ELSEVIER

Dyes and Pigments, Vol. 32, No. 1, pp. 25–42, 1996

PII: S0143-7208(96)00001-0 0143-7208/96 $15.00 + 0.00

Lipophilicity in Dye–Cellulose Fibre Binding

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

[a]Institute of Chemistry, Romanian Academy, Bul. Mihai Viteazul 24, 1900-, Timişoara, Romania
[b]Department of Biophysics, Department of Physiology, University of Medicine and Pharmacy, P-ta E. Murgu 2, 1900-, Timişoara, Romania

(Received 6 November 1995; accepted 11 December 1995)

ABSTRACT

Lipophilicity in cellulose dyeing has been modelled by correlations of dye structural parameter with dye affinity and lipophilic dye parameters. In a series of anionic azo dyes hydrophobic interactions in dye–fibre binding only appeared in two subseries of dyes having a coupling component with one or two sulphonic groups. Good correlations with dye affinity and lipophilic parameters were obtained by the Multiple Linear Regression (MLR) Analysis and Principal-Component-Regression (PCRA) approach. Minimum Steric Difference (MTD) calculations suggest the presence of mainly steric and possible electronic interactions in the adsorption of anionic azo dyes by cellulose fibre. Similar calculations were performed for a series of disperse azo dyes, suggesting possible electronic and hydrophobic effects in dye–fibre binding. Dye lipophilic properties also depend on polarity and bulk terms. Copyright © 1996 Elsevier Science Ltd

INTRODUCTION

Hydrophobicity interactions have been studied as a result of an apolar solute transfer in an aqueous medium,[1] being mainly of entropic nature. The equilibrium state of a system made of different molecule types is determined not only by the interaction enthalpies, but also by the enthalpy of that state, as has been mentioned by Zollinger.[2] It is supposed that hydrophobic parts of molecules dissolved in water cause an ice-like structure for water molecules

*Corresponding author.

by the closest proximity of hydrophobic molecules. Dye–dye aggregation, and also dye–polymer binding, is favoured by entropy effects associated to this water recognition.

When a dried fibre is dipped in water, it usually swells, except for hydrophobic fibres.[3] The wetting assumes rupture of some cohesion forces, yielding an increased average distance between the molecular chains. The mechanism of this process would result in water penetration of the polymer structure, followed by competition between fibre–fibre and fibre–water linkages. In cellulose dyeing, alternative weak binding forces (e.g. dispersive ones) occur because a hydration layer exists around the cellulose fibres.[3] Surface-active type dyes are therefore needed; these are characterised by an ability to concentrate at the water–cellulose interface, and are also capable of strong aggregation in the concentrated surface phase. Since most dyes have surface active properties, the latter factor is considered the most significant.[3]

Rattee[3] considers that the hydrophobic effects are entirely related to water structure, which is affected by the presence of ionic centres, and that the relative position of molecular hydrophobic and ionic regions could have considerable effects in the dyeing process.

Bach *et al.*[4] consider that dyes with substantivity are not chemically bonded to cellulose but that they aggregate in the fibre micelles. Dye molecules in the monomolecular state migrate into the textile fibre and are also eliminated from it in this form.

The lipophilic octanol–water partition coefficient (log P) and the hydrophobic substituent parameter (π-Hansch) are the most important and widely used physicochemical parameters in quantitative structure–activity relationships (QSAR).[5] The physicochemical significance of these parameters has been examined extensively, and it has been stated that these parameters could be expressed in terms of volumes, shapes, electronic and also hydrogen bonding factors. We have observed that only a few such quantitative correlations have been reported in the literature, as noted in the following.

Moriguchi related log P values to molecular volume and hydrophilic group effect (E_s), Kamlet correlated log P with the solvatochromic parameters π^* and β and Franke (as quoted in Ref. 6) expressed log P values as dependent on solute bulk, polar and hydrogen bonding effects. Ou *et al.*[6] correlated log P values with physicochemical parameters such as log MW (molecular weight), μ (dipole moment) and a hydrogen bonding parameter HB_2.

Lien[7] correlated log P and log t_R (retention time in HPLC) with log MW, the hydrogen bond forming ability of the substituents, and the substituent group dipole moment (μ).

It has been stated that lipophilicity can be expressed by a volume or cavity term accounting for hydrophobic and dispersion forces, and polarity terms which express electrostatic interactions.[8] The bulk (or cavity) term can be

described by the molar volume or molar refractivity;[9] the polarity terms are more difficult to express.

The lipophilic index R_M is defined[8] as in eqn (1)

$$R_M = \log(1/R_F - 1) \tag{1}$$

where R_F is the ratio of the distances covered by the solute and the solvent, derived from thin-layer chromatography and lipid-impregnated paper, and has been taken into account in order to express the lipophilicity.

R_M is considered to be proportional to the free energy accompanying the transfer of a mole of substance from a mobile phase to a stationary one under equilibrium conditions.[10] It has been found that R_M values derived from paper chromatography can be correlated to dye affinity for cellulose fibres.

