Modeling the Interactionof Small Organic Molecules with Biomacromolecules. IV. The In Vivo Interaction of 2- and 6-Substituted Purines with Murine Solid Tumor Adenocarcinoma CA

ARTICLE *in* EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY · APRIL 1991

Impact Factor: 3.45 · DOI: 10.1016/0223-5234(91)90063-S

5 AUTHORS, INCLUDING:

Danail Bonchev
Virginia Commonwealth University
263 PUBLICATIONS **4,230** CITATIONS
SEE PROFILE

Ivan Petkov Bangov
Konstantin Preslavsky University of Shumen
44 PUBLICATIONS **202** CITATIONS
SEE PROFILE

Eur J Med Chem (1991) 26, 305–312

Modelling the interaction of small organic molecules with biomacromolecules IV. The *in vivo* interaction of substituted purines with murine tumor adenocarcinoma CA 755

O Mekenyan[1], D Bonchev[1], DH Rouvray[2], D Peitchev[1], I Bangov[3]

[1]*Higher School of Chemical Technology, 8010 Burgas, Bulgaria;*
[2]*Department of Chemistry, University of Georgia, Athens, Georgia 30602, USA;*
[3]*Institute of Organic Chemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria*

(Received 27 April 1988; accepted 20 November 1989)

Summary — Following the approach adopted in earlier studies in the series, this work undertakes a QSAR investigation of the *in vivo* interaction of 2- and 6-substituted purine derivatives with murine solid tumour adenocarcinoma CA 755. The interaction is analyzed in terms of hydrophobic, electronic (orbital and electrostatic), and geometric (topological and steric) contributions. Exhaustive correlations with the bioactivity of a large number of different types of indices representing these various contributions have been made. It was concluded that the significant indices here are the electronic superdelocalizallity indices S^E_6, S^E_{10} and S^N_{10}, and the hydrophobic index π (6). Appropriate regression equations are presented, and these support the contention that our methodology is able to satisfactorily model biological interactions produced by sets of structurally related molecules.

Résumé — Modélisation de l'interaction des petites molécules organiques avec des biomacromolécules. IV. L'interaction *in vivo* de purines substituées avec la tumeur murine adénocarcinome CA 755. *Selon l'approche adoptée dans les études précédentes de la série, ce travail concerne une investigation QSAR de l'interaction* in vivo *des dérivés puriques 2- et 6-substitués avec la tumeur murine solide adénocarcinome CA 755. L'interaction est analysée en termes de contributions hydrophobes, électroniques (orbitalaire et électrostatique) et géométriques (topologique et stérique). Des corrélations exhaustives avec la bioactivité d'un grand nombre de types différents d'indices représentant les diverses contributions ont été établies. On peut conclure que les indices significatifs sont ici les indices S^E_6, S^E_{10} et S^N_{10}, de superdélocalisation électronique et l'indice hydrophobe π (6). Des équations de régression appropriées sont présentées, apportant la preuve que la méthodologie employée est apte à modéliser valablement les interactions produites par des séries de molécules structuralement apparentées .*

QSAR / purines / antitumor potency / OASIS approach

Introduction

In an endeavour to shed new light on the inner workings of biological responses, an innovative approach to the modelling of bioactivity has been outlined in previous parts of this series [1–3]. Our main focus of interest continues to be the bioactivity elicited when small organic molecules interact with biomacromolecules. The approach adopted so far has been successfully tried out on two specific classes of biologically active molecules, namely the substituted pyridines [1, 2], benzoates, phenols and amphetamines [3–6]. Here we shall extend the method to a QSAR study of substituted purine molecules interacting with murine solid tumour adenocarcinoma CA 755. The formalism employed is similar to that of Hansch [7, 8], and may in fact be viewed as a generalization of his methodology. The fundamental postulate upon which our model is based is that the net effect of any substituent on the (equilibrium) rate constant for a given biological interaction can always be factored into contributing effects arising from the hydrophobic, electronic, and geometric features of the parent compound. In the present context, both the topological and steric aspects of molecular structure will be grouped under the term geometric. To obtain estimates of the hydrophobic, electronic, and geometric contributions to blointeractions, Hansch [7, 8] made use of three indices, the *n*-octanol/water partition coefficient, P, the σ-constant of Hammett; and the steric parameter, E_s, of Taft respectively. By generalizing this approach, it is our objective to provide some new insights into the nature of biological interactions.

The novelty of our approach lies in the use of comprehensive statistical analyses to ascertain the relative significance of each of the known contributing factors to a given interaction. At the outset, individual assessments were made of the hydrophobic, electronic (including orbital and electrostatic components), and geometric (including topological and steric components) contributions. Those indices that are linearly independent are then determined by cluster analysis constructing an intercorrelation matrix for all indices found to correlate well with the bioactivity. This was accomplished by correlating the bioactivity against a variety of indices characterizing each of these contributions. Our mathematical model of the biological interaction under consideration is then constructed using the linearly independent set of indices characterizing all of the hydrophobic, electronic, and geometric component contributions. Since our method is, in essence, an optimized approach based on a structural indices set, it was designated for brevity by the acronym OASIS.

