Dyes and Pigments 24 (1994) 267–279

A QSAR Study of the Adsorption by Cellulose Fibre of Anthraquinone Vat Dyes

Simona Timofei,*[a] Walter Schmidt,[a] Ludovic Kurunczi[a]
Zeno Simon[b] & Aurora Salló[a]

[a] Institute of Chemistry, Romainian Academy,
Bul, Mihai Viteazul 24, 1900—Timisoara, Romania
[b] Department of Biophysics, University of Medicine and Pharmacy,
P-ta Eftimie Murgu 2, 1900—Timisoara, Romania

(Received 3 November 1993; accepted 8 December 1993)

ABSTRACT

The affinity of a series of 46 anthraquinone vat dyes to cellulose fibre was analysed by multiple regression analysis and the MTD method of QSAR. The main structural feature responsible for fibre binding is the number (n) *of bonds of the conjugated chain along the main axis of the molecules (correlation coefficient* r = 0·914*). Hydrophobic interactions and hydrogen bonding, especially by proton donor groups of the dye molecule are also important (*r = 0·943 *for a correlation with four structural parameters). The optimized receptor map of the MTD method suggests binding to both crystalline and amorphous regions of the cellulose fibre.*

1 INTRODUCTION

Dye adsorption by fibre, from the molecular point of view, is similar to biological ligand–receptor interaction, although it is possible that a textile fibre with both crystalline and amorphous regions will behave as a mixture of receptors of similar, but not identical, nature.

Correlations between technical properties and chemical dye structure have been made using the Free–Wilson analysis[1–4] and the PLS method.[4–7]

In this paper the influence of several structural factors on dye fixation

* To whom correspondence should be addressed.

Dyes and Pigments 0143-7208/94/$07.00

by fibres was studied by multiple regression analysis and the MTD method.[8] Three types of structural factors are considered to determine biological activity;[9] viz. (a) electronic factors connected with reactivity and intermolecular force characteristics; (b) steric factors connected with the fit of the ligand molecule in the receptor site; (c) partition effects connected especially with transport characteristics. These factors were used for the dye QSAR calculations.

2 METHODS

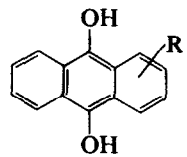
2.1 Definition of parameters used

The dye affinity for fibre (A), which is defined as the Gibbs free energy variation versus the dye content increase in one phase (e.g. fibre phase), can be used as a thermodynamic function to express the driving force of dyeing, similar to biological activity from QSAR calculations. Here we use the affinities of a set of 46 anthraquinone leuco vat dyes to cellulose fibre, in kcal/mole, as described by Giles and Hassan[10] (Table 1).

One of the structural parameters used was the length of the conjugated chain of the dye molecule (n), which Giles and Hassan[10] consider to be oriented along one cellulose chain and which determines the magnitude of the physical attraction between the π electronic system of the dye molecule and the fibre. The length of the conjugated chain was measured by the number of bonds between the extreme carbon atoms of the aromatic nuclei, or from the oxygen, nitrogen or halogen atoms of the terminal substitutents at either end of the axis of the molecule.

Hydrophobic effects (log P) were calculated for the leuco vat dyes as the sum of π-Hansch hydrophobicities[11] and/or the sum of the f-Rekker hydrophobicities,[12] with corrections for hydrogen bonding, expressed for the whole molecule and, separately, for the longest axis of the molecule (log P_n). As a hydrophobicity parameter we also tried the squared π-Hansch (or f-Rekker) hydrophobicities (log^2 P). These parameters also include electronic, steric or intermolecular effects between effector and receptor.[13] We are aware that such log P evaluations have a low degree of precision,[14] but we are actually interested in relative log P values. For a higher precision level of log P evaluations, molecular conformations are to be considered also, while for several of our dye molecules these conformations are not known. Grindea *et al.*[15] consider that the dye–cellulose interactions are mainly of a hydrophobic nature (especially van der Waals interactions) and minimize hydrogen bondings. The hydrophobic interactions are very closely related to the aggregation tendency

