Quantitative Structure–Activity Relationship of Phenyl *N*-Methylcarbamate Inhibition of Acetylcholinesterase

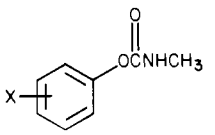
Amiram Goldblum,[1] Masafumi Yoshimoto,[2] and Corwin Hansch*

A quantitative structure–activity relationship (QSAR) has been formulated for a set of 269 phenyl *N*-methylcarbamates inhibiting fly head acetylcholinesterase. The data are from the extensive studies of Fukuto and Metcalf. The correlation equation obtained is $pI_{50} = 0.56MR_{3,4,5} + 1.56MR_2 - 0.61E_{s_2} - 0.94(\sum\sigma_{o,p}^- + \sigma_m)^2 + 1.43CHG - 0.23MR_2^2 - 5.24\mathcal{F}_{2,6}^2 + 3.46\mathcal{F}_{2,6} + 0.66RGMR - 0.62HB - 0.05MR_3^2 - 0.56E_{s_2}E_{s_6} + 3.46$ where $r^2 = 0.796$ and $s = 0.485$. In this expression MR refers to molar refractivity of substituents, E_s is the Taft steric parameter, $\mathcal{F}$ is an inductive parameter, RGMR refers to MR of certain parts of ring substituents, and CHG and HB are indicator variables for charged substituents and hydrogen bonding, respectively; r is the correlation coefficient and s is the standard deviation from the regression. The implication of this equation is discussed.

A large number of inhibition studies of acetylcholinesterase (AcChE) by various types of inhibitors have been performed in the past three decades, a great part of which were directed toward the rationalization of insecticide synthesis. These studies may provide us with basic information about the active sites of AcChE, their arrangement in space, and their forces of interaction with small molecules. Such information will be of value in the design of better inhibitors and possibly better insecticides.

In our discussion of the inhibition of cholinesterase we shall refer to it as AcChE when its source is housefly head (HFAcChE); any other source will be specifically noted (e.g., BAcChE = bovine erythrocyte AcChE).

In this paper we focus our investigation on a large set of phenyl *N*-methylcarbamates (PNMC) of structure I for



which I_{50} values (the molar concentration causing 50% inhibition of enzyme) were measured and documented by Metcalf and Fukuto with their collaborators (Metcalf and Fukuto, 1962, 1965a,b, 1967; Metcalf et al., 1960, 1962, 1963, 1964, 1965, 1966; Fukuto et al., 1967) for the inhibition of AcChE from homogenized fly heads.

It would be more desirable to have had k_i values (rate of inhibition) for direct reference to the mechanism since k_i values of the PNMC's are directly related to the proposed slow binding step:

$$E + I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EI \xrightarrow{k_2} EC \xrightarrow{k_3} E + C \qquad (1)$$

$$K_d = \frac{k_{-1}}{k_1} \qquad k_i = k_2/K_d$$

Abd Elroaf et al. (1977) and Nishioka et al. (1977) found k_2 to be more or less constant for a set of carbamates. The long reaction periods used in the I_{50} studies (Kolbezen et al., 1954) complicate the otherwise relatively straightforward relation of k_1 to I_{50} (O'Brien, 1976) because the decarbamylation rate constant, k_3, has a shorter $t_{1/2}$ than the span of time used for the I_{50} measurements. The parameter k_1 turns out to be nonsensitive to the concentration of inhibitor used (Nishioka et al., 1976) in the kinetic runs so it is of greater comparative value for results from different sources. Nevertheless, we believe that that data with a large set of congeners (even though they are less than ideal) can be of real value in enabling us to rationalize some of the salient events of the binding step and in exploring the characteristics of the binding site via physiocochemical correlation studies.

Since the previous quantitative structure–activity relationship (QSAR) of PNMC from this laboratory was proposed (Hansch and Deutsch, 1966), many more molecules have had their I_{50} values determined. Parameters for physicochemical correlations, as well as QSAR techniques, have been improved. The previous publication resulted in equations for three sets of monosubstituted PNMC's: a positive dependence of inhibition on π, a negative one on σ (for 3 and 4 substituents), and a positive dependence on σ for the ortho positions were found, this last position's activity being negatively influenced by steric effects.

Hetnarski and O'Brien (1972, 1973, 1975a,b) proposed a new charge transfer constant (C_T) obtained by the reaction of tetracyanoethylene with the PNMC. They formulated eq 2 and 3 (Hetnarski and O'Brien, 1975b) in para-substituted congeners

$$K_d = -3.48(\pm0.36)C_T - 2.64(\pm0.22)\pi + 5.88 \qquad (2)$$
$$n = 11 \quad r = 0.989 \quad s = 0.302$$

meta-substituted congeners

$$K_d = -1.23(\pm0.36)C_T - 1.96(\pm0.27)\pi + 2.44 \qquad (3)$$
$$n = 9 \quad r = 0.958 \quad s = 0.322$$

which n represents the number of congeners upon which the equation is based, r is the correlation coefficient, and s is the standard deviation from the regression equation. These are unconventional equations in that the authors have employed K_d instead of log K_d required in the usual Hammett treatment. In another publication (Hetnarski and O'Brien, 1972) a relationship of similar quality was drawn between the C_T constant and pI_{50} for HFAcChE.

Chiriac et al. (1975) used the recently devised (Simon and Szabadai, 1973) minimal steric difference parameter (MSD) in conjunction with other parameters (σ, MR, π, and C_T) and indicator variables (HB = hydrogen bonding; EC = charged substituents) to correlate pI_{50} for 97 PNMC's of the set reported by Metcalf and Fukuto et al. They derived eq 4:

$$pI_{50} = -0.54\sigma + 0.057MR + 0.64EC + 0.48HB + 0.058C_T - 0.16MSD + 5.16 \qquad (4)$$
$$n = 97 \quad r = 0.800 \quad s = 0.768$$

Department of Chemistry, Pomona College, Claremont, California 91711.

[1]Visiting scientist from the Hebrew University of Jerusalem, Israel.

[2]Visiting scientist from the Sankyo Co., Tokyo, Japan.

0021-8561/81/1429-0277$01.25/0

Their attempt to use π instead of MR gave an equation having $r = 0.713$ and $s = 0.891$ for the same 97 congeners. The confidence limits on the parameters of eq 4 were not reported so it is not clear just how important a role MSD plays. In this study Chiriac et al. (1975) selected the 3,5-diisopropyl congener as the reference molecule (MSD = 0). If the MSD term is significant, it implies a low sensitivity of PNMC to alterations in the molecular backbone of the 3,5-diisopropyl molecule. Though a low sensitivity might well be the case (since a multitude of different structures are able to interact with the AcChE as inhibitors), the use of MSD for structure–activity relations is still not well understood. It is built on the assumption that one can select an ideal "lead" molecule with minimal steric interactions which at best supplies us with information about the sensitivity of a system to structural differences. It cannot serve as a physicochemical parameter since, ab initio, it is not related to specific forces or energies.

In a recent elegant study of the binding of PNMC's by BAcChE enzyme, Nishioka et al. (1977) derived eq 5. In

$$\Delta \log 1/K = 1.40\pi_{2,3} + 0.31\pi_4 + 1.66\sigma_1^{\circ} - 1.78\sigma_2^{\circ} + 0.17E_s + 0.77\mathcal{F} + 1.36\mathrm{HB} + 0.07 \quad (5)$$

$$n = 53 \qquad r = 0.947 \qquad s = 0.238$$

this expression, $\pi_{2,3}$ refers to the sum of the π constants for substituents in the 2 and 3 positions, the σ_1° ($\rho > 0$) term is for all ortho substituents and for electron-withdrawing para substituents (NO_2, CN, COR, and SO_2CH_3) while σ_2° ($\rho < 0$) applies to all meta substituents and the rest of the para substituents. E_s and $\mathcal{F}$ are used only for ortho substituents, and the hydrogen-bonding term is an indicator variable taking the value of one for ortho OR and meta COR, CN, NO_2, and $N(CH_3)_2$. On the basis of this equation, they suggest a biphasic dependence of the transition state on σ for the slow binding step with a change in mechanism when electron-withdrawing groups are involved. Their postulate that electron donors at the ortho position do not participate in the mechanism involving a negative dependence on σ (carbonyl oxygen protonation) was explained by the steric repulsion found for that position. The proposed mechanism involves a tetrahedral transition state at the carbonyl carbon.

The structure of 3,5-di-CH_3-4-SCH_3PNMC was recently determined by X-ray crystallography (Takusagawa and Jacobson, 1977); they found the carbamate group to be perpendicular to the plane of the ring. No special intermolecular interactions were found to be present in the crystal. In a dipole moment study of some PNMC's (Exner and Bláha, 1977), the results conformed with the supposed noncoplanarity of the ester and aromatic ring for the 4-Cl-, 4-NO_2-, and unsubstituted PNMC. The results of log P measurements (Fujita et al., 1974) may indicate the PNMC's conformation in aqueous solution. All ortho-substituted compounds were found to have lower log P values than their meta and para analogues. It is often found that ortho substitution leads to higher log P values as a result of interference with hydration (Leo et al., 1971). To explain the finding of any opposite effect, we suggest that noncoplanarity of the carbamate side chain with the aromatic ring is the rule with ortho substituents; hence these substituents may not transmit an electronic effect to the carbonyl group by conjugation but may do so inductively or via a field effect with or without solvent aid.

In addition to the above-mentioned QSAR, many nonquantitative SAR studies of carbamate inhibitors of AcChE have been carried out. Kohn et al. (1965) noted the increasing influence of activity of α branching on meta and alkyl substituents and the negative effect of having too long a "tail" in that position. Kolbezen et al. (1954) pointed out the importance of "fit" and the higher activity of electron donors. Metcalf and Fukuto (1965a) stressed the importance of size, shape, and charge for high inhibitory activity. They used a plaster model of acetylcholine's extended conformation to "test" molecular models of PNMC for goodness of fit (lock and key theory) and concluded (Metcalf and Fukuto, 1967; Mahfouz et al., 1969) that bulkiness, hydrophobicity, and attraction to the anionic site (Metcalf et al., 1964) are operating to facilitate the attraction of PNMC's to AcChE.

Most of the workers studying the inhibition of AcChE from various sources rationalize their results by postulating hydrophobic binding. This point of view overlooks the fact that some effective PNMC's and many other powerful inhibitors carry *strongly* hydrophilic groups (N^+R_3, P^+R_3, and S^+R_2) as does the natural substrate which is an extremely hydrophilic molecule. In a recent review, O'Brien (1976) remarked that "Evidence of hydrophobic forces is not extensive". One must consider different sites or mechanisms of binding for the charged compounds as one must also do when the electronic effects of some substituents are not clearly understood (e.g., nitro-PNMC; Hastings et al., 1970; Hetnarski and O'Brien, 1973).

METHODS

The physicochemical constants in Table I were taken from our compilations (Hansch et al., 1973; Unger and Hansch, 1976; Hansch and Leo, 1979). MR values have been scaled by 0.1 to make them more nearly equiscalar with π. A number of values have not been experimentally determined; these were estimated from additivity principles (Hansch and Leo, 1979), assuming an increment of 0.50/CH_2 for π and 0.46/CH_2 for MR.

A number of E_s values have not yet been experimentally determined and, hence, have had to be estimated. For two substituents, $N(CH_3)_2$ and $P(Et)_2$, we have employed the isosteric principle advanced by Palm's group (Talvik and Palm, 1971); that is, we have used the E_s of $CH(CH_3)_2$ for $N(CH_3)_2$ and the E_s of $CH(Et)_2$ for $P(Et)_2$.

There are a number of OR and SR substituents for which measured E_s values are missing. To estimate these we used a principle of "isoincrements" wherein we considered the incremental values for E_s on these chains to be similar to their increments on carbon chains as follows:

Hydrocarbon E_s (Unger and Hansch, 1976)

CH_3	-1.24		OH	-0.55	
		-0.07			-0.07
CH_2CH_3	-1.31		OCH_3	-0.62	
		-0.29			-0.29
$CH_2CH_2CH_3$	-1.60		OCH_2CH_3	-0.91	
		-0.03			-0.03
$(CH_2)_3CH_3$	-1.63		$OCH_2CH_2CH_3$	-0.94	
		-0.01			-0.01
$(CH_2)_4CH_3$	-1.64		$O(CH_2)_3CH_3$	-0.95	
CH_3	-1.24		SH	-1.07	
		-0.07			-0.07
CH_2CH_3	-1.31		SCH_3	-1.14	
		-0.32			-0.32
$(CH_2)_3CH_3$	-1.63		$SCH_2CH_2CH_3$	-1.46	

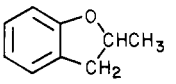
etc.