Considering dye adsorption by a textile fibre as being similar to biological ligand–receptor interaction, dye affinity for a fibre (the Gibbs free energy variation versus the dye content increase in the fibre phase) has been correlated to structural dye parameters. For a series of 46 anthraquinone vat dyes, dye affinity was related to electronic, steric and hydrophobic dye structural parameters.[11] In the final proposed models, besides other parameters, the hydrophobic ones (expressed by the octanol–water partition coefficient, calculated for the entire molecule and for the longest dye molecule axis) were reported.

A Principal-Component-Regression Analysis has been applied to a series of 49 anthraquinone vat dyes for cellulose fibres in order to express dye structure–fibre affinity relationships.[12] The octanol–water partition coefficient calculated for the longest dye molecule axis was found to make an important contribution to the second principal component resulting from PCRA calculations, and which appeared to be a measure of the hydrophobic effects in dye–fibre binding.

The objective of this paper is to evaluate lipophilicity in cellulose dyeing, as exemplified by hydrophobic dye–fibre interactions and also by lipophilic dye properties, using QSAR techniques.

METHODS

Multiple Linear Regression (MLR) Analysis

Multiple Linear Regression Analysis relates[13] one experimental variable (y_k) to one or several structural variables (x_i) by eqn (2)

$$y_k = b_o + \sum_i b_i \cdot X_{ik} + e_k \tag{2}$$

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

MLR calculations were performed with the MASCA package[14] at a general alpha level of 1% or less. As test statistics (TS), the largest root criterion was used as local criteria of identifying and choosing the most important individual parameter at a significance level of 5% or less. High leverage points were examined by the main diagonal elements of the so-called hat matrix, the influential points were tested by the likelihood function distance criterion and the outliers were tested by the externally Studentized residual.[10,11] The leave-one-out (similar to the leave-*n*-out[15]) cross validation procedure was applied in order to verify the reliability of our results. In this procedure one molecule is held out from the set, the correlation equation is computed for the rest of the molecules and the result is used to calculate the estimated affinity of the left-out molecule. These estimated affinities are then compared with the respective experimental values and 'predictive r^2' ($r2_{CV}$) thus obtained.

Principal-Component-Regression Analysis (PCRA)

The Principal-Component-Regression Analysis is based on a spectral decomposition of the correlation matrix (H) of the regressors X_l (l = 1, 2,..., c).[16,17] On the basis of this orthogonal transformation, principal components of the regressors are estimated and used instead of the original regressors. The ordinary least-squares regression is applied to estimate the regression matrix of the regressands Y_i (i = 1, 2,..., N) and component regressors Z_f (f = 1, 2,..., $s \leqslant c$) and finally, a simple re-transformation (rectification) lead to the desired estimators. The general PCRA equation is represented by eqn (3):

$$Y = h_0 + Z \cdot H \tag{3}$$

where $h_0 = \bar{y}$ is the intercept vector, which is equal to the vector $\bar{y}$ of the mean values of the regressands Y_i and Z is the matrix of the principal components. The PCRA calculations were performed with the MASCA package[14] at a general alpha level of 1% or less. Largest root criterion was used as local criteria of identifying and choosing the most important individual parameter at a significance level of 5% or less.

The Minimum Steric Difference (MTD) method

The hypermolecule which resulted from the MTD (Minimum Steric Difference) method[18] can be considered as a topological network. It represents the fibre receptor and is obtained by approximate non hydrogen atom per atom superposition of the whole set of molecules $M_i = 1, 2, \ldots, N$. The resultant vertices $j = 1, 2, \ldots, M$ of the hypermolecule correspond to the positions of these atoms. If molecule M_i occupies the vertex j, $x_{ij} = 1$ and 0 if this vertex is not occupied. The minimal steric difference MTD of the molecule i with respect to the receptor is calculated by the equation:

$$MTD_i = s + \sum \epsilon_j \cdot X_{ij} \tag{4}$$

with $\epsilon_j = -1$, 0 or $+1$ for the vertices attributed to the receptor cavity, exterior vertices or receptor walls and s the total number of cavity vertices.

The cross validation-like procedure[19] was applied in order to validate the proposed model and, also, a leave-one-out cross validation procedure applied for the MTD method similarly to that presented in the section on MLR analysis. The MTD calculations were carried out with a program developed from that of Ref. 18.

Definition of parameters used

The dye affinity (A), which expresses the dyeing driving force, was used in the same way as the biological activity from QSAR calculations as regressand. Affinity values were taken for the anionic azo dyes from Ref. 20 and for the disperse azo dyes, in accordance to Ref. 21 in kJ/mole.

The paper chromatographic R_M values were used as lipophilic parameter for the series of anionic azo dyes[20] and the sum of π-Hansch substituent terms ($\Sigma\pi$)[22] for a series of disperse azo dyes.[21]

The number of donor–acceptor hydrogen bonding groups (n_H) was estimated by the number of amine nitrogen, carbonyl oxygen and hydroxy oxygen atoms. Hydrogen bonding was also quantified through the number of H^+-donating hydrogen bonding groups (σ_D) and by the proton acceptor groups (σ_A).

The MTD parameter was estimated for the planar structure of the dye molecules (see the section on MTD results).