TABLE 1
Structural Parameters of the Anthraquinone Leuco Vat Dyes



No.	*R*	*A*[a] *(kcal/mole)*	*n*	*Log P*	*Log P_n*	σ_D	σ_{Dn}	n_H	*MTD*	*F*	*j for x_{ij} = 1 (from eqn (2))*
1	H	0	7	5·71	0·479	2	1	4	22	0	—
2	1-Cl	0	7	6·42	1·200	2	1	4	21	11·30	1
3	1-CH_3	0	7	6·27	0·477	2	1	4	21	11·60	1
4	1-OCH_3	0	7	5·04	0·538	2	1	3	21	17·20	1, 2
5	1-NH_2	0	7	5·13	−1·422	4	3	5	21	22·96	1
6	2-NH_2	0	7	6·43	−1·422	4	3	7	22	22·96	26
7	1-N(CH_3)($COCH_3$)	0	7	4·36	−0·771	3	2	4	19	18·52	1, 55–58
8	1-$NHCH_2C_6H_5$	0	7	3·84	−0·802	2	1	3	22	10·13	1, 2, 4–9
9	1-N(CH_3)(COC_6H_5) 4-N(CH_3)(COC_6H_5)	0	7	3·89	−0·802	2	1	2	22	25·33	1, 14, 55–63, 65–72, 64
10	1-NH_2; 4-NH_2	1·46	7[b]	4·55	−1·422	6	3	6	20	45·93	1, 14
11	1-$NHCH_3$	1·46	7[b]	5·24	−1·621	3	2	4	21	17·24	1, 2
12	1-$NHCOCH_2Cl$	1·46	9	2·60	−1·613	3	2	4	20	19·07	1–4, 9
13	1-NH_2; 8-NH_2	1·49	7[b]	5·85	−1·422	6	3	8	20	45·93	1, 33
14	1-$NHCOCH_3$	1·52	9	1·81	−1·613	3	2	4	20	21·34	1–4
15	2-NHCOCH	1·53	10	3·11	−1·257	3	2	6	20	21·34	2, 3, 26, 27
16	1-$NHCH_3$; 4-$NHCH_3$	1·55	8	4·77	−1·621	4	2	4	19	24·02	1, 2, 14, 15

continued

TABLE 1—*contd.*

No.	*R*	*A*[a] *(kcal/mole)*	*n*	*Log P*	*Log P*$_n$	σ_D	σ_{Dn}	n_H	*MTD*	*F*	*j for* $x_{ij} = 1$ *(from eqn (2))*
17	(1, 2)-$(CH_2)_4$	1·91	9	6·83	0·665	2	1	4	20	7·00	1–3, 26
18	2-$NHCOC_6H_5$	2·05	13	3·25	−0·032	3	2	5	18	19·17	2, 3, 26–32
19	1-NH_2; 5-NH_2	2·31	7[b]	5·85	−1·403	6	5	8	20	45·93	1, 42
20	1-$NHCOC_6H_5$	2·37	11	3·26	−0·744	3	2	3	18	19·17	1–9
21	1-NH_2; 5-$NHCOC_6H_5$	2·39	11	2·48	−1·950	5	4	4	18	15·53	1, 42–50
22	1-$NHCOC_6H_5$; 8-$NHCOC_6H_5$	2·60	11	2·85	−0·545	4	2	2	17	33·75	1–9, 33–41
23	1-[4-$CH_3C_6H_4CONH$—]	2·68	11	4·02	−0·545	3	2	3	17	13·20	1–10
24	1-[3-ClC_6H_4CONH—]	2·73	12	3·82	−0·013	3	2	3	17	13·12	1–9, 12
25	1-[3-$CH_3C_6H_4CONH$—]	2·74	11	3·32	−1·299	3	2	3	17	13·20	1–9, 12
26	1-[4-$OCH_3C_6H_4CONH$—]	2·90	12	2·18	−0·587	3	2	3	17	14·40	1–11
27	1-NH_2; 4-$NHCOC_6H_5$	2·94	11	1·83	−0·744	5	2	3	16	15·53	1, 14–20
28	1-[4-ClC_6H_4CONH—]	3·02	12	3·12	−1·299	3	2	3	17	13·12	1–10
29	1-$NHCOC_6H_5$; 5-$NHCOC_6H_5$	3·57	15	2·21	−1·059	4	3	6	15	33·75	1–9, 42–50
30	1-$NHCOC_6H_5$; 5-OCH_3	3·59	12	1·53	−0·481	3	2	2	16	13·87	1–9, 42, 43
31	1-$NHCOC_{10}H_7$—2'	3·61	12	3·30	−0·744	3	2	3	15	13·20	1–9, 12, 13, 28, 29
32	1-$NHCOC_6H_5$; 4-OCH_3	3·59	11	1·43	−0·744	3	2	2	16	13·87	1–9, 14, 15
33	1-[4-$CH_3C_6H_4CONH$—]; 5-[4-$CH_3C_6H_4CONH$—]	3·73	15	3·66	−3·263	4	3	2	15	31·06	1–10, 42–51
34	1-[3-$CH_3C_6H_4CONH$—]; 5-[3-$CH_3C_6H_4CONH$—]	3·77	15	3·54	−3·263	4	3	2	15	31·06	1–9, 12, 42–50, 53

