MOLECULAR PHARMACOLOGY, 10, 954–962

The Quantitative Structure-Activity Relationship of 9-(X-Phenyl)-guanines Reversibly Inhibiting Guanine Deaminase: Quantitative Comparison of Enzyme from Two Sources

CARLO SILIPO[1] AND CORWIN HANSCH

Department of Chemistry, Pomona College, Claremont, California 91711

(Received June 7, 1974)

SUMMARY

SILIPO, CARLO, AND HANSCH, CORWIN: The quantitative structure-activity relationship of 9-(X-phenyl)guanines reversibly inhibiting guanine deaminase: quantitative comparison of enzyme from two sources. *Mol. Pharmacol.* **10,** 954–962 (1974).

The structure-activity relationship for 9-(X-phenyl)guanines inhibiting guanine deaminase (EC 3.5.4.3) has been studied using substituent constants and regression analysis. The resulting quantitative relationship confirms what one might expect *a priori:* there are two types of space in or on an enzyme with which substituents might interact. Substituents in position 4 of the phenyl ring interact with hydrophobic space, while substituents in position 3 interact with polar space. The electronic effects of substituents on the phenyl ring appear to be of very minor importance. Steric effects of 2-functions are quite important. Equations were formulated for enzymes from two sources: rabbit liver and Walker 256 rat tumor. While different sets of congeners were studied with the two different enzymes, similar equations were obtained, showing the great similarity in enzyme from two different sources.

INTRODUCTION

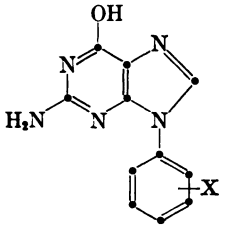
Guanine deaminase (EC 3.5.4.3) is an enzyme which converts guanine to xanthine. In his search for enzyme inhibitors of possible antitumor activity, B. R. Baker (1) was drawn to a study of this system because it had been proposed that the selective action of thioguanine on certain tumors might be due to the lack of guanine deaminase in these susceptible cells; that is, if thioguanine could not be converted to thioxanthine within the cell, death of the cell would occur. Baker first tried to discover highly effective reversible inhibitors of an enzyme (2). After the discovery of a highly active reversible inhibitor, a function (such as —SO_2F) active with nucleophiles could be attached to produce an irreversible inhibitor. His theory was that while there might not be much difference in the active sites of an enzyme obtained from different types of tissue, one could expect isoenzyme features to be found in the amino acid moieties near the active site. He hoped that an —SO_2F function could be placed on the inhibitor in such a way that it could react

This investigation was supported by Public Health Service Research Grant CA-11110 from the National Cancer Institute.

[1] North Atlantic Treaty Organization Fellow and Visiting Professor of Chemistry from the Institute of Pharmaceutical and Toxicological Chemistry, University of Naples, Italy.

with a nucleophile in a region outside the active site. It was Baker's expectation that differences in the nearby amino acids in the isoenzyme of the pathogenic cell and isoenzyme from the host cell could be exploited to produce inhibitors with a high therapeutic index. Unfortunately, Baker did not live long enough to explore this interesting idea completely.

Our concern in this report is with the reversible inhibitors of guanine deaminase of type I:



I

Baker and his students (3–6) studied the inhibition of guanine deaminase from two sources: rabbit liver and Walker 256 rat tumor. The concentrations of inhibitors producing 50% inhibition (as log $1/C$) of enzyme from rabbit liver, along with the substituent constants used in the regression analyses, are assembled in Table 1. The corresponding data for studies on enzyme from Walker tumor are listed in Table 2. The techniques used in the formulation of quantitative structure-activity relationships have received considerable discussion and are now understood moderately well (7–10).

METHODS

C in log $1/C$ of Table 1 represents the molar concentration for 50% reversible inhibition of guanine deaminase from rabbit liver when assayed at pH 7.4. C in Table 2 represents the molar concentration for 50% reversible inhibition of guanine deaminase from Walker 256 rat tumor when assayed under the same experimental conditions.

To estimate the hydrophobic interaction with the enzyme, π-2, π-3, and π-4 have been used for substituents in the *ortho*, *meta*, and *para* positions of the N-phenyl ring.