Since E_s is almost constant for chain lengths of propyl and longer, we do not believe that estimated E_s values will in any case introduce much error into our calculations.

For electronic effects we studied σ^- for ortho and para positions, σ for meta substituents, and $\mathcal{F}$ for each of these separately. $\Delta\sigma$ ($\sigma^- - \sigma$) for para substituents was also

considered. A number of indicator variables were examined; HB = 1 was employed for hydrogen bonding for cases where the donor atom was directly attached to the ring [OR and $N(CH_3)_2$]. Also checked were a charge indicator for salts (N^+R_3, S^+R_2, and P^+R_3) and a hydrolysis indicator (≡1) for all compounds having $(\sum\sigma_{o,p}^- + \sigma_m) \geq 0.8$.

For compounds having a ring structure connected in the 2–3 or 3–4 positions, we used MR for the two atoms attached to the ring separately from the rest of the ring, e.g.



The $>CHCH_3$ unit was parameterized under RGMR, O was included in MR_2 and CH_2 was included in MR_3. MR was obtained for charged groups by simple summation [e.g., $MR_{-N^+(CH_3)_3} = MR_{N(CH_3)_2} + MR_{CH_3}$]. No doubt the + charge will make the true value somewhat lower. However, this difference will not be very large and we assume it can be neglected for our purposes. Errors of this type will be taken care of by the indicator variable CHG.

Many combinations and cross-product terms of the parameters were investigated to detect special effects (i.e., $E_{s_2}\cdot E_{s_6}$, $E_{s_2}\cdot E_{s_3}$, etc.). It was especially convenient to carry out many of these studies using interactive APL. Composite terms in eq 1 such as $MR_{3,4,5}$ were constructed only after finding close correspondence in the susceptibility coefficients and confidence limits of terms of the individual variables. The steric and length constants of Verloop et al. (1976) and a variety of indicator variables for chain branching of complex meta substituents were tested but not found to be of value.

One of the more difficult decisions was that of numbering the positions for polysubstituted compounds. All 2,3-, 2,4-, and 3,4-disubstituted compounds and congeners with more than three substituents were numbered conventionally. For disubstituted compounds (2,5-; 2,6-; 3,5-) activity was compared with that of the mono analogues. The derivative having the *lower absolute value* was used to assign the lower position number to that substituent, assuming the other substituent to simply have a perturbation effect. For example, for no. **176** ($2\text{-}CH_3\text{-}5\text{-}NO_2$), 4.64 is compared with $2\text{-}CH_3$ (3.85) and $3\text{-}NO_2$ (3.70). Since activity is closer to 3.85, the numbering is $2\text{-}CH_3\text{-}5\text{-}NO_2$ rather than $3\text{-}NO_2\text{-}6\text{-}CH_3$. In the case of no. **209**, $3\text{-}CH(CH_3)_2\text{-}6\text{-}OCH_3$ (5.55), we compare it with $3\text{-}CH(CH_3)_2$ (6.47) and $2\text{-}OCH_3$ (4.43). Since 5.55 is closer to 6.47, this substituent is assigned to the 3 position rather than the 5 position. In one case [$3\text{-}C_6H_5\text{-}6\text{-}C(CH_3)_3$] we followed the results of the other two *tert*-butyl compounds since a $3\text{-}C_6H_5$ congener for reference was not tested.

Most of the nonsymmetrical 2,6-substituted compounds were not used in the formulation of eq 6 since these are mixtures of stereoisomers which may have quite different activities. Their deviations from eq 6 are quite large and we are unable to properly parameterize such structures at this time. A few other molecules proved to have large deviations and were also dropped in deriving eq 6. We have used 269 out of 288 data points for the formulation of eq 6. Of the 19 molecules dropped, all but six had substituents in the 2 and/or 6 positions. Ortho substitution is the most difficult to parameterize.

Our general approach (Fukunaga et al., 1976) has been to study all possible combinations of parameters rather than use stepwise regression analysis. The correlation matrix for the variables used in the final equation is presented in Table II. Each of the parameters used in the final equation was regressed against all possible permutations of the other parameters in order to search for special collinearities. The highest correlations found were

$$\mathcal{F}_{2,6} = 1.29\mathcal{F}_{2,6}^2 + 0.134\text{HB} + 0.018$$

$$n = 272 \quad r = 0.930 \quad s = 0.73$$

$$MR_2 = 0.21MR_2^2 - 0.17MR_{3,4,5} + 0.66$$

$$n = 272 \quad r = 0.945 \quad s = 0.333$$

The stepwise development of eq 6 is given in Table III and certain special equations are listed in Table IV.

RESULTS AND DISCUSSION

With the above results and ideas of previous studies in mind, we have undertaken an analysis of Fukuto and Metcalf's monumental investigation of PNMC's inhibiting fly head cholinesterase. Our best correlation equation derived from the data in Table I is

$$\begin{aligned} pI_{50} = {} & 0.557(\pm0.08)MR_{3,4,5} + 1.558(\pm0.20)MR_2 - \\ & 0.611(\pm0.09)E_s - 0.940(\pm0.19)(\textstyle\sum\sigma_{o,p}^- + \sigma_m) + \\ & 1.431(\pm0.31)\text{CHG} - 0.227(\pm0.04)MR_2^2 - \\ & 5.236(\pm1.27)\mathcal{F}_{2,6}^2 + 3.465(\pm0.90)\mathcal{F}_{2,6} + \\ & 0.659(\pm0.22)\text{RGMR} - 0.618(\pm0.22)\text{HB} - \\ & 0.052(\pm0.02)MR_3^2 - 0.563(\pm0.29)E_{s_2}\cdot E_{s_6} + 3.458(\pm0.21) \end{aligned} \quad (6)$$

$$n = 269 \quad r^2 = 0.796 \quad s = 0.485$$

$$\text{ideal } \mathcal{F}_{2,6} = 0.331 \ (0.295\text{–}0.368)$$

$$\text{ideal } MR_2 = 3.43 \ (3.16\text{–}3.79)$$

Terms in this equation have been arranged in order of decreasing importance (see Table III). The first three terms account for 47% out of 80% of the "explained" variance in pI_{50}. Considering the other terms in MR and E_s, most of the variance in pI_{50} seems related to the bulky effects of the substituents. We realize in eq 6 that E_s and especially MR are highly related to the volume of the substituent and, hence, may be accounting for substituents producing conformational changes in the enzyme.

As pointed out in the introduction, pI_{50} is not an ideal parameter for structure–activity work. This being so, how good is the result of eq 6? At first glance the standard deviation of 0.485 seems and is high; however, when one considers the range in pI_{50} (2.30–8.30), this is not unreasonable. On the average, we are able to predict the concentration causing 50% inhibition for 269 different molecules within a factor of ±3 (i.e., antilog of 0.485) for a concentration range of 1 000 000. Equation 6 is based on 22.4 data points/variable, on the average; hence it is highly unlikely that eq 6 is a chance correlation (Topliss and Edwards, 1979).

The HB term takes the value of 1 for hydrogen bond acceptors attached to the ortho position (OR and NR_2), and CHG is assigned the value of 1 for a charged substituent [e.g., $N^+(CH_3)_3$]. In one instance, two charged groups are present and CHG for this case is given the value of 2. The parameter RGMR is used for that portion of a ring attached between positions 2 and 3 or 3 and 4 with the exception of the two atoms directly attached to the rings. These two atoms are parameterized in MR_2 and $MR_{3,4,5}$. The σ terms have their usual connotation. The subscripts refer to the ortho, meta, and para positions, and σ^- is used to account for "through resonance" where a negative charge is delocalized by the substituent (e.g., $4\text{-}NO_2$).

The electronic effect of substituents correlated by Hammett-type parameters (σ and $\mathcal{F}$) show optima. In the case of $(\sum\sigma_{o,p}^- + \sigma_m)$, the optimum value is 0 (i.e., the best substituent is H). Empirically, either strong electron-at-