The novelty of our OASIS approach compared to that of Hansch consists of the following features:

1) The reaction sites are assumed to be located within the 'lead' molecule but not within its substituents. The site descriptors are determined by taking into account the molecular structure as a whole, analogously to the case of the global descriptors used for portraying nonspecific biological interactions.

2) Instead of including three specific structural parameters (one steric, one hydrophobic, and one electronic), our mathematical model of biological interaction rests upon an optimized procedure which selects the best structural indices within the groups of all known indices of this type. In addition, a fourth group, namely the group of topological indices, is also taken into account.

3) The early elimination of parameters that correlate poorly diminishes the possibility of chance correlations occurring. According to Topliss and Edwards [9] 'a good approach in correlation studies where a large number of potential variables could be considered would be to initially select for the correlation study, where possible, a limited group of preferred variables. Any correlation which emerged would then be unlikely to be clouded by chance factors'.

4) The large variety of structural parameters considered makes feasible a fuller elucidation than hitherto of the details of the underlying mechanism of the biological interaction. One index of particular importance in this regard appears to be the superdelocalizability of the electron donor or acceptor. This parameter identifies and locates any possible charge occurring between the interacting species.

5) The influence of a given substituent on the activity of the parent molecule is estimated directly in terms of the various structural characteristics of the substituted molecule, rather than indirectly by analogy with the substituent effects in a quite different class of compounds (benzoic acids).

Our method is applied in this work to the interaction of 2-and 6-substituted purines with murine solid tumor adenocarcinoma CA 755. Modelling of this particular biological system is of considerable importance, for it has been established [10, 11] that a number of purines are impressively carcinostatic and sometimes even carcinolytic. The theoretical prediction of their bioactivity by means of relatively sophisticated mathematical modelling would thus seem to be fully warranted. The model presented here, which is founded on an exhaustive statistical interpretation of the biological response of interest, may perhaps best be viewed as a valuable new weapon in the armory of practical pharmacology. In the past, quantitative structure-activity relationships (QSAR) involving the purines [12, 13] have employed as substituent constants [14] the hydrophobic parameter π; the Swain-Lupton field and resonance constants F and R as the electronic parameters; and the molar refractivity (MR) as a measure of the geometrical properties. The use of such a restricted set of parameters in the modelling procedure, however, results in models that are able to account for no more than a small fraction of the variance on the biological data [13]. Here we show that it is necessary to consider additional component terms for the hydrophobic, electronic and geometric contributions to account for the observed variance.

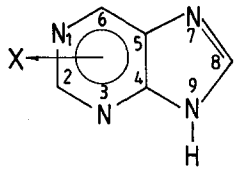
A test set of eighteen 2- and 6-mono- and disubstituted purines, listed in table I, was selected as the class of compounds for our study. This set is derived from table II in the work of Neiman and Quinn [13], although five of the derivatives listed by them have not been included here. We have omitted the iodo derivatives, as standard CNDO-type calculations have not yet been performed on these owing to a lack of reliable CNDO parameters for iodine. Moreover, for the two larger purine derivatives, no attempt has yet been made to perform the relevant CNDO calculations. The biological activities of the substituted purine compounds we consider are further presented. The activity is given in terms of log (1/C), where C is the concentration in mol·kg^{-1} which produces a tumor mass regression of 80%. Since there is no definitive measure available for the biological activity of the unsubstituted purine molecule, this particular molecule was not included in our test set. It was pointed out by Neiman and Quinn [13] that the parent purine molecule appears to be inactive.

The geometric descriptors

In accordance with our previously elaborated methodology [1–3], we now investigate in turn the various

hydrophobic, electronic, and geometric descriptors which have an influence on purine CA 755 activity. To quantify each of the component factors, use is made initially of a set of topological indices [15–28]. The indices selected for this purpose were: (i) the Zagreb group indices, M_1 and M_2; (ii) the polarity number, N_2; (iii) the Hosoya index, Z; (iv) the Hosoya information-theoretical analogue, I_z; (v) the Randic molecular connectivity index, ^1x; (vi) the Randic information-theoretical analogue, I_X; (vii) the Wiener index, W; (viii) the information-theoretical index based on distribution of the distance matrix entries; I^W_D; and (ix) the Balaban and Motoc mean square distance index, $D^{(2)}$. These indices were selected because they have been most widely employed to-date in QSAR correlations [15–28].

Table I. A listing of the test set of purine derivatives examined.