The hydrophobic constant π is from the benzene system (11); that is, π is defined as

$$\pi_X = \log P_X - \log P_H$$

where P_H is the partition coefficient of the parent molecule between 1-octanol and water, and P_X is that of the derivative X. Many of the π values of Tables 1 and 2 have been reported previously (11); others were estimated from additive-constitutive properties (12, 13).

For example, the calculated π value for the 3-$NHCOC_6H_4$-3′-Cl is the sum

$$\pi_{NHCOC_6H_5} + \pi_{Cl} = 0.49 + 0.71 = 1.20$$

The aliphatic value for π has been used for Br in aliphatic moieties. Thus π_{NHCOCH_2Br} is the sum

$$\pi_{NHCOCH_3} + \pi_{Br(aliphatic)} = -0.97 + 0.60 = -0.37$$

To estimate $\pi_{OCH_2CH_2CH_2C_6H_5}$, one must consider that $\pi_{OCH_3} \sim 0$; that is, π for the whole substituent $OCH_2CH_2CH_2C_6H_5$ is simply

$$\pi_{CH_2CH_2C_6H_5} = 2.66$$

In the case of the α- and β-naphthalene derivatives the π values for the *ortho*, *meta*, and *para* functions were estimated as

$$\tfrac{1}{2}\pi_{3,4\text{-}CH=CH-CH=CH-CH} = 0.66$$

It has also been assumed that the same value of π can be employed for *meta* and *para* substituents. Since both molar refraction and π are, to a certain extent, dependent on molar volume, there is a certain amount of correlation (14) between them. With the present data it is clear from Table 3 that the vectors π-3 and MR-3[2] are orthogonal; however, there is considerable collinearity between π-4 and MR-4. Hence there is some ambivalence in interpreting the role of 4-substituents with respect to these latter two variables.

The MR values have been calculated as previously reported (11) and have been scaled by 0.1. Taft's E_s parameter (15) is employed to correlate the steric effects of substituents in the *ortho* position of the N-phenyl ring. E_s values for the particular functions in the present paper were those obtained from the regression equation (16)

$$E_s = -1.839r_{vave} + 3.484$$

[2] The abbreviation used is: MR, molar refraction.

TABLE 1

Inhibition constants and physicochemical parameters for inhibition of guanine deaminase (from rabbit liver) by 9-(X-phenyl)guanine

C is the concentration which produces 50% inhibition. The calculated value of log $1/C$ is obtained using Eq. 1. π-4 is the hydrophobic constant for substituents in position 4 of the phenyl ring. MR-3 is the molecular refractivity of substituents in position 3. E_s-2 is the steric factor for substituents in position 2. D is an indicator variable for the presence of a 4-oxygen atom. σ-2,3,4 is the sum of σ constants for substituents on the phenyl ring.