Quantum-chemical parameters, like HOMO and LUMO orbital energies (E_{HOMO} and E_{LUMO}) (in β units), were used in order to study possible donor–acceptor interactions with the textile fibre. For the series of anionic azo dyes, the donor electron dye ability (ϕ_0), the sum of the total π electron

substituent charges (ΣQ), E_{HOMO} and E_{LUMO}, calculated by the Hückel method were taken from Ref. 23.

As sterical parameters, for the series of disperse azo dyes, the sum of substituent molar refractivity (ΣMR) and the sum of L and B_4 substituent Verloops STERIMOL parameters (ΣL and ΣB_4) and the sum of substituent steric constants by Charton (ΣE_{S-V}) were used according to Ref. 22.

Another electronic parameter used in the correlations of the disperse azo dyes is the Hammett constant,[22] calculated as the sum of the substituent contributions ($\Sigma\sigma$).

The van der Waals volumes (V_W) were calculated by the additivity of van der Waals volume increments.[24] For the sulphonic and azo groups, the increment values were taken from Ref. 25; for $-CH_2-CH_2-OH$ side groups attached to the amine nitrogen, increment values were appreciated in accordance to Ref. 26.

Structural dye parameters are presented for the series of anionic azo dyes in Table 1 and for disperse azo dyes in Table 2.

RESULTS AND DISCUSSION

Correlations with dye affinity

Fisichella *et al.* found good correlations between the chromatographic R_M value and the dye affinity values for a subseries of six anionic azo dyes having the same coupling component.[20]

We obtained good statistical results by correlation with dye affinity for a series of 12 anionic azo dyes (compounds 1, 2, 3, 4, 5, 7, 9, 12, 14, 15, 24 and 26 from Table 1):

$$\hat{A} = 25.83 + 19.59(\pm 8.71) R_M$$

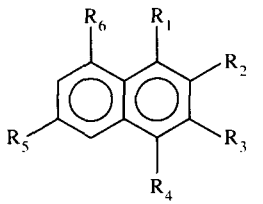
$$N = 12\, r = 0.849\, s = 2.13\, TS = 5.01\, r^2_{CV} = 0.565 \qquad (5)$$

The critical quantile equals 2.23 for the largest root criterion used as test statistics (TS). Neither high-leverage points nor outliers were observed, indicating that eqn (5) was a satisfactory model; its predictability was confirmed by the 'predictive r^2' value.

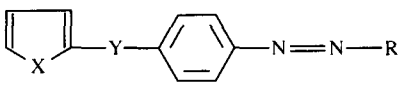
MLR analysis was applied to larger subseries of dyes in order to express the dependence of dye affinity upon R_M values. For compounds: 6, 8, 10, 11, 13, 16, 17, 18, 19, 20, 21, 22, 23, 25, 27, 28, 29 and 30 (see Table 1) the following equation was obtained:

TABLE 1

Coupling Components (I), Experimental Dye Affinities (A) and Structural Parameters of Anionic Azo Dyes (II)



I



II

No.	*X*	*Y*	R[a]	*A* (kJ/mole)	R_M	V_W (cm^3/mole)	ϕ_o (grd)	n_H	E_{HOMO} (ß units)	E_{LUMO} (ß units)
1	-S-	-CH=CH-	N	15.80	-0.52	203.14	40.80	2.00	0.26	0.32
2	-CH=CH-	-CH=CH-	N	14.25	-0.63	208.46	36.62	2.00	0.27	0.39
3	-S-	-CONH-	N	13.08	-0.72	205.51	38.89	5.00	0.32	0.42
4	-CH=CH-	-CONH-	N	12.00	-0.79	210.83	38.92	5.00	0.32	0.42
5	-S-	-CH=CH-	N-W	9.66	-0.83	199.59	38.70	2.00	0.25	0.33
6	-S-	-CH=CH-	H	9.45	0.31	229.57	40.31	4.00	0.26	0.32
7	-CH=CH-	-CH=CH-	N-W	9.20	-0.83	204.91	33.92	2.00	0.25	0.42
8	-S-	-CONH-	H	9.03	0.29	232.76	39.20	7.00	0.32	0.41
9	-S-	-CO-	N	8.78	-1.00	197.65	28.67	4.00	0.22	0.46
10	-CH=CH-	-CH=CH-	H	8.40	0.25	235.71	36.45	4.00	0.26	0.39
11	-CH=CH-	-CONH-	H	8.28	0.23	238.08	38.51	7.00	0.31	0.41
12	-S-	-CONH-	N-W	7.15	-0.87	201.96	35.65	5.00	0.31	0.47
13	-S-	-CH=CH-	C	7.06	0.14	226.84	13.60	4.00	0.25	0.32
14	-CH=CH-	-CO-	N	7.02	-1.12	202.97	28.66	4.00	0.22	0.46
15	-CH=CH-	-CONH-	N-W	6.52	-0.91	207.28	35.66	5.00	0.31	0.47
16	-S-	-CH=CH-	R′	6.27	0.12	222.37	37.50	2.00	0.23	0.32
17	-S-	-CONH-	C	6.23	0.12	229.21	37.50	7.00	0.31	0.44