No	*Purine derivative*	*Substituent X*
1	Purine	H
2	6-Chloro	6-Cl
3	6-Bromo	6-Br
4	6-Methoxy	6-OCH_3
5	6-Propoxy	6-$OCH_2CH_2CH_3$
6	6-Hydrazino	6-NH-NH_2
7	6-Methylthio	6-SCH_3
8	6-Ethylthio	6-SCH_2CH_3
9	6-Methylsulfonyl	6-SO_2CH_3
10	6-Sulfonamide	6-SO_2NH_2
11	2-Methyl-6-amino	2-CH_3, 6-NH_2
12	2-Chloro-6-methylamino	2-Cl, 6-$NHCH_3$
13	2-Bromo-6-amino	2-Br, 6-NH_2
14	2-Amino-6-chloro	2-NH_2, 6-Cl
15	2-Amino-6-bromo	2-NH_2, 6-Br
16	2-Amino-6-methylsulfonyl	2-NH_2, 6-SO_2CH_3
17	2,6-Bishydrazino	2,6-NH-NH_2
18	2-Fluorosulfonyl-6-chloro	2-SO_2F, 6-Cl

The steric aspects of the biological activity of the purines were studied by making use of the so-called 'sterimol' parameters of Verloop *et al* [29]. A quantitative characterization was attempted using the indices $L^{(2)}$, $B^{(2)}_1 - B^{(2)}_4$ and $L^{(6)}$, $B^{(6)}_1 - B^{(6)}_4$ for the directional orientation of each substituent at the purine ring positions 2 and 6 (table I). Neiman and Quinn [13] regard two positions as possible reaction centers. Use was also made of the steric indices *GW, GIW, GIE*, and L_{max}. The first three of these were introduced in part I of the present series [1] and represent, respectively, the uniformation analogues of the topological indices W, I^W_D and I^E_D, while the index L_{max} represents the maximum geometric distance in the molecule. These geometrical indices may be calculated by either including or neglecting the hydrogen atoms present in the chemical graph of the molecule.

In table II correlation coefficients for the best correlating indices against biological activity are presented for both first and second order polynomials, *ie* for linear and parabolic correlations. Both the topological and steric indices correlate poorly with the CA 755 biological activity of the substituted purines. In the case of the topological indices, no correlation coefficient exceeded 0.348 whereas the only steric index that shows a correlation coefficient higher than 0.7 is $B^{(6)}_1$. These results indicate that the biological interaction will not be affected by steric hindrances. One might therefore expect the purine molecules to fit into the receptor site with the five-membered ring at the bottom and the six-membered ring with its various substituents at the top, as illustrated in figure 1. With this alignment at the receptor, the substituents will have comparatively little impact on the docking process itself.

The hydrophobic and electronic descriptors

Correlations between the CA 755 activity of the purines and the hydrophobic index, π, were next

Table II. Correlation coefficients for the bioactivity of the purine species in table I against various hydrophobic, steric, and electronic indices.

Structural index	π (6)	B^6_1	q_6	S^E_6	S^E_8	S^E_{10}	S^N_1	S^N_8	S^N_{10}
Linear correlation	0.407	0.683	0.593	0.830	0.408	0.702	0.300	0.277	0.225
Parabolic correlation	0.725	0.716	0.728	0.831	0.716	0.798	0.756	0.716	0.865

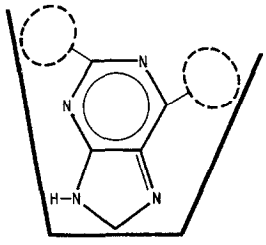


Fig 1. A possible disposition for purine derivatives in the receptor cavity.

established for both entire molecules and position 6 (table I) in the molecules. Because of the small number (10) of 2-substituted purine derivatives, no attempt was made to assess the role of the hydrophobic factor at position 2 (table I) of the purines. The values of the π-index were abstracted from the work of Neiman and Quinn [13]. They pointed out that although their π values were only obtained for aromatic systems, namely for benzene derivatives, the values might also be appropriate for investigation of the hydrophobic behaviour of purine species. The correlations for π (6) found for the purines are presented in table II. From these correlations it may be concluded that the hydrophobic characteristics pertaining to position 6 on the purines are of importance for the biological interaction. The correlation coefficient for the parabolic equation is 0.725 and that for the third order polynomial is 0.818.

The influence of the electronic structure of the purine species on their CA 755 activity was studied using the polyelectronic perturbation theory of Klopman and Hudson [30–32]. This theory postulates that the electrostatic and orbital components of the drug-receptor interaction can be characterized in terms of the net ($\sigma + \pi$) charge and the superdelocalizability of the molecule respectively. These indices are calculated by the standard CNDO procedure assuming that the molecules possess idealized geometry [33]. All positions on the purine rings were investigated as possible reaction centres. The α-atom of the substituents located at position 6 has also been included here and is numbered as position 10. Other possible reaction sites, which could be obtained from a superpositioning of the various purines, were not examined since the small correlation sample sizes would have led to statistically invalid results. The best correlation coefficients between log ($1/C$) and both the charges, q_1, and the donor and acceptor superdelocalizabilities, S^{E}_{1} and S^{N}_{1}, for several positions on the purine ring are also listed in table II.