Compound	X	Log $1/C$		$\lvert\Delta \log 1/C\rvert$	π-4[b]	MR-3[b]	E_s-2[b]	D	σ-2,3,4[b]
		Observed[a]	Calculated						
1	2-Br	3.92	3.906	0.01	0.0	0.103	0.08	0.0	0.71
2	2-Cl	4.00	4.115	0.11	0.0	0.103	0.27	0.0	0.67
3	4-$N(CH_3)_2$	4.47	5.249	0.78	0.18	0.103	1.24	0.0	−0.83
4	4-NH_2	4.58	4.713	0.13	−1.23	0.103	1.24	0.0	−0.66
5	2-F	4.68	4.675	0.00	0.0	0.103	0.78	0.0	0.47
6	2,3-CH=CH—CH=CH	4.85	4.745	0.10	0.0	0.873	0.36	0.0	0.04
7	4-$NHCOCH_2Br$	4.89	5.040	0.15	−0.37	0.103	1.24	0.0	−0.03
8	4-$CONH_2$	4.89	4.614	0.28	−1.49	0.103	1.24	0.0	0.36
9	H	5.00	5.181	0.18	0.0	0.103	1.24	0.0	0.0
10	4-$C(CH_3)_3$	5.03	5.933	0.90	1.98	0.103	1.24	0.0	−0.20
11[c]	3,4-$(OCH_3)_2$	5.15	6.981	1.83	−0.02	0.787	1.24	1.0	−0.15
12	3-NH_2	5.23	5.483	0.25	0.0	0.542	1.24	0.0	−0.16
13	4-CF_3	5.27	5.515	0.24	0.88	0.103	1.24	0.0	0.54
14	3-CF_3	5.35	5.456	0.11	0.0	0.502	1.24	0.0	0.43
15	3-OCH_3	5.42	5.652	0.23	0.0	0.787	1.24	0.0	0.12
16	4-Cl	5.42	5.451	0.03	0.71	0.103	1.24	0.0	0.23
17	4-CH_3	5.45	5.394	0.06	0.56	0.103	1.24	0.0	−0.17
18	3-Cl	5.60	5.525	0.07	0.0	0.603	1.24	0.0	0.37
19	3-CH_3	5.70	5.499	0.20	0.0	0.565	1.24	0.0	−0.07
20	4-C_2H_5	5.89	5.568	0.32	1.02	0.103	1.24	0.0	−0.15
21	4-$CH(CH_3)_2$	5.92	5.762	0.16	1.53	0.103	1.24	0.0	−0.15
22	4-OH	6.00	6.262	0.26	−0.67	0.103	1.24	1.0	−0.37
23	3,4-Cl_2	6.03	5.795	0.23	0.71	0.603	1.24	0.0	0.60
24	4-$(CH_2)_3CH_3$	6.19	5.941	0.25	2.00	0.103	1.24	0.0	−0.16
25	3-NHCHO	6.36	5.821	0.54	0.0	1.031	1.24	0.0	0.19
26	4-C_6H_5	6.47	5.926	0.54	1.96	0.103	1.24	0.0	−0.01
27	3-C_6H_5	6.62	6.859	0.24	0.0	2.536	1.24	0.0	0.06
28	3,4-CH=CH—CH=CH	6.66	5.963	0.70	0.66	0.873	1.24	0.0	0.04
29	4-OCH_3	6.70	6.509	0.19	−0.02	0.103	1.24	1.0	−0.27
30	3-$NHCOCH_2Br$	6.77	6.681	0.09	0.0	2.278	1.24	0.0	0.17
31	4-OC_2H_5	7.01	6.661	0.35	0.38	0.103	1.24	1.0	−0.24
32	4-$O(CH_2)_3C_6H_5$	7.25	7.528	0.28	2.66	0.103	1.24	1.0	−0.24
33	3-$NHCOC_6H_5$	7.30	7.499	0.20	0.0	3.464	1.24	0.0	0.02

[a] From refs. 3 and 4.

[b] See METHODS for sources of these constants.

[c] This point was not used in deriving Eqs. 1–5.

where $r_V ave$ is the group radius of Bondii (17). To obtain E_s for the 2,3-$(CH)_4$ moiety in the naphthyl moiety, we assumed that only the first carbon atom is involved in the steric effect. We have used the value of 1.70 for the carbon radius in the above regression equation for the calculation of E_s. This mode of calculating E_s from van der Waals radii is

TABLE 2

Inhibition constants and physicochemical parameters for inhibition of guanine deaminase (from Walker 256 rat tumor) by 9-(X-phenyl)guanine

C is the concentration with produces 50% inhibition. The calculated value of log $1/C$ is obtained using Eq. 6. π-4 is the hydrophobic constant for substituents in position 4 of the phenyl ring. MR-3 is the molecular refractivity of substituents in position 3. D is an indicator variable for the presence of a 4-oxygen atom. σ-3,4 is the sum of σ constants for substituents on the phenyl ring.