35	1-[3-$OCH_3C_6H_4CONH$—]; 5-[3-$OCH_3C_6H_4CONH$—]	3·86	17	0·70	−2·199	4	3	2	15	33·40	1–9, 12, 13, 42–50, 53, 54
36	1-[4-$OCH_3C_6H_4CONH$—]; 5-[4-$OCH_3C_6H_4CONH$—]	4·11	17	−0·70	−2·007	4	3	2	14	33·40	1–11, 42–52
37	1-[4-ClC_6H_4CONH—]; 5-OCH_3	4·06	13	1·34	0·343	4	2	3	17	12·95	1–10
38	1-[3-ClC_6H_4CONH—]; 5-[3-ClC_6H_4CONH—]	4·17	17	2·58	−3·462	4	3	2	15	30·84	1–9, 12, 42–50, 53
39	1-$NHCOC_6H_5$; 4-$NHCOC_6H_5$	4·29	15	4·13	−0·310	3	2	3	15	33·75	1–9, 14–22
40	1-[4-ClC_6H_4CONH—]; 5-[4-ClC_6H_4CONH—]	4·37	17	2·58	0·032	4	3	2	15	30·84	1–10, 42–51
41	1-[4-$OCH_3C_6H_4CONH$—]; 4-[4-$OCH_3C_6H_4CONH$—]	4·44	15	0·70	−0·700	4	2	2	14	33·40	1–11, 14–24
42	1-[3-$CH_3C_6H_4CONH$—]; 4-[3-$CH_3C_6H_4CONH$—]	4·56	15	4·33	−0·310	4	2	2	13	31·06	1–9, 12, 14–22, 25
43	1-$NHCOC_6H_5$; 4-$NHCOC_6H_5$; 5-$NHCOC_6H_5$	4·56	15	2·84	−3·269	5	3	1	12	52·93	1–9, 14–22, 42–50
44	1-[4-$CH_3C_6H_4CONH$—]; 4-[4-$CH_3C_6H_4CONH$—]	4·57	15	4·45	−0·310	4	2	2	13	31·06	1–10, 14–23
45	1-[3-ClC_6H_4CONH—]; 4-[3-ClC_6H_4CONH—]	4·98	15	2·58	0·440	4	2	2	13	30·84	1–9, 12, 14–22, 25
46	1-[4-ClC_6H_4CONH—]; 4-[4-ClC_6H_4CONH—]	5·20	17	2·58	1·160	4	2	2	13	30·84	1–10, 14–23

[a] The compounds are listed in increasing order of the A values, except numbers 31/32 and 36/37, in order to have balanced distribution of the vertices in the 'even-odd' cross-validation procedure for the MTD method.
[b] n-values from the original paper[10] were modified for steric reasons.

of dyes.[16] The number of H^+-acceptor hydrogen bonding groups (n_H) was calculated as the number of amine nitrogen atoms, carbonyl oxygen and hydroxy oxygen atoms. It has been considered that the availability of lone electrons of these atoms determines the dye affinity.[17]

The number of H^+-donating hydrogen bonding groups was also estimated by the number of proton donor groups for the entire dye molecule (σ_D) and for the longest axis of the molecule (σ_{Dn}), and by the proton acceptor groups, for the whole dye molecule (σ_A) and for the longest axis of the molecule (σ_{An}). Correlations were also tried with the sum of the donor and acceptor groups for the whole molecule (σ_{AD}) and for the longest axis of the molecule (σ_{ADn}).

