# Correlation Analysis. Its Application to the Structure–Activity Relationship of Triazines Inhibiting Dihydrofolate Reductase[1a]

**Carlo Silipo[1b] and Corwin Hansch***

*Contribution from the Department of Chemistry, Pomona College, Claremont, California 91711.* *Received November 6, 1974*

**Abstract:** Correlation analysis has been used to formulate equations relating the chemical structure of 4,6-diamino-1,2-dihydro-2,2-dimethyl-1-(X-phenyl)-*s*-triazines to the molar concentration causing 50% reversible inhibition of tumor dihydrofolate reductase. It is shown that for 244 congeners, substituents in the 3 position of the *N*-phenyl ring interact in hydrophobic enzymic space, while those in the 4 position interact with a different kind of enzymic space. This space may be generally polar in character. The use of indicator variables is demonstrated to greatly increase one's ability to formulate quantitative structure–activity relationships for large numbers of highly complex molecules interacting with a complex macromolecular system. In the present analysis, 76% of the variance is correlated by indicator variables and 85% by indicator variables plus hydrophobic and molar refractivity parameters. This approach to structure–activity relationships allows one to carry on objective discussions of massive amounts of chemical and biological data.

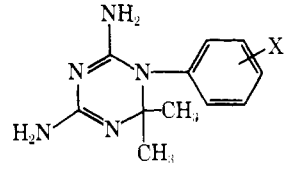
Hammett's classic book on physical organic chemistry[2] which appeared in 1940 can be taken as a turning point in the study of organic reactions. The now famous Hammett equation provided the means for translating the "English School" of chemists' qualitative thinking about the electronic effects of substituents on the dynamics of aromatic side-chain reactions into numerical terms. The use of $\sigma$ constants obtained from benzoic acids to correlate sets of rate or equilibrium constants came to be called the formulation of linear free-energy relationships.

Taft made a major advance by showing that linear free-energy relationships could be formulated for electronic and steric effects in aliphatic systems,[3] and Brown and his colleagues[4] extended the approach to the broad field of aromatic electrophilic substitution reactions. More recently, the technique has been extended to hydrophobic interactions in biochemical and medicinal chemical systems.[5]

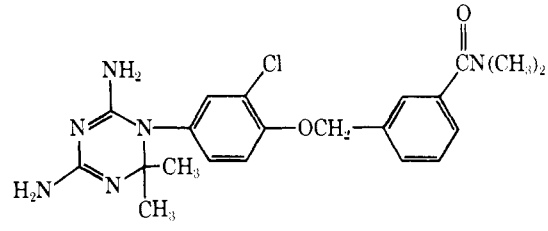
In an elegant treatment of work in this area, Leffler and Grunwald[6] in 1963 suggested the more encompassing term extrathermodynamic relationships and outlined a formal approach for development of such equations. Shorter[7] and Exner,[8] as well as Russian[8] workers, have now suggested the still broader term correlation analysis. The term extrathermodynamic implies relationships between thermodynamic quantities which do not obey the laws of thermodynamics. Since a large amount of structure–activity work is now carried out in which the quantities correlated are not thermodynamically based, correlation analysis seems to be a more appropriate term. This is especially true of work in bioorganic chemistry with which the present paper is concerned.

Before the general availability of computers, one often spoke of fitting data to an equation. This was a tedious and time-consuming process in which one simply could not consider many possibilities. The situation is now completely turned around and one can readily explore hundreds or thousands of possible equations in studying the interrelationship of sets of data. Today, one often speaks of fitting equations to data.[9] On the frontiers of chemical structure–activity relationships, especially in bio- and medicinal chemistry, so little solid theory is at hand on which to build that all kinds of purely empirical ideas need to be explored. Computerized statistical techniques promise to be of great help in sorting out important structure–activity features which can then be used to form more firmly based theory. Techniques such as pattern recognition,[10,11] discriminate analysis,[12] cluster analysis,[11,13] and regression analysis[5a] which have been developed and used heavily outside of chemistry are now beginning to be used by those working with structure–activity relationships.

In a brilliant burst of effort starting with a first publication in 1967, Baker and a few graduate students[14] synthesized variations of I to achieve, before Baker's death in 1971, II, a drug now in clinical trials against cancer. This



I



II

study appears to constitute the largest published set of biochemical congeners whose activity has been measured quantitatively in one laboratory. As such it presents a great challenge to those interested in structure–activity studies. In all, Baker's group synthesized about 260 variations of I and studied their inhibiting effect on dihydrofolate reductase isolated from Walker 256 and L1210 leukemia tumors. While Baker did not live long enough to achieve his goal of finding a derivative which would be highly selective for enzyme from tumor tissue but relatively inactive against enzyme from normal human tissue, he did demonstrate vividly that starting at the enzyme level rather than with whole animals constitutes a powerful technique for drug development. This approach has also been brilliantly exploited by Hitchings and his group[15,16] in the development of allopurinol for gout and the new antibacterial agent, trimethoprim.

Equation 1 was formulated[17] in a first attempt to place Baker's

$$\log 1/C = 0.89\ (\pm 0.14)\ (\pi\text{-}3) - 0.13\ (\pm 0.03)\ (\pi\text{-}3)^2 + 0.15\ (\pm 0.03)\ (\text{MR-}4) + 6.62\ (\pm 0.13) \quad (1)$$

| $n$ | $r$ | $s$ |
|---|---|---|
| 83 | 0.905 | 0.328 |

$$\text{ideal } \pi\text{-}3 = 3.5\ (3.0\text{-}4.3)$$

results in a quantitative context. $C$ in this equation is the molar concentration of inhibitor causing 50% *reversible* inhibition of enzyme. The figures in parentheses are the 95% confidence limits. Equation 1 is based on those derivatives tested on Walker tumor which do not contain an $SO_2F$ group. The correlation of this relatively simple set of 83 congeners enabled us to see the general outlines of the problem and start on the more difficult task of placing all derivatives in a single equation. $\pi$-3 represents hydrophobic constants for substituents in the 3 position of I and MR-4 represents molar refractivity of substituents in the 4 position of I. Equation 1 is based on 83 derivatives ($n$), $r$ represents the correlation coefficient, and $s$ is the standard deviation from the regression. This equation "explains" 82% of the variance in the data in terms of the two parameters $\pi$-3 and MR-4. It suggests that substituents in three-space (that portion of enzyme near the 3 position) are acting via hydrophobic forces. Substituents above a certain size set by the $(\pi\text{-}3)^2$ term appear to be too large to be effective in three-space. Substituents in the 4 position appear to be causing inhibition by a different mechanism more involved with the polarizability and dispersion forces of the substituent. While a rather large amount of experience[5a,18] has been obtained with the hydrophobic constants ($\pi$), relatively little use has been made of MR in quantitative terms.

Pauling was one of the first to emphasize the importance of MR in the interactions of biomacromolecules.[19] Agin et al.[20] developed an expression to relate MR to log $1/C$ for biochemical structure-activity studies. Experience in our own work shows MR to be an important parameter in enzymic quantitative structure-activity relationships (QSAR).

This report covers the development of a much more general correlation equation (eq 2) which accommodates essentially all of the triazines studied by Baker. Equation 2 (see Results) has been formulated from the data in Table I.

**Method.** $C$ in log $1/C$ in Table I represents the molar concentration which produces 50% *reversible* inhibition of dihydrofolate reductase obtained from either L1210/DF8 mouse leukemia or Walker 256 rat tumors when assayed with 6 $\mu M$ dihydrofolate[14] at pH 7.

$\pi$-2, $\pi$-3, and $\pi$-4 have been employed to estimate the hydrophobic interaction of substituents in the ortho, meta, and para positions of the $N$-phenyl ring of I. The $\pi$ constants derived from partition coefficients between 1-octanol and water[21] are from the benzene solute system. The $\pi$ values in Table I are from the literature[17,22] or have been estimated by additivity principles.[18,23] Two values of $\pi_{SO_F}$ were used; one was calculated as log $P_{CH_3\text{—}C_6H_4\text{—}4\text{—}SO_2F}$ − log $P_{CH_3C_6H_5}$ = 2.74 − 2.69 = 0.05, and the other as log $P_{CH_3CONH\text{—}C_6H_4\text{—}4\text{—}SO_2F}$ − log $P_{CH_3CONHC_6H_5}$ = 2.17 − 1.16 = 1.01. Whenever a strong electron-withdrawing group is attached to a benzene ring having a substituent with loosely held lone-pair electrons, $\pi$ is greatly altered. The following calculations illustrate the use of additivity

$$\pi_{CH=CHCONHC_6H_5\text{—}4'\text{—}SO_2F} = \pi_{SO_2F} + \pi_{CONHC_6H_5} + \Delta\pi = 1.01 + 0.49 + 0.49 = 1.99$$

$$\Delta\pi = \pi_{CH=CH} = \log P_{C_6H_5CH=CHCOCH_3} - \log P_{C_6H_5COCH_3} = 2.07 - 1.58 = 0.49$$

of log $P$ and $\pi$. $\pi_{CH_2CH(C_6H_4\text{—}X)CONHR}$ values were estimated from the appropriate $\pi_{CH_2CH(R')CONHR}$ and $\pi_{CONHR}$. For example, it is assumed that the value

$$\Delta\pi_{CH_2CH_2} = \log P_{C_6H_5CH_2CH_2CONH_2} - \log P_{C_6H_5CONH_2} = 0.91 - 0.64 = 0.27$$

is the same as log $P_{C_6H_5CH_2CH_2CON(R)_2}$ − log $P_{C_6H_5CON(R)_2}$. Thus

$$\pi_{CH_2CH(C_6H_4\text{—}3''\text{—}OMe)CONH\text{—}C_6H_4\text{—}4'\text{—}SO_2F} = \Delta\pi_{CH_2CH_2} + \Delta\pi_{\text{paraffin branch}} + \pi_{C_6H_5} + \pi_{OCH_3} + \pi_{CONHC_6H_4\text{—}4'\text{—}SO_2F} = 0.27 - 0.20 + 1.96 - 0.02 + 1.50 = 3.51$$

$$\pi_{OCH_2C_6H_4CONHC_6H_4SO_2F} = \pi_{OCH_2C_6H_5} + \pi_{CONHC_6H_4SO_2F} = 1.66 + 1.50 = 3.16$$

The values of the other substituents were estimated by the same method or measured values from the literature were used.[17,22a,c]

It has been assumed that the same value of $\pi$ can be employed for ortho, meta, and para substituents and no attempt was made to correct for groups adjacent to each other. $\Sigma\pi$ was also studied to explore the possibility that the overall lipophilicity of the inhibitors might play a significant role.

MR values have been taken from our recent compilation or calculated using these values.[22a] They have been scaled by 0.1 which makes the *apolar* functions essentially equiscalar with $\pi$. For example

| | | | |
|---|---|---|---|
| $\pi_{Cl}$ | 0.71 | $MR_{Cl}$ | 0.60 |
| $\pi_H$ | 0.00 | $MR_H$ | 0.10 |
| $\pi_{C_6H_5}$ | 2.13 | $MR_{C_6H_5}$ | 2.54 |
| $\pi_{CH_3}$ | 0.56 | $MR_{CH_3}$ | 0.56 |
| $\pi_{CF_3}$ | 0.88 | $MR_{CF_3}$ | 0.50 |

This collinearity between $\pi$ and MR does not hold for polar groups. Since both molar refractivity and $\pi$ depend to a certain extent on molar volume, the collinearity between these vectors may be high,[23] depending on the choice of substituents under consideration. The collinearity among the important vectors used in this study is displayed in Table II.