Table I. Parameters Used in the Formation of Equation 6

no.	group	pI_{50} obsd	pI_{50} calcd	$\lvert\Delta pI_{50}\rvert$	$\mathcal{F}_{2,6}$	MR_2	HB	CHG	RGMR	E_{s_3}	$MR_{3,4,5}$	$E_{s_2}\cdot E_{s_6}$	$\Sigma\sigma$
1	2-NO_2	2.30	3.17	0.88	0.67	0.74	0	0	0	0.0	0.30	0.0	1.24
2	2-$OCH_2C_6H_5$	3.52[a]	6.19	2.67	0.26	3.22	1	0	0	0.0	0.30	0.0	−0.16
3	H	3.70	3.78	0.08	0.0	0.10	0	0	0	0.0	0.30	0.0	0.00
4	2-CH_3	3.85	4.27	0.42	−0.04	0.57	0	0	0	0.0	0.30	0.0	−0.15
5	2-OCH_3	4.43	4.62	0.19	0.26	0.79	1	0	0	0.0	0.30	0.0	−0.16
6	2-OC_2H_5	4.80	5.08	0.28	0.22	1.25	1	0	0	0.0	0.30	0.0	−0.16
7	2-F	4.80	4.28	0.52	0.43	0.09	0	0	0	0.0	0.30	0.0	0.05
8	2-$Si(CH_3)_3$	4.89[a]	5.95	1.06	−0.04	2.50	0	0	0	0.0	0.30	0.0	0.08
9	2-C_2H_5	4.89	4.78	0.11	−0.05	1.03	0	0	0	0.0	0.30	0.0	−0.15
10	2-OC_4H_9	4.92	5.83	0.91	0.25	2.17	1	0	0	0.0	0.30	0.0	−0.16
11	2-*i*-C_5H_{11}	5.05	5.82	0.77	−0.06	2.43	0	0	0	0.0	0.30	0.0	−0.15
12	2-OC_3H_7	5.06	5.49	0.43	0.22	1.71	1	0	0	0.0	0.30	0.0	−0.16
13	2-$C(CH_3)_3$	5.22	5.52	0.30	−0.07	1.96	0	0	0	0.0	0.30	0.0	−0.13
14	2-D-*s*-C_4H_9	5.22	5.57	0.35	−0.06	1.96	0	0	0	0.0	0.30	0.0	−0.09
15	2-C_3H_7	5.22	5.20	0.02	−0.06	1.50	0	0	0	0.0	0.30	0.0	−0.15
16	2-*i*-C_3H_7	5.22	5.22	0.00	−0.06	1.50	0	0	0	0.0	0.30	0.0	−0.09
17	2-S-$C_{10}H_{21}$	5.28	5.78	0.50	0.23	5.54	0	0	0	0.0	0.30	0.0	0.17
18	2-Cl	5.30	4.95	0.35	0.41	0.60	0	0	0	0.0	0.30	0.0	0.27
19	2-$OCH_2C{\equiv}CH$	5.54	5.47	0.07	0.26	1.64	1	0	0	0.0	0.30	0.0	−0.16
20	2-$CH_2CH{=}CH_2$	5.47	5.23	0.24	−0.04	1.45	0	0	0	0.0	0.30	0.0	−0.18
21	2-*i*-C_4H_9	5.64	5.56	0.08	−0.06	1.96	0	0	0	0.0	0.30	0.0	−0.15
22	2-Br	5.66	5.27	0.39	0.44	0.89	0	0	0	0.0	0.30	0.0	0.28
23	2-$N(CH_3)_2$	5.70	5.16	0.53	0.10	1.56	1	0	0	0.0	0.30	0.0	−0.12
24	2-$CH_2C(CH_3){=}CH_2$	5.74	5.60	0.14	−0.04	1.91	0	0	0	0.0	0.30	0.0	−0.15
25	2-c-C_6H_{11}	5.85	5.96	0.11	−0.05	2.67	0	0	0	0.0	0.30	0.0	−0.15
26	2-$CH{=}C(CH_3)_2$	5.92	5.94	0.02	0.07	1.90	0	0	0	0.0	0.30	0.0	0.20
27	2-*s*-C_4H_9	5.96	5.60	0.36	−0.05	1.96	0	0	0	0.0	0.30	0.0	−0.15
28	2-c-C_5H_{11}	5.96	5.74	0.21	−0.05	2.20	0	0	0	0.0	0.30	0.0	−0.15
29	2-L-*s*-C_4H_9	6.00	5.60	0.40	−0.05	1.96	0	0	0	0.0	0.30	0.0	−0.15
30	2-SCH_3	6.05	5.80	0.25	0.20	1.38	0	0	0	0.0	0.30	0.0	0.17
31	2-I	6.10	5.81	0.29	0.40	1.39	0	0	0	0.0	0.30	0.0	0.30
32	2-S-*i*-C_5H_{11}	6.13	6.78	0.65	0.23	3.23	0	0	0	0.0	0.30	0.0	0.17
33	2-O-*i*-C_3H_7	6.16	5.55	0.61	0.30	1.71	1	0	0	0.0	0.30	0.0	−0.16
34	2-$SCH_2C_6H_4$-4-OCH_3	6.21	6.45	0.24	0.20	4.58	0	0	0	0.0	0.30	0.0	0.17
35	2-$CH(OCH_2CH_2O{-})$	6.30	5.92	0.38	0.10	1.73	0	0	0	0.0	0.30	0.0	0.10
36	2-O-c-C_5H_9	6.40	5.98	0.42	0.30	2.38	1	0	0	0.0	0.30	0.0	−0.15
37	2-SC_8H_{17}	6.40	6.46	0.06	0.23	4.62	0	0	0	0.0	0.30	0.0	0.17
38	2-SC_9H_{19}	6.44	6.17	0.27	0.23	5.08	0	0	0	0.0	0.30	0.0	0.17
39	2-$SCH_2C{\equiv}CH$	6.47	6.42	0.05	0.20	2.23	0	0	0	0.0	0.30	0.0	0.17
40	2-O-*s*-C_4H_9	6.51	5.86	0.65	0.30	2.17	1	0	0	0.0	0.30	0.0	−0.16
41	2-$SCH_2C_6H_4$-4-CH_3	6.55	6.59	0.04	0.20	4.28	0	0	0	0.0	0.30	0.0	0.17
42	2-$SCH_2CH{=}CH_2$	6.59	6.50	0.09	0.20	2.38	0	0	0	0.0	0.30	0.0	0.17
43	2-S-*i*-C_8H_{17}	6.62	6.47	0.15	0.23	4.62	0	0	0	0.0	0.30	0.0	0.17
44	2-$SCH_2C_6H_5$	6.66	6.72	0.06	0.20	3.82	0	0	0	0.0	0.30	0.0	0.17
45	2-SC_3H_7	6.74	6.50	0.24	0.23	2.31	0	0	0	0.0	0.30	0.0	0.17
46	2-SC_7H_{15}	6.68	6.67	0.01	0.23	4.15	0	0	0	0.0	0.30	0.0	0.17
47	2-SC_6H_{13}	6.70	6.77	0.07	0.23	3.69	0	0	0	0.0	0.30	0.0	0.17
48	2-SC_4H_9	6.80	6.69	0.11	0.23	2.77	0	0	0	0.0	0.30	0.0	0.17
49	2-S-*i*-C_3H_7	6.85	6.50	0.35	0.23	2.31	0	0	0	0.0	0.30	0.0	0.17
50	2-D-*s*-SC_4H_9	6.92	6.69	0.23	0.23	2.77	0	0	0	0.0	0.30	0.0	0.17
51	2-$SCH_2C_6H_4$-4-NO_2	7.00	6.51	0.49	0.20	4.45	0	0	0	0.0	0.30	0.0	0.17
52	2-L-*s*-SC_4H_9	7.10	6.69	0.41	0.23	2.77	0	0	0	0.0	0.30	0.0	0.17
53	3-SC_6H_{13}	3.22[a]	5.98	2.76	0.00	0.10	0	0	0	−1.52	3.90	0.0	0.15
54	3-$SCH_2C_6H_4$-4′-Cl	3.43[a]	6.02	2.59	0.00	0.10	0	0	0	−1.45	4.52	0.0	0.15
55	3-NO_2	3.70	4.25	0.55	0.00	0.10	0	0	0	−1.01	0.94	0.0	0.71
56	3-F	4.07	3.95	0.12	0.00	0.10	0	0	0	−0.46	0.30	0.0	0.34
57	3-Cl	4.30	4.51	0.21	0.00	0.10	0	0	0	−0.97	0.81	0.0	0.37
58	3-OCH_3	4.66	4.50	0.16	0.00	0.10	0	0	0	−0.62	1.00	0.0	0.12
59	3-O-*i*-C_4H_9	4.70	5.23	0.53	0.00	0.10	0	0	0	−0.90	2.38	0.0	0.10
60	3-OC_3H_7	4.80	5.09	0.29	0.00	0.10	0	0	0	−0.94	1.92	0.0	0.10
61	3-SF_5	4.85	5.65	0.80	0.00	0.10	0	0	0	−2.91	1.19	0.0	0.61
62	3-CH_3	4.85	4.78	0.07	0.00	0.10	0	0	0	−1.24	0.78	0.0	−0.07
63	3-Br	4.89	4.75	0.14	0.00	0.10	0	0	0	−1.16	1.10	0.0	0.39
64	3-OC_4H_9	5.03	5.26	0.23	0.00	0.10	0	0	0	−0.95	2.38	0.0	0.10
65	3-O-*i*-C_3H_7	5.04	5.42	0.38	0.00	0.10	0	0	0	−1.48	1.92	0.0	0.10
66	3-$N(CH_3)_2$	5.10	5.50	0.40	0.00	0.10	0	0	0	−1.71	1.77	0.0	−0.15
67	3-O-c-C_5H_9	5.10	5.70	0.60	0.00	0.10	0	0	0	−1.53	2.63	0.0	0.10
68	3-O-*s*-C_4H_9	5.15	5.61	0.46	0.00	0.10	0	0	0	−1.51	2.38	0.0	0.10
69	3-SCH_3	5.15	5.07	0.08	0.00	0.10	0	0	0	−1.14	1.59	0.0	0.15
70	3-I	5.15	5.14	0.01	0.00	0.10	0	0	0	−1.40	1.60	0.0	0.35
71	3-$SC_{10}H_{21}$	5.15[a]	6.07	0.92	0.00	0.10	0	0	0	−1.45	5.75	0.0	0.15
72	3-OC_5H_5	5.22	4.89	0.33	0.00	0.10	0	0	0	−0.91	1.46	0.0	0.10
73	3-C_2H_5	5.32	5.04	0.28	0.00	0.10	0	0	0	−1.31	1.24	0.0	−0.07
74	3-$OCH_2C{\equiv}CH$	5.40	5.00	0.40	0.00	0.10	0	0	0	−0.85	1.84	0.0	0.12
75	3-$SCH_2CH{=}CH_2$	5.44	5.57	0.13	0.00	0.10	0	0	0	−1.37	2.58	0.0	0.15
76	3-$SCH_2C{\equiv}CH$	5.62	5.52	0.10	0.00	0.10	0	0	0	−1.37	2.43	0.0	0.15

Table I (*Continued*)

