Inhibition of Dihydrofolate Reductase: Structure-Activity Correlations of Quinazolines Based upon Molecular Shape Analysis

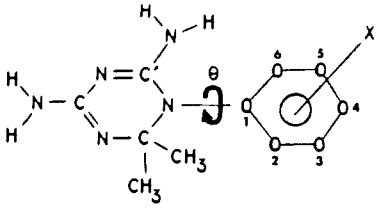
Carol Battershell, D. Malhotra, and A. J. Hopfinger*

Department of Macromolecular Science, Case Institute of Technology, Case Western Reserve University, Cleveland, Ohio 44106. *Received October 20, 1980*

A quantitative structure–activity relationship (QSAR) investigation of a set of 35 substituent-diverse quinazolines spanning an activity range of 4.16 log $(1/I_{50})$ units was carried out using molecular shape analysis (MSA). I_{50} is the molar concentration necessary for 50% inhibition of rat liver dihydrofolate reductase (DHFR). A correlation equation, analogous in descriptor form to those developed for two sets of 2,4-diaminotriazines which are also DHFR inhibitors, was constructed. The correlation coefficient, r, is 0.965, and the standard deviation, s, is 0.360. A second correlation equation was developed to explain the activities of the quinazolines on the basis of their shape similarity to a 2,4-diaminotriazine in its postulated active conformation as determined in a previous study. This correlation equation is again identical in descriptor form with those previously constructed for the triazines and quinazolines; $r = 0.945$ and $s = 0.451$ for this correlation equation. The ability to quantitatively explain activity in a congeneric set of compounds using a structurally diverse reference compound indicates the potential to design new lead compounds using MSA.

A theory of molecular shape analysis (MSA) for QSAR investigations was presented in an earlier paper.[1] The basis for comparing molecular shapes is to determine the common steric overlap volume, V_0, between pairs of molecules as a function of conformation and relative intermolecular geometry. V_0 physically measures how much steric space a pair of molecules share under a prescribed intermolecular relationship. An arbitrarily large set of molecules can be shape-compared to one another by selecting some common reference compound to which all molecules in the set are compared.

A set of Baker triazines (I) selected from a parent data



I

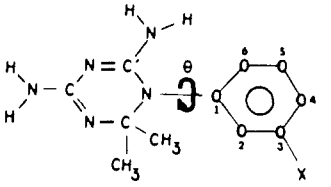
base constructed by Silipo and Hansch[2] was studied using MSA and reported as part of this earlier paper.[1] A relatively simple correlation equation (eq 1) was constructed

$$\log (1/I_{50}) = 1.384S_0 - 0.02127S_0^2 + 0.434\sum\pi - 0.574D_4 - 0.294D_4^2 - 15.66 \quad (1)$$

$n = 27$; $r = 0.953$; $s = 0.440$; ref X = 3,4-Cl_2

which accurately accounts for the activity of representative members of all seven conformational classes of Baker triazines, including compounds neglected in the QSAR analysis by Silipo and Hansch.[2] In eq 1, $S_0 = V_0^{2/3}$, that is, S_0 is power proportional to the common steric overlap volume, V_0, using the X = 3,4-Cl_2 compound as the reference shape molecule. The corresponding "active" conformation of the X = 3,4-Cl_2 compound is shown in a space-filling stereo model in Figure 1. This active conformation, with respect to θ, was deduced by comparing activity to preferred *intramolecular* conformer state.[1] $\sum\pi$ is the sum of water–octanol fragment constants[3] for 3 and 4 substituents. D_4 is needed to account for the low activity of X = 4-C_6H_5 and 4-CN, and it is assumed to measure a hypothetical intermolecular steric interaction involving a 4-X substituent directly along a vector defined by the C_4H_4 bond. I_{50} is the concentration needed to produce a 50% inhibition of dihydrofolate reductase, DHFR, from L1210 leukemia and Walker 256 tumor cells. The range in log $(1/I_{50})$ is 5.11.

In a subsequent MSA-based QSAR[4] of an alternate data base of 2,4-diaminotriazines[5] (II) having substitutions only



II

in the 3 position, correlation eq 2 and 3 were constructed.

$$\log (1/I_{50}) = 1.074S_0 - 0.01674S_0^2 + 0.604\pi_3 - 0.110\pi_3^2 - 0.297D_3 - 10.75 \quad (2)$$

$n = 31$; $r = 0.926$; $s = 0.25$; ref X = $C(CH_3)_3$; I_{50} values measured for bovine liver DHFR

$$\log (1/I_{50}) = 1.086S_0 - 0.01606S_0^2 + 0.723\pi_3 - 0.142\pi_3^2 - 0.535D_3 - 12.17 \quad (3)$$

$n = 20$; $r = 0.947$; $s = 0.23$; ref X = $C(CH_3)_3$; I_{50} values measured for rat liver DHFR

The forms of eq 2 and 3 are remarkably similar to eq 1. S_0 is again $V_0^{2/3}$ using X = $C(CH_3)_3$ as the shape reference compound which has the *identical "active" conformation*, with respect to θ, as the X = 3,4-Cl_2 compound used to develop eq 1. π_3 replaces the $\sum\pi$ in eq 1, since there is no 4 substituent. A parabolic dependence in π_3 is needed in eq 2 and 3 obstensibly because of a larger range in substituent lipophilicity in this data base as compared to that of ref 1. D_3 measures the maximum distance of a nonhydrogen atom of X *perpendicular* to a vector defined by the CH bond at position 3 on the ring. Like D_4, D_3 is assumed to account for an intermolecular steric interaction which diminishes activity as D_3 increases. D_3 neatly accounts for the dramatic loss in activities of X = *cis*-$COOC_2H_5$ and D-$CH(OH)C_6H_5$. Further, D_3 appears to correspond to a DHFR species specificity in that a D_3 term is not needed in eq 1 to account for the activity of

(1) A. J. Hopfinger, *J. Am. Chem. Soc.*, **102**, 7196 (1980).
(2) C. Silipo and C. Hansch, *J. Am. Chem. Soc.*, **97**, 6849 (1975).
(3) C. Hansch and A. Leo, "Substituent Constants for Correlation Analysis in Chemistry and Biology", Wiley-Interscience, New York, 1979.
(4) A. J. Hopfinger, *Arch. Biochem. Biophys.*, **206**, 153 (1981).
(5) S. W. Dietrich, R. N. Smith, J. Y. Fukunaga, M. Olney, and C. Hansch, *Arch. Biochem. Biophys.*, **194**, 600 (1979).