In addition to the continuous variables $\pi$ and MR, a number of discrete variables (indicator variables[24]) were studied. These variables take the value of 1 or 0 for structural features which could not be parameterized by $\pi$ and MR. The linear combination of such terms assumes that these structural features are additive properties independent of other changes in the system. Using the indicator variable I-1 with a value of 1 for Walker enzyme data and 0 for L1210 enzyme data allows the merging of data from the two test systems.

In our first study no attempt was made to include congeners with substituents in the ortho positions. We first attempted to include these by using the Taft steric parameter $E_s$. Although $E_s$-2 did result in considerable reduction in the variance, it was not quite as effective as the simple indicator variable I-2. It was thought that $E_s$-2 plus $\pi$-2, MR-2, $\mathfrak{F}$, or $\mathfrak{R}$ might model ortho effects better than I-2; however, this is not true.

I-3 was used for rigidity in groups attached to the $N$-phenyl ring. The direct attachment of $-C_6H_5$ or bridges such as -CONH- or -CH=CHCONH- from either the 3 or 4 position to a second phenyl ring has a very bad effect on activity. This is also true for the unit $-CH(C_6H_5)-$. These functions reduced inhibitory power by a factor of about 100. Rigid groups in the 3 position are not as well parameterized by I-3 as those in the 4 position; however, not enough 3 functions are available to justify an additional parameter.

The variable I-4 takes a value of 1 for those congeners having the highly active leaving group $-C_6H_4SO_2OC_6H_4X$.

Table I. Inhibition Constants and Physicochemical Parameters Used in Deriving Equation 2–11 for the Reversible Inhibition of Dihydrofolate Reductase by 4,6-Diamino-1,2-dihydro-2,2-dimethyl-1-(X-phenyl)-*s*-triazines