Another possible structural parameter for such correlations is the F parameter, which characterizes the polarity of a certain molecular region and estimates the electrostatic field on the surface of the van der Waals envelope of the corresponding functional group; F was calculated by the additivity of electrostatic field increments, with some corrections for the presence of benzene nuclei according to Náray-Szabó.[18]

The structural data used in the final correlations for the leuco vat dyes are presented in Table 1.

2.2 Multiple regression analysis

In this method one experimental variable y_k (usually the biological response in QSAR calculations) is correlated with one or several structural variables x_i by the equation:[19]

$$y_k = b_o + \sum_i b_i \, . \, x_{ik} + e_k \tag{1}$$

where b represents partial regression coefficients, and e the deviations and residuals.

For the calculation of the statistical parameters of the regression equation we used the SYSTAT program, from SYSTAT, Inc., Evenston, IL, USA.

2.3 The MTD method

The minimal steric (topological) difference (MTD)[8] is a measure of steric misfit between i molecules, $i = 1, 2, 3, \ldots, N$, and the receptor site. The hypermolecule considered as a steric and topological network is constructed through approximate atom per atom superpositions of the whole set of molecules (H atoms are neglected), having $j = 1, 2, 3, \ldots, M$ vertices. If the i molecule occupies the j vertex, then $x_{ij} = 1$, and $x_{ij} = 0$, if the vertex j is not occupied.

The MTD_i value of the i molecule, in comparison with the receptor is calculated by the equation:

$$MTD_i = S + \sum_{j=1}^{M} \varepsilon_j \cdot x_{ij} \tag{2}$$

(with S the total number of cavity vertices).

The j vertices denote the approximate positions of the atoms from the molecules and are assigned to the receptor cavity (annotation $\varepsilon_j = -1$), receptor walls ($\varepsilon_j = +1$), or to the region outside the receptor ($\varepsilon_j = 0$) through an optimization procedure. These vertex attributions $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_M$ constitute the receptor map; MTD_i is the number of wall vertices occupied by the i molecule plus the number of unoccupied cavity vertices. MTD calculations were carried out on a IBM—PC 486 computer with a program similar to that of Ref. 8.

2.4 Reliability tests for QSAR results

In this QSAR attempt, a rather large number of explicit and implicit (ε_j parameters for MTD) structural variables are used. Therefore, together with the usual statistical tests, we verified the reliability of our results by a cross validation procedure applied in two manners, viz. (1) The 'leave-one-out' procedure (similar to the 'leave-n-out' technique[20]) in which one molecule is held out from the set, the correlational equation is computed for the rest of $N - 1$ molecules and the result is used to calculate the estimated affinity of the left out molecule. These estimated affinities are compared with the respective experimental values. (2) In the second 'even-odd' procedure,[21] the molecules are arranged in decreasing order of experimental affinities; two subseries were set up, one for the molecules with even ordering numbers ($i = 2, 4, \ldots$), the second for the odd molecules ($i = 1, 3, \ldots$). Correlational equations were obtained for each subseries. Estimated affinities of the even number compounds were computed by the equation obtained for the odd subseries, and vice versa. The calculated values were compared with the experimental ones.

3 RESULTS

3.1 Multiple regression analysis

For the set of 46 compounds from Table 1, regression calculations (uniparametric correlations) were carried out with all the parameters mentioned in Section 2.1. Partial correlation and intercorrelation coefficients are listed in Table 2.