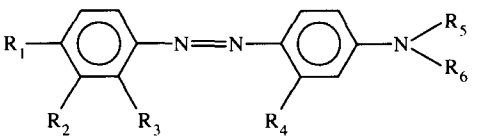
continued

TABLE 1—*continued*

No.	*X*	*Y*	R[a]	*A* (kJ/mole)	R_M	V_W (cm^3/mole)	ϕ_o (grd)	n_H	E_{HOMO} (ß units)	E_{LUMO} (ß units)
18	-CH=CH-	-CH=CH-	C	6.02	0.12	231.34	35.23	4.00	0.26	0.40
19	-CH=CH-	-CH=CH-	R′	5.81	0.10	227.69	32.85	2.00	0.23	0.40
20	-CH=CH-	-CONH-	C	5.18	0.09	234.53	37.52	7.00	0.31	0.44
21	-S-	-CONH-	R′	5.10	0.07	224.74	34.30	5.00	0.28	0.45
22	-S-	-CO-	H	4.64	0.09	224.90	29.09	6.00	0.21	0.44
23	-CH=CH-	-CONH-	R′	4.26	0.04	230.06	34.26	5.00	0.28	0.45
24	-S-	-CO-	N-W	4.22	-0.95	194.10	25.70	4.00	0.20	0.51
25	-CH=CH-	-CO-	H	4.10	0.04	230.25	29.09	6.00	0.21	0.44
26	-CH=CH-	-CO-	N-W	4.05	-0.95	199.42	25.60	4.00	0.20	0.51
27	-S-	-CO-	C	3.85	0.05	221.35	27.46	6.00	0.21	0.47
28	-CH=CH-	-CO-	C	3.43	0.04	226.67	27.41	6.00	0.21	0.47
29	-S-	-CO-	R′	3.22	-0.02	216.88	24.60	4.00	0.19	0.50
30	-CH=CH-	-CO-	R′	2.84	-0.04	222.20	24.37	4.00	0.19	0.51

[a]R represents the coupling component: N = naphthionic acid (I, $R_1 = NH_2$, $R_2 = H$, $R_3 = H$, $R_4 = SO_3H$, $R_5 = H$, $R_6 = H$); N-W = Nevile-Winter's acid (I, $R_1 = OH$, $R_2 = H$, $R_3 = H$, $R_4 = SO_3H$, $R_5 = H$, $R_6 = H$); R′ = R-acid (I, $R_1 = H$, $R_2 = OH$, $R_3 = SO_3H$, $R_4 = H$, $R_5 = SO_3H$, $R_6 = H$); H = H-acid (alkaline coupled) (I, $R_1 = NH_2$, $R_2 = H$, $R_3 = SO_3H$, $R_4 = H$, $R_5 = SO_3H$, $R_6 = OH$); C = chromotropic acid (I, $R_1 = OH$, $R_2 = H$, $R_3 = SO_3H$, $R_4 = H$, $R_5 = SO_3H$, $R_6 = OH$).

TABLE 2
Structural Parameters of Disperse Azo Dyes



No.	R_1	R_2	R_3	R_4	R_5	R_6	A (kJ/mole)	$\Sigma\pi$	$\Sigma\sigma$	ΣMR	σ_A
1	NO_2	-	-	-	C_2H_5	C_2H_4OH	11.06	-0.03	0.89	29.50	2.00
2	NO_2	-	-	CH_3	C_2H_4OH	C_2H_4OH	10.89	-1.26	1.03	36.69	4.00
3	NO_2	-	-	-	-	-	9.73	-0.28	0.78	7.36	0.00
4	Br	-	-	-	C_2H_4OH	C_2H_4OH	9.32	-0.68	0.65	32.56	4.00
5	NO_2	-	-	-	C_2H_4OH	C_2H_4OH	8.15	-1.82	1.20	31.04	4.00
6	Cl	-	-	-	C_2H_4OH	C_2H_4OH	8.15	-0.83	0.65	29.71	4.00
7	-	Cl	-	-	C_2H_4OH	C_2H_4OH	7.83	-0.83	0.79	29.71	4.00
8	-	-	-	-	C_2H_5	C_2H_4OH	7.27	0.25	0.11	22.14	2.00
9	-	-	-	CH_3	C_2H_4OH	C_2H_4OH	6.48	-0.98	0.35	29.33	4.00
10	CN	-	-	-	C_2H_4OH	C_2H_4OH	6.46	-2.11	1.08	30.01	4.00
11	-	NO_2	-	-	C_2H_4OH	C_2H_4OH	6.37	-1.82	1.13	31.04	4.00
12	OCH_3	-	-	-	C_2H_4OH	C_2H_4OH	6.14	-1.56	0.15	31.55	4.00
13	-	-	-	-	-	C_2H_4OH	6.03	-0.77	0.21	11.84	2.00
14	-	CH_3	-	-	C_2H_4OH	C_2H_4OH	5.94	-0.98	0.35	29.33	4.00
15	CH_3	-	-	-	C_2H_4OH	C_2H_4OH	5.93	-0.98	0.25	29.33	4.00
16	F	-	-	-	C_2H_4OH	C_2H_4OH	5.69	-0.98	0.29	29.33	4.00
17	-	-	CH_3	-	C_2H_4OH	C_2H_4OH	5.69	-1.40	0.48	24.60	4.00
18	-	-	-	-	-	-	5.29	0.00	0.00	2.06	0.00
19	-	-	-	-	C_2H_4OH	C_2H_4OH	4.61	-1.54	0.42	23.68	4.00
20	-	-	NO_2	-	C_2H_4OH	C_2H_4OH	3.14	-1.82	1.37	31.04	4.00

$$\hat{A} = 3.46 + 20.19(\pm 2.08)R_M$$

$$N = 18\, r = 0.979\, s = 0.40\, TS = 20.56\, r^2_{CV} = 0.952 \qquad (6)$$

The critical quantile for the largest root criterion test statistics was $c_0 = 2.12$. Neither high-leverage points, nor influential points or outliers were detected. Equation (6) indicates a good predictive model.

