Structure–Activity Analysis of Tetrahydrofolate Analogs Using Substituent Constants and Regression Analysis

By ELINOR MILLER and CORWIN HANSCH

The binding to dihydrofolate reductase of a series of substituted pyrimidines and triazines is shown to be dependent on the electron-contributing and lipophilic character of the substituents, as expressed by the parameters π and σ. The configuration of binding is variable, and is determined by the necessity of placing the most lipophilic substituent in a hydrophobic region of the enzyme. When two or more lipophilic substituents are present, competition between them for the hydrophobic site is dominated by the more lipophilic group, while the less lipophilic group contributes to hydrophobic bonding to a slight extent. Suggestions are made for the design of a new tetrahydrofolate analog with possible application as a nonclassical antimetabolite.

RECENT WORK has shown that biological response to congeneric drugs may be quantitatively correlated with molecular structure by means of substituent constants (1, 2). Two of the more useful parameters are Hammett's σ constant for electronic effects of substituents, and π, which is related (3) to the lipophilic character of a substituent ($\pi = \log P_X - \log P_H$ where P_X is the octanol–water partition coefficient of a derivative and P_H is that of the parent molecule). The combination of σ and π, as well as the steric parameter (4) of Taft (E_s), has been applied to a variety of compounds and systems, including the action of penicillins on bacteria (5), the relative sweetness of aromatic nitroamines (6), the binding of organic compounds by proteins (7, 8), the inhibition of photosynthesis by amides (9), the inhibitory activity of phosphate esters on cholinesterase (10), the microsomal demethylation of tertiary amines (11), and the interaction of enzymes and substrates (12).

In the present paper substituent constants and regression analysis are applied to series of substituted pyrimidines (13, 14) (Tables II and III) and triazines (15) (Table IV). These results come from the very interesting and extensive work of Baker and his co-workers on analogs of tetrahydrofolic acid designed to act as nonclassical antimetabolites (16) in the inhibition of dihydrofolate reductase. Baker's early studies were concerned primarily with steric and functional group requirements for enzyme binding (17, 18), but it has recently become apparent that hydrophobic bonding regions on the enzyme may also aid in binding the substrate (15). The purpose of this report is to show that many of Baker's deductions can be treated in a quantitative fashion, and in this way the relative importance of different substituent effects can be more precisely delineated.

Received August 19, 1966, from the Department of Chemistry, Pomona College, Claremont, Calif. 91713.

Accepted for publication September 27, 1966.

The work was supported by research grant GM-07492 from the National Institutes of Health, U. S. Public Health Service, Bethesda, Md.

The authors are indebted to Susan Anderson, Smith Kline & French Research Associate, for determining the partition coefficient of diphenhydramine.

METHOD

If ΔF°_{BR} represents the free-energy change in some biological response which can be attributed to a single physical or chemical reaction, then as a first approximation this free-energy change can be factored as follows:

$$\Delta F^\circ_{BR} = \Delta F^\circ_{L/H} + \Delta F^\circ_{\text{elect.}} + \Delta F^\circ_{\text{steric}} \sim \ln k_{BR} \quad \text{(Eq. 1)}$$

In Eq. 1, k_{BR} may represent a rate or equilibrium constant, while $\Delta F^\circ_{L/H}$ represents that portion of the free-energy change which can be attributed to hydrophobic bonding (19), $\Delta F^\circ_{\text{elect.}}$ represents the electronic contribution, and $\Delta F^\circ_{\text{steric}}$ represents the spatial demands of reactants and products on the free-energy change. The effect of substituents on the change in free energy of a reference molecule can then be represented as in Eq. 2:

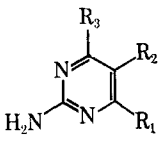
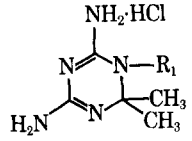
$$\delta_x \Delta F^\circ_{BR} = \delta_x \Delta F^\circ_{L/H} + \delta_x \Delta F^\circ_{\text{elect.}} + \delta_x \Delta F^\circ_{\text{steric}} \sim \delta_x \log k_{BR} \quad \text{(Eq. 2)}$$