TABLE 2
Partial Correlation Coefficients (r_p) and Intercorrelation Coefficients

	r_p	n	$log\ P$	$log^2\ P$	$log\ P_n$	n_H	σ_D	σ_{Dn}	σ_A	σ_{An}	σ_{AD}	σ_{ADn}	MTD	F
n	0·914	1												
log P	−0·615	−0·658	1											
$\log^2 P$	−0·617	−0·648	0·942	1										
log P_n	−0·149	−0·186	0·245	0·282	1									
n_H	−0·534	−0·599	0·518	0·543	0·010	1								
σ_D	0·392	0·239	−0·164	−0·148	−0·445	0·273	1							
σ_{Dn}	0·330	0·271	−0·183	−0·169	−0·613	0·310	0·822	1						
σ_A	0·729	0·808	−0·643	−0·676	−0·291	−0·590	0·202	0·197	1					
σ_{An}	0·852	0·944	−0·718	−0·731	−0·325	−0·590	0·222	0·306	0·841	1				
σ_{AD}	0·702	0·642	−0·493	−0·500	−0·482	−0·155	0·818	0·691	0·728	0·649	1			
σ_{ADn}	0·806	0·852	−0·637	−0·642	−0·519	−0·324	0·531	0·673	0·739	0·910	0·805	1		
MTD	−0·950	−0·919	0·601	0·625	0·159	0·612	−0·347	−0·268	−0·763	−0·872	−0·691	−0·795	1	
F	0·407	0·355	−0·136	−0·142	−0·449	0·077	0·741	0·613	0·438	0·353	0·776	0·541	−0·396	1

Although, as indicated by the calculated log P values (see Table 1), the dye molecules studied have a rather high lipophilicity, the A versus log P correlation yields a rather small and negative correlation coefficient, $r_p = -0{\cdot}615$. Our dye molecules seem to have log P values above the optimal lipophilicity for binding to cellulose fibre.

The affinity (A) was correlated first with the length (n) of the conjugated chain (one of the best r_p values from Table 2). The results are:

$$\begin{aligned} \hat{A} &= -2{\cdot}213\ (\pm 0{\cdot}332) + 0{\cdot}420\ (\pm 0{\cdot}028)n \\ N &= 46, \quad r = 0{\cdot}914, \quad s = 0{\cdot}667, \quad F(\text{Fischer}) = 223{\cdot}3 \end{aligned} \tag{3}$$

Other statistical test values for eqn (3) are $r^2_{\text{adjusted}} = 0{\cdot}832$, t(Student) = $-6{\cdot}67$ for the free term and 14·94 for n.

Attempts at biparametric correlations with n and other structural parameters (σ_D, log P_n, σ_{An}, F) did not yield satisfactory results from the point of view of the usual statistical tests.

We tried to improve the uniparametric correlation eqn (3) by adding various other structural parameters which do not intercorrelate with n and between themselves ($r^2_p < 0{\cdot}4$ in Table 2) in a multiparametric attempt. The best result is:

$$\begin{aligned} \hat{A} &= -2{\cdot}455\ (\pm 0{\cdot}508) + 0{\cdot}353\ (\pm 0{\cdot}035)n - 0{\cdot}158\ (\pm 0{\cdot}076)n_H \\ &\quad + 0{\cdot}483\ (\pm 0{\cdot}108)\sigma_D + 0{\cdot}197\ (\pm 0{\cdot}086)\ \log P_n \\ N &= 46, \quad r = 0{\cdot}943, \quad s = 0{\cdot}565, \quad F(\text{Fischer}) = 83{\cdot}0 \end{aligned} \tag{4}$$

We obtained $r^2_{\text{adjusted}} = 0{\cdot}879$ and t(Student) values for the free term and the four correlation coefficients: $-4{\cdot}83$; 10·13; $-2{\cdot}06$; 4·48 and 2·30, respectively. For the cross-validated r^2 value, the 'leave-one-out' procedure gives $r^2_{CV1} = 0{\cdot}863$, the 'even-odd' procedure $r^2_{CV2} = 0{\cdot}827$.