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2,5-$Cl_2$ | 3.43 | 4.220 | 0.79 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 2 | 2-$OCH_3$ | 3.68 | 4.220 | 0.54 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 3 | 2,4-$Cl_2$ | 3.82 | 4.326 | 0.51 | 0.00 | 0.71 | 0.60 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 4 | 2-$CH_3$ | 4.00 | 4.220 | 0.22 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 5 | 2-Cl | 4.15 | 4.220 | 0.07 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 6 | 2-Br | 4.25 | 4.220 | 0.03 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 7 | 2,4,5-$Cl_3$ | 4.38 | 4.326 | 0.05 | 0.00 | 0.71 | 0.60 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 8 | 2-I | 4.62 | 4.220 | 0.40 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 9 | 4-$CONHC_6H_4$-4'-$SO_2F$ | 4.68 | 5.272 | 0.59 | 0.00 | 1.50 | 4.23 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 100 | 14a |
| 10 | 4-$CONHC_6H_4$-3'-$SO_2F$ | 4.68 | 5.272 | 0.59 | 0.00 | 1.50 | 4.23 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0 | 14d |
| 11 | 4-$C_6H_5$ | 4.70 | 5.161 | 0.46 | 0.00 | 1.96 | 2.54 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 12 | 2-F | 4.74 | 4.220 | 0.52 | 0.00 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 13 | 3-$OCH_2CO$-N-$(CH_2CH_2)_2O$ | 4.85 | 5.576 | 0.73 | −1.39 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 14[e] | 4-CN | 5.14 | 6.862 | 1.72 | 0.00 | −0.57 | 0.63 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 15 | 4-CH=CHCONH-$C_6H_4$-4'-$SO_2F$ | 5.19 | 5.272 | 0.08 | 0.00 | 1.99 | 5.22 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0 | 14f |
| 16 | 3-$OCH_2CONMe_2$ | 5.44 | 5.606 | 0.17 | −1.36 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 17 | 4-CH(Ph)$CH_2CONH$-$C_6H_4$-4'-$SO_2F$ | 5.74 | 5.078 | 0.66 | 0.00 | 3.53 | 7.59 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0 | 14g |
| 18[e] | 4-Cl, 3-$(CH_2)_2C_6H_4$-4'-$SO_2F$ | 5.82 | 8.280 | 2.46 | 2.71 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 13 | 14z |
| 19 | 4-CH=CHCONH-$C_6H_4$-3'-$SO_2F$ | 5.89 | 5.272 | 0.62 | 0.00 | 1.99 | 5.22 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 55 | 14f |
| 20 | 3-$CONHC_6H_4$-4'-$SO_2F$ | 5.96 | 5.513 | 0.45 | 1.50 | 0.00 | 0.10 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0 | 14a |
| 21 | 3-$NHCOCH_2Br$, 4-$O(CH_2)_3C_6H_5$ | 6.11 | 6.993 | 0.88 | −0.37 | 2.66 | 4.15 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14c |
| 22 | 3-$CH_2NHCONEt_2$ | 6.11 | 6.305 | 0.20 | −0.29 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14m |
| 23 | 3-$OCH_3$ | 6.17 | 6.736 | 0.57 | −0.02 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 24[e] | 4-$OCH_2CON$-$(Me)C_6H_5$ | 6.17 | 7.268 | 1.10 | 0.00 | 0.12 | 4.55 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 25 | 4-$CH_2CH(CH_2CH_2Ph)$-$CONHC_6H_4$-4'-$SO_2F$ | 6.20 | 6.919 | 0.72 | 0.00 | 4.23 | 8.52 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14g |
| 26 | 3-$COCH_2Cl$ | 6.21 | 6.638 | 0.43 | −0.16 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 27 | 4-$CH_2CH(\alpha$-$C_{10}H_7)$-$CONHC_6H_4$-4'-$SO_2F$ | 6.24 | 6.559 | 0.32 | 0.00 | 5.02 | 9.13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 61 | 14i |
| 28 | 4-$OCH_2CONMe_2$ | 6.26 | 7.157 | 0.90 | 0.00 | −1.36 | 2.58 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 29 | 4-$CH_2CH$-(Ph-2"-$OCH_3$)$CONHC_6H_4$-4'-$SO_2F$ | 6.33 | 6.725 | 0.40 | 0.00 | 3.51 | 8.27 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 76 | 14i |
| 30 | 3-Cl, 4-$OCH_2C_6H_{10}$-$CH_2OC_6H_4$-4'-$SO_2F$ | 6.37 | 7.300 | 0.93 | 0.71 | 5.16 | 7.25 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 99 | 14q |
| 31[e] | 3-CH($CH_2NHCO$-$CH_2Br$)$(CH_2)_3C_6H_5$ | 6.37 | 7.729 | 1.36 | 2.94 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14c |
| 32[e] | 3-$CH_2NHCO$-$N(CH_2CH_2)_2O$ | 6.43 | 5.408 | 1.02 | −1.32 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14m |
| 33 | 4-$COCH_2Cl$ | 6.45 | 7.034 | 0.58 | 0.00 | −0.16 | 1.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 34 | 4-$CH_2CH$(Ph-3"-$OCH_3$)$CONHC_6H_4$-4'-$SO_2F$ | 6.46 | 6.725 | 0.27 | 0.00 | 3.51 | 8.27 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 68 | 14i |
| 35 | 4-CH($CH_2NHCO$-$CH_2Br$)$(CH_2)_3C_6H_5$ | 6.52 | 7.139 | 0.62 | 0.00 | 2.94 | 7.03 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14c |
| 36[e] | 2,3-$Cl_2$ | 6.52 | 4.643 | 1.88 | 0.71 | 0.00 | 0.10 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14z |
| 37 | 2-Cl, 4-$(CH_2)_4C_6H_5$ | 6.54 | 5.422 | 1.12 | 0.00 | 3.66 | 4.39 | 1.00 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14z |
| 38 | 3-Cl, 4-$O(CH_2)_4O$-$C_6H_4$-4'-$SO_3C_6H_4$-4"-Cl | 6.55 | 7.031 | 0.48 | 0.71 | 4.92 | 8.90 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 89 | 14q |
| 39 | 3-$CH_2NHCOCH_2Br$ | 6.58 | 6.364 | 0.22 | −0.52 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 40[e] | 3-$CONHC_6H_4$-3'-$SO_2F$ | 6.60 | 5.513 | 1.09 | 1.50 | 0.00 | 0.10 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0 | 14d |
| 41 | 4-$CH_2CONMe_2$ | 6.63 | 7.133 | 0.50 | 0.00 | −1.70 | 2.37 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 42 | 4-$OCH_2CO$-$N(CH_2)_4$ | 6.66 | 7.220 | 0.56 | 0.00 | −0.72 | 3.31 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 43 | 3-$OCH_2CON(Me)C_6H_5$ | 6.68 | 6.830 | 0.15 | 0.12 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 44 | 4-$OCH_2CONEt_2$ | 6.72 | 7.233 | 0.51 | 0.00 | −0.36 | 3.51 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 45 | 3-$CH_2CH(CH_2$-$NHCOCH_2Br)C_6H_5$ | 6.72 | 7.625 | 0.91 | 1.94 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14c |
| 46 | 4-Cl, 3-$O(CH_2)_5$-$OC_6H_4$-4'-$SO_2F$ | 6.72 | 7.314 | 0.59 | 4.43 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 49 | 14q |
| 47 | 4-$CH_2CONEt_2$ | 6.77 | 7.219 | 0.45 | 0.00 | −0.70 | 3.29 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 48[e] | 4-Cl, 3-$(CH_2)_4$-$C_6H_4$-4'-$SO_2F$ | 6.77 | 8.133 | 1.36 | 4.01 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 33 | 14z |
| 49 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$CH_2OC_6H_4$-4"-$SO_2F$ | 6.82 | 7.317 | 0.50 | 0.71 | 4.33 | 7.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 89 | 14q |
| 50 | 3-$OCH_2CONHC_6H_5$ | 6.85 | 7.115 | 0.27 | 0.60 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 51[e] | 3-$C_6H_5$ | 6.85 | 5.638 | 1.21 | 1.96 | 0.00 | 0.10 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 52 | 4-$CH_2CH(Ph)CONHC_6H_4$-3'-$SO_2F$ | 6.89 | 6.832 | 0.06 | 0.00 | 3.53 | 7.56 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14i |
| 53 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CONHC_6H_4$-4"-$SO_2F$ | 6.92 | 7.288 | 0.37 | 0.71 | 3.16 | 7.34 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 14s |
| 54 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$CONHC_6H_4$-4"-$SO_2F$ | 6.92 | 7.288 | 0.37 | 0.71 | 3.16 | 7.34 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 96 | 14s |
| 55 | 3-$OCH_2CONHC_6H_4$-4'-$SO_2F$ | 6.92 | 7.301 | 0.38 | 1.61 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 10 | 14s |
| 56 | 4-$CH_2CN$ | 6.92 | 6.934 | 0.01 | 0.00 | −0.57 | 1.01 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 57 | H | 6.92 | 6.75 | 0.17 | 0.00 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 58 | 3-$OCH_2C_6H_4$-3'-$NHCOCH_2Br$ | 6.92 | 7.431 | 0.51 | 1.29 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 59 | 4-$CH_2CON(Me)C_6H_5$ | 7.00 | 7.265 | 0.27 | 0.00 | −0.19 | 4.34 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 60 | 4-$(CH_2)_2CONMe_2$ | 7.05 | 7.181 | 0.13 | 0.00 | −1.20 | 2.83 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 61[e] | 3-Cl, 4-$(CH_2)_4C_6H_3$-5'-Cl, 2'-$SO_2F$ | 7.06 | 8.119 | 1.06 | 0.71 | 4.72 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 99 | 14z |
| 62 | 3-Cl, 4-$O(CH_2)_3$-$OC_6H_4$-4'-$SO_2F$ | 7.07 | 7.451 | 0.38 | 0.71 | 4.21 | 5.13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 81 | 14j |
| 63 | 3-$NO_2$ | 7.07 | 6.550 | 0.52 | −0.28 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 64 | 3-$(CH_2)_2COCH_2Cl$ | 7.10 | 6.881 | 0.22 | 0.20 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 65 | 3-$(CH_2)_4COCH_2Cl$ | 7.10 | 7.396 | 0.30 | 1.20 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 66 | 4-$OCH_2CO$-$N(CH_2)_5$ | 7.12 | 7.247 | 0.13 | 0.00 | −0.32 | 3.78 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 67 | 4-$CH_2CO$-$N(CH_2CH_2)_2O$ | 7.12 | 7.217 | 0.10 | 0.00 | −1.70 | 3.27 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 68 | 4-$(CH_2)_6C_6H_4$-4'-$SO_2F$ | 7.12 | 7.672 | 0.55 | 0.00 | 5.01 | 6.09 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 66 | 14z |
| 69 | 3-Cl, 4-$OCH(CH_3)$-$CONHC_6H_4$-4'-$SO_2F$ | 7.13 | 7.444 | 0.31 | 0.71 | 1.91 | 5.37 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 23 | 14s |
| 70 | 4-$CH_2CH(Ph)CONHC_6H_4$-4'-$SO_2F$ | 7.13 | 7.069 | 0.06 | 0.00 | 3.53 | 7.59 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 80 | 14g |
| 71 | 3-Cl, 4-$O(CH_2)_2$-$O(CH_2)_2OC_6H_4$-4'-$SO_2F$ | 7.14 | 7.426 | 0.29 | 0.71 | 3.38 | 5.80 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 76 | |
| 72 | 3-Cl, 4-$O(CH_2)_3$-$CONHC_6H_4$-4'-$SO_2F$ | 7.15 | 7.424 | 0.27 | 0.71 | 2.38 | 5.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 57 | 14s |
| 73 | 3-Cl, 4-$OCH_2CONMe_2$ | 7.16 | 7.581 | 0.42 | 0.71 | −1.36 | 2.58 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 74 | 3-Cl, 4-$O(CH_2)_3$-$CONHC_6H_4$-3'-$SO_2F$ | 7.17 | 7.424 | 0.25 | 0.71 | 2.38 | 5.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 48 | 14s |
| 75 | 4-Cl, 3-$O(CH_2)_4$-$OC_6H_4$-4'-$SO_2F$ | 7.17 | 7.47 | 0.30 | 3.92 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 51 | 14q |
| 76 | 4-$CH_2CH$(Ph-3"-Me)-$CONHC_6H_4$-4'-$SO_2F$ | 7.17 | 7.000 | 0.17 | 0.00 | 4.09 | 8.05 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 14g |
| 77[e] | 3-$(CH_2)_2CONHC_6H_4$-4'-$SO_2F$ | 7.19 | 8.287 | 1.10 | 1.77 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14a |
| 78 | 4-$CH_2CH$(Ph-4"-Me)-$CONHC_6H_4$-4'-$SO_2F$ | 7.24 | 7.000 | 0.24 | 0.00 | 4.09 | 8.05 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14g |
| 79 | 4-$CH_2CH$(Ph-2"-$CH_3$)-$CONHC_6H_4$-4'-$SO_2F$ | 7.24 | 7.000 | 0.24 | 0.00 | 4.09 | 8.05 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 80 | 14g |
| 80 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CONHC_6H_4$-3"-$SO_2F$ | 7.24 | 7.288 | 0.05 | 0.71 | 3.16 | 7.34 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 41 | 14s |
| 81 | 3-Cl, 4-$OCH_2C_6H_4$-2'-$CONHC_6H_4$-4"-$SO_2F$ | 7.24 | 7.288 | 0.05 | 0.71 | 3.16 | 7.34 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 12 | 14s |
| 82 | 3-Cl, 4-$O(CH_2)_4$-$CONHC_6H_4$-4'-$SO_2F$ | 7.24 | 7.394 | 0.15 | 0.71 | 2.88 | 6.30 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 89 | 14s |
| 83 | 3-Cl, 4-$OCH_2C_6H_3$-5'-Cl, 2'-$SO_2F$ | 7.27 | 7.453 | 0.18 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 73 | 14r |
| 84 | 4-Cl, 3-$O(CH_2)_2$-$OC_6H_4$-4'-$SO_2F$ | 7.27 | 7.596 | 0.33 | 3.00 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 29 | 14q |
| 85 | 3-$SO_2F$ | 7.27 | 6.783 | 0.49 | 0.05 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14d |
| 86 | 3-Cl, 4-$O(CH_2)_3NH$-$CONHC_6H_4$-3'-$SO_2F$ | 7.28 | 7.403 | 0.12 | 0.71 | 2.72 | 6.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 8 | 14o |
| 87 | 4-$(CH_2)_2CONEt_2$ | 7.28 | 7.246 | 0.03 | 0.00 | −0.21 | 3.76 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 88 | 3-Cl, 4-$OCH_2CO$-$N(CH_2)_4$ | 7.29 | 7.644 | 0.35 | 0.71 | −0.72 | 3.31 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 89 | 4-$OCH_2CO$-$N(CH_2CH_2)_2O$ | 7.29 | 7.231 | 0.06 | 0.00 | −1.39 | 3.49 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 90 | 4-$CH(CH_3)CH_2$-$CONHC_6H_4$-4'-$SO_2F$ | 7.29 | 7.249 | 0.04 | 0.00 | 2.07 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 50 | 14g |
| 91 | 4-$CH_2CON(Me)$-$CH_2C_6H_5$ | 7.30 | 7.268 | 0.03 | 0.00 | 0.43 | 4.80 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 92 | 4-$(CH_2)_2CON(Me)$-$CH_2C_6H_5$ | 7.31 | 7.261 | 0.05 | 0.00 | 0.93 | 5.27 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 93 | 4-$(CH_2)_2CO$-$N(CH_2CH_2)_2O$ | 7.32 | 7.245 | 0.08 | 0.00 | −1.20 | 3.74 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 94 | 4-$O(CH_2)_3NHCONH$-$C_6H_4$-3'-$SO_2F$ | 7.32 | 6.979 | 0.34 | 0.00 | 2.72 | 6.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 23 | 14o |
| 95 | 3-Cl, 4-$O(CH_2)_3$-$NHCOC_6H_4$-4'-$SO_2F$ | 7.34 | 7.424 | 0.08 | 0.71 | 1.42 | 5.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 69 | 14o |
| 96 | 3-$CH_2CONHC_6H_4$-4'-$SO_2F$ | 7.34 | 7.438 | 0.10 | 1.31 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14a |
| 97 | 4-$CH_2NHCONHC_6H_4$-4'-$SO_2F$ | 7.35 | 7.732 | 0.38 | 0.00 | 1.84 | 5.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 42 | 14n |
| 98 | 4-$(CH_2)_2CON(C_3H_7)_2$ | 7.35 | 7.269 | 0.08 | 0.00 | 0.80 | 4.69 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 99 | 3-Cl, 4-$OCH_2C_6H_3$-6'-Cl, 3'-$SO_2F$ | 7.38 | 7.453 | 0.07 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 87 | 14r |
| 100 | 3-Cl, 4-$OCH_2C_6H_3$-2'-$CH_3$, 4'-$SO_2F$ | 7.38 | 7.453 | 0.07 | 0.71 | 2.27 | 4.44 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 16 | 14r |
| 101 | 3-Cl, 4-$S(CH_2)_2$-$CONHC_6H_4$-4'-$SO_2F$ | 7.39 | 7.417 | 0.03 | 0.71 | 2.74 | 5.97 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 74 | 14s |
| 102 | 4-$(CH_2)_2C_6H_4$-4'-$SO_2F$ | 7.41 | 7.711 | 0.30 | 0.00 | 2.71 | 4.23 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0 | 14v |
| 103 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$CONHC_6H_4$-3"-$SO_2F$ | 7.41 | 7.288 | 0.12 | 0.71 | 3.16 | 7.34 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 14s |
| 104 | 4-$(CH_2)_2NHSO_2C_6H_4$-4'-$SO_2F$ | 7.41 | 7.255 | 0.16 | 0.00 | 1.01 | 5.48 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 38 | 14f |
| 105 | 3-Cl, 4-$SCH_2CONH$-$C_6H_4$-4'-$SO_2F$ | 7.42 | 7.440 | 0.02 | 0.71 | 2.24 | 5.51 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 86 | 14s |
| 106 | 3-Cl, 4-$OCH_2C_6H_3$-3'-Cl, 2'-$SO_2F$ | 7.42 | 7.453 | 0.03 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 19 | 14r |
| 107 | 3-Cl, 4-$OCH_2CONH$-$C_6H_4$-4'-$SO_2F$ | 7.43 | 7.454 | 0.02 | 0.71 | 1.61 | 4.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 76 | 14y |
| 108 | 3-Cl, 4-$OCH_2C_6H_4$-2'-$SO_2F$ | 7.43 | 7.441 | 0.01 | 0.71 | 1.71 | 3.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 36 | 14r |
| 109 | 3-Cl, 4-$OCH_2C_6H_3$-3'-Cl, 4'-$SO_2F$ | 7.43 | 7.453 | 0.02 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 97 | 14r |
| 110 | 3-Cl, 4-$OCH_2C_6H_3$-2'-Cl, 4'-$SO_2F$ | 7.44 | 7.453 | 0.01 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 75 | 14r |
| 111 | 3-Cl, 4-$O(CH_2)_2$-$OC_6H_4$-4'-$SO_2F$ | 7.44 | 7.454 | 0.01 | 0.71 | 3.00 | 4.66 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 97 | 14y |
| 112 | 3-$(CH_2)_4C_6H_3$-2',4'-$Cl_2$ | 7.45 | 7.843 | 0.39 | 5.08 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14e |
| 113 | 3-Cl, 4-$O(CH_2)_6$-$OC_6H_4$-4'-$SO_2F$ | 7.46 | 7.376 | 0.08 | 0.71 | 5.00 | 6.52 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 96 | 14q |
| 114 | 4-$(CH_2)_2CONHC_6H_3$-3'-OMe, 4'-$SO_2F$ | 7.46 | 7.942 | 0.48 | 0.00 | 1.75 | 5.84 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 95 | 14h |
| 115 | 3-Cl, 4-$OCH_2CON$-$(CH_3)C_6H_4$-4'-$SO_2F$ | 7.47 | 7.444 | 0.03 | 0.71 | 1.13 | 5.37 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 47 | 14j |
| 116 | 3-Cl, 4-$OCH_2CO$-$N(CH_2)_5$ | 7.47 | 7.670 | 0.20 | 0.71 | −0.32 | 3.78 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 117 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_2NMe_2$ | 7.48 | 7.685 | 0.21 | 0.71 | 0.88 | 5.27 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 118 | 3-Cl, 4-$OCH_2C_6H_3$-2'-Cl, 3'-$SO_2F$ | 7.49 | 7.453 | 0.04 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 79 | 14r |
| 119 | 3-$O(CH_2)_4OC_6H_4$-4'-$SO_2F$ | 7.49 | 7.344 | 0.15 | 4.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 48 | 14q |
| 120 | 4-Cl, 3-$O(CH_2)_3$-$OC_6H_4$-4'-$SO_2F$ | 7.51 | 7.553 | 0.04 | 3.50 | 0.71 | 0.60 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 26 | 14q |
| 121 | 3-Cl, 4-$OCH_2C_6H_4$-3'-CN | 7.51 | 7.669 | 0.16 | 0.71 | 1.09 | 3.75 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 122 | 3-Cl, 4-$OCH_2C_6H_5$ | 7.52 | 7.637 | 0.12 | 0.71 | 1.66 | 3.22 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 123 | 4-$SCH_2CONHC_6H_4$-4'-$SO_2F$ | 7.52 | 7.016 | 0.50 | 0.00 | 2.24 | 5.51 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 31 | 14s |
| 124 | 3-Cl, 4-$OCH_2C_6H_3$-4'-Cl, 2'$SO_2F$ | 7.52 | 7.453 | 0.07 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 60 | 14r |
| 125 | 3-$CH_2NHCONHC_6H_5$ | 7.52 | 7.699 | 0.18 | 0.83 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 126 | 4-$CH_2CH(Me)CONH$-$C_6H_4$-4'-$SO_2F$ | 7.55 | 7.249 | 0.30 | 0.00 | 2.07 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 32 | 14g |
| 127 | 3-$O(CH_2)_3OC_6H_4$-4'-$NHCOCH_2Br$ | 7.55 | 7.584 | 0.03 | 1.77 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 128 | 4-$(CH_2)_2CON(Me)$-$C_6H_5$ | 7.56 | 7.972 | 0.41 | 0.00 | 0.31 | 4.80 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14u |
| 129 | 3-Cl, 4-$O(CH_2)_4$-$OC_6H_4$-4'-$SO_2F$ | 7.57 | 7.436 | 0.13 | 0.71 | 4.00 | 5.59 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 90 | 14q |
| 130 | 3-Cl, 4-$O(CH_2)_5$-$OC_6H_4$-4'-$SO_2F$ | 7.57 | 7.412 | 0.16 | 0.71 | 4.50 | 6.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 25 | 14q |