Now substituent constants from model systems can be utilized in an extrathermodynamic approach (4) to evaluate substituent effects on k_{BR} as shown in Eq. 3.

$$\log BR \equiv \log \frac{1}{C_x} = k\pi + \rho\sigma + k'S + k'' \sim \delta_x \log k_{BR} \quad \text{(Eq. 3)}$$

In Eq. 3, C_x is the molar concentration of derivative x causing an equivalent biological response under standard conditions. π, σ, and S are extrathermodynamic constants for the respective substituent effects described in Eq. 2. The constants π and σ represent two parameters which are useful for hydrophobic and electronic effects; others can be used (3). This report will not attempt to evaluate steric effects included in S.

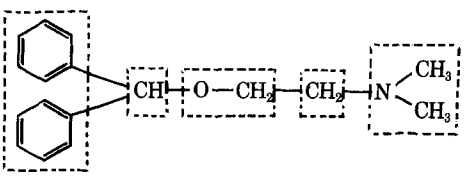
TABLE I.—STRUCTURE-ACTIVITY CORRELATIONS FOR TETRAHYDROFOLATE ANALOGS[a]

Pyrimidines	Compd.	Ref.	Eq.	n[b]	r[c]	S.D.[d]
	*1–16*	(13)	$\log \frac{1}{C}$[e] $= -5.002 - 5.162\sigma$ (Eq. 6)	16	0.760	1.076
	1–16	(13)	$\log \frac{1}{C} = -1.970 + 0.302\pi$ (Eq. 7)	16	0.328	1.565
	1–16	(13)	$\log \frac{1}{C} = -6.951 + 0.457\pi - 5.820\sigma$ (Eq. 8)	16	0.903	0.741
$R_3 = NH_2$	*5–16*	(13)	$\log \frac{1}{C} = -7.326 + 0.465\pi - 6.301\sigma$ (Eq. 9)	12	0.885	0.723
$R_3 = NH_2$, $\pi(R_1) = 0$	*5–16*	(13)	$\log \frac{1}{C} = -2.191 + 0.717\pi - 1.022\sigma$ (Eq. 10)	12	0.897	0.688
$R_3 = NH_2$, $\pi(R_2) = 0$	*5–16*	(13)	$\log \frac{1}{C} = -10.806 + 0.598\pi - 11.498\sigma$ (Eq. 11)	12	0.760	1.011
	5–16	(13)	$\log \frac{1}{C} = -8.144 + 0.800\pi_L$[f] $- 7.339\sigma$ (Eq. 12)	12	0.936	0.546
	5–16	(13)	$\log \frac{1}{C} = -6.686 + 0.512\pi_s$[g] $- 7.260\sigma$ (Eq. 12*a*)	12	0.735	1.054
	5–16	(13)	$\log \frac{1}{C} = -7.928 + 0.251\pi_s + 0.737\pi_L - 7.206\sigma$ (Eq. 13)	12	0.953	0.502
$R_1 = OH$, $R_3 = NH_2$	*17–26*	(14)	$\log \frac{1}{C} = -12.350 + 0.424\pi - 8.703\sigma$ (Eq. 14)	10	0.764	0.917
Omit R_2 = ethyl, *i*-amyl	*17–24*	(14)	$\log \frac{1}{C} = -17.864 + 0.103\pi - 13.724\sigma$ (Eq. 15)	8	0.899	0.523
Triazines						
						
pigeon liver	*28–39*	(15)	$\log \frac{1}{C} = -2.488 + 0.962\pi - 2.326\sigma$ (Eq. 16)	12	0.881	0.715
Omit R_1 = ethyl, *i*-amyl	*28–37*	(15)	$\log \frac{1}{C} = -2.215 + 0.861\pi - 1.724\sigma$ (Eq. 17)	10	0.930	0.499
E. coli	*27–36, 38, 39*	(15)	$\log \frac{1}{C} = -2.223 + 0.715\pi - 2.941\sigma$ (Eq. 18)	12	0.880	0.507