0022-2623/81/1824-0812$01.25/0

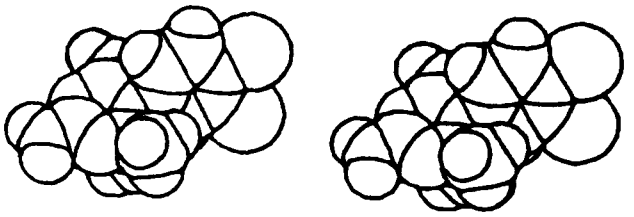


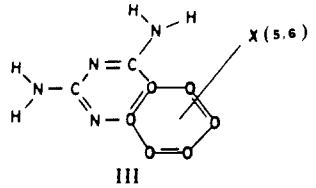
Figure 1. Space-filling stereo model of the 2,4-diaminotriazine having X = 3,4-Cl_2 in the postulated active conformation.

D-$CH(OH)C_6H_5$. All compounds reported by Dietrich et al.,[5] including ones deleted in constructing their QSAR, were used to derive eq 2 and 3.

If it is assumed that the amino group ortho to the substituted benzyl group hydrogen bonds to DHFR at the carbonyl of Ile-5, while the other amino group hydrogen bonds with O_γ of Thr-113, as in the *Escherichia coli* DHFR–methotrexate crystal complex,[6] then eq 1–3 are scalar representations of the three-dimensional binding modes of the respective triazines. Figure 1 correspondingly illustrates, at least parially, the preferred steric shape and lipophilic distribution of inhibitor volume necessary for reasonable inhibition activity. It should be noted that MSA studies will be independent of protonation at N1 which is the form of methotrexate observed in the crystal complex.[6] N1 protonation represents a constant feature over the data base with respect to MSA.

The specific reference shape compounds reported with eq 1–3 each yielded the most significant correlation equation of all compounds in each respective data base. This is the reason they were selected. An analysis of the choice of reference shape compound used in a QSAR is given in ref 1, and some additional discussion is presented under the Methods section of this paper.

The purpose of this paper is to report the results of applying MSA to develop a QSAR in a set of 2,4-diaminoquinazolines (III), which are also DHFR inhibitors.[7]



III

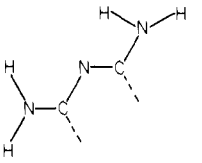
It will be shown that not only are the resultant correlation equations for the quinazolines of the same form and quality as eq 1–3, but also that the X = 3,4-Cl_2 compound of I, *in its postulated active conformation* (see Figure 1), can be used as a reference shape compound to accurately predict the activities of the quinazolines. Thus, QSAR based upon MSA may be useful in developing new lead compounds. This has already been demonstrated, to a limited extent, by predicting the activity of IV, which is noncongeneric to I and II, to be log $(1/I_{50})$ = 6.58 for bovine liver DHFR using eq 2.[4] The *subsequently* measured activity is log $(1/I_{50})$ = 6.73.[4]

Method

(1) **Theory of Shape Analysis.** Three additional features have recently been added to the theory of shape analysis given in ref 1. First, the multiple conformation problem has been considered. That is, how does one select a particular conformation of a molecule which can assume several possible, and probable, intramolecular conformer states? The answer is to include *all* calculated stable conformers in the pairwise shape matrix. Each conformer, therefore, becomes a unique entry in the shape analysis. The specific conformer finally selected is simply the result of optimizing the regression fit between activity and the correlation descriptors which obviously includes V_0 or its proportional powers. The optimization is achieved by considering all unitary combinations of conformers of each of the compounds in the data base. The relative intramolecular energy of each conformer can be included as a second set of features in this optimization process. Thus, it is ultimately the activities that prescribe molecular conformation. As already implicitly stated, this optimization is contained within the optimizations involving choice of shape-reference compound and shape independent molecular descriptors.

V_0 is not exactly calculated because it is based upon computing the overlap volumes of intersecting spheres (van der Waals spherical volumes of atoms). The general solution of the common intersection volume of multiple spheres has apparently not been solved. We have tried to gain some control over this source of error in an empirical way. A rectangular lattice of predescribed size and orientation is constructed around each pair of molecules (i,j) being shape compared. If atom k of i is in a particular lattice cell, then the only atoms of j which are considered to overlap with k are those residing in the same cell. By varying lattice size and relative orientation of the lattice with respect to i,j, it is possible to estimate the relative extent of error in the spherical volume overlaps. Further, within the empirical nature of QSAR analysis, the correlation fit can be optimized as a function of lattice size. The preferred lattice size, in turn, gives some information regarding which corresponding atoms in a set of compounds are most crucial to specifying activity.

Lastly, as part of the general shape analysis package, SHAPE, we have included a method of minimizing the sum of distances between arbitrarily specified atoms in molecules i and j as developed by Pensak.[8] Measures of the minimized sums between molecular pairs (i,j) may, in themselves, be useful QSAR shape descriptors. However, at this time we have used this procedure only to orient the quinazolines relative to the triazines in space by minimizing the sum of distances between the corresponding atoms of the structural fragment common to both classes of compounds.