Equations (5) and (6) show that the lipophilicity, as expressed by the R_M values, has an important contribution to the affinity for cellulose of dyes having coupling components containing one or two sulphonic acid groups.

For this last subseries of 18 dyes, dye affinity values were related to various structural parameters: R_M, ϕ_o, n_H, MTD, V_W, E_{HOMO} and E_{LUMO}. After a regressor selection based on the largest root criterion, the regressor eqn (7) and the rectified regression eqn (8) were obtained:

$$\hat{A} = 5.65 + 1.81(\pm 0.49)Z_1 - 0.78(\pm 0.49)Z_2$$

$$N = 17\, r = 0.949\, s = 0.71\, TS_1 = 10.15\, TS_2 = 4.36 \qquad (7)$$

The critical quantile for the largest root criterion test statistics was $c_0 = 2.73$.

$$\hat{A} = -2.08 + 8.08R_M + 0.16\phi_0 - 0.26n_H + 11.01E_{HOMO} \qquad (8)$$

PCRA applied to the series of 18 dyes gave: $r^2 = 0.79$ and $s = 0.99$, indicating compound 13 as outlier (checked by the leave-one-out resampling technique).

Equation (7) describes a robust model with predictive power. Both eqns (6) and (7) indicate good correlations with dye affinity and their predictability was checked by the 'leave-one-out' cross validation procedure (in the MLR analysis), respectively, using the rules outlined for the PCRA approach.[17]

Correlations between dye affinity and R_M values were performed for a series of 30 anionic azo dyes, but weaker statistical results were obtained ($r^2 = 0.10$ and $s = 3.21$). Attempts to correlate, with the PCRA approach, dye affinity with the following dye structural parameters: R_M, ϕ_0, n_H, MTD, V_W, E_{HOMO}, E_{LUMO} and ΣQ, for the same series of dyes, yielded the regression eqn (14) (see the section on MTD results) after a regressor selection procedure based on the largest root criterion as test statistic.

For the series of 20 disperse azo dyes (see Table 2), affinity values were related to the following structural parameters: $\Sigma\pi$, $\Sigma\sigma$, ΣMR, $\Sigma E_{S\text{-}V}$, ΣL and ΣB_4. Following a regressor selection procedure based on the largest root criterion (at a significance level of 5% or less), MLR analysis yields a robust predictive model:

$$\hat{A} = 6.66 + 2.07(\pm 1.13) \sum \pi + 4.53(\pm 1.96) \sum \sigma$$

$$N = 19\, r = 0.854\, s = 1.03\, TS_1 = 4.93\, TS_2 = 6.22\, r^2_{CV} = 0.590 \quad (9)$$

The critical quantile for the largest root criterion was $c_0 = 2.70$; neither high-leverage points, nor influential points or outliers were detected.

Equation (9) expresses a model without multicollinearity, in which the dye affinity is expressed as a function of hydrophobicity and electronic effects. MLR analysis applied to the whole series (20 compounds) yielded weaker statistical results ($r^2 = 0.450$, $s = 1.61$), indicating compound 20 as outlier.

Correlations with lipophilic parameters

R_M-values, as a measure of dye lipophilicity, were related to dye structural parameters for the series of anionic azo dyes.

For the subseries of 12 dyes (see the section on correlations with dye affinity), R_M values were correlated with the following structural parameters: ϕ_0, MTD, V_W, E_{HOMO}, E_{LUMO}. After a regressor selection procedure, based on the largest root criterion as test statistics for selecting the most important individual parameters, the following MLR equation was obtained:

$$R_M = -1.64 + 0.02(\pm 0.01)\phi_0$$

$$N = 12\, r = 0.775\, s = 0.11\, TS = 3.88\, r^2_{CV} = 0.402 \quad (10)$$

The critical quantile for the largest root criterion equals 2.23. Neither high-leverage points nor outliers were detected. It seems that the lipophilic term is expressed, in this case, by polarity effects. Correlations between R_M values and dye structural parameters were also performed for the subseries of 18 anionic azo dyes (see the section on correlations with dye affinity). A regressor selection procedure based on the largest root criterion (with the critical quantile $c_0 = 2.12$) for choosing the most important individual parameter was applied and the following equation was obtained:

$$R_M = 0.42 - 0.04(\pm 0.01)MTD$$

$$N = 18\, r = 0.831\, s = 0.06\, TS = 5.90\, r^2_{CV} = 0.602 \quad (11)$$

Neither high-leverage points nor outliers were detected. Statistical results indicate a satisfactory predictive model. As can be seen from eqn (11), the lipophilic R_M chromatographic parameter for this subseries strongly depends

upon steric effects, which are expressed by the MTD parameter (see the section on MTD results).