[a] All compounds are assayed in a pigeon liver enzyme system unless otherwise noted. [b] n is the number of points used in the regression. [c] r is the multiple correlation coefficient. [d] S.D., standard deviation. [e] C is the micromolar concentration causing 50% inhibition. [f] π_L is the π value of the most lipophilic substituent. [g] π_s is the π value of the second most lipophilic substituent.

The authors have taken σ from the compilation of Jaffé (20), making the assumption that although values for σ will change for substituents on a heterocyclic nucleus, they will do so in a parallel fashion. Using σ^+ or σ_m instead of σ_p did not give improved correlations.

While considerable evidence has accumulated to show that σ is an additive constant, an example illustrating the additive character (3) of log P and π is instructive. The experimental value of log P for a well-known antihistamine[1] is 3.07. Log P for diphenhydramine[1] can be calculated by taking advantage of the additivity of log P and π shown in I.



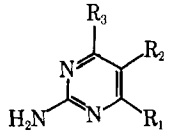
3.40 obs.
+ 4.26 + 0.30 −0.98 + 0.50 −0.95 = 3.13 calcd.

I

[1] Marketed as Benadryl by Parke-Davis Co., Detroit, Mich.

In I, log *P* for benzene (21) is taken as 2.13. Hydrogen is 0. The trisubstituted CH is given the value of 0.30 which is obtained by subtracting 0.50 for each methyl group in the isopropyl (π = 1.30) moiety. The value of CH_3 or CH_2 is 0.50. The value of −0.98 is that for an aliphatic methoxy

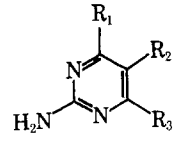
TABLE II.—DATA USED IN DERIVATION OF EQS. 6–11, 14–15 (TABLE I)