In doing this we assume that the two NH_2 groups of both types of compounds are involved in hydrogen bonding to DHFR as discussed above for methotrexate.[6]

(2) **Construction of the Data Base.** Fukunaga et al.[7] have put together a data base of 104 quinazolines which inhibit rat liver DHFR. They constructed a QSAR[7] from this data base as a follow-up study to their first triazine QSAR investigation[2] just as we are doing in this paper. The quinazoline data base was developed largely from the extensive work reported by Hynes and Ashton.[9] Thirty-three of the compounds in the data base of Fukunaga et al.[7] do not contain both the 2- and 4-NH_2 groups. These compounds were not included in our QSAR on the basis of what we now know about the DHFR–methotrexate crystal complex and because of their corresponding chemical dissimilarity to the 2,4-diaminotriazines. Those neglected compounds which contain proton donor groups in both the 2 and 4 positions should be represented in a later study where one is not focusing upon interrelating features of 2,4-diaminoquinazolines and 2,4-diaminotriazines.

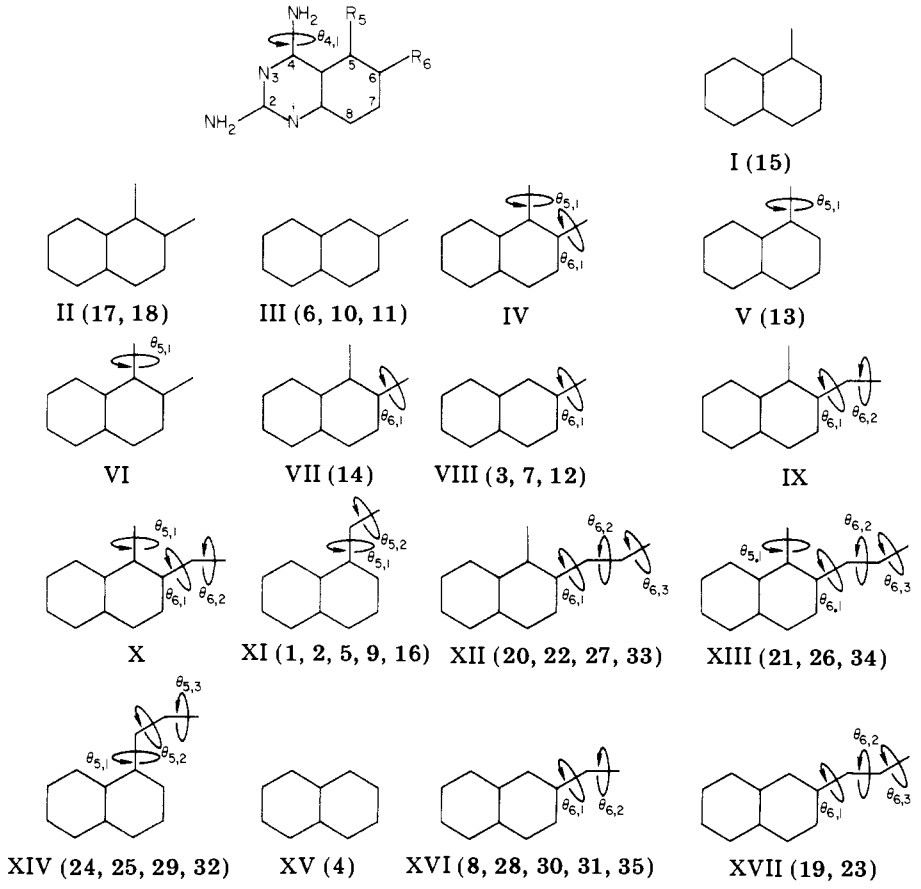
(6) D. A. Matthews, R. A. Alden, J. T. Bolin, D. J. Filman, S. T. Freer, R. Hamlin, W. G. J. Hol, R. L. Kisliuk, E. J. Pastore, L. T. Plante, N.-H. Xuong, and J. Kraut, *J. Biol. Chem.*, **253**, 6946 (1978).

(7) J. Y. Fukunaga, C. Hansch, and E. E. Steller, *J. Med. Chem.*, 19, 605 (1976).

(8) M. Pensak, unpublished software.

(9) (a) W. T. Ashton, F. C. Walker III, and J. B. Hynes, *J. Med. Chem.*, 16, 694 (1973); (b) W. T. Ashton and J. B. Hynes, *ibid.*, 16, 1233 (1973); (c) J. B. Hynes, W. T. Ashton, D. Bryansmith, and J. H. Freisheim, *ibid.*, 17, 1023 (1974); (d) J. B. Hynes and W. T. Ashton, *ibid.*, 18, 263 (1975); (e) J. B. Hynes and C. M. Garrett, *ibid.*, 18, 632 (1975).

Chart I. The 17 Conformational Classes of 2,4-Diaminoquinazolines in the Fukanaga et al. Data Base[7]



Like the triazines,[1] the quinazolines can be broken down into different conformational classes with respect to torsional degrees of freedom. For computational reasons, only compounds having three or fewer degrees of conformational freedom (torsional rotations) in a substituent were considered. This still leads to 17 different conformational classes which are defined in Chart I. The numbers in parentheses next to each conformational classification in Chart I indicate which compounds in the data base reported in Table I fall into that class. A maximum of five compounds are allowed to represent any single conformational classification, such that the five compounds span the broadest range in activity. It can be seen that some conformational classes are not represented, and other underrepresented, in the parent data base.[7]

The assumption made in reducing the number of compounds in a data base on the basis of conformational classification is that if representative members of a class fit the QSAR, any member of the class will fit. This is a reasonable, though unproven, simplification which makes it practical to introduce explicit three-dimensional shape features into a QSAR based upon regression analysis. The 35 compounds used in this investigation are given as part of Table I. They span an activity range of $\Delta[\log (1/I_{50})] = 4.16$.