| 131 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_2F$ | 7.58 | 7.441 | 0.14 | 0.71 | 1.71 | 3.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 91 | 14r |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 132 | 4-$(CH_2)_2CONC_6H_4$-4'-$SO_2F$ | 7.60 | 7.730 | 0.13 | 0.00 | 1.77 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 73 | 14y |
| 133 | 3-Cl, 4-$(CH_2)_2CONH$-$C_6H_4$-4'-$SO_2F$ | 7.62 | 8.154 | 0.53 | 0.71 | 1.77 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 95 | 14y |
| 134 | 3-$CH_2NHCONHC_6H_4$-3'-$SO_2F$ | 7.62 | 8.305 | 0.69 | 1.84 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14d |
| 135 | 4-$(CH_2)_2NHSO_2C_6H_4$-3'-$SO_2F$ | 7.64 | 7.255 | 0.39 | 0.00 | 1.01 | 5.48 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14f |
| 136 | 3-Cl, 4-$OCH_2CONEt_2$ | 7.64 | 7.656 | 0.02 | 0.71 | −0.36 | 3.51 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 137 | 3-$O(CH_2)_3OC_6H_4$-3'-$NHCOCH_2Br$ | 7.64 | 7.584 | 0.06 | 1.77 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 138 | 3-$O(CH_2)_2OC_6H_4$-2'-$NHCOCH_2Br$ | 7.66 | 7.423 | 0.24 | 1.27 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 139 | 3-$O(CH_2)_2OC_6H_4$-3'-$NHCOCH_2Br$ | 7.66 | 7.423 | 0.24 | 1.27 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 140 | 3-Cl, 4-$SCH_2CONH$-$C_6H_4$-3'-$SO_2F$ | 7.66 | 7.440 | 0.22 | 0.71 | 2.24 | 5.51 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 27 | 14s |
| 141 | 3-Cl, 4-$O(CH_2)_4$-$NHCOC_6H_4$-4'-$SO_2F$ | 7.66 | 7.401 | 0.26 | 0.71 | 1.92 | 6.21 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 100 | 14o |
| 142 | 4-$(CH_2)_3CONHC_6H_4$-4'-$SO_2F$ | 7.66 | 7.953 | 0.29 | 0.00 | 2.27 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 22 | 14d |
| 143 | 3-Cl, 4-$O(CH_2)_3NH$-$CONHC_6H_4$-4'-$SO_2F$ | 7.68 | 7.403 | 0.28 | 0.71 | 2.72 | 6.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 93 | 14o |
| 144 | 3-Cl, 4-$O(CH_2)_4NH$-$CONHC_6H_4$-3'-$SO_2F$ | 7.70 | 7.365 | 0.34 | 0.71 | 3.22 | 6.65 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 98 | 14o |
| 145 | 3-$(CH_2)_4C_6H_3$-3'-Cl, 4'-$SO_2F$ | 7.70 | 7.897 | 0.20 | 4.42 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 33 | 14z |
| 146 | 3-Cl, 4-$(CH_2)_4C_6H_3$-3'-Cl, 4'-$SO_2F$ | 7.70 | 8.119 | 0.42 | 0.71 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 35 | 14z |
| 147 | 4-$(CH_2)_4C_6H_4$-4'-$SO_2F$ | 7.70 | 7.712 | 0.12 | 0.00 | 3.71 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 75 | 14z |
| 148 | 4-$CH_2CONHC_6H_4$-4'-$SO_2F$ | 7.70 | 7.269 | 0.43 | 0.00 | 1.31 | 4.69 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14a |
| 149 | 4-$O(CH_2)_2OC_6H_4$-4'-$NHCOCH_2Br$ | 7.70 | 7.223 | 0.48 | 0.00 | 1.27 | 6.09 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 150 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CONMe_2$ | 7.72 | 7.690 | 0.03 | 0.71 | 0.15 | 5.02 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 151 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-Cl | 7.72 | 8.409 | 0.69 | 0.71 | 3.92 | 7.29 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 41 | 14t |
| 152 | 4-$OCH_2CONHC_6H_4$-4'-$SO_2F$ | 7.72 | 7.268 | 0.45 | 0.00 | 1.61 | 4.91 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14f |
| 153 | 3-Cl, 4-$OCH_2CONH$-$C_6H_4$-3'-$SO_2F$ | 7.72 | 7.454 | 0.27 | 0.71 | 1.61 | 4.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 61 | 14j |
| 154 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$SO_2F$ | 7.72 | 7.441 | 0.28 | 0.71 | 1.71 | 3.98 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 59 | 14r |
| 155 | 3-Cl, 4-$OCH_2C_6H_3$-6'-Cl, 2'-$SO_2F$ | 7.72 | 7.453 | 0.27 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 30 | 14r |
| 156 | 4-$CH_2NHCONHC_6H_3$-3'-Me, 4'-$SO_2F$ | 7.72 | 7.718 | 0.00 | 0.00 | 2.40 | 5.54 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 70 | 14k |
| 157 | 4-$(CH_2)_2CONHC_6H_4$-3'-$SO_2F$ | 7.74 | 7.730 | 0.01 | 0.00 | 1.77 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 49 | 14i |
| 158 | 3,5-$Cl_2$, 4-$OCH_2$-$CONHC_6H_4$-4'-$SO_2F$ | 7.74 | 7.454 | 0.29 | 0.71 | 1.62 | 4.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 65 | 14y |
| 159 | 3-Cl | 7.76 | 7.173 | 0.59 | 0.71 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 160 | 3-$CF_3$ | 7.76 | 7.257 | 0.50 | 0.88 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 161 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-4"-Cl | 7.77 | 8.409 | 0.64 | 0.71 | 3.92 | 7.29 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 39 | 14t |
| 162 | 3-$CH_2NHCONHC_6H_4$-3'-$CON(Me)_2$ | 7.77 | 6.699 | 1.07 | −0.68 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 163 | 3-Cl, 4-$(CH_2)_4C_6H_4$-2'-$SO_2F$ | 7.77 | 8.136 | 0.37 | 0.71 | 3.71 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 46 | 14z |
| 164 | 3-Cl, 4-$(CH_2)_4C_6H_3$-2'-Cl, 4'-$SO_2F$ | 7.77 | 8.119 | 0.35 | 0.71 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 94 | 14z |
| 165 | 3-Cl, 4-$CH_2NHCONH$-$C_6H_3$-3'-Me, 4'-$SO_2F$ | 7.80 | 8.142 | 0.34 | 0.71 | 2.40 | 5.54 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 57 | 14k |
| 166 | 4-$O(CH_2)_2OC_6H_4$-4'-$SO_2F$ | 7.80 | 7.268 | 0.53 | 0.00 | 3.00 | 4.67 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14f |
| 167 | 4-$(CH_2)_3CONHC_6H_4$-2'-$SO_2F$ | 7.80 | 7.953 | 0.15 | 0.00 | 2.27 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14d |
| 168 | 3-Cl, 4-$O(CH_2)_2$-$NHCONHC_6H_3$-3'-Me, 4'-$SO_2F$ | 7.82 | 7.403 | 0.42 | 0.71 | 2.78 | 6.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 90 | 14o |
| 169 | 3-$O(CH_2)_2OC_6H_4$-4'-$SO_2F$ | 7.82 | 7.490 | 0.33 | 3.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 25 | 14q |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 170 | 3-Cl, 4-$(CH_2)_4C_6H_3$-4'-Cl, 2'-$SO_2F$ | 7.82 | 8.119 | 0.30 | 0.71 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 93 | 14z |
| 171 | 3-Cl, 4-$(CH_2)_2C_6H_4$-4'-$SO_2F$ | 7.85 | 8.134 | 0.28 | 0.71 | 2.71 | 4.23 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 93 | 14z |
| 172 | 3-Cl, 4-$(CH_2)_2C_6H_3$-5'-Cl, 2'-$SO_2F$ | 7.85 | 8.140 | 0.29 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 31 | 14z |
| 173 | 3-Cl, 4-$(CH_2)_2C_6H_3$-3'-Cl, 4'-$SO_2F$ | 7.85 | 8.140 | 0.29 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 83 | 14z |
| 174 | 3-Cl, 4-$OCH_2CO$-$N(CH_2CH_2)_2O$ | 7.85 | 7.655 | 0.20 | 0.71 | −1.39 | 3.49 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 175 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CON(CH_2CH_2)_2O$ | 7.85 | 7.657 | 0.19 | 0.71 | 0.13 | 5.93 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 176 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CO$-$N(CH_2)_4$ | 7.85 | 7.666 | 0.18 | 0.71 | 0.80 | 5.75 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 177 | 3-Cl, 4-$OCH_2CON$-$(Me)C_6H_5$ | 7.89 | 7.691 | 0.20 | 0.71 | 0.12 | 4.55 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 178 | 4-$OCH_2CONHC_6H_5$ | 7.89 | 7.259 | 0.63 | 0.00 | 0.60 | 4.09 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 179 | 4-$(CH_2)_2C_6H_5$ | 7.89 | 7.678 | 0.21 | 0.00 | 2.66 | 3.47 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 180 | 4-$(CH_2)_2CONHC_6H_3$-3'-Me, 4'-$SO_2F$ | 7.89 | 7.953 | 0.06 | 0.00 | 2.33 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 100 | 14h |
| 181 | 3-Cl, 4-$CH_2NHCONH$-$C_6H_4$-4'-$SO_2F$ | 7.92 | 8.155 | 0.24 | 0.71 | 1.84 | 5.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 85 | 14y |
| 182 | 3-Cl, 4-$O(CH_2)_2NH$-$CONHC_6H_4$-4'-$SO_2F$ | 7.92 | 7.665 | 0.26 | 0.71 | 2.22 | 5.77 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 77 | 14y |
| 183 | 4-$(CH_2)_3CONHC_6H_4$-3'-$SO_2F$ | 7.92 | 7.953 | 0.03 | 0.00 | 2.27 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14d |
| 184 | 4-$(CH_2)_2COCH_2Cl$ | 7.92 | 7.145 | 0.78 | 0.00 | 0.20 | 2.47 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 185 | 3-$OC_6H_4$-4'-$NHCOCH_2Br$ | 7.92 | 7.568 | 0.35 | 1.71 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 186 | 3-Cl, 4-$(CH_2)_4C_6H_5$ | 7.92 | 8.138 | 0.22 | 0.71 | 3.66 | 4.39 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 187 | 4-$(CH_2)_4C_6H_3$-2',4'-$Cl_2$ | 7.92 | 7.706 | 0.21 | 0.00 | 5.08 | 5.39 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 188 | 3-Cl, 4-$(CH_2)_4C_6H_5$ | 7.96 | 8.375 | 0.42 | 0.71 | 4.13 | 4.39 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14e |
| 189 | 3-$O(CH_2)_3OC_6H_4$-4'-$SO_2F$ | 7.96 | 7.446 | 0.51 | 3.50 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 18 | 14q |
| 190 | 3-$(CH_2)_4C_6H_3$-5'-Cl, 2'-$SO_2F$ | 7.96 | 7.897 | 0.06 | 4.42 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 16 | 14z |
| 191 | 4-$(CH_2)_4C_6H_3$-2'-Cl, 4'-$SO_2F$ | 7.96 | 7.696 | 0.26 | 0.00 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 69 | 14z |
| 192 | 3-Cl, 4-$OCH_2C_6H_3$-4'-Cl, 3'-$SO_2F$ | 8.00 | 7.453 | 0.55 | 0.71 | 2.42 | 4.48 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 19 | 14r |
| 193 | 3-$(CH_2)_4C_6H_3$-2'-Cl, 4'-$SO_2F$ | 8.00 | 7.897 | 0.10 | 4.42 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 100 | 14z |
| 194 | 4-$OCH_2CONHC_6H_4$-3'-$SO_2F$ | 8.00 | 7.268 | 0.73 | 0.00 | 1.61 | 4.91 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14f |
| 195 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CONHC_6H_5$ | 8.00 | 7.613 | 0.39 | 0.71 | 2.15 | 6.53 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 196 | 3-$CH_2C_6H_5$ | 8.00 | 8.325 | 0.33 | 2.01 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14e |
| 197 | 4-$(CH_2)_4C_6H_5$ | 8.00 | 7.714 | 0.29 | 0.00 | 3.66 | 4.39 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 198 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CO$-$N(CH_2)_5$ | 8.02 | 7.638 | 0.38 | 0.71 | 1.20 | 6.21 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 199 | 3-$CH_2NHCONHC_6H_4$-3'-$OCH_3$ | 8.02 | 7.689 | 0.33 | 0.81 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 200 | 4-$(CH_2)_2CONHC_6H_3$-4'-Me, 3'-$SO_2F$ | 8.02 | 7.953 | 0.07 | 0.00 | 2.33 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 55 | 14h |
| 201 | 3-Cl, 4-$(CH_2)_4C_6H_4$-3'-$SO_2F$ | 8.03 | 8.136 | 0.11 | 0.71 | 3.71 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 100 | 14z |
| 202 | 3-$(CH_2)_4C_6H_3$-2',4'-$Cl_2$ | 8.03 | 7.606 | 0.42 | 5.08 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 203 | 4-$CH_2NHCONHC_6H_4$-3'-$SO_2F$ | 8.04 | 7.969 | 0.07 | 0.00 | 1.84 | 5.08 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14f |
| 204 | 4-$(CH_2)_2CON(Me)$-$C_6H_4$-4'-$SO_2F$ | 8.04 | 7.953 | 0.09 | 0.00 | 1.28 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14g |
| 205 | 3-Cl, 4-$(CH_2)_2C_6H_3$-4'-Cl, 2'-$SO_2F$ | 8.05 | 8.140 | 0.09 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 65 | 14z |
| 206 | 4-$CH_2C_6H_5$ | 8.05 | 7.882 | 0.17 | 0.00 | 2.01 | 3.00 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14e |
| 207 | 3-$CH_2NHCONHC_6H_4$-3'-Cl | 8.05 | 7.983 | 0.07 | 1.54 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 208 | 3-Cl, 4-$O(CH_2)_3NH$-$CONHC_6H_3$-4'-Me, 3'-$SO_2F$ | 8.06 | 7.365 | 0.70 | 0.71 | 3.28 | 6.64 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 87 | 14o |
| 209 | 4-$CH_2CONHC_6H_4$-3'-$SO_2F$ | 8.06 | 7.269 | 0.79 | 0.00 | 1.31 | 4.69 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0 | 14d |
| 210 | 4-$(CH_2)_2CONHC_6H_3$-6'-OMe, 3'-$SO_2F$ | 8.08 | 7.942 | 0.14 | 0.00 | 1.75 | 5.84 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 20 | 14h |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 211 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-$CF_3$ | 8.09 | 8.422 | 0.33 | 0.71 | 4.09 | 7.19 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 22 | 14t |
| 212 | 3-$CH_2NHCONHC_6H_4$-3'-$NO_2$ | 8.10 | 7.554 | 0.55 | 0.55 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 213 | 3-$(CH_2)_4C_6H_4$-4'-$SO_2F$ | 8.10 | 8.096 | 0.00 | 3.71 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 86 | 14z |
| 214 | 3-$(CH_2)_4C_6H_4$-3'-$SO_2F$ | 8.10 | 8.096 | 0.00 | 3.71 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 94 | 14z |
| 215 | 3-$(CH_2)_2C_6H_4$-4'-$SO_2F$ | 8.10 | 8.174 | 0.07 | 2.71 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 53 | 14v |
| 216 | 4-$(CH_2)_2NHCOC_6H_4$-4'-$SO_2F$ | 8.11 | 7.968 | 0.14 | 0.00 | 1.11 | 5.16 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 27 | 14f |
| 217 | 3-Cl, 4-$(CH_2)_4C_6H_3$-4'-Cl, 3'-$SO_2F$ | 8.11 | 8.119 | 0.01 | 0.71 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 97 | 14z |
| 218 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CON(Me)C_6H_5$ | 8.12 | 7.567 | 0.55 | 0.71 | 2.15 | 6.99 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 219 | 3-$O(CH_2)_2OC_6H_4$-4'-$NHCOCH_2Br$ | 8.13 | 7.423 | 0.71 | 1.27 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14b |
| 220 | 3-Cl, 4-$OCH_2C_6H_4$-3'-$CONEt_2$ | 8.14 | 7.655 | 0.49 | 0.71 | 1.15 | 5.95 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14u |
| 221 | 3-Cl, 4-$(CH_2)_4C_6H_4$-4'-$SO_2F$ | 8.14 | 8.136 | 0.00 | 0.71 | 3.71 | 5.16 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 100 | 14z |
| 222 | 3-Br, 4-$OCH_2CONH$-$C_6H_4$-4'-$SO_2F$ | 8.14 | 7.528 | 0.61 | 0.86 | 1.61 | 4.91 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 64 | 14p |
| 223 | 4-$(CH_2)_4OC_6H_4$-4'-$SO_2F$ | 8.14 | 7.944 | 0.20 | 0.00 | 4.62 | 5.37 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 66 | 14z |
| 224 | 3-$(CH_2)_2C_6H_5$ | 8.19 | 8.172 | 0.02 | 2.66 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 225 | 3-$CH_2NHCONHC_6H_4$-3'-CN | 8.19 | 7.385 | 0.81 | 0.26 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | 14m |
| 226 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_2OC_6H_5$ | 8.20 | 8.465 | 0.27 | 0.71 | 3.21 | 6.79 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 54 | 14t |
| 227 | 3-Cl, 4-$(CH_2)_4C_6H_3$-3'-Cl, 2'-$SO_2F$ | 8.20 | 8.119 | 0.08 | 0.71 | 4.42 | 5.66 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0 | 14z |
| 228 | 4-$(CH_2)_2CONHC_6H_3$-2'-Me, 4'-$SO_2F$ | 8.24 | 7.953 | 0.29 | 0.00 | 2.33 | 5.62 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 90 | 14h |
| 229 | 4-$(CH_2)_2CONHC_6H_3$-4'-OMe, 3'-$SO_2F$ | 8.24 | 7.942 | 0.30 | 0.00 | 1.75 | 5.84 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0 | 14h |
| 230 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-CN | 8.24 | 8.405 | 0.17 | 0.71 | 2.64 | 7.32 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 59 | 14t |
| 231 | 4-$(CH_2)_4OC_6H_5$ | 8.24 | 7.717 | 0.52 | 0.00 | 3.61 | 4.61 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 232 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_3$-3",4"-$Cl_2$ | 8.25 | 8.341 | 0.09 | 0.71 | 4.63 | 7.79 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 22 | 14t |
| 233 | 3-$(CH_2)_2C_6H_4$-4'-$NHCOCH_2Br$ | 8.26 | 8.374 | 0.11 | 2.29 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14b |
| 234 | 3-Cl, 4-$(CH_2)_2C_6H_3$-4'-Cl, 3'-$SO_2F$ | 8.27 | 8.140 | 0.13 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 76 | 14z |
| 235 | 3-Cl, 4-$(CH_2)_2C_6H_3$-3'-Cl, 2'-$SO_2F$ | 8.30 | 8.140 | 0.16 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0 | 14z |
| 236 | 3-Cl, 4-$(CH_2)_2C_6H_3$-2'-Cl, 4'-$SO_2F$ | 8.33 | 8.140 | 0.19 | 0.71 | 3.42 | 4.73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 76 | 14z |
| 237 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-$CF_3$ | 8.33 | 8.422 | 0.09 | 0.71 | 4.09 | 7.19 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 37 | 14t |
| 238 | 3-$(CH_2)_4OC_6H_5$ | 8.35 | 8.115 | 0.24 | 3.61 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 239 | 3-$(CH_2)_4C_6H_5$ | 8.35 | 8.106 | 0.24 | 3.66 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14v |
| 240 | 3-$(CH_2)_4C_6H_3$-4'-Cl, 3'-$SO_2F$ | 8.37 | 7.897 | 0.47 | 4.42 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 100 | 14z |
| 241 | 3-$(CH_2)_4C_6H_4$-4'-$NHCOCH_2Br$ | 8.38 | 8.400 | 0.02 | 3.24 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14b |
| 242 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-4"-CN | 8.39 | 8.405 | 0.02 | 0.71 | 2.64 | 7.32 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 71 | 14t |
| 243 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-4"-$OCH_3$ | 8.40 | 8.385 | 0.02 | 0.71 | 3.19 | 7.47 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 75 | 14t |
| 244 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-4"-F | 8.40 | 8.466 | 0.07 | 0.71 | 3.35 | 6.78 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 45 | 14t |
| 245 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-$OCH_3$ | 8.40 | 8.385 | 0.02 | 0.71 | 3.19 | 7.47 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 55 | 14t |
| 246 | 3-$(CH_2)_4C_6H_4$-3'-$NHCOCH_2Br$ | 8.41 | 8.400 | 0.01 | 3.24 | 0.00 | 0.10 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | | 14b |
| 247 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-$CH_3$ | 8.44 | 8.414 | 0.03 | 0.71 | 3.77 | 7.25 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 26 | 14t |
| 248 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-F | 8.46 | 8.466 | 0.01 | 0.71 | 3.35 | 6.78 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 46 | 14t |
| 249 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-$OCH_3$ | 8.52 | 8.385 | 0.14 | 0.71 | 3.19 | 7.47 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 38 | 14t |
| 250[e] | 3,4-$Cl_2$ | 8.54 | 7.279 | 1.26 | 0.71 | 0.71 | 0.60 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | 14e |
| 251 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-Cl | 8.62 | 8.409 | 0.21 | 0.71 | 3.92 | 7.29 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 48 | 14t |