The correlations presented, apart from that for q_6, indicate that none of the remaining changes at any ring position correlate well with the biological activity index, from which we may conclude that electrostatic interactions are of lesser importance in modelling this type of biological response. It is apparent that only the S^{E} indices at positions 6 and 10, as well as the S^{N} index at position 10, display satisfactory quadratic correlations with biological activity while for linear correlations this only holds true for S^{E}_{6}. Correlations between the biological activity and a number of global electronic indices were also examined. In addition to using quantum-chemical parameters, such as the total electronic energy, E, the energies of the frontier orbitals, E_{HOMO} and E_{LUMO}, and their difference, E_{HL}, we have also included the valence connectivity index, X^{v}, of Kier and Hall [23], the electropy index, ε, of I'Haya *et al* [34], and the molar refractivity of position 6 of the purine ring, MR [6]. Electronic indices characterizing the total electronic structure of the purines were found not to correlate significantly with their biological activity.

The results obtained provide support for the formulation of a preliminary hypothesis concerning the mechanism of the interaction. Based on our mathematical model [1], it is proposed that the reaction determining the biological activity of the purines is due to complex change transfers occurring at two points: from the biomacromolecule to the purine derivatives at position 10, and from position 6 of the purine to the biomacromolecule. These charge transfers thus take place in opposite directions, and the transfer from position 10 to the biomacromolecule is thought to be the less likely of the two to occur. Since these interactions would be orbitally controlled, any chemical bonds which are formed will possess covalent character.

Intercorrelation of hydrophobic and electronic descriptors

The next stage in our procedure is intercorrelation between the hydrophobic and electronic indices which have been found to correlate best ($r > 0.7$, table II) with the biological activity under consideration. The indices in question are: the hydrophobic index, π (6), the steric index B^{6}_{1}, and best electronic indices q_6, S^{E}_{6}, S^{E}_{10}, and S^{N}_{10}. The coefficients obtained for linear correlations of these indices are presented in table III. As seen from this table, the S^{E}_{6} electronic index best correlating with biological activity does not correlate with the hydrophobicity index π (6) while it correlates well with the significant steric index $B^{(6)}_{1}$. On the other hand, the remaining significant electronic indices (q_6, S^{E}_{10}, S^{N}_{10}) correlated well either with S^{E}_{6} or with π (6). In accordance with this finding, table IV reveals that the best models incorporate pairs of only poorly intercorrelating parameters. However, to avoid any possibility of excluding any significant factor by statistical artifact, in constructing our mathematical models of the biological interaction, we concluded our

procedure by examining all possible combinations of the six variables. The results of the best statistical correlations obtained (correlation coefficient, r, and standard deviation, s) are presented in table IV. Standard computer programmes were used throughout for the correlational analysis. For the linear correlations, a linear regression program was used whereas for the nonlinear correlations a multivariate correlation program was employed. All parameters were equally weighted and standardized by the well-known procedure described in [32].

From tables III and IV it can be concluded that the most successful combination of variables is that involving the indices S^{E}_{6} and π (6). The standardized electronic and hydrophobic indices which correlate best with purine CA 755 activity are listed in table V. Below we present the best mathematical equation obtained from our study together with the following statistics: the regression coefficients'confidence intervals according to Student's t-test at a 95% confidence limit, correlation coefficient, r, standard deviation, s, Fisher value, F, as well as mean square deviation between experimental and predicted values obtained after the 'leave-one-out' procedure, s':

$$\log (1/C) = 3.69 (\pm 0.14) + 0.51 (\pm 0.14)\, S^{E}_{6} + 0.24 (\pm 0.14)\, \pi (6) \quad (1)$$

$$(n = 17, r = 0.920, s = 0.265, F = 39.67, s' = 0.298)$$

Inclusion of the indices S^{E}_{10}, S^{N}_{10} and q_6 makes the statistics for our models worse (table IV). On the other hand, as seen from table IV, the parabolic equations including $(S^{E}_{6})^2$ produce higher correlation coefficients and lower standard deviations than the linear ones. However, the 95% confidence interval for the quadratic term coefficient is comparable in value to the coefficient itself, due to the relatively small correlation sample. Thus, for applicability of the nonlinear model, a larger number of purine derivatives needs to be examined.