(3) Conformational Analyses. The CHEMLAB package (part of which was earlier known as CAMSEQ-II)[10] was employed to determine the preferred conformations of III. The valence geometry of III with X = H was arrived at using Allinger's consistent-force-field method as formulated in program MMI.[11,12] This program has been partially parameterized in our laboratory and is a CHEMLAB component. The valence geometry of III with X = H was held fixed for all substitutions on the 5 and 6 positions. However, the 4-NH_2 group was allowed to rotate so as to minimize conformational energy when substituents in the 5 position were considered. The valence geometries of the substituent groups are based upon "standard" bond lengths and angles.[13]

Conformational space, that is, the $\theta_{5,j}$ and $\theta_{6,j}$ as defined in Chart I, were explored at an initial resolution of 30° increments. Each of the minima identified in this scan was then rigorously minimized using the minimization option, MINI, in CHEMLAB.[14] Torsional rotation of the 4-NH_2 group, $\theta_{4,1}$ in Chart I, was included in the scan and minimization procedures when 5-substituents were considered. Only solvent-independent intramolecular conformational energetics have been considered in this work.

(4) Molecular QSAR Descriptors. V_0 and S_0 were independently used as relative measures of molecular shape in formulating the QSAR. In addition, each ΔE_j, the difference in free-space, intramolecular conformational energy between the global energy minimum and the jth relative minimum for each compound, was considered in the regression analyses. The ΔE_js were represented by both linear and Boltzmann exponential terms in separate studies. Also, π_5, π_6, and their sum $(\pi_5 + \pi_6)$, where π_k is the water/octanol fragment constant[3] for a substituent at the kth position, were individually and collectively considered in generating QSAR regression equations. The consideration of the π descriptors follows from the triazine studies where they constitute important components in the QSARs.[1-5] Lastly, the torsional rotation angle, $\theta_{4,1}$, was included as a molecular descriptor, since it can represent changes in the hydrogen-bonding properties between the 4-NH_2 group and the enzyme.

Results

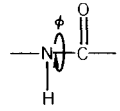
The most significant correlation equation realized using each of the quinazolines as the shape reference compound is eq 4. It is obvious that eq 4 is of the same form and

$$\log (1/I_{50}) = 0.3491S_0 - 0.0021S_0^2 + 0.4870\,(\pi_5 + \pi_6) - 0.0897\Delta\theta - 6.950 \quad (4)$$

$n = 35$; $r = 0.965$; $s = 0.360$; ref X = 6-SO_2-2-$C_{10}H_7$

(10) R. Potenzone, Jr., E. Cavicchi, H. J. R. Weintraub, and A. J. Hopfinger, *Comput. Chem.*, 1, 187 (1977).
(11) N. L. Allinger and J. T. Sprague, *J. Am. Chem. Soc.*, **95**, 3983 (1975).
(12) "Quantum Chemistry Program Exchange", program number 318, Chemistry Department, Indiana University.
(13) A. J. Hopfinger, "Conformational Properties of Macromolecules", Academic Press, New York, 1973.
(14) M. J. D. Powell, *Comp. J.*, **7**, 163 (1963).

Table I. The Set of 2,4-Diaminoquinazolines and the Shape Descriptors V_0 and S_0, Relative to Both the 3,4-Dichlorotriazine and Compound 35, and the Descriptors $\pi_5 + \pi_6$ and θ_4. The Observed, Predicted, and Difference Log $(1/I_{50})$ Values Are Also Reported