Table I. (*Continued*)

| No. | X | Log 1/C Obsd[a] | Log 1/C Calcd[b] | \|Δ log 1/C\| | π-3[c] | π-4[c] | MR-4[c] | I-1[c] | I-2[c] | I-3[c] | I-4[c] | I-5[c] | I-6[c] | % inactivation[d] | Ref |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 252 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-4"-$CON(CH_3)_2$ | 8.62 | 8.206 | 0.41 | 1.71 | 1.70 | 8.59 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 58 | 14t |
| 253 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-$CON(CH_3)_2$ | 8.63 | 8.206 | 0.42 | 0.71 | 1.70 | 8.59 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 71 | 14t |
| 254 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-CN | 8.70 | 8.405 | 0.30 | 0.71 | 2.64 | 7.32 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 54 | 14t |
| 255 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-2"-F | 8.74 | 8.466 | 0.27 | 0.71 | 3.35 | 6.78 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 56 | 14t |
| 256 | 3-Cl, 4-$OCH_2C_6H_4$-4'-$SO_3C_6H_4$-3"-$CON(CH_3)_2$ | 8.76 | 8.206 | 0.55 | 0.71 | 1.70 | 8.59 | 1.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 55 | 14t |

[a] See ref 14. [b] Calculated using eq 2. [c] See section on Methods for sources of these constants. [d] Percent irreversible inhibition. [e] These points not used in deriving eq 2–11.