Compound	X	Log $1/C$		$\lvert\Delta \log 1/C\rvert$	π-4[b]	MR-3[b]	D[b]	σ-3,4[b]
		Observed[a]	Calculated					
1	4-$CONH_2$	4.64	4.498	0.14	−1.49	0.103	0.0	0.36
2	4-NH_2	4.66	4.618	0.04	−1.23	0.103	0.0	−0.66
3	4-$NHCOCH_2Br$	5.00	5.012	0.01	−0.37	0.103	0.0	−0.03
4	3-NH_2	5.05	5.437	0.39	0.0	0.542	0.0	−0.16
5	4-CH_3	5.09	5.439	0.35	0.56	0.103	0.0	−0.17
6	4-CF_3	5.10	5.586	0.49	0.88	0.103	0.0	0.54
7	H	5.10	5.182	0.08	0.0	0.103	0.0	0.0
8	3-CH_3	5.58	5.450	0.13	0.0	0.565	0.0	−0.07
9	4-C_2H_5	5.64	5.650	0.01	1.02	0.103	0.0	−0.15
10	4-Cl	5.65	5.508	0.14	0.71	0.103	0.0	0.23
11	4-C_6H_5	6.27	6.082	0.19	1.96	0.103	0.0	−0.01
12	4-OCH_3	6.46	6.498	0.04	−0.02	0.103	1.0	−0.27
13	3,4-CH=CH—CH=CH	6.55	5.932	0.62	0.66	0.873	0.0	0.04
14	3-$NHCOCH_2Br$	6.70	6.445	0.25	0.0	2.278	0.0	0.17
15	4-OC_2H_5	6.72	6.682	0.04	0.38	0.103	1.0	−0.24
16	3-$NHCOC_6H_3$-2′,4′-Cl_2	7.33	7.713	0.38	0.0	4.464	0.0	0.02
17	3-$NHCOC_6H_4$-3′-Cl	7.44	7.423	0.02	0.0	3.964	0.0	0.02
18	3-$NHCOC_6H_4$-2′-Cl	7.60	7.423	0.18	0.0	3.964	0.0	0.02

[a] From refs. 5 and 6.

[b] See METHODS for sources of these constants.

TABLE 3

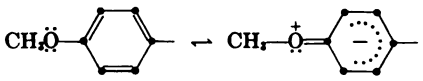
Squared correlation matrix

Numbers in this table show the percentage correlation (r^2) between each of the variables used in the correlation study of guanine deaminase from rabbit liver.

	π-4	π-3	π-2	MR-4	MR-3	MR-2	E_s-2	D	σ-2,3,4
π-4	1.00	0.00	0.02	0.45	0.03	0.02	0.02	0.01	0.01
π-3		1.00	0.00	0.01	0.13	0.01	0.00	0.01	0.03
π-2			1.00	0.06	0.01	0.95	0.98	0.02	0.19
MR-4				1.00	0.13	0.05	0.07	0.11	0.15
MR-3					1.00	0.00	0.01	0.04	0.01
MR-2						1.00	0.89	0.01	0.13
E_s-2							1.00	0.02	0.22
D								1.00	0.12
σ2,3,4									1.00

based on the work of M. Charton (18). The values of σ for *meta* and *para* substituents were taken from our recent compilation (11). We have taken Charton's values (19) from the anilium series for *ortho* functions. The "dummy" parameter D was used to account for the exaltation in activity of *para* OH and OR functions. The utility of σ^+, $\mathfrak{F}$, and $\mathfrak{R}$ was also explored. The modified Hammett constant σ^+ has been widely used to uncover

special direct resonance interactions between substituents and a reaction center in the molecule where the delocalization of a positive charge is important (7). In the present instance our concern was for interactions of the 4-OCH_3:



It was felt that an electron-deficient site in the enzyme might interact with the ring electrons in such a way that it would be greatly facilitated by resonance with the 4-OCH_3 function.

The parameters $\mathcal{F}$ and $\mathcal{R}$ have been formulated by Swain and Lupton (11) to account for the two types of electronic character of substituents, inductive ($\mathcal{F}$) and resonance ($\mathcal{R}$). *Sigma* constants have long been assumed to represent the resultant sum of these two forces: $\sigma_X = \mathcal{F}_X + \mathcal{R}_X$. Factoring σ in this fashion allows one to explore the two types of electronic substituent space independently.

π-2 were usually negative, indicating large groups lower the activity, it seems safe to conclude that a steric effect is involved, and E_s-2 is the parameter of choice. MR often appears to be a good measure of intermolecular interactions in biochemical processes (21–23). Four types of electronic effects were considered; however, σ, σ^+, $\mathcal{F}$, and $\mathcal{R}$ do not appear to have a significant role in the quantitative structure-activity relationships. In the preliminary studies with the aforementioned constants, which gave fair correlations, it was observed that congeners with alkoxy functions in position 4 were invariably much more active than predicted by regression equations. Therefore a dummy or indicator variable (24) was employed to account for this structural feature. This produced quite a significant improvement in correlation. In all the initial studies compound 11 (3,4-dimethoxy) of Table 1 was found to be much less active than expected. This data point was deleted from subsequent analyses. The "best" equation obtained is