		triazine, 3,4-Cl_2		compd 35, 6-SO_2-2-$C_{10}H_7$								
no.	structure[a]	V_0, Å^3	S_0, Å^2	V_0, Å^3	S_0, Å^2	$\pi_5 + \pi_6$	(b) $\Delta\theta$	log $(1/I_{50})$ obsd	log $(1/I_{50})$ pred, eq 4	diff (obsd − pred), eq 4	log $(1/I_{50})$ pred, eq 5	diff (obsd − pred), eq 5
1	5-SO_2-C_6H_3-3,4-Cl_2	387.1	53.12	506.2	63.53	1.12	29.4	4.24	4.56	−0.32	4.52	−0.28
2	5-SO-C_6H_3-3,4-Cl_2	372.6	51.79	490.3	62.18	1.07	30.5	4.26	4.33	−0.07	4.20	0.06
3	6-NH_2	279.0	42.71	405.3	54.78	−1.00	0	4.57	5.31	−0.74	4.75	−0.18
4	5,6-H_2	259.7	40.71	325.8	47.36	0.46	0	4.66	5.04	−0.38	5.12	−0.46
5	5-SO_2-2-$C_{10}H_7$	399.8	54.28	512.6	64.06	1.32	28.9	4.82	4.74	0.08	4.82	−0.00
6	6-CN	284.5	43.26	340.2	48.74	−0.34	0	4.92	4.85	0.07	5.18	−0.26
7	6-CHO	313.6	46.16	369.1	51.47	−0.19	0	5.00	5.30	−0.30	5.76	−0.76
8	6-CH_2NH_2	312.9	46.10	369.2	51.47	−0.71	0	5.03	5.04	−0.01	5.49	−0.46
9	5-*t*-CH=CH-2-$C_{10}H_7$	277.0	42.50	339.4	48.66	4.14	20.4	5.28	5.19	0.09	5.28	−0.00
10	6-Cl	298.9	44.70	388.4	53.23	0.92	0	5.40	6.06	−0.66	6.08	−0.68
11	6-Br	301.7	44.99	392.6	53.62	1.09	0	5.60	6.15	−0.55	6.22	−0.62
12	6-CH_3	291.7	43.98	348.3	49.51	0.78	0	5.66	5.50	0.16	5.88	−0.22
13	5-CH_3	281.6	42.96	342.1	48.92	0.78	0.3	6.09	5.39	0.70	5.67	0.42
14	5-Cl, 6-NH_2	303.6	45.18	424.9	56.52	−0.16	0	6.18	5.91	0.27	5.61	0.57
15	5-Cl	290.5	43.87	388.1	53.22	0.92	1.4	6.25	5.93	0.32	5.80	0.45
16	5-*c*-CH=CH-2-$C_{10}H_7$	278.4	42.64	379.2	52.43	4.14	14.7	5.52	6.21	0.31	5.88	0.64
17	5,6-Cl_2	323.5	47.13	438.0	57.69	1.42	1.4	6.70	6.68	0.02	6.60	0.10
18	5-Cl, 6-Br	328.9	47.65	451.9	58.89	1.57	1.4	7.00	6.88	0.12	6.76	0.24
19	6-S-C_6H_4-3-CF_3	310.2	45.83	445.7	58.36	3.13	0	7.44	7.71	−0.27	7.40	0.04
20	5-CH_3, 6-$NHCOCH_2$-C_6H_4-3-Br	338.1	48.53	478.8	61.20	1.76	0.6	7.47	7.26	0.21	7.65	−0.18
21	5-CH_3, 6-$NHCOCH_2$-C_6H_4-4-Cl	335.9	48.33	471.9	60.62	1.61	0.6	7.48	7.13	0.35	6.97	0.51
22	5-Cl, 6-$NHCOCH_2$-C_6H_3-3,4-Cl_2	336.4	48.37	482.9	61.55	2.53	1.6	7.62	7.57	0.05	7.35	0.27
23	6-$NHCOCH_2$-C_6H_4-4-Cl	363.9	50.98	493.3	62.44	1.33	0	7.64	7.21	0.43	7.29	0.35
24	5-CH_2S-C_6H_4-4-Cl	396.7	54.00	485.8	61.80	3.25	7.9	7.70	7.38	0.32	7.89	−0.19
25	5-CH_2CH_2-2-$C_{10}H_7$	351.2	49.79	440.2	57.89	4.67	5.9	7.70	7.88	−0.18	8.23	−0.53
26	5-CH_3, 6-$NHCOCH_2$-C_6H_3-3,4-Cl_2	334.6	48.20	474.7	60.85	1.49	0.6	7.77	7.10	0.67	6.89	0.88
27	5-Cl, 6-$NHCOCH_2$-C_6H_4-4-Br	349.2	49.59	501.3	63.12	1.98	1.6	7.89	7.44	0.45	7.25	0.64
28	6-S-C_6H_3-3,4-Cl_2	325.8	47.35	459.3	59.53	3.82	0	7.96	8.16	−0.20	8.01	−0.05
29	5-CH_2-S-2-$C_{10}H_7$	402.0	54.48	482.7	61.54	4.05	2.6	8.00	8.22	−0.22	8.90	−0.90
30	5-S-2-$C_{10}H_7$	403.2	54.59	638.4	74.14	3.52	0	8.15	8.97	−0.82	8.90	−0.75
31	6-SO-2-$C_{10}H_7$	409.8	55.17	708.5	79.48	1.37	0	8.15	8.04	0.11	7.87	0.28
32	6-$NHCOCH_2$-2-$C_{10}H_7$	405.5	54.79	650.9	75.11	2.13	0	8.30	8.32	−0.02	8.21	0.09
33	5-Cl, 6-$NHCOCH_2$-2-$C_{10}H_7$	389.5	53.34	644.5	74.63	2.79	1.6	8.30	8.48	−0.18	8.20	0.10
34	5-CH_3, 6-$NHCOCH_2$-2-$C_{10}H_7$	376.3	52.10	640.0	74.28	2.58	0.6	8.38	8.46	−0.08	8.03	0.35
35	6-SO_2-2-$C_{10}H_7$	412.2	54.40	784.2	85.06	1.54	0	8.40	8.12	0.28	7.86	0.54

[a] ϕ (see structure over column heads) is cis planar rigid for all substituents containing an –NHCO– group. [b] Closest difference to 0° or 90°. This is the rotation of the 4-NH_2 to minimize conformational energy. $\theta_{4,1} = 0°$ corresponds to the plane of the NH_2 being in same plane as the quinazoline ring.

Table II. Values of V_0 and S_0 for Cis and Trans X = 5-Cl, 6-$NHCOCH_2$-C_6H_3-3,4-Cl_2 As Well As the Predicted Activities Using Equations 4 and 5 and the Intramolecular Conformational Energies[a]

	eq 4			eq 5			
isomer	V_0, Å^3	S_0, Å^2	log $(1/I_{50})$	V_0, Å^3	S_0, Å^2	log $(1/I_{50})$	energy, kcal/mol
trans	606.0	71.62	8.37	399.6	54.25	8.32	−35.1
cis	482.9	61.55	7.57	336.4	48.37	7.34	−42.6

[a] Log $(1/I_{50})_{obsd}$ = 7.62.

quality as eq 1–3. The preferred shape descriptor is S_0 occurring in parabolic form as in eq 1–3, and $\pi_5 + \pi_6$ measures substituent lipohilicity and is analogous to $\sum\pi$ in eq 1 and π_3 in eq 2 and 3. $\Delta\theta$ is the smaller value of $|\theta_{4,1}|$ or $|90°-\theta_{4,1}|$ where, as a remider, $\theta_{4,1}$ is the torsional rotation of the 4-NH_2 group, such that $\theta_{4,1} = 0°$ corresponds to an NH_2 group being in the same "plane" as the quinazoline ring. The shape reference compound used to construct eq 4 is the most active compound in the data base as was the X = 3,4-Cl_2 compound used to construct eq 1. This is not evidence that the most active compound will be the preferred shape reference compound. Equations 2 and 3 are examples where a compound of average activity is the preferred shape reference compound. Further, several other quinazolines, when used as shape reference compounds, yield r values greater than 0.940.

The predicted activities and their differences with the observed activities using eq 4 are reported in Table I. The V_0 and S_0, as well as $\pi_5 + \pi_6$ and $\Delta\theta$, are also listed Table I. It is to be noted that compound **9**, X = 5-*trans*-CH=CH-2-$C_{10}H_7$, was an outlier in the study by Fukunaga et al.[7] and not included in construction of their correlation equation. It can be seen from Table I that the difference in the predicted and observed activity for this compound using eq 4 is only 0.09.