Our results indicate that the preliminary hypothesis postulating a charge transfer process between the purine and biomacromolecules at position 10 is perhaps invalid, although the large contribution of the S^{N}_{10} index in models 2 and 12 in table IV could be said to favour such an interaction. The essential electronic process taking place is most likely charge transfer from position 6 on the purine ring to the biomacromolecule. The large positive coefficient for the S^{E}_{6} term in equation (1) may be viewed as supportive of such a conclusion. Our results confirm speculation by Neiman and Quin [12] that potency in CA 755 activity is predominantly determined by the resonance effect of the substituents at position 6. In general, resonance effects are associated with charge transfer processes. Furthermore, table V shows that all purine derivatives having high values of the donor superdelocalizability index S^{E}_{6} possess high CA 755 potency. For instance, 6-chloro, 6-bromo, 6-methylthio, 6-ethylthio, 6-methylsulfonyl, 6-sulfonamide, 2-amino-6-methylsulfonyl substituted purines have a positive standardized value for S^{E}_{6} and all exhibit significant potency against murine solid tumor CA 755 (in each case log $(1/C) > 4.00$). One exception to this qualitative picture which must be admitted is that the potency of 2-amino-6-chloropurine is relatively high (log $1/C$ = 4.05), even though its standardized donor superdelocalizability index has a negative value ($S^{E}_{6} = -0.209$). All the other substituted purines investigated here having negative standardized S^{E}_{6} values display a relatively low potency against CA 755 (log $1/C < 4.00$).

Table III. The intercorrelation matrix for various electronic and hydrophobic indices.

	π (6)	B^{6}_{1}	q_6	S^{E}_{6}	S^{E}_{10}	S^{N}_{10}
π (6)	1					
B^{6}_{1}	0.0539	1				
q_6	0.3073	0.8079	1			
S^{E}_{6}	0.0120	0.7025	0.7753	1		
S^{E}_{10}	0.7247	0.0771	0.3475	0.0349	1	
S^{N}_{10}	0.7611	0.2266	0.5387	0.0851	0.9418	1

Table IV. Statistical correlations for combinations of variables against bioactivity of the purines using both linear and nonlinear mathematical models.

No [a]	*Variables*	*Correlation coefficient*	*Standard deviation*
1	S^{E}_{6}, π (6)	0.920	0.265
2	S^{N}_{10}, q_6	0.904	0.290
3	S^{E}_{6}, S^{E}_{10}	0.896	0.300
4	S^{E}_{6}, S^{N}_{10}	0.892	0.306
5	π (6), q_6	0.858	0.348
6	S^{E}_{6}, B^{6}_{1}	0.823	0.385
7	π (6), B^{6}_{1}	0.791	0.414
8	S^{N}_{10}, B^{6}_{1}	0.782	0.422
9	S^{E}_{6}, q_6	0.782	0.422
10	S^{E}_{6}, π (6)	0.943	0.235
11	S^{E}_{6}, S^{E}_{10}	0.921	0.273
12	S^{N}_{10}, q_6	0.917	0.281
13	S^{E}_{6}, S^{N}_{10}	0.915	0.283
14	S^{N}_{10}, S^{E}_{10}	0.913	0.286
15	S^{E}_{10}, π (6)	0.889	0.327
16	S^{N}_{10}, π (6)	0.885	0.328

[a]The number of empirical coefficients equals 3 (**1–9**) and 4 (**10–16**).

Table V. Electronic and hydrophobic indices displaying the best correlations with the bioactivity of purine species.

Purine derivative	S^{E}_{6}	S^{N}_{10}	S^{E}_{10}	$\pi(6)$
6-Chloro	0.7948	− 1.0625	0.7379	0.7683
6-Bromo	0.8716	− 0.9572	1.1609	0.9218
6-Methoxy	− 0.5452	− 0.4884	− 0.1946	0.0216
6-Propoxy	− 0.4669	− 0.5135	− 0.1682	1.1161
6-Hydrazino	− 0.4684	0.3505	− 0.1895	− 0.8580
6-Methylthio	1.0235	− 0.9970	1.4860	0.6660
6-Ethylthio	1.0450	− 1.0072	1.4975	1.1365
6-Methylsulfonyl	1.5301	1.5642	− 1.3795	− 1.6251
6-Sulfonamide	1.4917	1.6931	− 1.4768	− 1.8194
2-Methyl-6-amino	− 0.6634	0.9197	− 1.0334	− 1.2160
2-Chloro-6-methylamino	− 1.5199	0.4989	− 0.3542	0.6149
2-Bromo-6-amino	− 0.5406	1.0247	− 1.0973	− 1.2160
2-Amino-6-chloro	− 0.2090	− 1.0546	0.7458	0.7683
2-Amino-6-bromo	− 0.1323	− 0.9465	1.1629	0.9218
2-Amino-6-methylsulfonyl	0.6858	1.5403	− 1.4075	0.6149
2,6-Bishydrazino	− 1.8944	0.1197	0.0912	− 0.8580
2-Fluorosulfonyl-6-chloro	− 1.0026	− 0.6843	0.4191	0.0421

The smaller coefficients for the π (6) terms in equation (1) support the contention that hydrophobic interactions play a subordinate role in determining CA 755 activity. The individual correlations of purine CA 755 activity with π (6) are comparatively low (for polynomials of second and third order, the respective correlation coefficients are 0.725 and 0.818), although the results shown in table IV indicate that our mathematical models are only improved from the statistical standpoint when the S^{E}_{6} variable is combined with π (6). Bearing in mind that correlation is not in itself proof of causality [36], a reasonable speculation based on these findings might be that, before any type of electronic (change transfer) interaction can take place, the purine ring needs to be attached to the receptor site by means of an initial hydrophobic interaction occurring at position 6 of the ring.

