STUDIA UNIVERSITATIS BABEȘ-BOLYAI, CHEMIA, XLVII, 1-2, 2002

Dedicated to Professor Ionel Haiduc
on the occasion of his 65th birthday

STUDY OF LIPOPHILICITY OF SOME 4,4'-DIAMINOBENZANILIDE BASED DIRECT DYES IN THE CELLULOSE DYEING

SIMONA FUNAR-TIMOFEI[*], **GEORGETA SIMU**, **EUGENIA SĂRĂNDAN**

** Institute of Chemistry, Romanian Academy, Bul. Mihai Viteazul 24,*
1900 Timisoara, Romania

ABSTRACT. This paper presents an experimental evaluation of chromatographic mobilities i.e. R_f values of a series of 16 direct dyes containing 4,4'-diaminobenzanilide as a middle component. The R_f values were obtained by reverse-phase thin-layer chromatography and were quantitatively analysed by MLR (Multiple Linear Regression) analysis. Good correlation with the scores of the R_f values obtained by principal component analysis and predictability were noticed. Hydrophobicity of the direct dyes influences significantly their lipophilic behaviour.

Keywords: direct dyes, RP-TLC (reverse-phase thin-layer chromatography), lipophilicity, PCA (Principal Component Analysis), MLRA (Multiple Linear Regression Analysis)

INTRODUCTION

Lipophilicity is an important characteristic of dyestuffs, which can affect their affinity for the textile substrates [1-2]. Other interactions also influenced by lipophilicity include dye-dye aggregation [3-5] and dye-surfactant complexation [6,7]. Furthermore, toxicity can correlate with lipophilicity [8].

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 [9]. The polarity terms are more difficult to be expressed.

An experimental procedure used for assessing lipophilicity is the reverse-phase thin-layer chromatography (RP-TLC) [10]. Chromatographic mobility in such systems has been extensively used as a measure of the lipophilic character, though not apparently by dyestuffs or textile workers [11].

In the present work, chromatographic mobilities of a series of disazoic and trisazoic direct dyes with 4,4'-diaminobenzanilide as middle component were qualitatively analyzed by PCRA (Principal Component Regression Analysis). The experimental R_f values were obtained by RP-TLC analysis and are presented in Table 1.

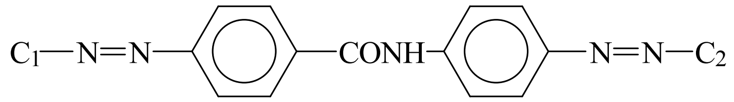
EXPERIMENTAL

Materials and RP-TLC measurements

The investigated dyestuffs were asymmetric disazoic and trisazoic direct dyes, of general formula (I).

---

[*] *Corresponding author, e-mail: timofei@acad-tim.utt.ro*



(I)

where for dye: 1-10, 12, 15, 16 C_1 represents the first coupling component (salicylic acid) and C_2 - the second coupling component, as follows: the monoazoic compound 2-amino-phenol-4-sulphoneamide→resorcine (dye 1), γ acid (dye 2), resorcine (dye 3), the monoazoic compound sulphanilic acid→2,4-diaminotoluene (dye 4), p-aminobenzoyl I acid (dye 5), benzoyl I acid (dye 6), the monoazoic compound 5-nitro-anthranilic acid→H acid (dye 7) and the monoazoic compound 5-nitro-anthranilic acid→I acid (dye 8), 2-amino-phenol-5-sulphone-amide→ resorcine (dye 9), salicylic acid (dye 10), 4 nitro-aniline→H acid (dye 12), I acid (dye 15), benzoyl γ acid (dye 16). For dye 11 the coupling components are: γ acid (C_1), respectively H acid (C_2); for dye 13 the coupling component C_1 is the monoazoic compound aniline →acid H and C_2 is m-phenylenediamine; for dye 14 the coupling component C_1 is the Schäffer acid and C_2 is the monoazoic compound 5-nitro-anthranilic acid→I acid. These dyes were synthesized and purified at the Institute of Chemistry of Timișoara [12].

The measurement of the chromatographic R_f values was performed by RP-TLC analysis. The chromatographic plates used were of Merck, DC Fertigplatten Kieselgel 60 type. They were pre-developed during 24 hours in a hexane:paraffin oil of 95:5 (v:v) mixture. The eluting solvent, which is imposed to be a binary system, was isopropanol:NH_3 25 %, at 80-60 % ratios, which varied with an increment of 5% of organic phase.