Perhaps the most far reaching result of this study is that correlation eq 5 could be generated for the set of quinazolines in Table I using the X = 3,4-dichlorotriazine, shown in Figure 1, as the shape reference compound. Equation 5 is again identical in form and quality with the previous four correlation equations. It is quite significant, however, that a structurally noncongeneric compound can be used to accurately explain activity in a different set of congeneric compounds on the basis of relative shape. This raises the prospect of designing new lead compounds using correlation equations based upon MSA.

$$\log (1/I_{50}) = 0.3735S_0 - 0.0022S_0^2 + 0.5115(\pi_5 + \pi_6) - 0.1008\Delta\theta - 6.596 \quad (5)$$

n = 35; r = 0.945; s = 0.451; ref X = 3,4-dichloro-2,4-diaminotriazine

The predicted activites and differences in predicted and observed activites using eq 5 are also reported in Table I. A qualitative understanding of the shape similarity between the X = 3,4-dichloro-2,4-diaminotriazine and the X = 6-sulfonyl-2-naphthylquinazoline can be inferred by visually comparing their space-filling molecular representations to the least-active quinazoline in our data base, X = 5-SO_2-C_6H_3-3,4-Cl_2. Figures 2 and 3 show stereo models of the most and least active quinazolines, respectively. It is clear that the -C_6H_3-3,4-Cl_2 group of the triazine intersects the 6-SO_2-2-$C_{10}H_7$ structure to a much larger extent than it does the 5-SO_2-C_6H_3-3,4-Cl_2 substituent. However, no visual representation, including the so-called receptor-site mapping methods, quantifies the shape similarity relationship *and* also allows shape to be treated on an equal basis to other physicochemical descriptors.

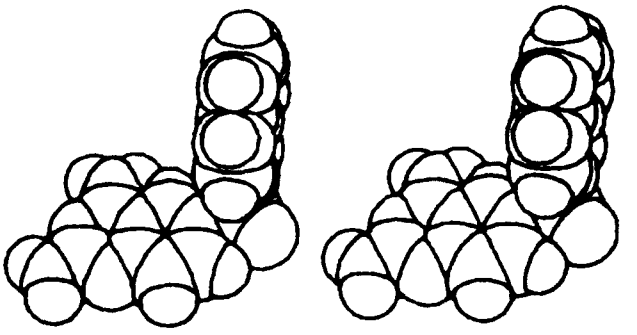


Figure 2. Compound **35** in Chart I, the most active 2,4-diaminoquinazoline in the conformation which best fits the activity scale. The external view is the same as in Figure 1.

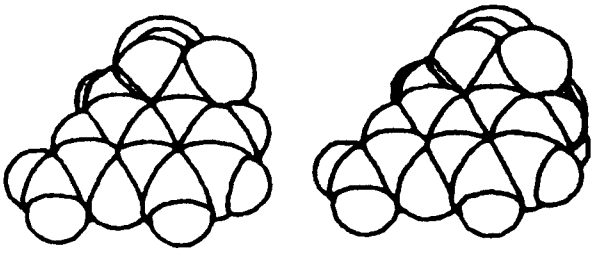
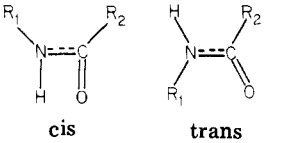


Figure 3. Compound **1** in Table I, the least active 2,4-diaminoquinazoline in the conformation which best fits the activity scale. The external view is the same as in Figures 1 and 2. It can be seen that the 4-NH_2 group is rotated out of the plane of the quinazoline ring.

The regression fit, to some extent, dictates conformer choice for compounds that can adopt multiple stable intramolecular conformations. Surprisingly, values of V_0 in this data base were not too sensitive to conformer state. One particular case to the contrary, however, involves compounds containing an amide group. The problem arises as to whether the groups on either side of the amide bond are cis or trans planar to one another,



owing to the partial double-bond character of the amide bond. It was found that only the cis isomers yielded V_0 values which led to a meaningful QSAR. Quite nicely the intramolecular conformational energies of the cis isomers are 6, or more, kcal/mol more stable than the trans isomers. This is illustrated in Table II for X = 5-Cl 6-$NHCOCH_2$-C_6H_3-3,4-Cl_2.

Lastly, the structure–activity correlation equations are optimized for lattice cell lengths of 5 Å, and the optimizations are independent of the length of the lattice cell used to compute V_0 for lengths greater than 5 Å. We believe this indicates that the overestimation of V_0, in terms of the sum of pairs of intersecting van der Waals spheres, is relatively uniform over the entire data base and,

Table III. Observed vs. Predicted Activities of Four Compounds Not Used in Developing Equations 4 and 5

compd	log (1/I_{50})$_{obsd}$	$\pi_5 + \pi_6$	$\Delta\theta$	eq 4			eq 5		
				V_0, Å^3	S_0, Å^2	log (1/I_{50})	V_0, Å^3	S_0, Å^2	log (1/I_{50})
6-$CH_2NHCOCH_2C_6H_3$-3,4-Cl_2	5.82	1.15	0	363.9	50.98	5.95	300.5	44.87	6.28
5-S-2-C_6H_3-3,4-Cl_2	6.70	3.92	2.0	349.6	49.63	5.98	271.3	41.92	6.93
6-$NHCH_2$-C_6H_3-3,4-Cl_2	7.74	2.44	0	478.8	61.32	7.75	321.9	46.98	7.33
5-Cl, 6-$NHCOCH_2$-C_6H_4-3-Br	8.12	1.98	1.6	539.6	66.29	7.78	378.8	52.36	7.78

consequently, does not significantly effect the results.