no.	group	pI_{50} obsd	pI_{50} calcd	$\lvert\Delta pI_{50}\rvert$	$\mathcal{F}_{2,6}$	MR_2	HB	CHG	RGMR	E_{S_3}	$MR_{3,4,5}$	$E_{S_2}\cdot E_{S_6}$	$\Sigma\sigma$
77	3-$SCH_2C_6H_4$-4′-OCH_3	5.70	6.05	0.35	0.00	0.10	0	0	0	−1.45	4.78	0.0	0.15
78	3-$SCH_2C_6H_4$-4′-CH_3	5.70	6.02	0.32	0.00	0.10	0	0	0	−1.45	4.48	0.0	0.15
79	3-c-C_6H_{11}	5.70	6.07	0.37	0.00	0.10	0	0	0	−2.03	2.87	0.0	−0.07
80	3-S-*i*-C_3H_7	5.74	5.93	0.19	0.00	0.10	0	0	0	−2.00	2.51	0.0	0.15
81	3-S-*i*-C_8H_{17}	5.82	6.09	0.27	0.00	0.10	0	0	0	−1.52	4.82	0.0	0.15
82	3-c-C_5H_{11}	5.82	5.76	0.06	0.00	0.10	0	0	0	−1.75	2.40	0.0	−0.07
83	3-SC_8H_{17}	5.92	6.09	0.17	0.00	0.10	0	0	0	−1.52	4.82	0.0	0.15
84	3-SC_3H_7	5.96	5.60	0.36	0.00	0.10	0	0	0	−1.46	2.51	0.0	0.15
85	3-SC_7H_{15}	6.03	6.05	0.02	0.00	0.10	0	0	0	−1.52	4.36	0.0	0.15
86	3-S-*i*-C_5H_{11}	5.06	5.91	0.15	0.00	0.10	0	0	0	−1.52	3.54	0.0	0.15
87	3-SC_4H_9	6.11	5.77	0.34	0.00	0.10	0	0	0	−1.47	3.07	0.0	0.15
88	3-$P(C_2H_5)_2$	6.13	6.90	0.77	0.00	0.10	0	0	0	−3.22	3.25	0.0	0.03
89	3-$CH_2C_6H_5$	6.13	5.90	0.23	0.00	0.10	0	0	0	−1.61	3.20	0.0	−0.08
90	3-$Si(CH_3)_3$	6.15	6.84	0.69	0.00	0.10	0	0	0	−3.36	2.70	0.0	−0.04
91	3-$C(CH_3)_3$	6.40	6.30	0.10	0.00	0.10	0	0	0	−2.78	2.16	0.0	−0.10
92	3-*i*-C_3H_7	6.47	5.48	0.99	0.00	0.10	0	0	0	−1.71	1.70	0.0	−0.07
93	3-SC_9H_{19}	6.54	6.11	0.43	0.00	0.10	0	0	0	−1.52	5.28	0.0	0.15
94	3-$SCH_2C_6H_4$-4′-NO_2	6.68	6.03	0.65	0.00	0.10	0	0	0	−1.45	4.65	0.0	0.15
95	3-*s*-C_4H_9	6.80	6.06	0.74	0.00	0.10	0	0	0	−2.37	2.16	0.0	−0.08
96	3-$CH(C_2H_5)_2$	6.96	6.73	0.23	0.00	0.10	0	0	0	−3.22	2.62	0.0	−0.08
97	4-NO_2	2.52	2.69	0.17	0.00	0.10	0	0	0	0.00	0.94	0.0	1.24
98	4-$N(CH_3)_2$	3.62	4.58	0.96	0.00	0.10	0	0	0	0.00	1.76	0.0	−0.12
99	4-Cl	3.62	3.99	0.37	0.00	0.10	0	0	0	0.00	0.80	0.0	0.27
100	4-F	3.64	3.77	0.13	0.00	0.10	0	0	0	0.00	0.29	0.0	0.05
101	4-CN	3.77	3.13	0.64	0.00	0.10	0	0	0	0.00	0.83	0.0	1.00
102	4-$C(CH_3)_3$	3.82	4.80	0.98	0.00	0.10	0	0	0	0.00	2.16	0.0	−0.13
103	4-OC_3H_7	3.96	4.65	0.69	0.00	0.10	0	0	0	0.00	1.91	0.0	−0.16
104	4-Br	4.00	4.14	0.14	0.00	0.10	0	0	0	0.00	1.09	0.0	0.28
105	4-CH_3	4.00	4.02	0.02	0.00	0.10	0	0	0	0.00	0.77	0.0	−0.15
106	4-SO_2CH_3	4.00	3.57	0.43	0.00	0.10	0	0	0	0.00	1.55	0.0	0.98
107	4-I	4.06	4.42	0.36	0.00	0.10	0	0	0	0.00	1.60	0.0	0.30
108	4-O-*i*-C_3H_7	4.06	4.65	0.59	0.00	0.10	0	0	0	0.00	1.91	0.0	−0.16
109	4-OCH_3	4.10	4.14	0.04	0.00	0.10	0	0	0	0.00	0.99	0.0	−0.16
110	4-OC_2H_5	4.15	4.40	0.25	0.00	0.10	0	0	0	0.00	1.45	0.0	−0.16
111	4-*i*-C_3H_7	4.15	4.55	0.40	0.00	0.10	0	0	0	0.00	1.70	0.0	−0.09
112	4-$OCH_2C{\equiv}CH$	4.21	4.61	0.40	0.00	0.10	0	0	0	0.00	1.84	0.0	−0.16
113	4-SC_2H_5	4.25	4.72	0.47	0.00	0.10	0	0	0	0.00	2.04	0.0	0.17
114	4-C_2H_5	4.42	4.27	0.15	0.00	0.10	0	0	0	0.00	1.23	0.0	−0.15
115	4-SCH_3	4.47	4.46	0.01	0.00	0.10	0	0	0	0.00	1.58	0.0	0.17
116	4-O-*s*-C_4H_9	4.49	4.91	0.42	0.00	0.10	0	0	0	0.00	2.37	0.0	−0.16
117	4-c-C_5H_9	4.57	4.92	0.35	0.00	0.10	0	0	0	0.00	2.41	0.0	−0.18
118	4-C_6H_5	4.59	5.13	0.54	0.00	0.10	0	0	0	0.00	2.74	0.0	0.10
119	4-OC_4H_9	4.70	4.91	0.21	0.00	0.10	0	0	0	0.00	2.37	0.0	−0.16
120	4-O-c-C_5H_9	4.77	5.05	0.28	0.00	0.10	0	0	0	0.00	2.62	0.0	−0.16
121	4-$SCH_2C{\equiv}CH$	4.77	4.94	0.17	0.00	0.10	0	0	0	0.00	2.43	0.0	0.17
122	4-$SOCH_3$	4.80	4.12	0.68	0.00	0.10	0	0	0	0.00	1.57	0.0	0.62
123	4-$Si(CH_3)_3$	4.81	5.11	0.30	0.00	0.10	0	0	0	0.00	2.70	0.0	0.08
124	4-S-*i*-C_8H_{17}	4.89[a]	6.27	1.38	0.00	0.10	0	0	0	0.00	4.82	0.0	0.17
125	4-SC_8H_{17}	4.89[a]	6.27	1.38	0.00	0.10	0	0	0	0.00	4.82	0.0	0.17
126	4-SC_3H_7	4.92	4.98	0.06	0.00	0.10	0	0	0	0.00	2.51	0.0	0.17
127	4-*i*-C_5H_{11}	5.05	5.06	0.01	0.00	0.10	0	0	0	0.00	2.63	0.0	−0.15
128	4-c-C_6H_{11}	5.05	5.19	0.14	0.00	0.10	0	0	0	0.00	2.87	0.0	−0.15
129	4-S-*i*-C_3H_7	5.05	4.98	0.07	0.00	0.10	0	0	0	0.00	2.51	0.0	0.17
130	4-$SCH_2CH{=}CH_2$	5.17	5.02	0.15	0.00	0.10	0	0	0	0.00	2.58	0.0	0.17
131	4-$SCH_2C_6H_4$-4′-OCH_3	5.36	6.25	0.89	0.00	0.10	0	0	0	0.00	4.78	0.0	0.17
132	4-SC_6H_{13}	5.40	5.75	0.35	0.00	0.10	0	0	0	0.00	3.89	0.0	0.17
133	4-$OCH_2C_6H_5$	5.46	5.49	0.03	0.00	0.10	0	0	0	0.00	3.42	0.0	−0.16
134	4-$C(CH_3)_2CH_2C(CH_3)_2CH_3$	5.52	5.83	0.31	0.00	0.10	0	0	0	0.00	4.02	0.0	−0.15
135	4-SC_4H_9	5.52	5.24	0.28	0.00	0.10	0	0	0	0.00	2.97	0.0	0.17
136	4-$SCH_2C_6H_4$-4′-CH_3	5.72	6.08	0.36	0.00	0.10	0	0	0	0.00	4.48	0.0	0.17
137	4-*s*-C_4H_9	5.74	4.79	0.95	0.00	0.10	0	0	0	0.00	2.16	0.0	−0.15
138	4-SC_5H_{11}	5.74	5.50	0.24	0.00	0.10	0	0	0	0.00	3.43	0.0	0.17
139	4-S-*i*-C_5H_{11}	5.80	5.50	0.30	0.00	0.10	0	0	0	0.00	3.43	0.0	0.17
140	4-$CH_2C_6H_5$	5.82	5.38	0.44	0.00	0.10	0	0	0	0.00	3.20	0.0	−0.12
141	4-$SCH_2C_6H_5$	5.89	5.82	0.07	0.00	0.10	0	0	0	0.00	4.02	0.0	0.17
142	4-$C(CH_3)_2C_6H_5$	6.12	5.90	0.22	0.00	0.10	0	0	0	0.00	4.13	0.0	−0.12
143	4-$SCH_2C_6H_4$-4′-Cl	6.37	6.10	0.27	0.00	0.10	0	0	0	0.00	4.52	0.0	0.17
144	4-$SCH_2C_6H_4$-4′-NO_2	6.70	6.17	0.53	0.00	0.10	0	0	0	0.00	4.65	0.0	0.17
145	2-NO_2-3-CH_3	3.70	4.34	0.64	0.67	0.74	0	0	0	−1.24	0.77	0.0	1.17
146	2,3-Cl_2	4.32[a]	5.49	1.17	0.41	0.60	0	0	0	−0.97	0.80	0.0	0.64
147	2,3-$(OCH_3)_2$	4.85	5.37	0.52	0.26	0.79	1	0	0	−0.62	0.99	0.0	−0.04
148	2,3-$(CH_3)_2$	5.09	5.25	0.16	−0.04	0.57	0	0	0	−1.24	0.77	0.0	−0.22
149	2-SCH_3-3-*i*-C_3H_7	6.74	7.52	0.78	0.20	1.38	0	0	0	−1.71	1.70	0.0	0.10
150	2,4-$(CH_3)_2$	3.89	4.47	0.58	−0.04	0.57	0	0	0	0.00	0.77	0.0	−0.30

Table I (*Continued*)

no.	group	pI_{50} obsd	pI_{50} calcd	$\lvert\Delta pI_{50}\rvert$	$\mathcal{F}_{2,6}$	MR_2	HB	CHG	RGMR	E_{s_3}	$MR_{3,4,5}$	$E_{s_2}{\cdot}E_{s_6}$	$\Sigma\sigma$
151	2-CH_3-4-Cl	4.32	4.56	0.24	−0.04	0.57	0	0	0	0.00	0.80	0.0	0.12
152	2-CH_3-4-SCH_3	4.76	5.01	0.25	−0.04	0.57	0	0	0	0.00	1.58	0.0	0.02
153	2-*i*-C_3H_7-4-CH_3	5.36	5.43	0.07	−0.06	1.50	0	0	0	0.00	0.77	0.0	−0.24
154	2,4-*i*-C_3H_7	5.85	5.97	0.12	−0.06	1.50	0	0	0	0.00	1.70	0.0	−0.18
155	2,4-Cl_2	4.85	5.02	0.17	0.41	0.60	0	0	0	0.00	0.80	0.0	0.54
156	2-Cl-4-CH_3	5.10	5.27	0.17	0.41	0.60	0	0	0	0.00	0.77	0.0	0.12
157	2,4-$(OCH_3)_2$	4.55	4.93	0.38	0.26	0.79	1	0	0	0.00	0.99	0.0	−0.32
158	2-SCH_3-4-CH_3	5.72	6.09	0.37	0.20	1.38	0	0	0	0.00	0.77	0.0	0.02
159	2-SCH_3-4-CH_3O	5.85	6.21	0.36	0.20	1.38	0	0	0	0.00	0.99	0.0	0.01
160	2-NO_2-4-CH_3	3.89	3.77	0.12	0.67	0.74	0	0	0	0.00	0.77	0.0	1.09
161	3-CH_3-4-NO_2	3.49	3.85	0.36	0.00	0.10	0	0	0	−1.24	1.41	0.0	1.17
162	3-CH_3-4-Cl	4.54	5.02	0.48	0.00	0.10	0	0	0	−1.24	1.27	0.0	0.20
163	3,4-$(CH_3)_2$	4.59	5.00	0.41	0.00	0.10	0	0	0	−1.24	1.24	0.0	−0.22
164	3-CH_3-4-SCH_3	5.89	5.49	0.40	0.00	0.10	0	0	0	−1.24	2.05	0.0	0.10
165	3-CH_3-4-*i*-C_3H_7	5.96	5.54	0.42	0.00	0.10	0	0	0	−1.24	2.17	0.0	−0.16
166	3-C_2H_5-4-NO_2	3.70	4.11	0.41	0.00	0.10	0	0	0	−1.31	1.87	0.0	1.17
167	3-*i*-C_3H_7-4-F	5.24	5.48	0.24	0.00	0.10	0	0	0	−1.71	1.69	0.0	−0.02
168	3-*i*-C_3H_7-4-NO_2	5.55	4.56	0.99	0.00	0.10	0	0	0	−1.71	2.34	0.0	1.17
169	3-*i*-C_3H_7-4-$N(CH_3)_2$	6.82	6.27	0.55	0.00	0.10	0	0	0	−1.71	3.16	0.0	−0.19
170	3-*i*-C_3H_7-4-SCH_3	7.00	6.19	0.81	0.00	0.10	0	0	0	−1.71	2.98	0.0	0.10
171	3,4-Cl_2	4.74	4.52	0.22	0.00	0.10	0	0	0	−0.97	1.30	0.0	0.64
172	3,4-$(OCH_3)_2$	4.72	4.89	0.17	0.00	0.10	0	0	0	−0.62	1.68	0.0	−0.04
173	3-SCH_3-4-CH_3O	5.58	5.47	0.11	0.00	0.10	0	0	0	−1.14	2.27	0.0	−0.01
174	3-NO_2-4-CH_3	4.80	5.61	0.81	0.00	0.10	0	0	0	−2.52	1.41	0.0	0.56
175	2-NO_2-5-*i*-C_3H_7	3.64	4.12	0.48	0.67	0.74	0	0	0	0.00	1.70	0.0	1.17
176	2-CH_3-5-NO_2	4.64	4.35	0.29	−0.04	0.57	0	0	0	0.00	0.94	0.0	0.56
177	2-*i*-C_3H_7-5-CH_3	5.85	5.46	0.39	−0.06	1.50	0	0	0	0.00	0.77	0.0	−0.16
178	2-c-C_6H_{11}-5-OCH_3	6.57	6.36	0.21	−0.05	2.67	0	0	0	0.00	0.99	0.0	−0.06
179	2-F-5-*i*-C_3H_7	5.26	5.06	0.20	0.43	0.09	0	0	0	0.00	1.70	0.0	−0.02
180	2-OCH_3-5-$C(CH_3)_3$	5.38	5.61	0.23	0.26	0.79	1	0	0	0.00	2.16	0.0	−0.26
181	2-O-*i*-C_3H_7-5-OCH_3	6.25	5.96	0.29	0.30	1.71	1	0	0	0.00	0.99	0.0	−0.04
182	2-O-*i*-C_3H_7-5-CH_3	6.55	5.79	0.76	0.30	1.71	1	0	0	0.00	0.77	0.0	−0.23
183	2,6-$(CH_3)_2$	2.00[a]	3.18	1.18	−0.08	0.57	0	0	0	0.00	0.30	1.54	−0.30
184	2-CH_3-6-Cl	3.72	4.31	0.59	0.37	0.57	0	0	0	0.00	0.30	1.20	0.12
185	2,6-Cl_2	2.89	2.99	0.10	0.82	0.60	0	0	0	0.00	0.30	0.94	0.54
186	2,6-$(OCH_3)_2$	2.68	3.55	0.87	0.52	0.79	2	0	0	0.00	0.30	0.38	−0.32
187	3,5-$(CH_3)_2$	5.22	5.02	0.20	0.00	0.10	0	0	0	−1.24	1.24	0.0	−0.14
188	3-C_2H_5-5-CH_3	5.32	5.29	0.03	0.00	0.10	0	0	0	−1.31	1.70	0.0	−0.14
189	3-*i*-C_3H_7-5-$N(CH_3)_2$	6.72	6.25	0.47	0.00	0.10	0	0	0	−1.71	3.16	0.0	−0.22
190	3-*i*-C_3H_7-5-CH_3	7.25	5.73	1.52	0.00	0.10	0	0	0	−1.71	2.17	0.0	−0.14
191	3,5-(*i*-$C_3H_7)_2$	7.48	6.25	1.23	0.00	0.10	0	0	0	−1.71	3.10	0.0	−0.14
192	3-$C(CH_3)_3$-5-CH_3	6.77	6.55	0.22	0.00	0.10	0	0	0	−2.78	2.63	0.0	−0.17
193	3,5-$[C(CH_3)_3]_2$	7.11	7.31	0.20	0.00	0.10	0	0	0	−2.78	4.02	0.0	−0.20
194	3-OCH_3-5-CH_3	4.77	4.77	0.00	0.00	0.10	0	0	0	−0.62	1.46	0.0	0.05
195	3,5-Cl_2	4.92	4.40	0.52	0.00	0.10	0	0	0	−0.97	1.30	0.0	0.74
196	3,5-$(OCH_3)_2$	5.10	4.84	0.26	0.00	0.10	0	0	0	−0.62	1.68	0.0	0.24
197	3,5-$(OC_2H_5)_2$	4.96	5.49	0.53	0.00	0.10	0	0	0	−0.91	2.60	0.0	0.22
198	3,5-di-$N(CH_3)_2$	5.59	6.24	0.65	0.00	0.10	0	0	0	−1.71	3.22	0.0	−0.30
199	3-CH_3-6-$C(CH_3)_3$	3.10[a]	4.48	1.38	−0.07	0.10	0	0	0	−1.24	0.77	0.0	−0.20
200	3-CH_3-6-NO_2	3.60	3.47	0.13	0.67	0.10	0	0	0	−1.24	0.77	0.0	1.17
201	3-OCH_3-6-$C(CH_3)_3$	3.72	4.24	0.52	−0.07	0.10	0	0	0	−0.62	0.99	0.0	−0.01
202	3,6-Cl_2	4.30	4.79	0.49	0.41	0.10	0	0	0	−0.97	0.80	0.0	0.64
203	3-$OCH(CH_3)_2$-6-OCH_3	4.82	5.35	0.53	0.26	0.10	1	0	0	−1.48	1.91	0.0	−0.06
204	3,6-$(OCH_3)_2$	4.89	4.44	0.45	0.26	0.10	1	0	0	−0.62	0.99	0.0	−0.04
205	3,6-$(CH_3)_2$	5.05	4.59	0.46	−0.04	0.10	0	0	0	−1.24	0.77	0.0	−0.22
206	3-OCH_3-6-CH_3	5.07	4.36	0.71	−0.04	0.10	0	0	0	−0.62	0.99	0.0	−0.03
207	3-CH_3-6-OCH_3	5.07	5.22	0.15	0.26	0.10	1	0	0	−1.24	0.77	0.0	−0.23
208	3-C_6H_5-6-$C(CH_3)_3$	5.27[a]	6.86	1.59	−0.07	0.10	0	0	0	−3.82	2.74	0.0	−0.07
209	3-*i*-C_3H_7-6-OCH_3	5.55	5.37	0.18	0.26	0.10	1	0	0	−1.71	1.70	0.0	−0.23
210	3-*i*-C_3H_7-6-CH_3	5.70	5.29	0.41	−0.04	0.10	0	0	0	−1.71	1.70	0.0	−0.22
211	3-*i*-C_3H_7-6-*i*-C_3H_7	6.37	5.39	0.98	0.30	0.10	1	0	0	−1.71	1.70	0.0	−0.23
212	3-*i*-C_3H_7-6-SCH_3	6.74	5.96	0.78	0.20	0.10	0	0	0	−1.71	1.70	0.0	0.10
213	3-*i*-C_3H_7-6-Cl	7.11[a]	5.99	1.12	0.41	0.10	0	0	0	−1.71	1.70	0.0	0.20
214	3-*i*-C_3H_7-6-Br	7.14[a]	5.96	1.18	0.44	0.10	0	0	0	−1.71	1.70	0.0	0.21
215	2,3,4-$(CH_3)_3$	4.82	5.43	0.61	−0.04	0.57	0	0	0	−1.24	1.24	0.0	−0.37
216	2,3-$(CH_3)_2$-4-Cl	4.96	5.57	0.61	−0.04	0.57	0	0	0	−1.24	1.27	0.0	0.05
217	2,3,5-$(CH_3)_3$	5.22	5.48	0.26	−0.04	0.57	0	0	0	−1.24	1.24	0.0	−0.29
218	2,3,6-$(CH_3)_3$	4.30	4.14	0.16	−0.08	0.57	0	0	0	−1.24	0.77	1.54	−0.37
219	2,4,5-Cl_3	4.85	4.80	0.06	0.41	0.60	0	0	0	0.00	1.30	0.0	0.91
220	2,4-Cl_2-5-CH_3	4.96	5.35	0.39	0.41	0.60	0	0	0	0.00	1.27	0.0	0.47
221	2-Cl-4,5-$(CH_3)_2$	5.82	5.54	0.28	0.41	0.60	0	0	0	0.00	1.24	0.0	0.05
222	2,4,5-$(CH_3)_3$	5.07	4.69	0.38	−0.04	0.57	0	0	0	0.00	1.24	0.0	−0.37
223	2,5-$(CH_3)_2$-4-Cl	5.22	4.83	0.39	−0.04	0.57	0	0	0	0.00	1.27	0.0	0.05
224	2-CH_3-4-$N(CH_3)_2$-5-*i*-C_3H_7	6.41	5.78	0.63	−0.04	0.57	0	0	0	0.00	3.16	0.0	−0.34
225	2-*i*-C_3H_7-4-$N(CH_3)_2$-5-CH_3	5.89	6.22	0.33	−0.06	1.50	0	0	0	0.00	2.23	0.0	−0.28