$$\log 1/C = 0.380\ (\pm 0.15)\ \pi\text{-}4 + 0.690\ (\pm 0.17)\text{MR-}3 + 1.100\ (\pm 0.45)E_s\text{-}2 + 1.336\ (\pm 0.41)D + 3.747\ (\pm 0.51) \quad (1)$$

n	r	s
32	0.928	0.363

RESULTS

In studying the structure-activity relationship of the 9-phenylguanine inhibitors of guanine deaminase from rabbit liver, the following parameters (11) for the substituents were considered: π-2, π-3, π-4, E_s-2, MR-2, MR-3, MR-4, ΣMR, $\Sigma\pi$, $\Sigma\pi^2$, $\Sigma\sigma$, $\Sigma\sigma^+$, $\Sigma\mathcal{F}$, $\mathcal{R}$-3, $\mathcal{R}$-4, and D. There is considerable evidence to support the directional nature of the hydrophobic effect of substituents (20); hence π has been examined for each of the three positions (*ortho*, *meta*, and *para*) on the phenyl ring (π-2, etc.) as well as in the composite sense ($\Sigma\pi$). Four of the substituents on the phenyl ring are in position 2. Since Taft's steric constants are available for these, E_s-2 was studied. It was found to model substituent effects of *ortho* functions better than MR-2 or π-2, although there is high collinearity among these three vectors (Table 3). Since the signs of the coefficients in the better correlations using MR-2 and

In Eq. 1 C is the molar concentration of inhibitor causing 50% inhibition, n represent the number of data points used in formulats ing the equation, r is the correlation coefficient, s is the standard deviation, and the figures in parentheses are the 95% confidence intervals. By "best" equation is meant the equation with the largest number of terms, all of which are justified by the stepwise application of the F test. Equation 1 is highly significant ($F_{4,27} = 42$; $F_{4,27,\alpha.005} = 5.4$). Moreover, there are eight data points per variable. In order to be sure that no significant linear combinations of the parameters mentioned above would give a better correlation than Eq. 1, 11 of the parameters (MR-2, σ^+, $\mathcal{F}$, and $\mathcal{R}$ were omitted) were used to derive all possible equations ($2^{11} - 1 = 2047$).

All linear combinations of the 11 variables were made. The general equation for such

combinations is

$$\text{No. of combinations} = \frac{n!}{(n-i)!\,i!}$$

where n represents the number of variables and i the number taken at one time. For example, the number of three-variable equations is 11!/(11-3)!3! = 165. This general approach to regression analysis, in which one considers all possible linear combinations of the independent variables, has been discussed by Daniel and Wood (24). Actually, because of some collinearity in the variables resulting in singular matrices, only 1920 equations resulted. Of these, the equation with the lowest standard deviation was

$$\log 1/C = 0.38\pi\text{-}4 + 0.67\text{MR-}3 + 1.28E_s\text{-}2 + 1.43D + 0.352\Sigma\sigma + 3.52 \qquad n = 32,\ r = 0.935,\ s = 0.352 \qquad (2)$$

Although the standard deviation of Eq. 2 is lower than Eq. 1, the additional term in $\Sigma\sigma$ is not significant ($F_{1,26} = 2.72$; $F_{1,26\,\alpha.10} = 2.91$). A better selection of substituents might reveal a small positive effect for electron-withdrawing functions.

Looking at the structure of Eq. 1 stepwise, we find the following order:

$$\log 1/C = 1.65E_s\text{-}2 + 3.79 \qquad n = 32,\ r = 0.553,\ s = 0.783 \qquad (3)$$

$$\log 1/C = 1.52E_s\text{-}2 + 0.49\text{MR-}3 + 3.80 \qquad n = 32,\ r = 0.698,\ s = 0.674 \qquad (4)$$

$$\log 1/C = 1.27E_s\text{-}2 + 0.61\text{MR-}3 + 1.38D + 3.72 \qquad n = 32,\ r = 0.853,\ s = 0.499 \qquad (5)$$

E_s-2 is the most important single variable; it stands out because of its highly deleterious effect on activity. The next most important variable is MR-3, followed by D and then π-4 (Eq. 1). Except for the coefficients with E_s-2, one observes a constancy in the coefficients in each of the terms and intercepts in Eqs. 1–5 which testifies to the independence in the variables, as is of course apparent in the correlation matrix of Table 3.