Conclusion

In terms of its statistics, the mathematical model proposed here in equation (1) for the porency of substituted purines against murine adenocarcinoma CA 755 is significantly better than that put forward by Neiman and Quinn [13]. For comparison purposes, we now give the correlation equation obtained by the latter workers:

$$\log (1/C) = 4.26 - 0.47\, MR\ (2) + 1.18\, R\ (6) \qquad (2)$$

$(n = 22, r = 0.815, s = 0.372)$

In attempting to reproduce their model, we obtained a slightly different version of the equation (2), namely

$$\log (1/C) = 4.30 - 0.58\, MR\ (2) + 1.11\, R\ (6) \qquad (3)$$

$(n = 22, r = 0.811, s = 0.376)$

Then, using the same 17 compounds we have employed in our study, we recalculated the result for the Neiman and Quinn [13] model thus:

$$\log (1/C) = 4.33\ (\pm 0.32) - 0.82\ (\pm 0.68)\, MR\ (2) + 1.51\ (\pm 0.63)\, R\ (6) \qquad (4)$$

$(n = 17, r = 0.836, s = 0.371, F = 19.51)$

On comparing our equation (1) with the above result, we are of the conclusion that the higher correlation coefficients and lower standard deviations from the OASIS model are not due to the smaller correlation sample (17 as against 22) but are rather a consequence of the optimized selection of the structural parameters. In fact, with the same set of 17 compounds we were able to improve the correlation coefficient from 0.836 to 0.920 for the linear equation. At the same time, the standard deviation was reduced from 0.371 to 0.265. Our best model [1] thus explains 85% of the variance in the selected data set. We would also mention, following Topliss and Edwards' report [9], that a careful preliminary selection of the structural parameters employed in the modelling makes the probability of a chance correlation occurring virtually zero.

For further validation of the proposed model its predictive ability was examined. More specifically, the potency of each compound was predicted in turn from the equation derived from scratch, including variable selection, in its absence. The 17 equations thus obtained are given in table VI along with model [1] presented as number 0. All the equations show remarkable convergence, with the respective deviation intervals of the 3 coefficients a, b, and c being 3.69 ± 0.03, 0.51 ± 0.04, and 0.24 ± 0.05, *ie* they are much smaller than the confidence intervals of model (1): $\pm$ 0.14. On the other hand, the coincidence in the predicted and observed potencies supports the predictive power of our model.

The satisfactory results achieved by us provide further evidence that the OASIS approach can be employed to model biological interactions of sets of structurally related molecules. Moreover, our methodology also confirms the conceptualization of the nature of biological interaction expounded in Part II of the series [2]. Therein, argumentation was advanced based on the now well-established requirement for bioactivity that there be optimal geometric [7, 37] and hydrophobic [38, 39] correspondence between the pharmacophore and its receptor site.

Table VI. Predicted CA 755 potencies of the purine derivatives obtained by linear regressions excluding the respective compound. The numbers 1 to 17 correspond to the purine derivatives 2 to 18 in table I.

No	a	b	c	r	s	*Predicted*	*Observed*	*Calculated by model 1*
0	3.6906	0.5070	0.2440	0.920	0.265	–	–	–
1	3.6940	0.5097	0.2466	0.917	0.274	4.29	4.23	4.28
2	3.6807	0.4985	0.2350	0.913	0.271	4.33	4.50	4.36
3	3.7173	0.4925	0.2448	0.931	0.246	3.45	3.00	3.42
4	3.7117	0.4969	0.2677	0.929	0.259	3.78	3.42	3.73
5	3.6922	0.5063	0.2427	0.917	0.275	3.25	3.22	3.24
6	3.6811	0.4974	0.2378	0.912	0.271	4.35	4.51	4.37
7	3.6941	0.5106	0.2480	0.912	0.274	4.51	4.45	4.50
8	3.6933	0.5112	0.2396	0.919	0.274	4.09	4.04	4.07
9	3.6659	0.4696	0.2894	0.925	0.259	3.84	4.26	4.00
10	3.6784	0.5150	0.2588	0.921	0.270	3.02	3.23	3.06
11	3.6680	0.5415	0.2297	0.929	0.258	2.99	3.37	3.07
12	3.7299	0.4864	0.1965	0.938	0.216	3.23	2.56	3.12
13	3.6725	0.5110	0.2301	0.926	0.262	3.74	4.05	3.77
14	3.6865	0.5076	0.2403	0.920	0.274	3.84	3.91	3.85
15	3.7004	0.5137	0.2500	0.921	0.271	4.21	4.04	4.19
16	3.6693	0.5471	0.2618	0.916	0.262	2.41	2.77	2.52
17	3.6914	0.5062	0.2441	0.917	0.275	3.19	3.18	3.19