Table I (*Continued*)

no.	group	pI_{50} obsd	pI_{50} calcd	$\lvert\Delta pI_{50}\rvert$	$\mathcal{F}_{2,6}$	MR_2	HB	CHG	RGMR	E_{S_3}	$MR_{3,4,5}$	$E_{S_2} \cdot E_{S_6}$	$\Sigma\sigma$
226	2-*i*-C_3H_7-4-Cl-5-CH_3	6.28	5.75	0.53	−0.06	1.50	0	0	0	0.00	1.27	0.0	0.11
227	2,4,6-$(CH_3)_3$	3.10	3.33	0.23	−0.08	0.57	0	0	0	0.00	0.77	1.54	−0.45
228	2,6-$(CH_3)_2$-4-Cl	4.48	3.54	0.94	−0.08	0.57	0	0	0	0.00	0.80	1.54	−0.03
229	2,4-Cl_2-6-CH_3	4.39	4.50	0.11	0.37	0.60	0	0	0	0.00	0.80	1.20	0.39
230	2,6-Cl_2-4-CH_3	4.03	3.39	0.64	0.82	0.60	0	0	0	0.00	0.77	0.94	0.39
231	2,4,6-Cl_3	4.77[a]	2.93	1.84	0.82	0.60	0	0	0	0.00	0.80	0.94	0.81
232	2,4-$(NO_2)_2$-6-CH_3	2.40[a]	−1.75	4.15	0.63	0.74	0	0	0	0.00	0.94	3.12	2.33
233	3,5-$(CH_3)_2$-4-NO	4.08	3.27	0.81	0.00	0.10	0	0	0	−1.24	1.66	0.0	1.46
234	3,5-$(CH_3)_2$-4-CN	4.32	4.64	0.32	0.00	0.10	0	0	0	−1.24	1.77	0.0	0.86
235	3,5-$(CH_3)_2$-4-SO_2CH_3	4.68	5.08	0.40	0.00	0.10	0	0	0	−1.24	2.49	0.0	0.84
236	3,5-$(CH_3)_2$-4-NO_2	5.38	4.26	1.12	0.00	0.10	0	0	0	−1.24	1.88	0.0	1.10
237	3,5-$(CH_3)_2$-4-$N(CH_3)_2$	5.48	5.79	0.31	0.00	0.10	0	0	0	−1.24	2.70	0.0	−0.26
238	3,5-$(CH_3)_2$-4-Cl	5.70	5.31	0.39	0.00	0.10	0	0	0	−1.24	1.74	0.0	0.13
239	3,4,5-$(CH_3)_3$	5.72	5.23	0.49	0.00	0.10	0	0	0	−1.24	1.71	0.0	−0.29
240	3,5-$(CH_3)_2$-4-$SOCH_3$	5.74	5.53	0.21	0.00	0.10	0	0	0	−1.24	2.51	0.0	0.48
241	3,5-$(CH_3)_2$-4-SCH_3	5.92	5.76	0.16	0.00	0.10	0	0	0	−1.24	2.52	0.0	0.03
242	3,5-$(CH_3)_2$-4-OCH_3	5.96	5.34	0.62	0.00	0.10	0	0	0	−1.24	1.93	0.0	−0.30
243	3,5-$(CH_3)_2$-4-SCN	6.10	5.54	0.56	0.00	0.10	0	0	0	−1.24	2.48	0.0	0.46
244	3,4,5-$(OCH_3)_3$	5.19	5.27	0.08	0.00	0.10	0	0	0	−0.62	2.37	0.0	0.08
245	2-CH_3-3,4,6-Cl_3	3.11[a]	4.92	1.81	0.37	0.57	0	0	0	−0.97	1.30	1.20	0.76
246	2-*i*-C_3H_7-4,6-Cl_2-5-CH_3	4.55[a]	5.49	0.94	0.36	1.50	0	0	0	0.00	1.27	1.66	0.38
247	2,6-di-CH_3O-3,5-di-NO_2	4.15	4.73	0.58	0.52	0.79	2	0	0	−2.52	1.58	0.38	1.10
248	2,3,4,5,6-F_5	3.32	2.37	0.95	0.86	0.09	0	0	0	−0.46	0.27	0.21	0.83
249	2,4,6-Cl_3-3,5-$(CH_3)_2$	3.72	4.39	0.67	0.82	0.60	0	0	0	−1.24	1.74	0.94	0.67
250	2,4-Cl_2-3,5,6-$(CH_3)_3$	3.74[a]	5.85	2.11	0.37	0.60	0	0	0	−1.24	1.74	1.20	0.25
251	2,3-[$OC(CH_3)_2CO$]	4.89	5.87	0.98	0.26	0.18	1	0	1.39	−1.50	0.78	0.0	0.22
252	2,3-[$OCH(CH_3)CH_2$]	5.61	5.65	0.04	0.26	0.18	1	0	1.39	−1.24	0.66	0.0	−0.23
253	2,3-[$OC(CH_3)_2CH(OH)$]	5.85	5.73	0.12	0.26	0.18	1	0	1.39	−1.21	0.82	0.0	−0.16
254	2,3-[$OC(CH_3)_2CH(OCOCH_3)$]	6.00	6.40	0.40	0.26	0.18	1	0	1.39	−1.63	1.74	0.0	−0.16
255	2,3-[$OC(CH_3)_2CH(OCH_3)$]	6.19	6.10	0.09	0.26	0.18	1	0	1.39	−1.43	1.31	0.0	−0.16
256	2,3-[$OC(CH_3)_2CH_2$]	6.69	5.65	0.95	0.26	0.18	1	0	1.39	−1.24	0.66	0.0	−0.23
257	2,3-(N=CH—CH=CH)	4.37	5.25	0.88	0.09	0.34	1	0	0.88	−1.24	0.69	0.0	0.26
258	2,3-(CH_2CH=$CHCH_2$)	5.68	5.66	0.02	−0.04	0.42	0	0	1	−1.31	0.62	0.0	−0.25
259	2,3-(CH=CH—CH=CF)	5.82	6.05	0.23	0.03	0.49	0	0	0.88	−1.48	0.60	0.0	0.17
260	2,3-$(CH=CH)_2$	6.05	6.05	0.00	0.03	0.49	0	0	1	−1.24	0.69	0.0	0.09
261	2,3-[CH=CH—CH=C(OH)]	6.30	6.00	0.30	0.03	0.49	0	0	0.88	−1.21	0.79	0.0	0.09
262	2,3-(S—CH=CH)	6.52	6.43	0.09	0.20	0.78	0	0	0.50	−1.24	0.69	0.0	0.21
263	2,3-(CH=CH—S)	6.60	5.75	0.85	0.03	0.49	0	0	0.50	−1.14	0.97	0.0	0.23
264	2,3-$(CH_2)_3$	5.62	5.78	0.16	−0.04	0.70	0	0	0.46	−1.24	0.90	0.0	−0.22
265	2,3-$(CH_2)_4$	5.85	6.51	0.66	−0.04	0.93	0	0	0.93	−1.31	1.13	0.0	−0.22
266	2,3-[$OC(CH_3)_2CH_2$]-5-Cl	7.00	5.96	1.04	0.26	0.18	1	0	1.39	−1.24	1.16	0.0	0.14
267	3,4-(OCH_2O)	3.89	4.55	0.66	0.00	0.10	0	0	0.46	−0.62	0.46	0.0	0.00
268	3,4-$(CH_2)_3$	4.74	5.19	0.45	0.00	0.10	0	0	0.46	−1.24	1.02	0.0	−0.22
269	3,4-$(CH_2)_4$	5.08	6.03	0.95	0.00	0.10	0	0	0.93	−1.31	1.96	0.0	−0.22
270	3,4-$(CH=CH)_2$	4.85[a]	5.93	1.08	0.00	0.10	0	0	0.88	−1.24	1.84	0.0	0.09
271	2,3-$(CH=CH)_2$-4-OH	5.28	6.08	0.80	0.03	0.49	0	0	0.88	−1.24	0.88	0.0	−0.07
272	2,3-$(CH=CH)_2$-4-Cl	5.89	6.13	0.24	0.03	0.49	0	0	0.88	−1.24	1.19	0.0	0.36
273	2,3-$(CH=CH)_2$-4-F	5.92	5.95	0.03	0.03	0.49	0	0	0.88	−1.24	0.68	0.0	0.14
274	2,3-[$OC(CH_3)_2CH_2$]-4-Cl	6.55	5.98	0.57	0.26	0.18	1	0	1.39	−1.24	1.16	0.0	0.04
275	2,3-[$OC(CH_3)_2CH_2$]-4-CH_3	6.80	5.82	0.98	0.26	0.18	1	0	1.39	−1.24	1.13	0.0	−0.38
276	4-$S^+(CH_3)_2$	4.96	4.97	0.01	0.00	0.10	0	1	0	0.00	2.15	0.0	1.16
277	4-$N^+(CH_3)_3$	5.46	5.87	0.41	0.00	0.10	0	1	0	0.00	2.32	0.0	0.70
278	3-$S^+(CH_3)_2$	6.19	6.15	0.04	0.00	0.10	0	1	0	−1.71	2.15	0.0	1.00
279	3-$P^+(C_2H_5)_2CH_3$	7.44	7.32	0.12	0.00	0.10	0	1	0	−3.36	3.81	0.0	1.14
280	3-$N^+(CH_3)_3$	7.74	7.11	0.63	0.00	0.10	0	1	0	−2.84	2.32	0.0	0.88
281	3,5-di-$N^+(CH_3)_3$	6.92	7.48	0.56	0.00	0.10	0	2	0	−2.84	4.34	0.0	1.76
282	3-$N^+(CH_3)_3$-5-$N(CH_3)_2$	7.43	8.15	0.72	0.00	0.10	0	1	0	−2.84	3.78	0.0	0.73
283	3-$N^+(CH_3)_3$-5-*i*-C_3H_7	8.15	8.00	0.15	0.00	0.10	0	1	0	−2.84	3.72	0.0	0.81
284	2-$S^+(CH_3)_2$	4.82	4.05	0.77	1.02	1.95	0	1	0	0.00	0.30	0.0	1.16
285	2-$N^+(CH_3)_3$	5.00	5.81	0.81	0.89	2.12	0	1	0	0.00	0.30	0.0	0.70
286	2-CH_3-4-$N^+(CH_3)_3$-5-*i*-C_3H_7	7.96	7.41	0.55	−0.04	0.57	0	1	0	0.00	3.72	0.0	0.48
287	2-*i*-C_3H_7-4-$N^+(CH_3)_3$-5-CH_3	8.00	7.81	0.19	−0.05	1.50	0	1	0	0.00	2.79	0.0	0.54
288	3-*i*-C_3H_7-4-$N^+(CH_3)_3$	8.30	7.67	0.63	0.00	0.10	0	1	0	−1.71	3.72	0.0	0.63