Equation 6 has been formulated from the data in Table 2 for inhibition of Walker 256 rat tumor enzyme for comparison with Eq. 1. Equation 6 does not contain an E_s term, since none of the 2-substituted derivatives were tested on Walker enzyme. When the confidence limits are taken into consideration, the coefficients of Eq. 6 do not appear to differ significantly from those of Eq. 1. Subtracting 1.24 × 1.1 (E_s for H) from the intercept of Eq. 6 gave a value of 3.76 compared with 3.75 for Eq. 1. Equation 6 is highly significant ($F_{3,14} = 57$; $F_{3,14\,\alpha.005} = 6.7$).

$$\log 1/C = 0.459\ (\pm 0.20)\pi\text{-}4 + 0.580\ (\pm 0.10)\text{MR-}3 + 1.325\ (\pm 0.49)D + 5.122\ (\pm 0.20) \qquad n = 18,\ r = 0.961,\ s = 0.297 \qquad (6)$$

DISCUSSION

In view of the large variations in substituents of the compounds of Tables 1 and 2, Eqs. 1 and 6 are very good correlations. The standard deviation of about 0.3 indicates that the I_{50} concentration is predictable within a factor of ±2. There is a 2500-fold range in concentrations in Table 1 and a 1000-fold range in concentrations in Table 2. One could expect better results if K_i had been determined instead of I_{50}; however, the results do show that I_{50} is a valuable constant in the study of structure-activity relationships.

Baker was well aware that 2-substituents greatly reduced inhibitory activity, and he attributed this to the necessity that the 9-phenyl ring be coplanar with the heterocyclic ring. While this steric effect is quantitatively described by the E_s-2 term in Eq. 1, there is no way of knowing exactly what is involved. Baker was also aware of the activating effect of the oxygen atom in position 4, whose effect we have quantified by the use of D. He considered it likely that the oxygen lone pair electrons might be interacting with a phenyl ring of the enzyme in a kind of charge transfer complex. This seems less likely to us than the possibility of an interaction with a positive center such as an $^{+}NH_3$. At first it was thought that a special resonance interaction might be responsible

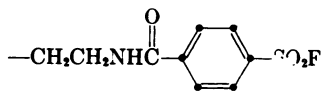
for this increased activity. For this reason the use of σ^+, $\mathcal{F}$, and $\mathcal{R}$ was explored, but without success. Since this effect of OH and OR in position 4 cannot be attributed to an electronic, steric, or hydrophobic effect on the ring, little is left but to postulate a special interaction of the lone pair electrons with some electron-deficient center of the enzyme. The use of the dummy parameter establishes the fact that as far as MR-4 goes, these 4-substituents behave like all other 4-substituents. It is only through careful regression analysis that such interactions can be delineated with some degree of certainty.

It is surprising that neither the 4-NH_2 nor the 4-$N(CH_3)_2$ shows this great increase in activity. This may be due to the much more hydrophilic NH_2 seeking out a slightly different binding point.

One data point (compound 11 in Table 1) was not included in the regression analysis. This derivative (3,4-di-OCH_3) is about 70 times less active than Eq. 1 predicts. In this instance one does not see the usual exaltation shown by 4-OR functions. The reason for this is not obvious, but it may be that the strong inductive-field effects of the 3-oxygen atom reduce the availability of the lone pair electrons on the 4-oxygen atom.

A most interesting feature of Eq. 1 is the different parameters, π-4 and MR-3, for substituents in the *meta* and *para* positions. This clearly signals the difference in *meta* and *para* space; that is, the space adjacent to those two positions. Since there can be considerable collinearity between the π and MR vectors (14), it is important to note in Table 3 that π-3 and MR-3 are essentially independent. The message from the MR-3 term is that *meta* space in the enzyme is not hydrophobic. Substituents in *meta* space cause inhibition by distortion of the enzyme or by bonding through dispersion forces or a combination of both. This appears to be a qualitatively different process from the desolvation process which is so important in hydrophobic interactions. Of course, dispersion forces are involved in both π and MR; however, MR effects appear to occur in polar space.