Compd.	R_1	R_2	R_3	$\Sigma\pi$	$\Sigma\sigma$	$\Sigma\pi$ $\pi(R_1)=0$	$\Sigma\pi$ $\pi(R_2)=0$	Obs.[a] log $\frac{1}{C}$	Calcd. log $\frac{1}{C}$
1	CH_3	$C_6H_5NH(CH_2)_3$	OH	2.29	−0.70	...	...	−2.90[b]	−1.83[c]
2	CH_3	$C_6H_5NH(CH_2)_3$	H	2.90	−0.34	...	...	−2.68	−3.65
3	CH_3	$C_6H_5(CH_2)_4$	OH	4.02	−0.70	...	...	−1.47	−1.04
4	C_6H_5	$C_6H_5(CH_2)_3$	H	5.76	−0.17	...	...	−3.58	−3.33
5	CH_3	$C_6H_5(CH_2)_4$	NH_2	3.48	−1.00	2.98	−0.65	1.57	0.46
6	CH_3	$C_6H_5NH(CH_2)_3$	NH_2	1.75	−1.00	1.25	−0.65	−0.34	−0.33
7	C_6H_5	$C_6H_5(CH_2)_4$	NH_2	5.11	−0.83	2.98	0.98	−0.04	0.22
8	C_6H_5	$C_6H_5NH(CH_2)_3$	NH_2	3.38	−0.83	1.25	0.98	0.06	−0.58
9	CH_3	H	NH_2	−0.65	−0.83	−1.15	−0.65	−3.04	−2.42
10	C_6H_5	H	NH_2	0.98	−0.66	−1.15	0.98	−2.20	−2.66
11	CH_3	*n*-C_4H_9	NH_2	1.35	−0.99	0.85	−0.65	−0.30	−0.57
12	C_6H_5	*n*-C_4H_9	NH_2	2.98	−0.82	0.85	0.98	−1.46	−0.82
13	*p*-$C_6H_5C_6H_4$	H	NH_2	3.11	−0.66	−1.15	3.11	−1.46	−1.67
14	*n*-C_3H_7	$C_6H_5(CH_2)_4$	NH_2	4.48	−0.96	2.98	0.35	1.68	0.69
15	$C_6H_5CH_2$	$C_6H_5NH(CH_2)_3$	NH_2	3.88	−0.83	1.25	1.48	−0.58	0.24
16	$C_6H_5CH_2$	$C_6H_5(CH_2)_4$	NH_2	5.61	−0.83	2.98	1.48	0.47	1.03
17	OH	$C_6H_5(CH_2)_4$	NH_2	2.37	−1.19	...	...	−0.92[d]	−1.28[e]
18	OH	$C_6H_5(CH_2)_3$	NH_2	1.87	−1.19	...	...	−1.78	−1.34
19	OH	H	NH_2	−1.76	−1.02	...	...	−4.11	−4.05
20	OH	*n*-C_8H_{17}	NH_2	−0.26	−1.15	...	...	−2.65	−2.11
21	OH	*n*-C_4H_9	NH_2	+0.24	−1.18	...	...	−1.60	−1.64
22	OH	*n*-C_6H_{13}	NH_2	1.24	−1.18	...	...	−1.48	−1.54
23	OH	*n*-C_8H_{17}	NH_2	2.24	−1.18	...	...	−1.68	−1.44
24	OH	$(CH_3)_2C{=}CHCH_2$	NH_2	0.24	−1.15[f]	...	...	−1.23	−2.06
25	OH	C_2H_5	NH_2	−0.76	−1.15	...	...	−4.23	−2.67
26	OH	*i*-C_5H_{11}	NH_2	0.54	−1.18	...	...	−0.60	−1.95

[a] *C* is the micromolar concentration causing 50% inhibition. [b] Calculated from data in *Reference 13*. [c] Calculated from Eq. 8. [d] Calculated from data in *Reference 14*. [e] Calculated from Eq. 15. [f] Value estimated.

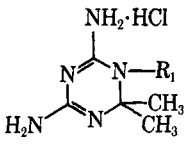
TABLE III.—DATA USED IN DERIVATION OF EQS. 12, 12*a*, AND 13 (TABLE I)



Compd.	R_1	R_2	R_3	Most Lipophilic R	π_L[a]	Obs. log $\frac{1}{C}$	Calcd. log $\frac{1}{C}$
5	CH_3	$C_6H_5(CH_2)_4$	NH_2	R_2	4.13	1.57	1.38
6	CH_3	$C_6H_5NH(CH_2)_3$	NH_2	R_2	2.40	−0.34	0.06
7	C_6H_5	$C_6H_5(CH_2)_4$	NH_2	R_2	4.13	−0.04	0.17
8	C_6H_5	$C_6H_5NH(CH_2)_3$	NH_2	R_2	2.40	0.06	−1.15
9	CH_3	H	NH_2	R_1	0.50	−3.04	−2.59
10	C_6H_5	H	NH_2	R_1	2.13	−2.20	−2.56
11	CH_3	*n*-C_4H_9	NH_2	R_2	2.0	−0.30	−0.32
12	C_6H_5	*n*-C_4H_9	NH_2	R_1	2.13	−1.46	−1.42
13	*p*-$C_6H_5C_6H_4$	H	NH_2	R_1	4.26	−1.46	−0.94
14	*n*-C_3H_7	$C_6H_5(CH_2)_4$	NH_2	R_2	4.13	1.68	1.09
15	$C_6H_5CH_2$	$C_6H_5NH(CH_2)_3$	NH_2	R_1	2.63	−0.58	−0.26
16	$C_6H_5CH_2$	$C_6H_5(CH_2)_4$	NH_2	R_2	4.13	0.47	0.88

[a] π_L = π value of most lipophilic substituent; *i.e.*, the largest π value. [b] *C* is the micromolar concentration causing 50% inhibition predicted by Eq. 12.