With the MTD procedure (Section 3.2) the correlation of affinity versus MTD values yields eqn (5):

$$\begin{aligned} \hat{A} &= 11{\cdot}743\ (\pm 0{\cdot}465) - 0{\cdot}527\ (\pm 0{\cdot}026)\text{MTD} \\ N &= 46, \quad r = 0{\cdot}950, \quad s = 0{\cdot}515, \quad F(\text{Fischer}) = 404{\cdot}0 \end{aligned} \tag{5}$$

Other statistical test values for eqn (5) are: $r^2_{\text{adjusted}} = 0{\cdot}900$, t(Student) = 25·26 for the free term and t(Student) = $-20{\cdot}10$ for MTD.

With a second structural variable besides MTD (log P, n_H, σ_D, log^2 P, σ_{Dn}, log P_n, F) we obtained regression equations which did not improve the significance from the statistical point of view. The multiparametric correlation, including MTD, with the highest r and F(Fischer) values, is:

$$\begin{aligned} \hat{A} &= 10{\cdot}837\ (\pm 0{\cdot}580) - 0{\cdot}486\ (\pm 0{\cdot}032)\ \text{MTD} - 0{\cdot}084\ (\pm 0{\cdot}056)\ \log P \\ &\quad + 0{\cdot}275\ (\pm 0{\cdot}118)\sigma_{Dn} + 0{\cdot}140\ (\pm 0{\cdot}086)\ \log P_n \\ N &= 46, \quad r = 0{\cdot}957, \quad s = 0{\cdot}493, \quad F(\text{Fischer}) = 111{\cdot}9 \end{aligned} \tag{6}$$

The lengths of conjugated chain (n) and MTD values could not be used in the same regression equation because of their high intercorrelation coefficient ($r = -0.919$, Table 2).

The introduction of additional parameters in eqn (4) with respect to eqn (3) produced a significant increase in the correlation coefficient for the whole equation, but the additional parameters in eqn (6) as compared with eqn (5) produced only a moderate increase in r.

3.2 MTD calculations

The hypermolecule obtained by the superposition of the 46 dye molecules, considered planar, is presented in Fig. 1.

The optimized receptor map was obtained by trial and error and verified as optimal by the regression coefficient criterion.[8] It consists of the following vertex attributions:

$$S^* \begin{cases} j(\varepsilon = -1)\text{: 1, 3, 6, 7, 10, 12, 14, 15, 17–20, 23, 25, 27–29, 33, 42–44, 52} \\ j(\varepsilon = +1)\text{: 16, 21, 22, 24, 51, 53, 55, 59} \\ j(\varepsilon = 0)\text{: 2, 4, 5, 8, 9, 11, 13, 26, 30–32, 34–41, 45–50, 54, 56–58, 60–72} \end{cases}$$

The MTD values calculated with the optimized map are listed in Table 1 and with them eqns (5) and (6) were obtained.

A 'even-odd' cross-validation procedure was applied for the MTD method (see Section 2.4). For the odd set of compounds the following equation was obtained:

$$\hat{A}_{odd} = 11.072 - 0.507\ MTD \qquad N = 23, \quad r = 0.990 \tag{7}$$

with the optimized receptor map:

$$S_1 \begin{cases} j(\varepsilon = -1)\text{: 1, 3–7, 10, 12–15, 18–20, 23, 27–29, 33, 42–44} \\ j(\varepsilon = +1)\text{: 8, 21, 22, 24, 34, 52, 53, 55, 59} \\ j(\varepsilon = 0)\text{: 2, 9, 11, 16, 17, 25, 26, 30–32, 35–41, 45–51, 54, 56–58, 60–72} \end{cases}$$

For the even set of compounds the result is:

$$\hat{A}_{even} = 12.028 - 0.517\ MTD \qquad N = 23, \quad r = 0.925 \tag{8}$$

with the optimized receptor map:

$$S_2 \begin{cases} j(\varepsilon = -1)\text{: 1, 3, 5–7, 10, 12, 14, 15, 18–20, 23, 25–29, 33, 42–44, 52} \\ j(\varepsilon = +1)\text{: 8, 21, 22, 24, 51, 53, 55, 59} \\ j(\varepsilon = 0)\text{: 2, 4, 9, 11, 13, 16, 17, 30–32, 34–41, 45–50, 54, 56–58, 60–72} \end{cases}$$