[a] These compounds were not used in the derivation of eq 6.

tracting or -releasing substituents make poorer inhibitors. In the case of ortho substituents there is superimposed on σ_o^- an inductive effect modeled by $\mathcal{F}$. The optimum value of this is 0.33; that is, weak electron withdrawal does appear to increase activity. Otherwise, strong release or withdrawal is similar to the meta and para positions.

We believe that strong electron-attracting groups cause loss of carbamate via hydrolysis. This has been well established for the nitro groups by Fukuto et al. (1967); however, other groups besides NO_2 can also cause significant loss via hydrolysis. There are many groups in the present set having $|\sum\sigma^-| \geq 0.72$, the value of σ leading to a predicted decrease in activity larger than the standard deviation. This predicted loss of activity can occur only with electron-attracting groups since donors in this set, with the exception of one (2,4,6-tri-CH_3, $\sum\sigma^- = -0.45$), do

Table II. Squared Correlation Matrix (r^2) for Collinearity among Variables of Equation 6

	$\Sigma\sigma^a$	$\mathcal{F}_{2,6}$	HB	CHG	MR_2	MR_3	$MR_{3,4,5}$	RGMR	E_{S_2}	E_{S_3}	E_{S_6}
$\Sigma\sigma$	1.00	0.99	0.03	0.21	0.01	0.00	0.00	0.02	0.00	0.02	0.00
$F_{2,6}$		1.00	0.11	0.00	0.06	0.04	0.13	0.01	0.13	0.03	0.06
HB			1.00	0.01	0.00	0.01	0.04	0.14	0.01	0.00	0.01
CHG				1.00	0.00	0.02	0.05	0.00	0.00	0.04	0.00
MR_2					1.00	0.11	0.26	0.01	0.49	0.17	0.01
MR_3						1.00	0.36	0.00	0.17	0.52	0.00
$MR_{3,4,5}$							1.00	0.02	0.36	0.15	0.02
RGMR								1.00	0.03	0.04	0.01
E_{S_2}									1.00	0.17	0.00
E_{S_3}										1.00	0.00
E_{S_6}											1.00

^{a} $\Sigma\sigma = (\Sigma\sigma_{o,p}^- + \sigma_m)$.

not have σ^- values higher than −0.4 so that their predicted depressing activity is smaller than −0.15 (pI_{50}).

Another way of dealing with the detrimental effect of high $\sum\sigma$ on activity is to replace $(\sum\sigma_{o,p}^- + \sigma_m)^2$ with an indicator variable (=1) when the value of this term is ≥0.8 for any given congener (see eq 10). Making this substitution yields essentially the same equation (compare coefficients). This can also be illustrated by dropping all the nitro compounds, leaving the set with weaker attracting groups (see eq 9). Of course, one obtains a poorer correlation with the indicator variable since it does not take into account increments in $\sum\sigma$ which affect hydrolysis in a continuous fashion. Equations 8 and 12 in Table IV predict a decrease of ~1.45 (in pI_{50}) for a 4-NO_2 group so that its "intrinsic" activity for inhibition, were it not hydrolyzed, is $pI_{50} \sim 4.2$; that is, it should be more active than the parent molecule as Nishioka et al. (1977) indeed found (BAcChE).

The set of carbamates in Table I, large as it is, lacks variety in σ^- and $\mathcal{F}$ values. Since the low activity of electron attractors was noticed at the initial experimental stage (Kolbezen et al., 1954), but only later was related (Fukuto et al., 1967) to reduction of PNMC concentration during inhibition measurements, not many more molecules of this type were tested.

The electronic effect of substituents appears complex for groups which have rather large (positive or negative) σ values, but for those with values near zero, the effects are not very significant. The term $(\sum\sigma_{o,p}^- + \sigma_m)^2$ does not become significant until it reaches a value of ~0.5. The prediction of best activity in terms of $\mathcal{F}_{2,6}$ for 2-OR, 2-SR, and 2-X correlates with the observation (Metcalf and Fukuto, 1965a, 1967; Mahfouz et al., 1969) that most of these substituents, contrary to alkyl groups, are more active in the ortho than in the meta and para positions.

The role of the "hydrogen-bonding" variable in eq 6 cannot be specifically defined. Its negative coefficient brings out a deleterious effect on inhibition. Whether this is really due to hydrogen bonding or some other characteristic of these substituents cannot be stated at this time. It may be that the hydrogen-bonding variable is in some way a correction factor on the $\mathcal{F}_{2,6}$ terms. Nishioka et al. (1977) (eq 5) found an enhancing hydrogen-bonding effect on the binding of PNMC inhibitors to BAcChE. Their parameter is quite different from ours in that they postulate that 2-OR and 3-CN, 3-NO_2, and 3-COR are all interacting in the same way with the same hydrogen-bond donor on the enzyme.

Charge effects have also been considered central in binding studies of AcChE (Gearien and Mede, 1976), and the so-called anionic site of the enzyme is a permanent reminder of such an interaction. There is no indication of a separate binding site for charged PNMC as far as we can tell from our analysis. Virtually no change from the original equation is found when all charged compounds are not used for the regression analysis (eq 6). Our feeling (see below) is that since MR correlates binding better than π, binding may be occurring largely in a region composed of polar amino acid residues. If this is true, there would be many possibilities for dipolar interactions not only with charged compounds but also with "neutral" ones. There is a permanent polarization in the PNMC's due to the presence of the phenyl carbamate group. The unsubstituted molecule has a dipole amount of 3.10 D (Exner and Blâha, 1977).

Various authors have drawn attention to the possible importance of charged substituents being located in a highly specific way with respect to the ester or carbamate linkage. In parameterizing charge, we have not found it to be advantageous to make this constant position dependent. Some of the early work appeared to suggest that when groups such as $N^+(CH_3)_3$ are in the meta position, the most active compounds are obtained. It is seen from Table I that the most active compounds (286–288), except for compound 283, carry the charged groups in the 4 position.

Since much of the SAR analysis of PNMC is concerned with substituent effects, the importance of the phenyl ring to binding is prone to be overlooked. Metcalf and Fukuto (1967) found the cyclohexyl and benzyl analogues of PNMC to have considerably lower activity. The need for a "transmission effect" of electronic effects of substituents might be brought up as an explanation; however, this is complicated by the finding that benzyl carbamates are bad carbamylators of BAcChE. We are not able to parameterize these molecules in the framework of this analysis. A separate QSAR study of these compounds would be desirable.

Charge transfer is most likely ruled out as an important mechanism of binding to AcChE unless one invokes different regions for electron-attracting substituents. The finding (Millner and Purcell, 1976) that pyridine carbamates do not seem to bind better than their nonnitrogeneous analogues also discounts charge-transfer interactions. [We compared their data for EEAcChE vs. data from Nishioka et al. (1977) for BAcChE on a common I_{50} basis.] Steinberg et al. (1975) found that π binding was of no importance in the binding of molecules related to tetrahydroaminoacridine to EEAcChE.

Hydrophobic vs. Polar Interaction. The parameter that "dominates" the QSAR of eq 6 is MR (molar refractivity). The experimental values of MR are obtained from the expression $MR = [(n^2 - 1)/(n^2 + 2)](M_r/d)$ where n = refractive index, M_r = molecular weight, and d = density. MR is also linearly proportional to the mean polarizability $\bar{\alpha}$: $MR = {}^3/_4\pi N\alpha$; through the first of the above relations it is connected to "bulk" effects in molecular interactions; through the second it is related to van der Waals free

Table III. Development of Equation 6

no. variables	intercept	$MR_{3,4,5}$	MR_2	E_{s_3}	$(\Sigma\sigma)^2$ [a]	CHG	MR_2^2	$\mathcal{I}_{2,6}^2$	$\mathcal{I}_{2,6}$	RGMR	HB[b]	MR_3^2	$E_{s_2} \cdot E_{s_6}$ [c]	r^2	s	F_1 [d]
1	4.98	0.25												0.082	1.007	24.0
2	4.16	0.52	0.594											0.334	0.860	100
3	3.84	0.43	0.718	−0.52										0.471	0.767	68.7
4	3.94	0.42	0.699	−0.57	−0.65									0.527	0.727	31.6
5	4.16	0.33	0.628	−0.53	−1.12	1.57								0.615	0.658	59.0
6	3.27	0.41	1.298	−0.59	−1.07	1.39	−0.16							0.661	0.618	35.9
7	3.85	0.38	1.343	−0.57	−0.93	1.41	−0.17	−1.26						0.681	0.601	16.4
8	3.74	0.42	1.324	−0.555	−0.93	1.57	−0.18	−4.76	2.68					0.721	0.563	37.4
9	3.63	0.47	1.380	−0.50	−0.86	1.51	−0.18	−4.15	2.22	0.56				0.741	0.542	21.0
10	3.60	0.47	1.473	−0.51	−0.88	1.53	−0.21	−5.71	3.53	0.72	−0.62			0.766	0.517	26.5
11	3.40	0.57	1.574	−0.61	−0.91	1.45	−0.23	−5.69	3.62	0.69	−0.63	−0.05		0.783	0.498	21.4
12	3.46	0.56	1.558	−0.61	−0.94	1.43	−0.23	−5.24	3.46	0.66	−0.62	−0.05	−0.56	0.796	0.485	15.0

[a] $\Sigma\sigma \equiv (\Sigma\sigma_{o,p}^{-} + \sigma_m)$. [b] Hydrogen bonding of OR, NR_2 ortho substituents, HB = 1; otherwise HB = 0. [c] $E_{s_2} \cdot E_{s_6}$ = cross term of E_{s_2} and E_{s_6} (always ≥ 0). [d] $F_{1,120,\alpha 0.001} = 11.38$.