TABLE IV.—DATA USED IN DERIVATION OF EQS. 16–18 (TABLE I)



Compd.	R_1	log $\frac{1}{C}$[a] Pigeon Liver Obs.[b]	Pigeon Liver Calcd.[c]	*E. coli* Obs.[b]	*E. coli* Calcd.[d]	π	σ^*
27	C_6H_5	...	...	−0.48	−1.00	2.13	0.10[e]
28	$C_6H_5CH_2$	−0.52	−0.33	−1.53	−0.99	2.63	0.22
29	$C_6H_5(CH_2)_2$	0.15	0.34	−0.58	−0.22	3.13	0.08
30	$C_6H_5(CH_2)_3$	1.52	0.88	1.22	0.31	3.63	0.02
31	$C_6H_5(CH_2)_4$	1.40	1.31	0.68	0.67	4.13	0.02[e]
32	CH_3	−1.87	−1.78	−1.68	−1.87	0.50	0.00
33	n-C_3H_7	−1.04	−0.72	−0.67	−0.80	1.50	−0.12
34	n-C_4H_9	0.44	−0.27	−0.32	−0.41	2.00	−0.13
35	n-C_6H_{13}	0.50	0.59	0.04	0.30	3.00	−0.13[e]
36	n-C_8H_{17}	0.85	1.45	0.57	1.02	4.00	−0.13[e]
37	$C_6H_4(C_4H_9$-$n)$-p	1.22	1.17	...	...	4.13	0.10[e]
38	C_2H_5	−2.34	1.18	−1.79	−1.22	1.00	−0.10
39	i-C_5H_{11}	1.22	−0.01	0.16	−0.20	2.30	−0.13[e]

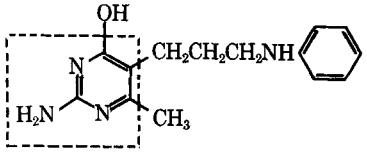
[a] C is the micromolar concentration causing 50% inhibition. [b] Calculated from data in *Reference 15*. [c] Calculated from Eq. 17. [d] Calculated from Eq. 18. [e] Value estimated.

group (3). The figure of −0.95 is that for an aliphatic dimethylamino fragment (3). The good agreement between observed and calculated log P exemplifies the additive character of π.

For the work in hand one does not have or need to use log P for the pyrimidine or triazine unit. Since π is an additive constant when strong group interactions are absent, partition coefficients may be estimated from previously determined values for functional groups.

$$\log P = \pi_1 + \pi_2 \ldots + \pi_i$$

where π_i represents logs of partition values for the components of a molecule. In series of compounds where a portion of the molecule is held constant, that portion may be disregarded and $\Sigma\pi$ for the substituents is then used for the correlation. For example, $\Sigma\pi$ for the molecule in II (number 1 in Table II) was calculated as shown in II and Eqs. 4 and 5.



constant

II

$$\Sigma\pi = \pi OH + \log P\ C_6H_5NH_2 + 3\pi CH_2 + \pi CH_3 \quad \text{(Eq. 4)}$$

$$\Sigma\pi = -0.61 + 0.90 + 3(0.50) + 0.50 = 2.29 \quad \text{(Eq. 5)}$$

$\Sigma\sigma$ was calculated for each molecule by adding the individual σ values. π for the H_2N group was taken as −1.15. For groups such as $C_6H_5CH_2CH_2CH_2$— and $C_6H_5NHCH_2CH_2CH_2$— it was assumed that the phenyl and phenylamino functions would be sufficiently insulated from the ring so that they played no electronic role. Taft's σ^* values (4) were used for the nonaromatic substituents in Table IV. The use of σ from the benzene system is an approximation which is particularly bad for the OH group since in the nitrogen heterocycles keto–enol tautomerism is so important. From the data in Tables II, III, and IV the equations in Table I have been derived by the method of least squares using an IBM 360/40 computer.