For the series of 20 disperse azo dyes (see Table 2) lipophilic dye effects were expressed by the sum of π-Hansch substituent terms ($\Sigma\pi$), which were related to different structural parameters. With the following set of initial structural parameters: V_W, $\Sigma\sigma$, ΣMR, MTD,[27] σ_A and σ_D, a Principal-Component-Regression Analysis was applied to the series of 20 dyes after a regressor selection based on the largest root criterion as test statistics. The following component regressor (12) and rectified (13) regression equations were obtained, with a critical quantile $c_0 = 2.10$ for the largest root criterion used as test statistics (TS):

$$\sum \pi = -1.02 - 0.46(\pm 0.24)Z_1$$

$$N = 20\, r = 0.693\, s = 0.49\, TS = 4.11 \tag{12}$$

$$\sum \pi = 0.27 - 0.15\sigma_A - 0.33 \sum \sigma - 0.02 \sum MR \tag{13}$$

Neither high-leverage points nor outliers were detected. Evaluation of the non least-squares regression eqn (13) for lipophilicity suggests the inclusion of both polarity effects, as expressed by the sum of the Hammett σ-constants, and the number of proton acceptor groups, which could be involved in hydrogen bonding, as well as of bulk parameters, as expressed by the sum of the substituent molar refractivity.

MTD results

MTD calculations were applied to the series of anionic azo dyes in order to model dye–cellulose fibre interactions. The hypermolecule obtained from molecule superposition for the 30 anionic azo dyes, considered planar, is depicted in Fig. 1.

The vertex attributions are presented in Table 3. The start standard was chosen on the basis of chemical reasons:

$$S_o \begin{cases} j(\epsilon = 0) : 1 - 4,\ 6 - 13,\ 23,\ 26 - 31 \\ j(\epsilon = +1) : 5,\ 14 - 22,\ 24,\ 25 \end{cases}$$

The optimization procedure was applied by trial and error, and verified as optimal by the regression coefficient criterion;[18] the following optimized receptor map and correlation equation are thus obtained:

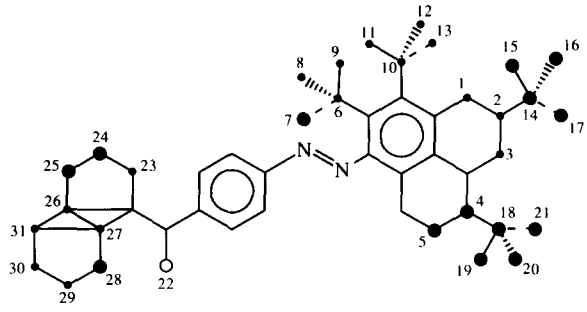


Fig. 1. Hypermolecule derived from MTD calculations with lipophilic R_M values as parameter with numbering of vertices for anionic azo dyes. Beneficial vertices ($\epsilon_j = -1$) are marked by white circles, detrimental vertices ($\epsilon_j = +1$), by black circles, irrelevant vertices ($\epsilon_j = 0$), by dots.

$$S*_1 \begin{cases} j(\epsilon = -1) : 7 \\ j(\epsilon =) : 1-4,\ 6,\ 8-13,\ 28-31 \\ j(\epsilon = +1) : 5,\ 14-27 \end{cases}$$

$$\hat{A} = 13.657(\pm 0.938) - 1.020(\pm 0.136)MTD$$

$$N = 30\ r = 0.817\ s = 1.960\ F = 56.067\ r^2_{CV1} = 0.527 \qquad r^2_{CV2} = 0.560 \tag{14}$$

The final MTD values are listed in Table 3. The 'predicted r^2' values obtained from the 'even-odd' (r^2_{CV1}) and leave-one-out (r^2_{CV2}) cross validation procedures indicate an acceptable predictability of the proposed model. In the cross validation-like ('even–odd') procedure, the dye molecules were numbered in increasing affinity order and separated into two subseries: ODD, for $i = 1, 3, 5, \ldots, 29$ and EVEN, for $i = 2, 4, 6, \ldots, 30$. Four inversions: $i = 3/4$, $i = 10/11$, $i = 13/14$ and $i = 28/29$ were performed for a balanced vertex distribution in both subseries. As start standard S^0-vertex attribution was used.

For the ODD and EVEN subseries, the following optimized receptor maps (S_{ODD}, respectively, S_{EVEN}) and correlation equations were obtained:

$$S_{ODD} \begin{cases} j(\epsilon = 0) : 2-4,\ 6-14;\ 22,\ 28-31 \\ j(\epsilon = +1) : 1,\ 5,\ 15-21,\ 23-27 \end{cases}$$

TABLE 3
MTD Results[a]