Table IV. Sensitivity of the Equation to the Omission of Certain Data Points

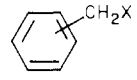
test	eq no.	intercept	$MR_{3,4,5}$	MR_2	Es_3	$(\Sigma\sigma)^2$ [a]	CHG	MR_2^2	$\mathcal{F}_{2,6}^2$	$\mathcal{F}_{2,6}$	RGMR[a]	HB_2[a]	MR_3^2	$E_{s_2} \cdot E_{s_6}$ [a]	ind pull[b]	n	r	s
no charged compds	7	3.46	0.56	1.55	−0.61	−0.90		−0.23	−5.67	3.69	0.66	−0.64	−0.052	−0.53		256	0.878	0.482
no hydrogen-bonded compds	8	3.46	0.55	1.58	−0.63	−0.92	1.39	−0.23	−4.84	3.13	0.23		−0.053	−0.53		239	0.903	0.466
no nitro compds	9	3.42	0.56	1.58	−0.63	−0.79	1.26	−0.23	−5.13	3.55	0.63	−0.60	−0.054	−0.60		255	0.886	0.473
indicator variable instead of $(\Sigma\sigma)^2$	10	3.38	0.57	1.59	−0.59		1.05	−0.23	−5.24	3.30	0.70	−0.59	−0.047	−0.59	−0.867	269	0.873[c]	0.524
no hydrolysis-related variable	11	3.27	0.59	1.69	−0.54		0.72	−0.24	−5.71	3.20	0.82	−0.57	−0.044	−0.47		269	0.848	0.568
full eq	12	3.46	0.56	1.56	−0.61	−0.94	1.43	−0.23	−5.24	3.46	0.66	−0.62	−0.052	−0.56		269	0.892[c]	0.485

[a] These constants are depicted in Table I. [b] ind pull = 1 for all $\Sigma\sigma \geq 0.8$. [c] F test for comparison with eq 5 not having any hydrolysis term; for eq 11, $F_{1,X} = 45.44$; for eq 13, $F_{7,4} = 96.21$; $F_{7,120,\alpha 0.001} = 11.38$.

energies—dispersion, orientation, and induction (all three proportional to the inverse sixth power of the interaction distance)—and higher order energies (Israelachvili, 1974). Orientation energies are of major importance when highly polar molecules are involved, while dispersion energies are dominant for nonpolar and medium polar molecules. MR values are additive as are dispersion forces in sets of congeners (Meyer and Hotz, 1976). The energies involved with dispersion are in the range of hydrogen-bond energies (for atom to atom at relatively short distances). In the meta and para directions "MR space" is less crucial to binding (parameterzied by $MR_{3,4,5}$), and in the 4 position a limit to such interactions (in congener set of Table I) has not been reached. The MR parameter for the 2,3 and 3,4 ring structures has comparable susceptibility coefficients to the $MR_{3,4,5}$ one. Since some fairly long substituents exist in those positions (no. **93** and **136**, Table I), the active site seems to be of large dimensions. Hellenbrand and Krupka (1975) commented on the broad hydrophobic region for binding in AcChE and defined it as being an "open field". A planar structure for the binding site of EEAcChE was proposed by Steinberg et al. (1975). Another recent study of EEAcChE (Abou-Donia et al., 1976) suggested a planar or slightly curved hydrophobic subsite with a radius of >10 Å and a second area designated as a van der Waals subsite. Many suggestions from various sources are found for more than one "anionic site" in AcChE of a variety of kinds (Triggle and Triggle, 1976; O'Brien, 1976; Krupka and Hellenbrand, 1974). The importance of hydrophobic interactions is agreed upon by most of the workers in the field and is thought to be of the utmost importance for the fly head enzyme (Hellenbrand and Krupka, 1975), increasing the inhibition with "bulk and nonpolarity". Since bulk refers to the molecular volume, it does not have to be connected with nonpolarity; in fact, it most obviously is not with groups of the type $N^+(CH_3)_3$, $S^+(CH_3)_2$, etc. There are few examples in the present data set where one can make simple comparisons to aid in deciding the type of surface with which substituents are interacting. Since substituents have complex properties, it is impossible to find two substituents that differ only in π and MR, other factors (σ, E_s, and HB) being constant. The two best possibilities are $N(CH_3)_2$ and $CH(CH_3)_2$; their respective log $(1/C)$ values in position 3 are 5.1 and 6.5, while they are 3.6 and 4.1 in position 4. The respective π values are 0.18 and 1.53. Their σ_m values are quite close (−0.16 and −0.07). Since 3-$CH(CH_3)_2$ is ~30 times as active as 3-$N(CH_3)_2$, one might be tempted to call 3-space hydrophobic. The value of $\Delta\pi$ between $N(CH_3)_2$ and $CH(CH_3)_2$ is 1.35, while ΔMR is only 0.4. A reasonable coefficient to expect with a π term is ~1. Thus this could account for 1.35 out of 1.40 of Δ log $(1/C)$ between two such meta substituents. This *simple* test does suggest 3-space to be hydrophobic; the same is not true for 4-space because in this instance difference in log $(1/C)$ for the two substituents is the same within experimental error. The binding process involved with each of these terms is the focus of disagreement in the field. Belleau and Lacasse (1964), assuming the enzyme's surface to be essentially nonpolar, argued strongly in favor of the solubility of BAcChE substrates in the nonpolar active site, suggesting that van der Waals forces may be operating only as soon as the maximal dissolution in the active site has taken place. This was proposed as an outcome of the distance specificity of the van der Waals forces. The calculations Belleau performed to prove his point were based on Salem's (1961) planar model for hydrocarbon chain induction which may be underestimating the three-dimensional effects. Nishioka et al. (1977) proposed a similar mechanism for PNMC inhibition of BAcChE. They view the binding step as involving the engulfment of substituents into the binding site. Their mechanism is based on the high correlation found with π constants (eq 5). In another study (EEAcChE), Milner and Purcell (1976) suggested the importance of the surface area of ortho substituents in pyridine carbamates as inducing "structure breaking" and expulsion of water from the active site. They could not explain their results in terms of log P.

Belleau's clear-cut distinction between van der Waals and "transfer" energy is at odds with the Wolfenden and Lewis (1976) view, explaining the distribution of hydrocarbons between water and nonpolar solvents as the outcome of the variations of free energies of interaction with the nonpolar solvent which are clearly of the van der Waals type.

The present evidence for the character of the binding areas in AcChE does not permit us to make a clear choice. MR gave better results than π in our preliminary calculations; hence we elected to use this parameter. However, the two variables for the present data set are so highly collinear for the 3 position that we cannot say MR is superior to π in the interactions of substituents at all positions with all of the corresponding enzymic space. A better set of substituents needs to be studied. Unfortunately, many of the polar substituents that one could use to break off colinearity between π and MR have either high or low σ constants inhibiting activity. Insulation of such groups from the ring, e.g.



could alleviate this problem to a certain extent.

The parameter E_s has been shown by Fujita and Nishioka (1976) to be effective in correlating the interactions of ortho substituents on benzene rings with reaction centers. It has also been shown to be of value in rationalizing intermolecular interactions occurring in biochemical systems (Unger and Hansch, 1976). There are two E_s terms in eq 6; the negative slope of E_s indicates that substituents with large van der Waals radii increase activity (the larger the substituent, the more negative its E_s value; E_s for H = 0). The E_s value of substituents increases rapidly in going from methyl to *n*-propyl but then remains essentially constant for larger groups. Hence the negative E_s term in eq 6 shows that, other factors being equal, branching on the X substituent near its attachment to the phenyl ring enhances inhibitory power. E_s is a crude descriptor of this critical steric effect which probably cannot be sharply parameterized with a continuous variable.

The negative coefficient with the cross-product term $E_{s_2} \cdot E_{s_6}$ means that 2,6 disubstitution has more than a simple additive detrimental effect on inhibitory potency. The E_s term accounts for what has been called the "branching" effect (Kohn et al., 1965) of substitution on the α-carbon atom. This is especially clear from the work of Kohn et al. (1965) on meta substitution. This special "volume effect" in the vicinity of the ring at the meta position is not ideally accounted for by eq 6 since some of the very active substituents (no. **92**, **199**, and **208**, Table I) are poorly predicted.

Another difficulty with eq 6 is that of squaring it with the evidence presented in the introduction on the probable positioning of the side-chain carbamate group with the use of σ^- for conjugated substituents. A better selection of 2-X-4-Y derivatives where X has little electronic effect (e.g., 2-CH_3) and Y has an exalted σ value (e.g., 4-CN,

4-SO_2CH_3, etc.) might show that the use of σ^- would only apply when the ortho positions are unsubstituted. An extremely poorly fit compound (no. **232**) has two conjugated NO_2 groups. Compound no. 1 is also poorly fit but in the opposite direction. Clearly, a more careful study of the role of σ and σ^- is called for.

In summary, it can be said that eq 6 brings order to a very large set of chemical structures and inhibition constants. One can quickly see from this equation and the data in Table I what has been explored and what remains to be explored in terms of substituents and parameters.

There are now many examples (Kim et al., 1979) where it has been shown that correlation equations have accurate predictive value. This is to be expected when predictions are made for new congeners having values of activity (in this case, pI_{50}; 2.30–8.30) in the range already studied and having parameter values within the range studied (explored data space). Although correlation equations have also been shown (Kim et al., 1979) to make accurate predictions beyond explored data space, one cannot count on this. The region beyond that explored is truly unknown, and its nautre cannot be forecast anymore than one can forecast the future. Nevertheless, eq 6 is a useful guide for trying to find cogeners with more or less activity than those studied. Being able to predict the activity of congeners within explored data space is no trivial accomplishment. For example, 67 different monovalent substituents and 17 different divalent substituents have been employed in making the congeners of Table I. If we considered making only congeners with the monovalent substituents in positions 2, 3, and 4, this would constitute 67^3 or 300 763 possible compounds. Actually, 67 is a rather small portion of a recently published (Hansch and Leo, 1979) list of well-characterized substituents. If this group of 166 were used in the three positions, it would mean a total of 4 574 296 possible compounds. While many of these would not be worth the trouble to make, one could easily design thousands that could be made without undue synthetic effort. Equation 6 would be an excellent guide to avoid redundance in such an undertaking. One might ask why, since Table I contains inhibitors with pI_{50} of over 8, one would want to make more of such carbamates. In the search for better insecticides, one might want compounds more resistant to metabolism or with a better balance of lipophilic/hydrophilic character or simply a compound with novel substituents for patent purposes. One could easily select from the current list (Hansch and Leo, 1979) of well-characterized substituents many novel substituents that are not present in Table I and, with the help of eq 6, produce highly potent inhibitors having other properties of a different nature and still not step out of explored substituent space.

The real value of correlation equations such as eq 6 does not rest only with their predictive value, important as that is; rather, it is the insight such equations provide us about how small organic compounds interact with macromolecules and macromolecular systems to influence living processes.

ACKNOWLEDGMENT

We thank T. R. Fukuto for several helpful discussions. To Nicholas Dreyer we are indebted for help in the use of APL.