DISCUSSION

Rather good correlations were found to result using simple two-parameter equations employing only π and σ. In no instance did a significant improvement in correlation result when either a π^2 or σ^2 term or both π^2 and σ^2 were added. This is what one would normally expect (12) in an *in vitro* system using purified enzymes.

Baker and his colleagues have concluded (13) that the dihydrofolate reductase inhibitors are bound to the enzyme by the interaction of ring electrons with an electron-deficient site, and by hydrophobic interaction of one or more side chains. The equations in Table I strongly support these conclusions. All of the equations have positive coefficients for π, indicating that the more lipophilic the side chain, the more effective the compound is as an inhibitor. The negative coefficient associated with σ means that the more electron-releasing the substituent (the more it tends to increase the electron density on the ring), the more effective the derivative is as an inhibitor. The importance of the use of the two parameters is best seen by considering Eqs. 6–8 in Table I. The two single-parameter equations account for only 58% ($r^2 = 0.58$) and 11% of the variance in the data, while the two-parameter equation accounts for more than 81% of the vari-

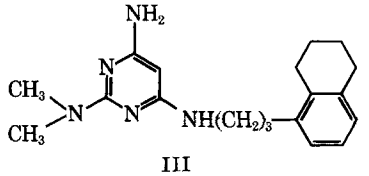
ance. In deriving these equations, only one compound [R_3=OH, R_2=$C_6H_5(CH_2)_4$, R_1=n-C_3H_7] was not used because it gave a poor fit in relation to the other data. The correlation obtained for the other 16 different derivatives in which changes were made in three different substituents is surprisingly good, especially when one considers no account has been taken of steric interactions of substrate and enzyme which classical enzyme theory has characterized as being so highly important (lock and key theory).

In order to evaluate more precisely the role of hydrophobic bonding, the 12 derivatives were taken in which changes were made only in R_1 and R_2. Equation 9 gives the result, treating the substituents R_1 and R_2 in simple additive fashion. First the hypothesis was tested that substituents in only one of the positions (R_1 or R_2) could fit into the hydrophobic area of the enzyme and thus contribute to inhibitory power of the derivative. This was done by setting R_1 equal to zero in Eq. 10, and R_2 equal to zero in Eq. 11. The higher correlation obtained with Eq. 10 would seem to confirm Baker's feeling (18) that the 6-position substituent is unable to contact the hydrophobic bonding area. However, the fact that one obtains essentially the same correlation with Eqs. 9 and 10 seems inconsistent. If, in fact, 6-position substituents contribute nothing to binding, an equation which includes their π values would include a varying and meaningless addition to π and should show a poor correlation. This inconsistency makes doubtful the theory that only one conformation is possible, but the relatively high correlations obtained suggested that hydrophobic binding may be *primarily* controlled by the larger lipophilic substituent. Inspection of the groups in positions 5 and 6, as shown in Table II, reveals that position 5 has more of the larger substituents. This could account for the fact that Eq. 10 gives a better correlation than Eq. 11. Equation 12 gives the best correlation. The hypothesis underlying Eq. 12 implies that substituted pyrimidines bind to dihydrofolate reductase in any of several conformers and that the choice of conformer is determined by the necessity of placing the most lipophilic group in contact with the hydrophobic bonding region of the enzyme. The most lipophilic group thus provides the dominating force in determining the binding configuration. Equation 12 confirms this hypothesis by rationalizing 87% of the variance in the data. Equation 12*a* shows that the contribution to the binding by the second most lipophilic substituent correlates about 50% of the variance in the data. However, when these two terms are included in a single equation (Eq. 13), the slight improvement in correlation is not statistically significant (F test).