Table II. Squared Correlation Matrix for Parameters Used in the Correlation Study[a]

| | π-3 | π-4 | MR-3 | MR-4 | σ-sum | I-1 | I-2 | I-3 | I-4 | I-5 | I-6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| π-3 | 1.00 | 0.05 | 0.54 | 0.06 | 0.00 | 0.09 | 0.02 | 0.01 | 0.14 | 0.02 | 0.00 |
| π-4 | | 1.00 | 0.20 | 0.58 | 0.00 | 0.12 | 0.02 | 0.00 | 0.07 | 0.04 | 0.00 |
| MR-3 | | | 1.00 | 0.48 | 0.00 | 0.01 | 0.02 | 0.00 | 0.01 | 0.04 | 0.00 |
| MR-4 | | | | 1.00 | 0.05 | 0.01 | 0.07 | 0.00 | 0.14 | 0.02 | 0.00 |
| σ-sum | | | | | 1.00 | 0.00 | 0.29 | 0.00 | 0.00 | 0.01 | 0.06 |
| I-1 | | | | | | 1.00 | 0.03 | 0.02 | 0.07 | 0.11 | 0.00 |
| I-2 | | | | | | | 1.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| I-3 | | | | | | | | 1.00 | 0.00 | 0.01 | 0.00 |
| I-4 | | | | | | | | | 1.00 | 0.02 | 0.01 |
| I-5 | | | | | | | | | | 1.00 | 0.03 |
| I-6 | | | | | | | | | | | 1.00 |

[a] Numbers in Table II (×100) show the percent correlation ($r^2$) between each of the variables.

Except for compound **38**, these derivatives are about eight times as active as the continuous variables alone would predict. Only one of these compounds (**38**) has a long bridge and this congener does not show this special activity. Rather than drop it, we have given I-4 a value of 0 and find that it is well fit. This suggests that this leaving group must be properly placed for I-4-type activity. It seems likely that Baker was not able to separate reversible from irreversible activity for these sulfonate functions.

No special activity could be found for the nucleophilic function $SO_2F$. Indicator variables for $SO_2F$ in the 3 and 4 positions were explored separately and combined but no significant reduction in the variance was observed when such terms were used. Other nucleophilically active groups, $NHCOCH_2Br$ and $COCH_2Cl$, with which Baker hoped to covalently bond inhibitors to the enzyme also showed no special activity.

I-5 takes a value of 1 for flexible bridges such as $-CH_2-$, $-CH_2CH_2-$, $-(CH_2)_4-$, $-(CH_2)_6-$, and $-(CH_2)_4O-$ between the *N*-phenyl moiety and a second phenyl ring. This variable does not apply to those examples where the attachment to *N*-phenyl is through oxygen.

I-6 takes the value of 1 for bridges of the type $CH_2NHCONHC_6H_4-X$, $-CH_2CH_2C(=O)N(R)C_6H_4-X$, and $-CH_2CH_2CH_2C(=O)N(R)C_6H_4-X$ (R = H or Me) when these groups are attached to either the 3 or 4 position of the *N*-phenyl ring.

The search for meaningful constellations or patterns of atoms is as old as organic chemistry itself. New efforts are now underway to use computerized techniques to find groups which have special meaning in bio- and medicinal chemical processes.[10,11] The use of indicator variables can be a powerful tool in this search.[22b,c,25]

Our general approach to the formulation of indicator variables has been to first formulate the best possible correlation equations using the established parameters π, MR, σ, and $E_s$. When further reduction in variance cannot be made, a study of the residuals at this point often uncovers certain groupings of atoms with special activity. One can also factor a large data set into subsets and study these independently. From a comparison of the correlations one can determine which sets can be merged and what kind of indicator variables will be necessary.

In the present study two continuous variables (π-3 and MR-4) account for four terms; one indicator is necessary for the two enzyme systems so that five terms are used to account for special structural features. A number of other indicator variables were examined. Baker was impressed with the value of Cl in the 3 position. Testing this idea for the complete set of congeners revealed no special effect for this atom. In the few cases where a 4-Cl was present, it seemed to be less active than expected; however, the results are too few and too scattered to justify an additional term. The possibility that a single group in the 3 or 4 position might have significance was considered but none was found. The positional importance of $SO_2F$ in the second or third ring was tested for differences between the 3 and 4 positions. No difference was found despite the fact that it is quite clear in Table I that such positional effects are very important for irreversible reaction of this function with the enzyme. Thus it would seem that Baker was successful in separating reversible from irreversible inhibition. The $SO_2F$ function in the 2 position of the last ring is, with one exception, well fit by eq 2. The cross product terms 3-Cl·MR-4 and π-3·MR-4 were examined and found to be without influence. Also, the use of σ or $\sigma^+$ did not uncover any electronic effects of substituents on inhibition.

There are three examples in Table I (compounds **1**, **7**,

Table III. Correlation Equations for the Inhibition of Dihydrofolate Reductase by Triazines

| I-2 | I-3 | I-4 | I-5 | I-6 | $\pi$-3 | $(\pi\text{-}3)^2$ | MR-4 | $(\text{MR-4})^2$ | I-1 | $s$ | $r$ | $F_{n,k}$ | $F_{1,k}$ | Eq. no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| −3.15 | | | | | | | | | | 0.727 | 0.654 | 180 | 180 | 3 |
| −3.22 | −2.32 | | | | | | | | | 0.615 | 0.769 | 174 | 97 | 4 |
| −3.15 | −2.25 | 0.88 | | | | | | | | 0.568 | 0.809 | 162 | 43.4 | 5 |
| −3.07 | −2.11 | 1.02 | 0.67 | | | | | | | 0.508 | 0.851 | 157 | 60.9 | 6 |
| −2.99 | −2.01 | 1.11 | 0.76 | 0.58 | | | | | | 0.476 | 0.871 | 149 | 33.6 | 7 |
| −2.90 | −2.01 | 0.88 | 0.63 | | 0.51 | −0.11 | | | | 0.465 | 0.878 | 130 | | 8 |
| −2.77 | −1.88 | 0.98 | 0.72 | 0.68 | 0.57 | −0.13 | | | | 0.416 | 0.904 | 150 | 64 | 9 |
| −2.63 | −1.87 | 0.86 | 0.70 | 0.69 | 0.59 | −0.12 | 0.04 | | | 0.410 | 0.907 | 137 | 8.49 | 10 |
| −2.52 | −1.90 | 1.05 | 0.62 | 0.68 | 0.60 | −0.11 | 0.21 | −0.024 | | 0.388 | 0.918 | 139 | 27.7 | 11 |
| −2.53 | −1.99 | 0.88 | 0.69 | 0.70 | 0.68 | −0.12 | 0.23 | −0.024 | 0.24 | 0.377 | 0.923 | 134 | 15.3 | 2 |

158) where a Cl atom is in the 5 position. Without any parameter for 5-substituents, two of these are reasonably well fit and one is poorly fit. Two explanations for the well-fit points are possible. Cl could project into the surrounding solution from the 5 position so that it does not contact the enzyme or, if it does contact the enzyme, the interaction may be of the type modeled by MR-4. MR-4 for Cl is 0.6 but H is already present; hence, the increment which 5-Cl would add is $0.5 \times 0.23 = 0.11$. This is too small to detect with three data points and our present level of resolution.

In Table I, 12 congeners marked by the footnote *e* have not been employed in formulating eq 2-11. Ten of these derivatives are mispredicted by a factor of about 10, one by a factor of 50, and one by a factor of 250. We have taken the factor of 10 as the point of failure even though this is not so bad when it is considered that the concentration range of $1/C$ in Table I is over 200000-fold. However, these 12 points were not used in the derivation of eq 2-11. Ten of them can be included without greatly affecting $r$ or the values of the parameters of eq 2.

Baker found that dihydrofolate reductase could be irreversibly inhibited by some, but by no means all, of the congeners which contained the $SO_2F$ or $-C_6H_4SO_2OC_6H_4X$ functions. The derivatives in Table I are marked with the degree of irreversible inhibition they produce. It is of interest to note that the congeners capable of irreversible inhibition are as well fit by eq 2 as the strictly reversible inhibitors.

In all, eq 2 contains six indicator variables. Actually, it is possible to fit $n$ discrete changes with $n - 1$ indicator variables. However, for simplicity we have elected to use one variable for each discrete property of the system.

## Results

Equations 2-11 were generated via the method of least squares (see Table II)

$$\log 1/C = 0.680\ (\pm 0.12)\ (\pi\text{-}3) - 0.118\ (\pm 0.03)\ (\pi\text{-}3)^2 + 0.230\ (\pm 0.07)\ (\text{MR-4}) - 0.0243\ (\pm 0.009)\ (\text{MR-4})^2 + 0.238\ (\pm 0.12)\ (\text{I-1}) - 2.530\ (\pm 0.27)\ (\text{I-2}) - 1.991\ (\pm 0.29)\ (\text{I-3}) + 0.877\ (\pm 0.23)\ (\text{I-4}) + 0.686\ (\pm 0.14)\ (\text{I-5}) + 0.704\ (\pm 0.16)\ (\text{I-6}) + 6.489\ (\pm 0.16) \quad (2)$$

| $n$ | $r$ | $s$ |
|---|---|---|
| 244 | 0.923 | 0.377 |

ideal MR-4 = 4.7 (4.2-5.6)

ideal $\pi$-3 = 2.9 (2.6-3.3)

from the data in Table I. The figures in parentheses in this equation are the 95% confidence limits. The optimum values for MR-4 and $\pi$-3 are obtained from the partial derivatives of eq 2. We did not employ an automatic stepwise regression program; instead, we developed eq 2 by trial and error. After no further reduction in the variance could be obtained by indicator variables or cross product terms (see Method), all possible equations were generated from the linear combination of 11 variables; that is, the ten variables of eq 2 plus an indicator variable for the presence or absence of $SO_2F$ were used. This yielded the theoretical number ($2^n - 1$ or 2047) of regression equations.