Para space appears to be hydrophobic, although there is some uncertainty because of some collinearity between π-4 and MR-4. From Table 3, $r^2 = 0.45$ or $r = 0.67$ for the correlation between these two vectors. Hence the angle separating the vectors is arcosine 0.67 or 48°. Ideally, one would like something approaching 90°. That π-4 is much the superior parameter for 4-functions can be seen by substituting MR-4 for π-4 in Eq. 1. This yields an equation with r of 0.879 compared to 0.928 for Eq. 1. Since Eq. 5 has an r of 0.853, MR-4 contributes little to reduction in the variance in the data. This suggests that *para* space is hydrophobic. It is of unusual interest that Baker arrived at similar conclusions about the nature of *meta* and *para* space by quite different reasoning (5). He noted that when an $-SO_2F$ moiety was projected into *para* space, the compound was generally found to be a reversible inhibitor, not an irreversible inhibitor. All but one of the irreversible inhibitors were cases in which the $-SO_2F$ function was projected into *meta* space (5). The one exception was the *para* substituent



It is possible that this long group extends beyond hydrophobic *para* space. Baker rationalized that $-SO_2F$ in para space failed to yield irreversible inhibitors because of the paucity of nucleophilic groups in this apolar region. In *meta* space, which Baker concluded was polar, there would be a much greater probability that an $-SO_2F$ function could find a suitably positioned nucleophile for the covalent bond formation necessary for a truly irreversible inhibitor. In the development of Eq. 1 we paid no attention to Baker's reasoning since, for us, most of it seemed difficult to follow. However, our mathematical expression turns out to square with all of Baker's qualitative thinking! Moreover, it shows the importance of Baker's modus operandi of first studying the reversible inhibitors of an enzyme before attempting to design irreversible inhibitors. Equation 1 shows how one can establish the polar or apolar nature of substituent space and in this way make a better decision about the placement of functions suitable for covalently attaching inhibitor to enzyme.

It must be remembered that any signifi-

cant amount of space in an enzyme cannot be truly homogeneous. For the present all that can be said is that two general types of space are known to be present in proteins and that π and MR appear to model substituent interactions for the limiting cases. In-between situations, in which a substituent is partly in hydrophobic space and partly in polar space, may be modeled more or less equally well by π or MR, depending on the percentage of each type of space involved.

Since the addition of a term in $(\pi\text{-}4)^2$ to Eq. 1 did not give a significant improvement in correlation, one should be able to make more potent inhibitors by increasing the hydrophobicity of 4-substituents. The use of cross-product interaction terms such as $(\pi\text{-}4 \cdot \text{MR-}3)$ in the correlation did not appear to be significant. This is not surprising, since there are few 3,4-disubstituted congeners.

Although Baker studied a variety of congeners containing an $-SO_2F$ group, we have not included them in this analysis because of the uncertainty about their reversible or irreversible action. We are now working with these derivatives.

The assertion is commonly made that adding more terms to a regression equation will in no case lower the correlation and will often improve it. While this is true as far as the correlation coefficient goes, it is not true of the more rigorous test of improved correlation, the standard deviation. The fact that generating over 1900 equations did not produce a more significant correlation than Eq. 1 shows that when one has a sufficient number of data points per variable and employs parameters with real chemical meaning in a good mathematical model, it is by no means easy to obtain a higher correlation by the random addition of irrelevant parameters. In the present case there are four important structural features which must be accounted for to formulate Eq. 1. When these features are taken care of, the addition of 1900 combinations of seven different meaningless (for the present structure-activity relationship) variables in no case leads to a further significant reduction in variance.

Equation 6 establishes the fact in quantitative terms that for reversible inhibitors, enzymes from rabbit liver and Walker 256 tumor are, to a first approximation, identical. Since no 2-substituted derivatives were tested against Walker enzyme, no E_s-2 term appears in Eq. 6. Another interesting point is that congeners 16–18 of Table 2 were not tested against the rabbit liver enzyme. Nevertheless, the coefficients with the variables in Eq. 6 are, within the confidence limits, the same as in Eq. 1.

One would not be justified in using the indicator variable D for the set of data in Table 2 taken alone, since only two data points have features modeled by this term. However, since the coefficient with D in Eq. 1 is much better established (four data points), the fact that the two coefficients are the same establishes the similar role of 4-OR functions in each type of enzyme with some degree of certainty.