Equations 14 and 15 correlate a series of trisubstituted pyrimidines which differ from the previous group chiefly in having an OH group in place of the lipophilic function at R_1. They appear to be intrinsically less active (the intercepts are more negative). Although this series appears very similar in structure to that considered in Eqs. 6–13, combining the two sets of data resulted in a very poor correlation (r = 0.776), indicating that the two series have different modes of binding. The coefficients of Eq. 15 indicate an increased dependence on σ and a decreased importance of π, which might result from a greater dependence on hydrogen bonding. In Eq. 15 the ethyl and isoamyl groups which were poorly correlated have been omitted. The ethyl group was about 30 times less active than predicted and the isoamyl was about 10 times more active than predicted. Why these groups are not adequately accommodated by the correlation is not clear. However, a reasonable suggestion by Baker (14) is that a conformational change may be caused in the enzyme in proceeding from ethyl to n-butyl. This brings out one of the great advantages of regression analysis, however, since the spotting of such exceptional structures indicates areas for further research. This same characteristic of the ethyl and isoamyl groups is apparent again in the substituted triazine series.

The coefficients in the best equations for pyrimidines (Eq. 12) and for the triazines (Eq. 17) are quite similar to each other. Hydrophobic bonding in the triazines seems to be somewhat more important, as indicated by the larger coefficient of π, and electronic bonding is somewhat less important as indicated by the smaller coefficient of σ. In general, the 90% confidence limits on the σ terms are much greater, no doubt because the differences in σ values are small compared to the experimental error in measuring σ. F tests also indicate lower significance for σ terms in the triazines. Of the more than 20 derivatives tested by Baker, the authors have not been able to treat those in which a phenyl group was attached directly to the ring nitrogen because of the uncertainty in electronic interaction. Delocalization of the lone pair electrons on nitrogen by the phenyl ring can greatly affect the lipophilic character (21) of the molecule, as well as the electron density on the ring which appears to react with an electron-deficient species. For the substituents which were considered, species differences in enzymes appear to be rather small (*cf.* Eqs. 16 and 18), although in some instances Baker has found quite marked differences. The species differences mentioned by Baker are, however, supported by the authors' finding that response for isoamyl and ethyl substituents are moderately well predicted for the *E. coli* enzyme (Eq. 18) but not for the enzyme from pigeon liver (Eqs. 14–17).

The equations in Table I should enable one to design more active inhibitors. Greater activity should result from compounds having larger positive $\Sigma\pi$ and more negative $\Sigma\sigma$. To increase the electron density on the ring, one could move the amino group Baker has incorporated into the side chain to a ring position in the 5- or 6-position and attach the largest possible lipophilic substituent to it. The possibility in III seems most logical.



III

In the variation of Baker's general type of inhibitor (III), the authors have added the quite lipophilic cyclohexyl moiety to the phenyl ring, hoping, of course, that the enzymic hydrophobic area is large enough to accommodate it. Another approach

would be to use a pentamethyl- or pentachlorophenyl group on the side chain. Placing methyl groups on one of the amino groups would lower affinity for the aqueous phase, even though hydrophobic bonding in the direct sense might not be possible at this position. We have indicated attachment of the side chain at the 6-position to make maximum use of the electron-releasing effect of the amino function. Since a better correlation was obtained using σ_p rather than σ_m, it is assumed that it is the high electron density on the 1,3,5-positions which is most important. This may not be the case, however. Since so few strong electron-releasing groups were tested, it is not possible to isolate this factor with confidence. In any case, it would be advantageous to separate the two functions so that steric interactions do not twist either of the nitrogen atoms out of the ring-plane and thus prevent effective overlapping of p-orbitals.