Discussion

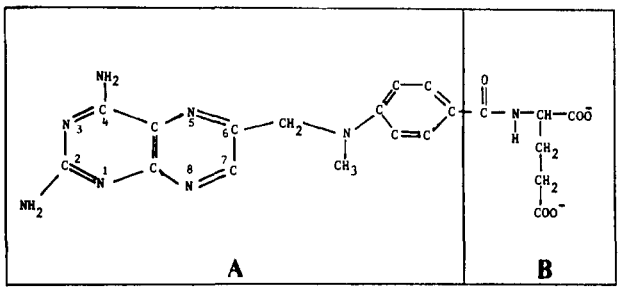
Equations 4 and 5 were tested for reliability by using them to predict the activity of four compounds not used in their derivations but given in the parent data base of Fukunaga et al.[7] The results are presented in Table III. Compound X = 6-$CH_2NHCOCH_2$-C_6H_3-3,4-Cl_2 was an outlier in the QSAR developed by Fugunaga et al.[7] Therefore, we felt it important to see if it fit our QSAR even though there are 4 torsional degrees of freedom in the substituent. From Table III it is seen that this compound, as well as the other three, are meaningfully explained by eq 4 and 5.

The descriptor $\Delta\theta$ deserves some discussion. Unlike D_4 and D_3 of eq 1 and 2, respectively, $\Delta\theta$ is necessary because of intramolecular, as opposed to intermolecular, interactions. Substitutions in the 5 position of the ring sterically force the 4-NH_2 from its preferred conformation planar to the ring. This intramolecular interaction probably modifies the intermolecular hydrogen bonding capacity of the 4-NH_2 group with the enzyme. From the way in which $\Delta\theta$ had to be constructed to yield the QSAR, we are forced to conclude that optimum intermolecular hydrogen bonding can be realized when the 4-NH_2 is either parallel or perpendicular to the ring. The only intermolecular hydrogen-bond geometry consistent with this conclusion requires two separate acceptor sites, which are in such close proximity that if C–O bonds are involved they are directed normal to each other.

Some perspectives should be put upon the molecular shape descriptors with regard to physical meaning and predictive potential for specifying new DHFR inhibitors. Firstly, each of the structure–activity correlation equations can be optimized for a particular value of S_0(opt): eq 1, 32.4; eq 2, 32.1; eq 3, 33.9; eq 4, 83.1; eq 5, 80.4. The first three values differ markedly from the last two because the overlap volume of the 2,4-diaminodimethylated ring with itself was not included in the calculation of V_0 (S_0) for the triazines. When this overlap volume (S_0 = 47.2) is added as a constant to the S_0(opt) of eq 1–3, the "normalized" S_0(opt) values are 79.6, 79.3, and 81.2, respectively. Thus, the optimum values of S_0 in all five equations are amazingly similar, varying only from 79.3 to 83.1. But what does it mean? Our interpretation is that the combined restrictions in shape flexibility of the triazines and quinazolines, along with the geometry of the active site, require the optimum inhibitor in both series to occupy about 710 Å^3.

The substituent groups selected to comprise part of this prescribed volume should be hydrophobic. D_3 and D_4 represent the presence of "receptor walls" in close proximity to the triazines bound to the enzyme in the "active" conformation illustrated for the 3,4-Cl_2 compound in Figure 1. The importance of intermolecular hydrogen bonding between the 4-NH_2 and the enzyme upon activity is reflected by $\Delta\theta$.

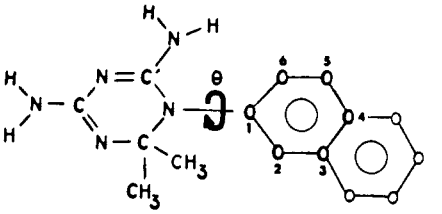
The design of more active in vitro inhibitors of DHFR must be thought of in terms of a two-site binding model.



V

Figure 4. The structure of methotrexate with the A and B DHFR binding sites defined.

This is perhaps illustrated best by looking at the methotrexate molecule (V) shown in Figure 4. The part of the molecule labeled A corresponds to the sole binding site present in both the triazines and quinazolines. Most of the compounds in our triazine and quinazoline data bases do not contain substituents "long" enough to mimic the second binding site of methotrexate as denoted by B in Figure 4. Those substituents which are long enough do not contain realistic end groups to strongly interact with the enzyme. Thus, eq 1–5 are prescriptions for optimizing binding to DHFR at one site. In this light, structure IV,



IV

which already has been predicted and tested, represents a good new lead. Its bovine liver DHFR activity [log (1/I_{50}) = 6.73] is greater than both the corresponding unsubstituted 2,4-diaminotriazine [log (1/I_{50}) = 4.66] and unsubstituted quinazoline [log (1/I_{50}) = 6.33]. This suggests that IV is intrinsically more active than both the triazines and quinazolines.

It is also to be noted that the respective "active" conformers of the triazine, quinazoline, and pyrimidine (discussed in the following paper in this issue) DHFR inhibitors were fit into the DHFR–methotrexate crystal structure.[6] The obvious assumption made was that each of these inhibitors replaces methotrexate. All the compounds fit with space left over. That is, as mentioned above, the compounds in our data base are generally smaller than methotrexate, and it is difficult to draw conclusions regarding binding or fit. Moreover, the uncertainty in the coordinates of the crystal structure, conformational flexibility of the active site, and enzyme species specificity with respect to amino acid sequence limit the reliable information that can be gleaned from such studies.

Nevertheless, the next logical step in our studies would be to cheat a little and use the methotrexate structure

bound to DHFR[6] as the shape reference compound for designing substituents on IV when in its active conformation. Further, some of the compounds in the original triazine data base[2] might be employed in such a study for purposes of evaluating polar/nonpolar binding constraints imposed on possible substituents by the second binding site. Such an investigation is now under consideration.

Acknowledgment. The theory and corresponding computer programs to calculate molecular shape descriptors have been developed under private funding. The application of MSA to the quinazolines was funded under a contract from the National Cancer Institute (Contract NO1-CP-75927) and a grant from the National Science Foundation (Grant ENV77-24061). The authors appreciate helpful discussions with Dr. J. Y. Fukunaga of Schering-Plough Corp. and R. Pearlstein and S. K. Tripathy of our laboratory during the course of this study.