Equations 1 and 6 can be used to develop more potent inhibitors, especially for use *in vivo*. Since adding a term in $(\pi\text{-}4)^2$ did not result in an improved correlation, the indication is that more hydrophobic groups could be placed in position 4. However, over-all hydrophobicity is of the utmost importance for activity *in vivo* (25). Low over-all hydrophobicity could be maintained by offsetting large lipophilic 4-substituents with very hydrophilic 3-substituents. Such molecules can now be designed in advance, using the extensive sets of known substituent constants (11). Since Baker was able to make irreversible inhibitors which were selective in their attack on the enzymes from two sources, it would seem likely that by proper regulation of the over-all lipophilic character of congeners, better antitumor potentiators for thioguanine could be developed. The crucial importance of over-all lipophilicity for antitumor drugs has been demonstrated recently (26).

REFERENCES

1. Baker, B. R. (1967) *J. Med. Chem.*, **10**, 59–61.
2. Baker, B. R. (1967) *Design of Active-Site-Directed Irreversible Enzyme Inhibitors*, Wiley, New York.
3. Baker, B. R. & Wood, W. F. (1967) *J. Med. Chem.*, **10**, 1101–1105.
4. Baker, B. R. & Wood, W. F. (1968) *J. Med. Chem.*, **11**, 644–649.

5. Baker, B. R. & Wood, W. F. (1969) *J. Med. Chem.*, **12**, 216–220.
6. Baker, B. R. & Siebeneick, H. U. (1971) *J. Med. Chem.*, **14**, 802–805.
7. Hansch, C. (1973) in *Structure-Activity Relationships* (Cavallito, C. J., ed.), Vol. 1, pp. 75–165, *International Encyclopedia of Pharmacology and Therapeutics*, Sect. 5, Pergamon Press, Oxford.
8. Tute, M. S. (1971) *Adv. Drug Res.*, **6**, 1–77.
9. Verloop, A. (1972) in *Drug Design* (Ariëns, E. J., ed.), Vol. 3, pp. 133–187, Academic Press, New York.
10. Franke, R. & Oehme, P. (1973) *Pharmazie*, **28**, 489–508.
11. Hansch, C., Leo, A., Unger, S. H., Kim, K. H., Nikaitani, D. & Lien, E. J. (1973) *J. Med. Chem.*, **16**, 1207–1216.
12. Leo, A., Hansch, C. & Elkins, D. (1971) *Chem. Rev.*, **71**, 525–616.
13. Hansch, C., Leo, A. & Nikaitani, D. (1972) *J. Org. Chem.*, **37**, 3090–3092.
14. Hansch, C., Unger, S. H. & Forsythe, A. B. (1973) *J. Med. Chem.*, **16**, 1217–1222.
15. Taft, R. W., Jr. (1956) in *Steric Effects in Organic Chemistry* (Newman, M. S., ed.), pp. 556–675, Wiley, New York.
16. Kutter, E. & Hansch, C. (1969) *J. Med. Chem.*, **12**, 647–652.
17. Bondii, A. (1966) *J. Phys. Chem.*, **68**, 441–451.
18. Charton, M. (1968) *J. Am. Chem. Soc.*, **91**, 615–618.
19. Charton, M. (1971) in *Progress in Physical Organic Chemistry* (Streitwieser, A., Jr. & Taft, R. W., Jr., eds.), Vol. 8, pp. 239–241, Wiley/Interscience, New York.
20. Hansch, C., Kim, K. H. & Sarma, R. H. (1973) *J. Am. Chem. Soc.*, **95**, 6447–6449.
21. Agin, D., Hersh, L. & Holtzman, D. (1965) *Proc. Natl. Acad. Sci. U. S. A.*, **53**, 952–958.
22. Hansch, C. & Coats, E. (1970) *J. Pharm. Sci.*, **59**, 731–743.
23. Hansch, C. & Silipo, C. (1974) *J. Med. Chem.*, **17**, 661–667.
24. Daniel, C. & Wood, F. S. (1971) *Fitting Equations to Data*, pp. 55, 169, 203, Wiley/Interscience, New York.
25. Hansch, C. & Clayton, J. M. (1973) *J. Pharm. Sci.* **62**, 1–21.
26. Hansch, C., Smith, N., Engle, R. & Wood, H. (1972) *Cancer Chemother. Rep.* **56**, 443–456.