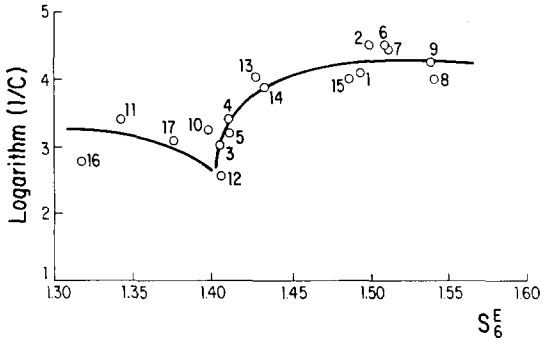


Fig 2. Plot of the CA 755 bioactivity of purine derivatives *vs* the donor superdelocalizability at position 6 on the purine fragment.

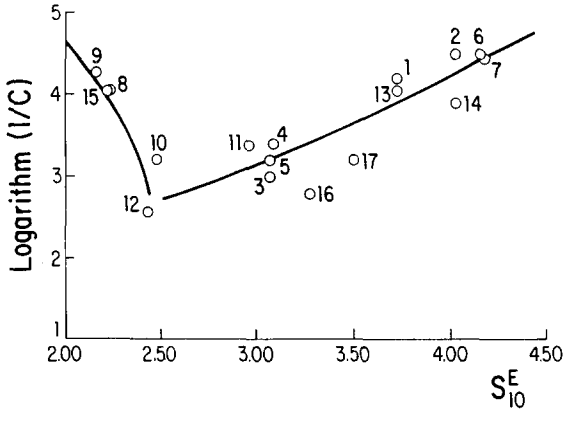


Fig 3. Plot of the CA 755 bioactivity of purine derivatives *vs* the donor superdelocalizability at position 10 on the purine fragment.

An additional requirement that we imposed had the effect of producing a parabolic dependence between the biological activity of chemical species and the relevant electronic parameters. This requirement is based on the biological interaction complexity, due to the large number of competing target and side reactions. It was shown [40] that this fact increases the probability for electronic structure/activity correlations manifesting extrema in biological activity. The great diversity of substituents in our test class of purine derivatives, together with the relatively low precision of the experimentally determined CA 755 potency, augurs well for the observation of a parabolic correlation between the bioactivity and those electronic indices which correlate best, namely S^E_6, S^E_{10}, S^N_{10}, and q_6 (table II). Correlations displaying extrema are clearly in evidence in figures 2, 3, 4 and 5. The minimum in figure 2 is less pronounced and this explains why the correlation coefficients for S^E_6 and log (1/*C*) are the same for both the linear and parabolic equations (table II). It is of interest to note that these extrema are located in the vicinity of compound 12, which is the one having minimum biological potency. The expectation of a parabolic dependence between the biological CA 755 potency and the hydrophobic index π (6) is also realized, although not illustrated here.

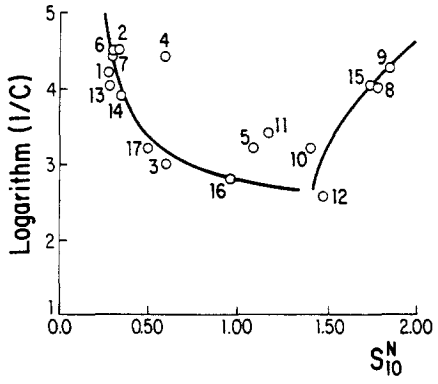


Fig 4. Plot of the CA 755 bioactivity of purine derivatives *vs* the acceptor superdelocalizability at position 10 on the purine fragment.

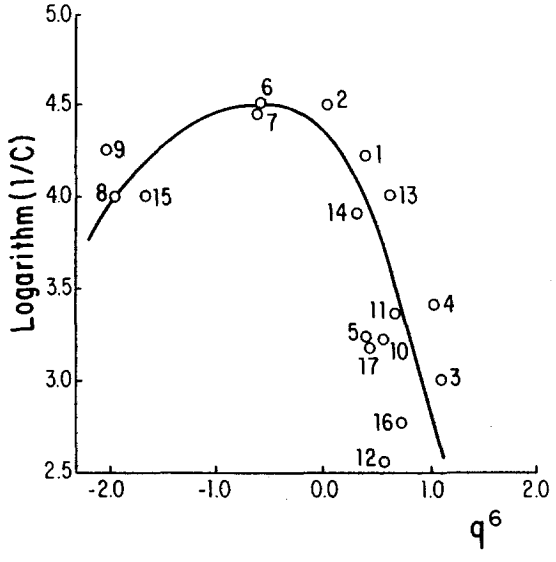


Fig 5. Plot of the CA 755 bioactivity of purine derivatives *vs* the electronic charge at position 6 on the purine fragment.