No.[b]	$j(x_{ij} = 1)$	$\hat{A}^*$	MTD_0	MTD*	MTD_1	Cross $\hat{A}_1$	valid. MTD_2	$\hat{A}_2$
1	1-3, 10-13, 27-31	11.62	0	2	2	11.23	3	11.23
2	1-3, 10-13, 26-31	10.59	0	3	3	10.76	3	10.18
4	1-3, 10-13, 23, 26-31	9.58	0	4	3	10.20	6	9.42
3	1-3, 10-13, 23, 27-31	10.59	0	3	3	10.76	4	10.29
5	1-3, 10-13, 27-31	11.62	0	2	2	11.23	3	12.05
6	1-4, 6-9, 14-17, 27-31	8.56	4	5	5	8.26	5	8.69
7	1-3, 10-13, 26-31	10.59	0	3	2	11.23	3	10.73
8	1-4, 6-9, 14-17, 23, 27-31	7.54	4	6	6	7.01	7	7.29
9	1-3, 10-13, 22-26	7.54	3	6	7	6.09	6	6.64
11	1-4, 6-9, 14-17, 23, 26-31	6.52	4	7	7	5.76	8	6.17
10	1-4, 6-9, 14-17, 26-31	7.54	4	6	5	8.15	6	7.48
12	1-3, 10-13, 23, 27-31	10.59	0	3	3	10.76	3	10.93
14	1-3, 10-13, 22-27	6.52	3	7	8	5.06	8	6.32
13	1-4, 6-9, 14-17, 27-31	8.56	4	5	5	8.26	5	8.79
15	1-3, 10-13, 23, 26-31	9.58	0	4	3	10.20	3	9.82
16	4-6, 10-13, 18-21, 27-31	6.52	5	7	6	7.01	6	7.63
17	1-4, 6-9, 14-17, 23, 27-31	7.54	4	6	6	7.12	6	7.86
18	1-4, 6-9, 14-17, 26-31	7.54	4	6	6	7.01	6	7.86
19	4-6, 10-13, 18-21, 26-31	5.49	5	8	5	8.15	8	5.48
20	1-4, 6-9, 14-17, 23, 26-31	6.52	4	7	7	5.76	8	6.95
21	4-6, 10-13, 18-21, 23, 27-31	5.49	5	8	6	7.12	9	5.46

continued

TABLE 3—*continued*

22	1-4, 6-9, 14-17, 22-26	4.48	7	9	8	4.51	11	4.21
23	4-6, 10-13, 18-21, 23, 26-31	4.48	5	9	6	7.12	9	5.12
24	1-3, 10-13, 22-26	7.54	3	6	5	8.26	6	7.65
25	1-4, 6-9, 14-17, 22-27	3.46	7	10	11	1.98	11	3.05
26	1-3, 10-13, 22-27	6.52	3	7	6	7.01	5	7.39
27	1-4, 6-9, 14-17, 22-26	4.48	7	9	10	3.01	9	4.52
28	1-4, 6-9, 14-17, 22-27	3.45	7	10	11	1.98	10	3.35
29	4-6, 10-13, 18-26	2.44	8	11	9	3.26	11	2.98
30	4-6, 10-13, 18-27	1.42	8	12	10	2.01	13	1.09

[a]MTD_0-initial MTD values with S^0 as standard; MTD^*-values calculated with the optimized receptor map S^* as standard; $\hat{A}^*$ values calculated with eqn (16); $\hat{A}_1$, MTD_1-values calculated with the 'even–odd' cross validation procedure; $\hat{A}_2$, MTD_2-values calculated with the leave-one-out cross validation procedure.

[b]Compound numbering is in agreement with the inversions mentioned in the section on MTD results.

$$\hat{A}_{\mathrm{ODD}} = 14.510 - 1.250MTD \tag{15}$$

$$N = 15\ r = 0.861$$

$$S_{\mathrm{EVEN}}\begin{cases} j(\epsilon = -1) : 4, 28, 29 \\ j(\epsilon = 0) : 1-3, 5-13, 25, 30, 31 \\ j(\epsilon = +1) : 14-25, 27 \end{cases}$$

$$\hat{A}_{\mathrm{EVEN}} = 13.286 - 1.028MTD \tag{16}$$

$$N = 15\ r = 0.812$$

With the MTD method one can assess especially the behaviour of the substituents attached to a parent compound, assumed to include an attractive region for the receptor site. As has already been mentioned, the MTD results include not only steric, but also polarity and possible hydrophobic effects. In order to have a better insight into the possible hydrophobic interactions in dye–cellulose binding and to improve the correlation with dye affinity, besides MTD, the lipophilic R_M values were introduced as a supplemental parameter. The same start standard (S°) as in previous calculations was used. By the same optimization procedure, the following optimized receptor map (see also Fig. 1) and regression equation were obtained:

$$S*_2\begin{cases} j(\epsilon = -1) : 22 \\ j(\epsilon = 0) : 1-3, 6, 8-13, 23, 26, 27, 29-31 \\ j(\epsilon = +1) : 4, 5, 7, 14-21, 24, 25, 28 \end{cases}$$

$$\hat{A} = 33.512(\pm 2.340) + 19.797(\pm 1.962)R_M - 3.753(\pm 0.324)MTD \tag{17}$$

$$N = 30\ r = 0.922\ s = 1.338\ F = 76.669\ r^2_{\mathrm{CV1}} = 0.647\ r^2_{\mathrm{CV2}} = 0.796$$

From inspection of the optimized receptor map, the presence of a single cavity vertex is evident, in addition to many irrelevant and detrimental vertices, especially in positions occupied by bulky substituents (such as sulphonic acid groups). Mainly steric, rather than hydrophobic effects, seem to be important in dye–cellulose fibre interactions for the series of the 30 anionic azo dyes investigated. Cavity vertices presumed to characterize the parent compound and vertex 22 indicate the binding region along the

molecule axis, as has been observed in a previous study of anthraquinone dye adsorption by cellulose fibres.[11]

Statistical results indicate a good correlation with dye affinity and acceptable predictability of the MTD proposed models for this series of dyes.