We have summarized the equations which have the lowest standard deviation for each class in Table III; that is, of all possible three-variable equations, eq 5 had the lowest standard deviation; of all possible four-variable equations, eq 6 had the lowest standard deviation, etc. The regression coefficients for each equation are listed under the variable headings. Standard deviations and correlation coefficients are listed under $s$ and $r$. These equations are all based on 244 data points. The value of the overall $F$ statistic (i.e., for eq 2, $F_{n,k} = F_{10,233}$) for each equation is shown under $F_{n,k}$. The $F$ statistic for the addition of each successive term (except for eq 8 where the $F$ test cannot be applied) is listed under $F_{1,k}$. All of the single-term additions are highly significant. The $F$ value closest to $F_{233}$ is $F_{120}$ whose values are $F_{1,120\ (\alpha\ 0.001)} = 11.38$ and $F_{1,120\ (\alpha\ 0.005)} = 8.18$. At no point in the development in Table III did the indicator variable for $SO_2F$ appear. Its addition to eq 2 did not reduce the standard deviation.

One sees the relative importance in Table III of the different terms in eq 2. Ortho substituents have an extremely bad effect on inhibitory power. The coefficient with I-2 indicates that these derivatives are 300 times less active than one would expect from the other terms of eq 2 alone; hence, I-2 accounts for a very large amount of variance. The same applies to I-3 whose coefficient indicates that the congeners with rigid attachment to the *N*-phenyl moiety are, on the average, 100 times less active than expected. The coefficient with I-4 indicates that the $SO_2OC_6H_4$-X function confers a special activity on the average eight times more active than congeners lacking this group. The flexible carbon side chain accounted for by I-5 contributes five times the activity of the less flexible side chains. The special amide moiety $CH_2CH_2CONHC_6H_4$-X has a special fivefold-activating ability. The least significant variable is I-1 which accounts for the difference in the two different types of enzyme used in the assay.

Except for the discontinuity of eq 8, the variables fall into place in regular order. The two $\pi$ terms displace I-6 in going from eq 7 to eq 8; however, I-6 comes back in eq 9.

Equation 2 is essentially an extension of eq 1. The parameters common to the two equations are in quite reasonable agreement even though eq 1 was based on one-third the data and on congeners having less complex structural features. The greatest difference is in MR-4. In our first study we did not anticipate the unusual activity of the $SO_2OC_6H_4X$ group. The larger coefficient with MR-4 in eq 1 accounted for this rather well. This effect became clear only after we worked with groups having large MR-4 values.

Using the two physicochemical parameters $\pi$ and MR, eq 1 with 83 derivatives accounted for 82% of the variance in

the data. Adding 161 new derivatives of more complex structure resulted in a huge increase in variance so that it is found in eq 2 that 76% of the variance is now accounted for by the indicator variables. This clearly demonstrates the importance of this technique in dealing with truly gross structural modifications in a parent molecule.

## Discussion

However one may view eq 2, it is an objective means for discussing the structures and activities in Table I. One of the greatest values of eq 2 is that it constitutes a means for keeping account of the course of an extensive structure–activity study. One can see what ground has been covered in terms of substituent constants and the special features accounted for by the indicator variables and plot the course of new experiments to explore enzymic space into which substituents fall. The "substituents" in Table I, which are often larger than the parent structure, are so varied and so complex that even after long study one finds it impossible to hold all structures along with their significant deviations in mind. Correlation equations such as eq 2 give one an immediate bird's eye view of the general problem and at once suggest avenues for exploration.

Besides structuring the data, correlation equations have important predictive value. One of their great values lies in suggesting which new derivatives under consideration should *not* be made. For example, if substituent space in terms of the $\pi$ vector has been well covered by substituents in a given position having values in, say, the range $-1.0$ to $+1.0$, then one should avoid derivatives with $\pi$ in this range and instead make new derivatives with values in the unexplored region. Once $\pi_0$ (ideal $\pi$) has been established with some confidence, it becomes much less interesting to make derivatives with superoptimal $\pi$ values. This same philosophy holds for other continuous variables.

Sooner or later, a large enough change can be made in the properties of a substituent so that the correlation equation will not hold. Correlation equations should not be thought of as invariant permanent expressions, but rather as developing expressions to follow the mapping of substituent space about the region of a specific site in a macromolecule or macromolecular complex. By moving along a vector through increasing or decreasing values of a property of substituents, one is moving into the complete unknown. One cannot be sure what kind of substituent space will turn up without knowledge of the complete structure of the macromolecule. A correlation equation is the best guide for the design of new derivatives; however, one should move with some caution along a given vector. Gross changes may lead to results which are extremely difficult to incorporate into the developing correlation equation.

In our first analysis of the triazine data (eq 1), congeners with groups large enough in the 4 position to test the linear limit of MR-4 were not included. It is necessary in the present study to introduce the quadratic MR-4 term because of groups having superoptimal MR values. Now, from the partial derivative of eq 2 with respect to MR-4, the ideal value of 4.7 can be obtained for MR-4. In our first study (eq 1) essentially all derivatives having MR-4 values significantly greater than 5 contained the $SO_2OC_6H_4X$ function. This is still true in Table I. However, when I-3 is introduced to account for the special activity of the sulfonate group, these compounds are well predicted. It seems likely that Baker's test procedure was not fine enough to precisely delineate the boundary between reversible and irreversible inhibition by these most active congeners.

Although the completely reversible inhibitors tend to be among the less active derivatives, they are in fact dispersed throughout Table I. With the exception of two molecules (compounds **227** and **235**), no congener having a Cl in the 3 position and an $SO_2F$ function is completely reversible. The function of the 3-Cl appears to position the inhibitor on the enzyme so that the $SO_2F$ function can react more readily with a nucleophilic group of the enzyme.

The dependence of log $1/C$ on $\pi$-3 is the same in eq 2 as in eq 1 within the limits of the precision of the analysis. The initial slope with $\pi$-3 is much greater than that of MR-4 so that although there is considerable collinearity between $\pi$-3 and MR-3 and $\pi$-4 and MR-4, substituent space around the 3 position is significantly different from the 4 position.

There are two kinds of space which, in a general way, one can imagine to exist within and on the surface of an enzyme. A large amount of experience establishes the existence of hydrophobic pockets or pools more or less composed of apolar moieties.[26] The rest of enzymic space or surface is more or less polar. Our hypothesis is that $\pi$ models hydrophobic bonding between enzyme and ligand while MR models binding in polar space. This inference can only be supported by further studies with sets of congeners for which care has been taken at the start to design a set of substituents in which $\pi$ and MR are orthogonal; this is not difficult to accomplish.[11,13] It is realized that any section of reasonable size in or on an enzyme will not be completely homogeneous; we are talking about the predominant characteristic of a given volume or area. The many hundreds of successful correlations with $\pi$ and a growing number with MR support this rough division of enzymic space.

It is unfortunate that enough collinearity exists between $\pi$ and MR (see Table II) that we cannot say with certainty that three-space is hydrophobic and four-space is otherwise. We can say with certainty that three-space is quite different from four-space and that since poorer correlations are obtained using MR-3 and $\pi$-4, $\Sigma$MR or $\Sigma\pi$, three-space is probably generally hydrophobic. The situation in four-space is more ambivalent. $\pi$-4 gives almost as good a correlation as MR-4. The high degree of collinearity between these two variables precludes a clean decision about four-space. The much lower initial slope with MR-4 or $\pi$-4 than with $\pi$-3 indicates that four-space is indeed quite different and substituent interactions appear to be less strong. This suggests the weaker forces of polarizability and dispersion rather than hydrophobic interactions. An alternative explanation is that substituents in four-space are interacting hydrophobically, but only on a surface; that is, desolvation is occurring on only half of a substituent.

In considering the limits of $\pi$-3 space set by $(\pi\text{-}3)^2$ and MR-4 space set by $(\text{MR-4})^2$, it seems most likely that it is the bulk tolerance of these regions that eventually sets a maximum value on the advantage to be gained by placing large apolar groups in these positions. Another explanation is that such very large molecules are hindered in their movement through crude enzyme to the site of inhibition so that this random walk process might become rate limiting.

The variable I-2 for ortho substituents is of special interest. Although $E_s$-2 correlates variance in 2-substituents rather well, it is not as good as I-2. A thorough, but unsuccessful, attempt was made to improve the fit of 2-substituents by using $E_s$-2, $\pi$-2, MR-2, $\mathfrak{F}$-2, and $\mathfrak{R}$-2. As Baker pointed out, something more than a simple steric effect is involved. He suspected an electronic contribution but we could not substantiate this. It may be that some hydrogen bonding interaction with an amino group of the triazine is also involved.

Indicator variables I-3 and I-5 for rigid and flexible bridges stem from more or less self-evident steric problems. If large groups are held too rigidly on the *N*-phenyl ring, they do not fit readily into the nearby enzymic space. In the case of the very flexible bridges, the bulky end groups have

much better opportunity to locate a site for maximum interaction.

I-6 is a more difficult structural feature to understand. It is not surprising that an amide moiety should have a strong interaction with enzymic space; amides are one of the most common functions to turn up in drugs. What is unusual about I-6 is that it is highly specific; it holds for the groups $-CH_2CH_2CONHC_6H_4X$, $-CH_2CH_2CH_2CONHC_6H_4X$, and $-CH_2NHCONHC_6H_4X$ but not for the groups $-O(CH_2)_nCONHC_6H_4X$, $-O(CH_2)_nNHCONHC_6H_4X$, $CH_2CONHC_6H_4X$, and $CH_2CH(R)CONHC_6H_4X$. This parameter appears to be somewhat arbitrary and, in fact, may be covering two kinds of effects which happen to be varying in about the same way. The oxygen atom seems to offset the effect of the amide moiety in two of the general structures. There may not be enough flexibility in the bridge in the other two type structures for the amide to show its extra effect.

We have found in approaching a complex biochemical correlation problem that a good idea can be obtained as to whether the subsets are behaving in the same or different fashion by first factoring the data into subsets. It soon became clear from such studies that test results from the two different types of enzymes were essentially the same. The positive coefficient with I-1 brings out the fact that the Walker 256 test is, on the average, 1.7 times as sensitive as the L1210/DF8 test system.

In Table I, 12 points with the footnote *e* have not been used in the formulation of eq 2-11. All but one of these are among the less active derivatives. In most of these cases activity turns out to be less than expected. These congeners have little in common and, in general, no ready explanation can be offered for their deviant behavior. It is surprising that compound **14** with the simple CN group is mispredicted by a factor of 50; compound **63** with the strong electron-withdrawing $NO_2$ is only mispredicted by a factor of 3. It is possible that clerical or testing errors could occur in handling such a large amount of test data. Compounds with unusually complex branches positioned in three-space (compounds **31, 35,** and **45**) are also poorly predicted, although only one of these was dropped. The variable I-4 for rigidity next to the *N*-phenyl ring works well for 4-substituents but not so well for the rather few 3-substituents having this character (compounds **40** and **51**).

In our first study, compound **187** was poorly fit. In the present instance we have used test data from the L1210 enzyme instead of Walker enzyme and found this compound to be well fit.

To us, the most difficult example to understand is compound **250,** the 3,4-$(Cl)_2$ derivative which is almost 20 times more active than expected. Equation 1 gives almost the same poor result. Compounds **36, 159,** and **160** [2,3-$(Cl)_2$, 3-Cl, and 3-$CF_3$, respectively] are also more active than expected. This does suggest that enzymic space near the 3 position is quite sensitive to relatively small symmetrical groups. Studies of groups such as I and $C(CH_3)_3$ in the 3 position are now in progress.

The problem of recognizing special structural features in a highly complex set of congeners such as those in Table I requires a large amount of patient study of the data. Many hundreds of equations were examined in the formulation of eq 1 and this provided the foundation from which to attack the whole problem with the ultimate formulation of eq 2.

