate on (a) the electronic mechanism of action of these compounds and (b) the structure of the receptor.

Electronic molecular mechanism of action. Without invoking any particular conceptualization of drug-induced receptor alteration,

we may propose that a drug molecule moves by diffusion in the biophase into equilibrium position for the formation of three, not necessarily simultaneous, weak molecular bonds at sites [*A*], [*D*], and [*E*] (Fig. 2). The molecule now occupies the receptor.

The receptor, however, normally activated by Ca^{++}, may only be composed of sites [*A*] and [*E*]. Sites [*A*] and [*E*] are chosen as likely centers for the normal Ca^{++} receptor because the distance between [*A*] and [*E*] is 3.5 A. The diameter of the singly hydrated Ca^{++} is 3.7 A (28). This interesting similarity may, of course, be of no more than coincidental interest, and does not affect the following discussion. The distance between sites [*D*] and either [*A*] or [*E*] is much greater, however, and it may be that [*D*] serves as the accessory site permitting stabilization of the receptor occupation.

With the drug molecule in 3-point attachment to the receptor, an electron or fraction thereof is donated to the receptor from the drug. With this event the charge distribution in the drug molecule changes so as to decrease markedly the binding at [*A*]. Binding at [*D*] and [*E*] is slightly enhanced by this exchange, but not sufficiently to overcome the effective lack of binding at [*D*]. The molecule is now easily disengaged from the receptor either by Ca^{++} or solvent forces, or by the changed charge characteristics of the receptor itself.

If this hypothesis is correct, at least in general outline, 4-methylated compounds, which possess very low $E^{[A]}$ values, should be inactive. This has been verified experimentally. Compounds substituted at position 3 with highly negatively charged substituents such as trifluoromethyl should also be inactive, again in keeping with experimental data. In accordance with predictions, compounds devoid of oxygen at position 1 are inactive, but this result is open to question because of the very poor solubility characteristics of such compounds.

The hypothesis does not account satisfactorily for the observation that 3-oxo derivatives possess activity, but in this case the 2-nitrogen carries a hydrogen atom and this structural difference complicates interpretation of the laboratory finding.

Structure of the receptor. This extrapolation is on at least as tenuous ground as the preceding discussion, but may also be open to test.

The "receptor" for the benzothiadiazines is thought to differ in possessing one more binding site, at [*D*], than the number required for activation by Ca^{++}. Sites [*A*] and [*E*] may be two anionic centers such as would be present in an organic phosphate. Site [*D*], which would seem logically to be a cationic center by virtue of its proposed interaction with oxygen on position 1 of the drug molecule, also appears to behave as an anionic site, according to the coefficients (positive) and the sign of $E^{[D]}$ (negative).

The above assignments are based on the regression equation data exclusively, but have interesting laboratory correlates. The 3-oxo compounds briefly discussed above are far more potent than inspection, based on the negatively charged 3-substituent, would indicate. Yet the EHT computation on such a compound indicates a much decreased negative charge on the 2-nitrogen, such that $E^{[D]}$ would decrease sizably in negative character and account for virtually all of the observed potency in view of the very large contribution made by it, as in Eq. IV.

Negatively charged 5-substituents, such as nitro, whose group net charge (EHT) is approximately -1.0, should and do increase the positive character of position 4 quite markedly while decreasing the electron-donating ability of the molecule; i.e., the compound should not act like other members of the series. Experimental observations confirm this supposition, in that 5-nitro-7-chloro-3-methylbenzothiadiazine 1, 1-dioxide stimulates the receptor in low doses and blocks it in higher doses, consistent with the hypothesis put forth.

It may not be possible to identify the receptor in more specific terms, but work is proceeding along these lines, utilizing both experimental findings and intermolecular

EHT computations along the lines suggested by Adam *et al.* (29), and will be reported in subsequent publications in this series.

CONCLUSION

Molecular orbital methods, combined with accurate laboratory data obtained under conditions in which structure and mechanism can be controlled and kept constant, permit predictive correlations as well as useful extrapolations of the pharmacological effects within a series of drugs. Limitations of the procedure are the requirements that (a) as broad a set of molecular orbital variables be included in regression analyses as is possible and (b) the compound series must be as wide as is consistent with synthetic capabilities. Within these limitations, highly accurate regression equations may be developed both for optimizing potency within a compound series and for understanding the underlying mechanism of action.