From eqns (7) and (8) $r^2_{CV} = 0.859$ was calculated. This together with the

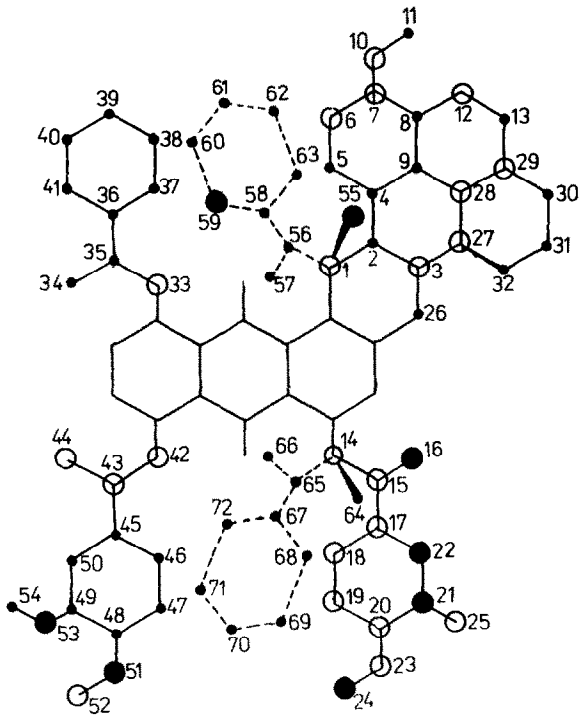


Fig. 1. Hypermolecule for anthraquinone vat dyes with numbering of vertices. Beneficial vertices ($\varepsilon_j = -1$) are marked by white circles (○), detrimental vertices ($\varepsilon_j = 1$), by black circles (●), irrelevant vertices ($\varepsilon_j = 0$), by dots.

high degree of similarity of the three optimized receptor maps (whole series, odd series, even series) indicates a significant predictive power for the MTD results.

4 DISCUSSION

From the multiple regression analysis results the most significant one is eqn (4); its r^2_{CV} values for the 'leave-one-out' and 'even-odd' procedure indicate the validity of the proposed model. There are no significant outliers in the data set, in accord with the standardized residual test.[22] This equation suggests the presence of electronic factors through the number of hydrogen bonds (n_H) and the number of proton donor groups (σ_D); the presence of steric variables, through the length of the conjugated chain (n), and hydrophobicity factors through the hydrophobicity of the axis of the conjugated chain (log P_n) in dye–fibre interactions. The effects of these structural factors cannot be exactly separated. The positive contribution of n, σ_D and log P_n is remarkable, which confirms the con-

tribution of steric, electronic and hydrophobicity effects to dye fixation on cellulose.

The significant parallelism between affinity (A) and length (n), $r_p = 0{\cdot}914$ in Table 2, indicates the rather uniform presence, along the dye molecule perimeter (length of conjugated chain), of groups allowing attractive interactions with the cellulose fibre. This length (n) of the conjugated chain can be considered as a parameter with electronic, steric and hydrophobic (van der Waals interactions) implications.

Hydrophobic interactions probably take place along the direction of the conjugated chain axis of the dye molecule. This is reflected by the hydrophobicity values divided by the length of conjugated chain, calculated for the more polar compounds; high values were obtained especially for compounds having benzamido substituents.

The presence of amino and amido nitrogen atoms, and also of hydroxy and amido oxygen atoms, could explain the occurrence of hydrogen bonds with the cellulose chains. The proton donor group of the dye molecules seem to be more important in this respect.

These correlational QSAR type results should allow predictions of new anthraquinone dyes from the same series, with increased affinity for cellulose fibres.

For the MTD results, the optimized receptor map suggests the existence of few steric compulsions—only 8 'wall' ($\varepsilon_j = +1$) vertices out of a total of 72. The 'cavity' ($\varepsilon_j = -1$) vertices should correspond to attractive interactions of various types: hydrophobic interactions, hydrogen bonds, etc. The positioning of the beneficial 'cavity' vertices not only along the longest axis of the dye molecules, but also in some lateral pockets, suggests binding of dye molecules not only in the crystalline, but also in the amorphous regions of the cellulose fibre. The positioning of detrimental 'wall' vertices suggests some steric hindrance to binding, by some lateral aromatic rings and by bulkier substituents, like chlorine and methoxy, on such rings.