CONCLUSIONS

Hydrophobicity in azo dye–cellulose binding and dye lipophilic properties were modelled by QSAR techniques through correlations of structural parameters with dye affinity and dye lipophilicity parameters.

Adsorption of anionic azo dyes by cellulose seem to depend mainly on steric, and to a lesser extent, on electronic interactions. Hydrophobic effects are involved only in a subseries of dyes having a coupling component with one or two sulphonic acid groups. Dye lipophilicity can be correlated with electronic or steric parameters. The MTD calculations indicate similar results for the series of anionic azo dyes: in dye–fibre binding, mainly steric interactions are involved and, perhaps, electronic effects. The latter could express the hydrogen bonding ability in dye–cellulose interactions.

The disperse azo dyes are involved not only in hydrophobic, but also in electronic interactions, in dye–fibre binding; their lipophilicity correlates well with polarity and bulk terms.

ACKNOWLEDGEMENTS

The authors thank Dr Peter P. Mager from the University of Leipzig (Germany) for the permission of access to the MASCA package and for useful suggestions throughout this work.

REFERENCES

1. Dill, K. A., *Science*, **250** (1990) 297.
2. Zollinger, H. *Color Chemistry. Synthesis, Properties and Applications of Organic Dyes and Pigments*, pp. 215–233. VCH, Weinheim, 1987.
3. Rattee, I. D., *J. Soc. Dyers Colour*, **90** (1974) 367.
4. Bach, H., Pfeil, E., Philipper, W. & Reich, M., *Angew. Chem.*, **75** (1963) 407.
5. Yang, G., Lien, E. J. & Guo, Z., *Quant. Struct. Act. Relat.*, **5** (1986) 12.
6. Ou, X., Ouyang, Y. & Lien, E. J., *J. Mol. Sci. (Wuhan, China)*, **4** (1986) 89.
7. Lien, E. J., Gao, H. & Prabhakar, H., *J. Pharm. Sci.*, **80** (1991) 517.
8. Van de Waterbeemd, H. and Testa, B. In: *Advances in Drug Research*, ed. B. Testa, chapter 16, pp. 85–225. Academic Press, London, 1987.

9. Gryllaki, M., Van de Waterbeemd, H., Testa, B., Tayar, N. E., Mayer, J. M. & Carrupt, P. A., *Int. J. Pharm.*, **51** (1989) 95.
10. Alberti, G. & De Giorgi, M. R., *Ann. Chim. (Rome)*, **73** (1983) 315.
11. Timofei, S., Schmidt, W., Kurunczi, L., Simon, Z. & Salló, A., *Dyes and Pigments*, **24** (1994) 267.
12. Timofei, S., Kurunczi, L., Schmidt, W., Fabian, W. M. F. & Simon, Z., *Quant. Struct. Act. Relat.*, in press.
13. Wold, S. & Dunn, III, W. J., *J. Chem. Inf. Comput. Sci.*, **23** (1983) 6.
14. Mager, P. P., Rothe, H., Mager, H. & Werner, H. In: *QSAR in Design of Bioactive Compounds*, pp. 131–182. J. R. Prous Science Publishers, 1992.
15. Franke, R., *Theoretical Drug Design Methods*, p. 227. Akademie-Verlag, Berlin, 1984.
16. Mager, P. P. & Rothe, H., *Pharmazie*, **45** (1990) 758.
17. Mager, P. P., *Med. Res. Rev.*, **14** (1994) 533.
18. Simon, Z., Chiriac, A., Holban, S., Ciubotariu, D. & Mihalas, G. I., *Minimum Steric Difference*. Res. Stud. Press, Letchwoorts and John Wiley, New York, 1984.
19. Ciubotariu, D., Deretey, E., Oprea, T. I., Sulea, T., Simon, Z., Kurunczi, L. & Chiriac, A., *Quant. Struct. Act. Relat.*, **12** (1993) 367.
20. Fisichella, S., Scarlata, G. & Torre, M., *J. Soc. Dyers Colour*, **94** (1978) 521.
21. Shibusawa, T. & Uchida, T., *Sen'i Gakkaishi*, **42** (1986) 70.
22. Hansch, C. & Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*. John Wiley, New York, 1979.
23. Osik, Yu. I., Kachkovskij, A. D. & Saribekov, G. S., *Ukrain. Khim. Zh.*, **51** (1985) 1071.
24. Bondi, A., *J. Phys. Chem.*, **68** (1964) 441.
25. Shibusawa, T., *Sen'i Gakkaishi*, **43** (1987) 401.
26. Ito, T. & Seta, J., *Sen'i Gakkaishi*, **38** (1986) 71.
27. Timofei, S., Kurunczi, L., Schmidt, W. & Simon, Z., *Dyes and Pigments*, in press.