ACKNOWLEDGMENTS

The author is deeply indebted to Miss Lorraine M. Hausler and Mrs. Emma Eynon for painstaking collection of pA_2 values, to Dr. J. G. Topliss for synthesis of compounds, to Mrs. Margaret Hanks for needed help in interpretation of the regression analyses, and to Mrs. Barbara Uhrmacher for programming assistance. Neither this nor the preceding paper would have been possible without the help of the Quantum Chemical Program Exchange of Indiana University, which supplied the EHT program (QCPE 64) and the program for calculation of Cartesian atomic coordinates as input to the EHT program as PROXYZ (QCPE 94).

Special acknowledgments are due Mr. Robert Meese of the Singer-General Precision Corporate Data Center, for help with the operations and scheduling of SRU 1108 computer time.

REFERENCES

1. C. Hansch and T. Fujita, *J. Amer. Chem. Soc.* **86**, 1616 (1964).
2. A. Cammarata and R. L. Stein, *J. Med. Chem.* **11**, 829 (1968).
3. W. B. Neely, H. C. White and A. Rudzik, *J. Pharm. Sci.* **57**, 1176 (1968).
4. A. Cammarata, *J. Med. Chem.* **11**, 1111 (1968).
5. G. Klopman and R. F. Hudson, *Theor. Chim. Acta* **8**, 165 (1967).
6. G. Klopman, *J. Amer. Chem. Soc.* **90**, 223 (1968).
7. A. J. Wohl, *Mol. Pharmacol.* **6**, 189 (1970).
8. J. G. Topliss, L. M. Konzelman, E. P. Shapiro, N. Sperber and F. E. Roth, *J. Med. Chem.* **7**, 269 (1964).
9. A. A. Rubin, F. E. Roth, R. M. Taylor and H. Rosenkilde, *J. Pharmacol. Exp. Ther.* **136**, 344 (1962).
10. A. Wohl, L. M. Hausler and F. E. Roth, *J. Pharmacol. Exp. Ther.* **158**, 531 (1967).
11. A. J. Wohl, L. M. Hausler and F. E. Roth, *Life Sci.* 7 (Pt. I), 381 (1968).
12. A. J. Wohl, L. M. Hausler and F. E. Roth, *J. Pharmacol. Exp. Ther.* **162**, 109 (1968).
13. A. Streitweiser, Jr., *in* "Quantum Aspects of Heterocyclic Compounds in Chemistry and Biochemistry," Second Jerusalem Symposium, Jerusalem, Israel. In press.
14. A. Streitweiser, Jr., "Molecular Orbital Theory for Organic Chemists." Wiley, New York, 1961.
15. R. Hoffmann and R. B. Woodward, *J. Amer. Chem. Soc.* **87**, 2046 (1965).
16. B. Pullman and A. Pullman, *Proc. Nat. Acad. Sci. U. S. A.* **44**, 1197 (1958).
17. C. R. Merril, S. H. Snyder and D. F. Bradley, *Biochim. Biophys. Acta* **118**, 316 (1966).
18. K. Fukui, T. Yonezawa and C. Nagata, *J. Chem. Phys.* **27**, 1247 (1957).
19. K. Fukui, C. Nagata and A. Imamura, *Science* **132**, 87 (1960).
20. K. Fukui, K. Morokuma, C. Nagata and A. Imamura, *Bull. Chem. Soc. Jap.* **34**, 1224 (1961).
21. T. Yonezawa, I. Muro and H. Kato, *Mol. Pharmacol.* **5**, 446 (1969).
22. B. H. Chirgwin and C. A. Coulson, *Proc. Roy. Soc. Ser. A Phys. Sci.* **291**, 196 (1950).
23. H. DeVoe and I. Tinoco, Jr., *J. Mol. Biol.* **4**, 500 (1962).
24. R. Rein, P. Claverie and M. Pollak, *Int. J. Quantum Chem.* **2**, 129 (1968).
25. O. Arunlakshana and H. O. Schild, *Brit. J. Pharmacol. Chemother.* **14**, 48 (1959).
26. J. M. Van Rossum, *Arch. Int. Pharmacodyn.* **143**, 299 (1963).
27. J. C. R. Li, "Statistical Inference," Vol. 2, p. 120. Edwards, Ann Arbor, Mich., 1964.
28. L. J. Mullins, *Ann. N. Y. Acad. Sci.* **94**, 390 (1961).
29. W. Adam, A. Grimison, R. Hoffmann and C. Zuazaga de Ortiz, *J. Amer. Chem. Soc.* **90**, 1509 (1968).