In this study, we assumed a planar structure of our dye molecules, as a first approximation. Conformational analysis of these molecules could possibly offer new insights in their binding mechanism to the cellulose fibre.

ACKNOWLEDGEMENT

The authors thank FORSLYS (Formal Study of Living Systems) group, PO Box 1120, Timisoara 14, R-1900 Romania, for permitting access to the SYSTAT program.

REFERENCES

1. Grecu, R., Pieroni, M. & Carpignano, R., *Dyes and Pigments*, **2** (1981) 305–18.
2. Carpignano, R., Barni, E., Di Modica, G., Grecu, R. & Bottaccio, G., *Dyes and Pigments*, **4** (1983) 195–211.
3. Carpignano, R., Savarino, P., Di Modica, G. & Scavia, G., *Tinctoria*, **4** (1984) 97–110.
4. Carpignano, R., Savarino, P., Barni, E., Di Modica, G. & Papa, S. S., *J. Soc. Dyers Colour.*, **101** (1985) 270–6.
5. Carpignano, R., Savarini, P., Barni, E., Clementi, S. & Giulietti, G., *Dyes and Pigments*, **6** (1985) 189–212.
6. Garpignano, R., Savarini, P., Barni, E., Viscardi, G., Clementi, S. & Giulietti, G., *Anal. Chim. Acta*, **191** (1986) 445–50.
7. Carpignano, R., Savarino, P., Viscardi, G., Barni, E., Clementi, S. & Cruciani, G., *Melliand Textilber.*, **70** (1989) 276–8.
8. Simon, Z., Chiriac, A., Holban, S., Ciubotariu, D. & Mihalas, G. I., *Minimum Steric Difference*. Res. Stud. Press, Letchwoorts and John Wiley, New York, 1984.
9. Apetroae, M., Singer, C., Bozga, R., Cilianu, S., Ceclan, R., Bânzaru, T. & Simon, Z., *Rev. Roumaine Chim.*, **32** (1987) 583–90.
10. Giles, C. H. & Hassan, A. S. A., *J. Soc. Dyers Colour.*, **74** (1958) 846–57.
11. Hansch, C. & Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*. John Wiley, New York, 1979.
12. Rekker, R. F., *The Hydrophobic Fragmental Constant, Vol. 1*. Elsevier Science Publishers, Amsterdam, 1977.
13. Franke, R., In *QSAR and Strategics in the Design of Bioactive Compounds: Proceedings of the Fifth European Symposium on Quantitative Structure–Activity Relationships, Bad Segeberg, 1984*, ed. J. K. Seydel. VCH Verlagsgesellshaft, Weinheim, 1985, pp. 59–78.
14. Leo, A. J., *Chem. Rev.*, **93** (1993) 1281–306.
15. Grindea, M., Forst, T. & Hanganu, A., *Technologia vopsirii si imprimării textilelor*. Editura Technică, Bucuresti, 1983.
16. Rattee, I. D. & Breuer, M. M., *The Physical Chemistry of Dye Adsorption*. Academic Press, London, 1974.
17. Venkataraman, K. *The Chemistry of Synthetic Dyes, Vol. 2*, Academic Press, New York, 1952.
18. Náray-Szabó, G., In *Trends in Medicinal Chemistry '88*, ed. H. van der Goot, G. Dománi, L. Pallos & H. Timmerman. Elsevier Science Publishers, Amsterdam, 1989, pp. 29–41.
19. Wold, S. & Dunn III, W. J., *J. Chem. Inf. Comput. Sci.*, **23** (1983) 6–13.
20. Franke, R., *Theoretical Drug Design Methods*. Akademie—Verlag, Berlin, 1984, p. 227.
21. Simon, Z. & Bohl, M., *Quant. Struct.—Act. Relat.*, **11** (1992) 23–8.
22. McPherson, G., *Statistics in Scientific Investigation. Its Basis, Application, and Interpretation*. Springer-Verlag, New York, 1990, p. 533.