LITERATURE CITED

Abd Elroaf, T. K.; Fahmy, M. A. H.; Fukuto, T. R.; Elsebae, A. H. *J. Econ. Entomol.* **1977**, *70*, 78.

Abou-Donia, M. B.; Rosen, G. M.; Paxton, J. *Int. J. Biochem.* **1976**, *7*, 371.

Belleau, B.; Lacasse, G. *J. Med. Chem.* **1964**, *7*, 768.

Chiriac, A.; Ciubotaru, D.; Szabadai, Z.; Vilceanu, R.; Simon, Z. *Rev. Roum. Biochim.* **1975**, *12*, 143.

Exner, O.; Bláha, K. *Collect. Czech. Chem. Commun.* **1977**, *42*, 2379.

Fujita, T.; Kamoshita, K.; Nishioka, T.; Nakajima, M. *Agric. Biol. Chem.* **1974**, *38*, 1521.

Fujita, T.; Nishioka, T. *Prog. Phys. Org. Chem.* **1976**, *12*, 49.

Fukunaga, J. Y.; Hansch, C.; Steller, E. E. *J. Med. Chem.* **1976**, *19*, 605.

Fukuto, T. R.; Fahmy, M. A. H.; Metcalf, R. L. *J. Agric. Food Chem.* **1967**, *15*, 273.

Gearien, J. E.; Mede, K. A. In "Principles of Medicinal Chemistry"; Foye, W. O., Ed.; Lea and Febiger: Philadelphia, 1976; p 322.

Hansch, C.; Deutsch, E. W. *Biochim. Biophys. Acta* **1966**, *126*, 117.

Hansch, C.; Leo, A. "Substituent Constants for Correlation Analysis in Chemistry and Biology"; Wiley-Interscience: New York, 1979.

Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207.

Hastings, F. L.; Main, A. R.; Iverson, F. *J. Agric. Food Chem.* **1970**, *18*, 497.

Hellenbrand, K.; Krupka, R. M. *Croat. Chem. Acta* **1975**, *47*, 345.

Hetnarski, B.; O'Brien, R. D. *Pestic. Biochim. Physiol.* **1972**, *2*, 132.

Hetnarski, B.; O'Brien, R. D. *Biochemistry* **1973**, *12*, 3883.

Hetnarski, B.; O'Brien, R. D. *J. Agric. Food Chem.* **1975a**, *23*, 709.

Hetnarski, B.; O'Brien, R. D. *J. Med. Chem.* **1975b**, *18*, 29.

Israelachvili, J. N. *Q. Rev. Biophys.* **1974**, *6*, 341.

Kim, K. H.; Hansch, C.; Fukunaga, J. Y.; Steller, E. E.; Jow, P. Y. C.; Craig, P. N.; Page, J. *J. Med. Chem.* **1979**, *22*, 366.

Kohn, G. K.; Ospenson, J. N.; Moore, J. E. *J. Agric. Food Chem.* **1965**, *13*, 232.

Kolbezen, M. J.; Metcalf, R. L.; Fukuto, T. R. *J. Agric. Food Chem.* **1954**, *2*, 864.

Krupka, R. M.; Hellenbrand, K. *Biochim. Biophys. Acta* **1974**, *370*, 208.

Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.

Mahfouz, A. M. M.; Metcalf, R. L.; Fukuto, T. R. *J. Agric. Food Chem.* **1969**, *17*, 917.

Metcalf, R. L.; Fuertes-Polo, C.; Fukuto, T. R. *J. Econ. Entomol.* **1963**, *56*, 862.

Metcalf, R. L.; Fukuto, T. R. *J. Econ. Entomol.* **1962**, *55*, 340.

Metcalf, R. L.; Fukuto, T. R. *J. Agric. Food Chem.* **1965a**, *13*, 220.

Metcalf, R. L.; Fukuto, T. R. *J. Econ. Entomol.* **1965b**, *58*, 1151.

Metcalf, R. L.; Fukuto, T. R. *J. Agric. Food Chem.* **1967**, *15*, 1022.

Metcalf, R. L.; Fukuto, T. R.; Frederickson, M. *J. Agric. Food Chem.* **1964**, *12*, 231.

Metcalf, R. L.; Fukuto, T. R.; Frederickson, M.; Peak, L. *J. Agric. Food Chem.* **1965**, *13*, 473.

Metcalf, R. L.; Fukuto, T. R.; Wilkinson, C.; Fahmy, M. H.; Abd El-Aziz, S.; Metcalf, E. R. *J. Agric. Food Chem.* **1966**, *14*, 555.

Metcalf, R. L.; Fukuto, T. R.; Winton, M. Y. *J. Econ. Entomol.* **1960**, *53*, 828.

Metcalf, R. L.; Fukuto, T. R.; Winton, M. Y. *J. Econ. Entomol.* **1962**, *55*, 345.

Meyer, E. F.; Hotz, C. A. *J. Chem. Eng. Data*, **1976**, *21*, 274.

Millner, O. E., Jr.; Purcell, W. P., Jr. *J. Pharm. Sci.* **1976**, *65*, 910.

Nishioka, T.; Fujita, T.; Kamoshita, K.; Nakajima, M. *Pestic. Biochem. Physiol.* **1977**, *7*, 107.

Nishioka, T.; Kitamura, K.; Fujita, T.; Nakajima, M. *Pestic. Biochem. Physiol.* **1976**, *6*, 320.

O'Brien, R. E. In "Insecticide Biochemistry and Physiology"; Wilkinson, C. F., Ed.; Plenum Press: New York, 1976, p 271.

Salem, L. *Can. J. Biochem. Physiol.* **1961**, *40*, 1287.

Simon, Z.; Szabadai, Z. *Stud. Biophys.* **1973**, *39*, 123.

Steinberg, G. M.; Mednick, M. L.; Maddox, J.; Rice, R.; Cramer, J. *J. Med. Chem.* **1975**, *18*, 1056.

Takusagawa, F.; Jacobson, R. A. *J. Agric. Food Chem.* **1977**, *25*, 329.

Talvik, I. V.; Palm, V. A. *Reakts. Sposobn. Org. Soedin.* **1971**, *8*, 445.

Topliss, J. G.; Edwards, R. P. *J. Med. Chem.* **1979**, *22*, 1238.
Triggle, D. J.; Triggle, C. R. In "Chemical Pharmacology of the Synapse"; Academic Press: New York, 1976.
Unger, S. H.; Hansch, C. *Prog. Phys. Org. Chem.* **1976**, *12*, 91.
Verloop, A.; Hoogenstraaten, W.; Tipker, J. "Drug Design"; Academic Press: Amsterdam, 1976; Vol. VII, p 176.
Wolfenden, R.; Lewis, C. A., Jr. *J. Theor. Biol.* **1976**, *59*, 231.

Received for review August 2, 1979. Accepted May 5, 1980. A.G. thanks the Hebrew University of Jerusalem for financial support for this work. This research was supported by Grant CA-11110 from the National Institutes of Health.

Reduction of Polychlorinated Biphenyl Toxicity and Uptake of Carbon-14 Activity by Plants through the Use of Activated Carbon

Harry J. Strek, Jerome B. Weber,* Patrick J. Shea, Edward Mrozek, Jr., and Michael R. Overcash

The use of soil-applied activated carbon in reducing the phytotoxicity and uptake from soil of polychlorinated biphenyls (PCB's) by a variety of crop plants was investigated. Reductions in growth parameters resulted at the highest rate of PCB for soybean [*Glycine max* (L.) Merr.] and beet (*Beta vulgaris* (L.)]. Growth parameters taken at harvest showed no apparent inhibition of corn (*Zea mays* L.) and sorghum [*Sorghum bicolor* (L.) Moensch] by PCB. The activated carbon treatment substantially reduced growth inhibition caused by PCB. Treatment with soil-applied activated carbon reduced ^{14}C uptake into foliage of beet, corn, sorghum, and peanut (*Arachis hypogaea* L.) in studies using a mixture of U-^{14}C-labeled polychlorinated biphenyls mixed with unlabeled PCB and applied to soil at 20 ppm (total PCB). Activated carbon continued to be effective in reducing ^{14}C uptake over three croppings of fescue (*Festuca arundinacea* Schrib).

Land application of municipal and industrial sludge is becoming an increasingly acceptable method of disposal. Application to agricultural land is viewed as an inexpensive method to exploit the nitrogen and other nutrients of these sludges for use as fertilizer (Overcash and Pal, 1979). However, concern has been voiced over the effect of the organic chemical content of these sludges on crops, and the potential for their uptake by crops needs to be determined (Weber, 1978; Pahren et al., 1979). The PCB content of sludges has been reported to range from <0.01 to 1700 ppm dry weight (Furr et al., 1976; Lawrence and Tosine, 1976; Bergh and Peoples, 1977; Pahren et al., 1979). With application to agricultural land of 2250–22 500 kg of sludge hectare^{-1} as dictatd by nitrogen content (Black and Kronis, 1974), an application range of ≪1 to 1500 kg of PCB ha^{-1} year^{-1} could potentially occur. Since PCB's are resistant to degradation, in particular those which are highly chlorinated (Gustafson, 1970; Kalmaz and Kalmaz, 1979; Pal et al., 1980), and do not leach readily in soils (Tucker et al., 1975; Scharpenseel et al., 1977a; Moza et al., 1979a; Weber, 1980), sludge application over a 3-year period could result in PCB levels ranging from ≪1 to 1000 ppm in the upper 7.6 cm of soil. At more reasonable application rates of 11 200–44 800 kg/ha/year of sludges containing 1–200 ppm of PCB's, one could anticipate a range of <1 to 24 ppm accumulating in the upper 7.6 cm of soil over a 3-year period. PCB levels in soils receiving dried sludge have been reported to range from 0 to >50 ppm (Bergh and Peoples, 1977), and PCB levels as high as 1200 ppm have been reported in some Japanese agricultural soils (Fujiwara, 1975).

Weed Science Center, Crop Science Department (H.J.S., J.B.W., and P.J.S.), Department of Botany (E.M.), and Department of Biological and Agricultural Engineering (M.R.O.), North Carolina State University, Raleigh, North Carolina 27607.

Uptake of PCBs by various crops and weeds from soils containing low amounts of PCB's has been reported for carrots (Iwata et al., 1974), carrots and radishes (Wallnöfer et al., 1975), carrots and sugar beets (Moza et al., 1976, 1979a,c), soybean (Suzuki et al., 1977), soybean and fescue (Weber and Mrozek, 1979), and pigweed and panicum (Strek and Weber, 1980). PCB concentrations in the soils ranged from 0.05 to 100 ppm and levels in the crops reached a maximum mean (average of five isomers) of 13.9 ppm (fresh weight) in carrots grown outdoors for 72 days in soil treated with 100 ppm of Aroclor 1254 (Iwata et al., 1974). Uptake over 4 years of growth by spruce trees from soil fortified with [^{14}C]PCB-treated sludge totaled 0.401 and 0.231 ppm (dry weight) in the needles and stems, respectively (Moza et al., 1979b), showing that plants can take up PCB's from contaminated sludge applied to soil. Disruption of growth in plants has been reported for an aquatic plant *Spirodela oligorrhiza* (Mahanty and Fineran, 1976; Mahanty and McWha, 1976) and for soybean and fescue (Weber and Mrozek, 1979). Aroclor 1242 is reported to inhibit photosynthesis in isolated spinach chloroplasts (Sinclair et al., 1977).

Volatilization appears to be an important route for the loss of PCB's from the soil, despite the moderate to low volatilities of 7×10^{-3} to 4×10^{-5} mmHg at 25 °C reported for Aroclors 1221–1260 (Pal et al., 1980). Losses from a woodland soil ranged from 79.2 to 41.5% of that applied in a single cropping season under outdoor conditions for single isomers possessing two to five chlorines (Moza et al., 1976, 1979a,c). The amount of loss through volatilization appears to depend upon the degree of chlorination, with the more highly chlorinated PCB's being lost to a lesser degree (Iwata et al., 1974; Kilzer et al., 1979) and also more readily adsorbed by the soil (Haque et al., 1974).

The effectiveness of activated carbon as an adsorbent of organic molecules has long been recognized (Mattson and Mark, 1971; Cheremisinoff and Ellerbusch, 1978). Although early investigation of this adsorptive property

0021-8561/81/1429-0288$01.25/0