While the above design might be very effective in *in vitro* work, such a highly lipophilic compound would probably not make the best drug. For *in vivo* use, one would first have to determine π_0 (2) and then work close to this value, making maximum use of σ, in designing derivatives. The problem of penetration of the blood-brain barrier which appears to be a limiting factor in the clinical use of the tetrahydrofolate analogs, aminopterin and amethopterin (22), could be circumvented by proper lipophilic design of the drug (23).

The good correlations obtained for the structure-activity relationships of these varied series of compounds show that the additive nature of π makes it a useful and reliable parameter in drug design. The large reduction in variance obtained by using the linear combination of σ with π confirms the value of the use of σ constants in heterocyclic systems, even though the values were obtained from benzene. Although it is likely better correlations would be obtained if pKa values for the drugs in question could be used, the values of Jaffé's σ are probably good enough for practical purposes. If one wished to isolate as precisely as possible the steric factors, it would then be essential to determine pKa values. The rather good mathematical correlation the authors have obtained with the wide variety of bulky substituents used by Baker is further evidence for the importance of the role of flexibility in enzyme action (24) which in recent years has become so apparent. The authors' correlations provide a mathematical basis for the essential conclusions of Baker's studies and reinforce these conclusions by deriving them in an independent way.

REFERENCES

(1) Hansch, C., and Fujita, T., *J. Am. Chem. Soc.*, **86,** 1616(1964).
(2) Hansch, C., Steward, A. R., Iwasa, J., and Deutsch, E. W., *Mol. Pharmacol.*, **1,** 205(1965).
(3) Iwasa, J., Fujita, T., and Hansch, C., *J. Med. Chem.*, **8,** 150(1965).
(4) Leffler, J. E., and Grunwald, E., "Rates and Equilibria of Organic Reactions," John Wiley & Sons, Inc., New York, N. Y., 1963.
(5) Hansch, C., and Deutsch, E. W., *J. Med. Chem.*, **8,** 705(1965).
(6) Deutsch, E. W., and Hansch, C., *Nature*, **211,** 75(1966).
(7) Hansch, C., Kiehs, K., and Lawrence, G., *J. Am. Chem. Soc.*, **87,** 5770(1965).
(8) Kiehs, K., Hansch, C., and Moore, L., *Biochemistry*, **5,** 2602(1966).
(9) Hansch, C., and Deutsch, E. W., *Biochim. Biophys. Acta*, **112,** 381(1966).
(10) *Ibid.*, **126,** 117(1966).
(11) Hansch, C., Steward, A. R., and Iwasa, J., *J. Med. Chem.*, **8,** 868(1965).
(12) Hansch, C., Deutsch, E. W., and Smith, R. N., *J. Am. Chem. Soc.*, **87,** 2738(1965).
(13) Baker, B. R., and Shapiro, H., *J. Pharm. Sci.*, **55,** 308(1966).
(14) Baker, B. R., Ho, B.-T., and Santi, D. V., *ibid.*, **54,** 1415(1965).
(15) Baker, B. R., and Ho, B.-T., *ibid.*, **55,** 470(1966).
(16) Baker, B. R., *ibid.*, **53,** 347(1964).
(17) Baker, B. R., Santi, D. V., Almaula, P., and Werkheiser, W. C., *J. Med. Chem.*, **7,** 25(1964).
(18) Baker, B. R., and Shapiro, H., *ibid.*, **6,** 664(1963).
(19) Kauzmann, W., *Advan. Protein Chem.*, **14,** 37(1959).
(20) Jaffé, H. H., *Chem. Rev.*, **53,** 191(1953).
(21) Fujita, T., Iwasa, J., and Hansch, C., *J. Am. Chem. Soc.*, **86,** 5175(1964).
(22) Freireich, E., *Ann. Internal Med.*, **59,** 939(1963).
(23) Hansch, C., Steward, A. R., and Iwasa, J., *Mol. Pharmacol.*, **1,** 87(1965).
(24) Koshland, D. E., *Cold Spring Harbor Symp. Quant. Biol.*, **27,** 473(1963).