Inhibition of Dihydrofolate Reductase: Structure–Activity Correlations of 2,4-Diamino-5-benzylpyrimidines Based upon Molecular Shape Analysis

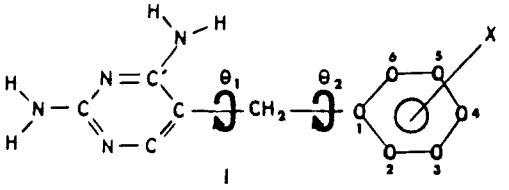
A. J. Hopfinger

Department of Macromolecular Science, Case Institute of Technology, Case Western Reserve University, Cleveland, Ohio 44106. *Received October 20, 1980*

A quantitative structure–activity rleationship (QSAR) investigation of a set of 23 substituted 2,4-diamino-5-benzylpyrimidines spanning an activity range of 1.8 log $(1/C)$ units was carried out using molecular shape analysis (MSA). C is the molar concentration necessary for 50% inhibition of bovine liver dihydrofolate reductase (DHFR). The "active" shape of these compounds was deduced by comparing the change in conformational state to the activity of four compounds outside the data base described above. A correlation equation, analogous in descriptor form to those developed earlier for DHFR inhibition by substituted 2,4-diaminotriazines and -quinazolines, was constructed. The correlation coefficient, r, was 0.931 and the standard deviation of fit, s, was 0.137. The results suggest that these pyrimidines bind to DHFR with shape features different from both the triazines and quinazolines. It is postulated from the "active" shape of the pyrimidines that it is preferable to substitute at the meta position of the benzyl ring rather than at the para position.

In the preceding paper of this issue we developed a QSAR to describe the dihydrofolate reductase (DHFR) inhibition activity of a set of 2,4-diaminoquinazolines based upon molecular shape analysis (MSA).[1] Equivalent QSARs, based upon MSA, were also constructed for sets of 2,4-diaminotriazines which inhibit DHFR.[2,3] A methodology breakthrough in the potential design of new lead structures was realized in the investigation of the quinazolines. It was shown that the activities of these compounds could be explained by their shape similarity to a 2,4-diaminotriazine in its postulated active conformation.[1]

In this paper we report the results of a QSAR investigation, based upon MSA, of yet another set of DHFR inhibitors, substituted 5-benzyl-2,4-diaminopyrimidines (I). The importance of trimethoprim [2,4-diamino-5-



I

(3,4,5-trimethoxybenzyl)pyrimidine] as a broad-spectrum antibacterial agent[4-7] has spurred an interest in this class of compounds.

Method

(1) The Data Base. Blaney et al.[8] have put together a set of 23 5-(X-benzyl)-2,4-diaminopyrimidines whose inhibitory effect on bovine liver DHFR has been measured. The activity is expressed as log $(1/C)$, where C is the molar concentration of inhibitor which produces 50% inhibition. The compounds and corresponding activities are given as part of Table I. Blaney et al.[8] formulated correlation eq 1 to explain the DHFR inhibition

$$\log (1/C) = 0.622\pi_3 + 0.322\sum\sigma^+ + 4.99 \quad (1)$$

$$n = 23;\ r = 0.931;\ s = 0.146$$

activity. In eq 1, π_3 is the hydrophobic constant for substituents in position 3, and $\sum\sigma^+$ represents the summed electronic effect of 3, 4, and 5 substituents on position 1. $\sum\sigma^+$ is a somewhat surprising feature to find related to activity, especially since it does not appear in any other QSARs for DHFR inhibition developed by Hansch and co-workers for the triazines[9,10] and quinazolines.[11] However, the authors do point out that the $\sum\sigma^+$ term in eq 1 does not play too significant a role.[8]

(2) Conformational Analysis. A fixed valence geometry intramolecular conformational analysis of I was carried out using CHEMLAB (part of which was formerly known as CAMSEQ-II)[12] components. The bond rotation angles were explored at 30° resolution in the initial scan. The identified minima were then rigorously energy minimized.

(3) Molecular Shape Analysis. An "active" conformation of I was deduced with respect to the principle torsional degress of freedom θ_1 and θ_2. The manner of doing this is presented under Results. Each of the compounds in Table I can adopt the active

(1) C. Battershell, D. Malhotra, and A. J. Hopfinger, *J. Med. Chem.*, preceding paper in this issue.
(2) A. J. Hopfinger, *J. Am. Chem. Soc.*, **102**, 7196 (1980).
(3) A. J. Hopfinger, *Arch. Biochem. Biophys.*, **206**, 153 (1981).
(4) S. R. M. Bushby and G. H. Hitchings, *Dr. J. Pharmacol. Chemother.*, **33**, 72 (1968).
(5) J. J. Burchall and G. H. Hitchings, *Mol. Pharmacol.*, 1, 126 (1965).
(6) *Drugs*, 1, 7 (1971), and references therein.
(7) *J. Infect. Dis.*, **128** (Suppl) (1973).
(8) J. M. Blaney, S. W. Dietrich, M. A. Reynolds, and C. Hansch, *J. Med. Chem.*, **22**, 614 (1979).
(9) C. Silipo and C. Hansch, *J. Am. Chem. Soc.*, **97**, 6849 (1975).
(10) S. W. Dietrich, R. N. Smith, J. Y. Fukunaga, M. Olney, and C. Hansch, *Arch. Biochem. Biophys.*, **194**, 600 (1979).
(11) J. Y. Fukunaga, C. Hansch, and E. E. Steller, *J. Med. Chem.*, **19**, 605 (1976).
(12) R. Potenzone, Jr., E. Cavicchi, H. J. R. Weintraub, and A. J. Hopfinger, *Comput. Chem.*, 1, 187 (1977).

0022-2623/81/1824-0818$01.25/0