It is generally found that once some kind of reasonable correlation is obtained with the relatively nonspecific parameters $\pi$ and/or MR, it is possible to spot groups of congeners which show exceptional activity through an examination of the residuals. What general help our experience with indicator variables has to offer those interested in computerized pattern recognition is not evident. It must be emphasized that the use of indicator variables in correlation equations is of the utmost importance for making progress in structure-activity studies. For example, in eq 1 we had reached the limit of correlation using continuous functions. It is only by the use of indicator variables that we were able to see that 161 other congeners were behaving in the same way with respect to $\pi$-3 and MR-4. One must approach biochemical correlations in a different way from those in the homogeneous systems studied by physical organic chemists. The heterogeneous character of enzymes means that continuous variables are bound to fail sooner or later. The steric effects of substituents and bridges constitute the most difficult structural features to model in numerical terms. The use of indicator variables[22b,c,25] opens up an exciting pathway through this difficult jungle.

It is worth examining our results in the light of Baker's philosophy of drug research. It was his thesis[27] that one should select an enzyme critical for a pathogen and then, by making *gross* changes in the normal substrate, develop potent *reversible* inhibitors. He believed that portions of the inhibitor should extend considerably beyond the active site and one should next introduce functions such as $-NHCOCH_2Br$, $SO_2F$, and $SO_2OC_6H_4X$ on those positions of the most potent reversible inhibitors which could covalently anchor the inhibitor to the enzyme. Baker's reason for having the covalent bonding occur outside the active site was because one might expect to find the greatest differences between enzyme from host and pathogen in this part of enzymic space. One can hope to devise a truly effective drug only by building selectivity in an inhibitor. Baker accepted current biochemical thinking which postulates that the structure of an active site will be strongly conserved in evolutionary processes and that differences in isozymes will most likely be found outside the active sites.

Baker notes in his early work with the triazine inhibitors that functions such as $NHCOCH_2Br$ and $COCH_2Cl$ yield reversible inhibitors.[14a] He points out that the $SO_2F$ function is relatively inert and, for example, does not react with pyridine, hot ethanol, or hot acid solution. It does react with OH groups in cellulose when attached via a bound dye molecule. It is noteworthy to find in Table I that while many congeners with the $SO_2F$ function do form irreversible inhibitors, many do not. Baker is right in that the positioning of this group in or on the enzyme has a critical effect on its ability to form a covalent bond. It is a surprisingly selective reagent for nucleophilic groups. Nevertheless, as used in the congeners of Table I, a derivative having this function did not eventually reach clinical trials. One reason is that serum, especially mouse serum, contains an enzyme which hydrolyzes $SO_2F$.[28] Although it proved to be possible to make inhibitors which showed resistance to the mouse serum hydrolases, Baker did not have time to thoroughly explore this lead.[14h] The fact that a metabolic study of the $SO_2F$ function showed that none of it was excreted in the urine as such (it appeared only in the hydrolyzed form[14x]), coupled with the fact that reversible inhibitors appeared to be just as potent in vivo if not more so, dampened Baker's zest in the search for an irreversible inhibitor. Later, Baker and Ashton[14t] concluded that $SO_2OC_6H_4$-X functions were too easily hydrolyzed in vivo to be suitable as leaving groups for covalent bond formation.

What lessons does this study offer for the design of better antitumor drugs? Many of Baker's compounds, although highly active against dihydrofolate reductase in vitro, were inactive in vivo. Baker assumed that, to a considerable extent, this was the result of poor penetration of the drugs in living systems. This is to be expected with the large very lipophilic drugs of Table I. It is now well established[29] that

each congeneric series of drugs has an optimum degree of lipophilic character and large deviations from this (either positive or negative) lead to inactive compounds. No systematic attempt has been made to determine this figure for the action of these inhibitors in animals. Recently, Skeel et al.[30] presented evidence to show that the reason triazine II is much less effective against L1210 leukemia than against Walker 256 in vivo is due to differences in transport. Proper modulation of the lipophilic character of Baker's compounds will be an easy task because of the great nonspecificity of the binding site in the enzyme. Our present approach to this problem is to place groups having optimum $\pi$ values in the 3 position and then offset the overall lipophilicity by placing polar groups in the 4 position.

In summary, we can say that eq 2 is a highly significant correlation from many points of view. It fits very well with our first attempt (eq 1) to structure Baker's triazine work. All terms are highly significant with respect to the $F$ statistic. With the exception of I-2 and I-3 which are based on 12 and 9 data points, respectively, a large number of data points support each term, on the average, 24 per term. This precludes the possibility of chance correlation.[31] Actually, one can include all but three data points (compounds **14, 18,** and **36**) and derive an equation with essentially the same coefficients as eq 2 with $r = 0.898$ and $s = 0.428$. Equation 2 is a robust expression covering a great range of substituent space.

We believe that correlation analysis using indicator variables will be of enormous assistance to those trying to structure the very large data bases which are resulting from the massive efforts in drug research as well as those in homogeneous physical-organic reaction studies. Correlation equations are not to be thought of as finished products; they are a means for structuring a developing area of research. They will be most valuable in the complex decision making involved in new moves in an ongoing problem.

It is our belief that Baker's triazine study in Table I will remain a classic example for some time because of its extensive nature and because eq 2 clearly brings out the large degree of self-consistency. This data set is an excellent testing ground for further new approaches to structure-activity analysis.

## References and Notes


(1) (a) This investigation was supported by U.S. Public Health Service Research Grant No. CA-11110 from the National Cancer Institute; (b) NATO Fellow and Visiting Professor of Chemistry from the Institute of Pharmaceutical and Toxicological Chemistry, University of Naples.
(2) L. P. Hammett, "Physical Organic Chemistry", 2nd ed, McGraw-Hill, New York, N.Y., 1970.
(3) R. W. Taft in "Steric Effects in Organic Chemistry", M. S. Newman, Ed., Wiley, New York, N.Y., 1956, p 556.
(4) H. C. Brown and Y. Okampto, *J. Am. Chem. Soc.*, **80**, 4979 (1958).
(5) (a) C. Hansch, *Acc. Chem. Res.*, **2**, 232 (1969); (b) W. J. Dunn, III, and C. Hansch, *Chem.-Biol. Interact.*, **9**, 75 (1974).
(6) J. E. Leffler and E. Grunwald, "Rates and Equilibria of Organic Reactions", Wiley, New York, N.Y., 1963.
(7) J. Shorter, "Correlation Analysis in Organic Chemistry. An Introduction to Linear Free-Energy Relationships", Clarendon Press, Oxford, 1973.
(8) O. Exner in "Advances in Linear-Free Energy Relationships", N. B. Chapman and J. Shorter, Ed., Plenum Press, London, 1972, p 2.
(9) C. Daniel and F. S. Wood, "Fitting Equations to Data", Wiley-Interscience, New York, N.Y., 1971.
(10) B. R. Kowalski and C. F. Bender, *J. Am. Chem. Soc.*, **96**, 918 (1974).
(11) S. S. Schiffman, *Science*, **185**, 112 (1974).
(12) Y. C. Martin, J. B. Holland, C. H. Jarboe, and N. Plotnikoff, *J. Med. Chem.*, **17**, 409 (1974).
(13) C. Hansch, S. H. Unger, and A. B. Forsythe, *J. Med. Chem.*, **16**, 1217 (1973).
(14) (a) B. R. Baker and G. J. Lourens, *J. Med. Chem.*, **10**, 1113 (1967); (b) *ibid.*, **11**, 26 (1968); (c) *ibid.*, **11**, 34 (1968); (d) *ibid.*, **11**, 38 (1968); (e) B. R. Baker, *ibid.*, **11**, 483 (1968); (f) B. R. Baker and G. J. Lourens, *ibid.*, **11**, 666 (1968); (g) *ibid.*, **11**, 672 (1968); (h) *ibid.*, **11**, 677 (1968); (i) *ibid.*, **12**, 92 (1969); (j) *ibid.*, **12**, 95 (1969); (k) *ibid.*, **12**, 101 (1969); (l) B. R. Baker and M. A. Johnson, *ibid.*, **11**, 486 (1968); (m) W. T. Ashton, L. L. Kirk, and B. R. Baker, *ibid.*, **16**, 453 (1973); (n) B. R. Baker, G. J. Lourens, R. B. Meyer, Jr., and N. M. J. Vermeulen, *ibid.*, **12**, 67 (1969); (o) B. R. Baker and E. E. Janson, *ibid.*, **12**, 672 (1969); (p) B. R. Baker and W. T. Ashton, *ibid.*, **12**, 894 (1969); (q) *ibid.*, **13**, 1149 (1970); (r) *ibid.*, **13**, 1161 (1970); (s) *ibid.*, **13**, 1165 (1970); (t) *ibid.*, **15**, 945 (1972); (u) *ibid.*, **16**, 209 (1973); (v) B. R. Baker, E. E. Janson, and N. M. J. Vermeulen, *ibid.*, **12**, 898 (1969); (w) B. R. Baker, N. M. J. Vermeulen, W. T. Ashton, and A. J. Ryan, *ibid.*, **13**, 1130 (1970); (x) A. J. Ryan, N. M. J. Vermeulen, and B. R. Baker, *ibid.*, **13**, 1140 (1970); (y) B. R. Baker and N. M. J. Vermeulen, *ibid.*, **13**, 1143 (1970); (z) *ibid.*, **13**, 1154 (1970).
(15) G. H. Hitchings, *J. Infect. Dis.*, **128**, S433 (1973).
(16) J. J. Burchall, *J. Infect. Dis.*, **128**, S437 (1973).
(17) C. Hansch and C. Silipo, *J. Med. Chem.*, **17**, 661 (1974).
(18) A. Leo, C. Hansch, and D. Elkins, *Chem. Rev.*, **71**, 525 (1971).
(19) L. Pauling and D. Pressman, *J. Am. Chem. Soc.*, **67**, 1003 (1945).
(20) D. Agin, L. Hersh, and D. Holtzman, *Proc. Natl. Acad. Sci. U.S.A.*, **53**, 952 (1965).
(21) T. Fujita, J. Iwasa, and C. Hansch, *J. Am. Chem. Soc.*, **86**, 5175 (1964).
(22) (a) C. Hansch, A. Leo, S. H. Unger, K. H. Kim, D. Nikaitani, and E. J. Lien, *J. Med. Chem.*, **16**, 1207 (1973); (b) C. Hansch and M. Yoshimoto, *ibid.*, **17**, 1160 (1974); (c) M. Yoshimoto, C. Hansch, and P. Y. C. Jow, *Chem. Pharm. Bull.*, **23**, 437 (1975).
(23) C. Hansch, A. Leo, and D. Nikaitani, *J. Org. Chem.*, 37, 3090 (1972).
(24) (a) N. R. Draper and H. Smith, "Applied Regression Analysis", Wiley, New York, N.Y., 1966, pp 134–141; (b) ref 9, pp 55, 169, 203.
(25) (a) C. Silipo and C. Hansch, *Mol. Pharmacol.*, **10**, 954 (1974); (b) C. Silipo and C. Hansch, *Farmaco, Ed. Sci.*, **30**, 35 (1975).
(26) I. D. Kuntz, *J. Am. Chem. Soc.*, **94**, 8568 (1972).
(27) B. R. Baker, "Design of Active-Site-Directed Irreversible Enzyme Inhibitors", Wiley, New York, N.Y., 1967.
(28) E. Fölsch and J. R. Bertino, *Mol. Pharmacol.*, **6**, 93 (1970).
(29) C. Hansch and J. M. Clayton, *J. Pharm. Sci.*, **62**, 1 (1973).
(30) R. T. Skeel, W. L. Sawicki, A. R. Cashmore, and J. R. Bertino, *Cancer Res.*, **33**, 2972 (1973).
(31) J. G. Topliss and R. J. Costello, *J. Med. Chem.*, **15**, 1066 (1973).