References

1 Mekenyan O, Peitchev D, Bonchev D, Trinajstić N, Bangov I (1986) *Drug Res* 36, 176
2 Mekenyan O, Bonchev D, Trinajstić N, Peitchev D (1986) *Drug Res* 36, 421
3 Mekenyan O, Peitchev D, Bonchev D, Trinajstić N, Dimitrova J (1986) *Drug Res* 36, 629
4 Mekenyan O, Bonchev D (1986) *Acta Pharm Jugosl* 36, 225
5 Mekenyan O, Bonchev D, Enchev V (1988) *Quant Struct Act Relat* 7, 240
6 Mekenyan O, Karabunarliev S, Bonchev D (1990) *J Math Chem* 4, 207
7 Hansch C (1969) *Acts Chem Res* 2, 232
8 Hansch C, Schaeffer J, Kerley R (1972) *J Biol Chem* 247, 4703
9 Topliss JG, Edwards RP (1979) *J Med Chem* 22, 1238
10 Skipper HE, Montgomery JA, Thompson JR, Schabel FM (1958) *Cancer Res* 19, 425
11 Geran RI, Greenberg NE, McDonald MM, Schumacher AM, Abbott BJ (1972) *Cancer Chemother Rept* Part 3, 3, 1
12 Neiman Z, Quinn FR (1981) *J Pharm Sci* 70, 425
13 Neiman Z, Quinn FR (1982) *J Pharm Sci* 71, 618
14 Hansch C, Leo A (1979) *Substituent Constants for Correlation Analysis in Chemistry and Biology*. Wiley, New York, Appendix I
15 Sabljić A, Trinajstić N (1981) *Acta Pharm Jugosl* 31, 189
16 Balaban AT, Chiriac A, Motoc I, Simon Z (1980) *Lect Notes Chem* 15, Springer, Berlin
17 Trinajstić N (1983) *Chemical Gr-aph Theory I and II*. CRC Press, Boca Raton, Florida
18 Bonchev D (1983) *Information-Theoretic Indices for Characterization of Chemical Structures*. Research Studies Press, Chichester, UK
19 Rouvray DH (1983) *Stud Theor Phys Chem* 28, 159
20 Balaban AT, Motoc I, Bonchev D, Mekenyan O (1984) *Topics Curr Chem* 114, 21
21 Dubois JE (1972) *Man Comput* 309
22 Mercier C, Sobel Y, Dubois JE (1981) *Eur J Med Chem* 16, 473
23 Kier LB, Hall LH (1976) *Molecular Connectivity in Chemistry and Drug Research*. Academic Press, New York
24 Kier LB, Hall LH (1986) *Molecular Connectivity in Structure-Activity Analysis*. Research Studies Press, Letchworth, England
25 Rouvray DH (1986) *In: Mathematics and Computational Concepts in Chemistry* (Trinajstić N, ed) Ellis Horwood, Chichester, England, p 295
26 Rouvray DH (1986) *Sci Am* 254, 40
27 Rouvray DH (1987) *J Comput Chem* 8, 470
28 Stankevich MI, Stankevich IV, Zefirov NS (1988) *Usp Khim* 57, 337
29 Verloop A, Hoogenstraaten W, Tipker J (1976) *Drug Design* 7, 165
30 Klopman GK (1967) *J Am Chem Soc* 90, 223
31 Klopman G, Hudson RF (1967) *Theor Chim Acta* 8, 165
32 Berges J, Peradejordi F (1977) *In: Molecular and Quantum Pharmacology* (Bergman E, Pullman B, eds) Reidel, Dordrecht, The Netherlands
33 Sutton LE (1965) *Tables of Interatomic Distances and Configuration in Molecules and Ions* The Chemical Society, London (Suppl)
34 Yee WT, Sakamoto K, I'Haya YJ (1976) *Rept Univ Electro-Commun* 27, 53; *ibid* (1977) 27, 227
35 Hansch C, Unger SH, Forsythe AB (1973) *J Med Chem* 16, 1217
36 Klopman G, Kalos AN (1985) *J Comput Chem* 6, 492
37 Higuchi T, Davis S (1970) *J Pharm Sci* 59, 1376
38 Lien E, Hansch C, Anderson S (1968) *J Med Chem* 11, 430
39 Martin YC (1978) *Quantitative Drug Design*. Marcel Dekker, New York
40 Mekenyan O, Bonchev D (1989) *Eur J Med Chem* 24, 141