To ascertain the reproducibility of the experimental chromatographic mobilities, the dyes were each spotted ten times onto the RP-TLC plates. Average R_f values were obtained by successive experiments.

METHODS

Definition of parameters

As independent variable the scores obtained by Principal Component Analysis (PCA) from chromatographic R_f values performed by RP-TLC analysis were used. They are presented in Table 1.

Molecular dye structures were built by the ChemOffice package [13], energy minimized using the MM2 force field. The minimum energy structures found in this way for each compound were then used for energy minimization by the semiempirical continuum-solvation model COSMO-AM1 [14] and employed for further analyses. Molecular surface area (A_{COSMO}) and volume (V_{COSMO}) refered to solution-phase (COSMO-AM1) structures were calculated using MST van der Waals radii [15].

Several quantum-chemical parameters were considered, like: the maximum (QMAX) and minimum (QMIN) atom charge in the dye molecules, the atomic formal charge on the most positive hydrogen atom (which reflects the electrostatic part of the acidity term and expresses the hydrogen-bond acceptor basicity, Q+).

The difference between the maximum and minimum atomic charge in a molecule (MAXMIN) was used as a submolecular polarity parameter. The HOMO and LUMO molecular orbital energies (E_{HOMO} and E_{LUMO}) were used in order to study possible interactions of the dye molecule with the textile fibre. Also, from the same quantum-mechanical calculations, dipole moments (μ) were considered in order to express possible dipole effects in dye-fibre interactions.

Table 1.

R_f scores (t[1]) and structural descriptors of the direct azo dyes

Dye	t[1]	A_{COSMO} [$Å^2$]	E_{HOMO} [eV]	E_{LUMO} [eV]	μ [D]	MAXMIN	ClogP	MR
1	1.45	581.77	-9.14	-1.42	9.91	4.13	6.25	19.16
2	1.94	513.9	-8.49	-1.58	22.99	4.13	4.26	17.16
3	1.72	435.51	-8.85	-1.48	12.23	0.99	6.48	14.23
4	1.62	590.16	-8.35	-1.3	4.77	4.17	5.00	19.68
5	1.33	612.45	-8.51	-1.63	17.19	3.92	5.08	20.54
6	1.09	600.49	-8.53	-1.52	13.65	4.17	5.08	20.54
7	-3.09	650.36	-8.46	-1.83	17.33	4.12	4.73	22.02
8	0.32	677.25	-8.35	-2.14	15.95	4.14	1.95	23.04
9	-1.32	577.75	-8.97	-1.52	14.57	4.15	6.25	19.16
10	2.84	454.59	-8.94	-1.52	22.99	0.99	7.57	14.73
11	-5.62	605.67	-8.19	-1.76	22.29	4.20	-3.84	20.43
12	-2.42	659.32	-8.52	-2.03	13.9	4.13	2.18	22.39
13	-0.69	626.10	-8.33	-1.96	20.56	4.12	1.77	21.65
14	-3.63	696.46	-8.53	-1.82	14.78	4.14	2.61	24.08
15	-0.29	513.47	-8.53	-1.54	7.54	4.15	4.26	17.16
16	1.42	602.05	-8.93	-1.6	14.42	4.12	5.99	20.17

Connolly Solvent Accessible Surface Area (SAS), Connolly Molecular Surface Area (MS) and Connolly Solvent-Excluded Volume (SEV) [16] and the hydrophobicity (ClogP) parameter (the logarithm of octanol/water partition coefficient) were performed by the ChemOffice package.

The molecular weight (MW) and molar refraction (MR) were used as a measure of the bulkiness of the azo dyes.

Principal Component Analysis (PCA)

The purpose of PCA is to find the simplest mathematical model able to describe satisfactorily a given set of data. In the latent variable analysis, the main objective is to find the possible relations between one or more dependent variables and a potential number of explanatory variables [17]. During this analysis the data matrix for a given set of objects are examined. The variables are described as linear combinations of the new variables, called principal components (PC), which are orthogonal vectors.

The latent variables directions in the M-dimensional space are specified by the "p" vectors and the location of the scores along this axis are given by the "t" vectors. The "t" vectors are describing the direction of the principal component axis, while the components of "p" represent the cosines of the angles between the principal component and the co-ordinates axis. The levelheaded coefficients are also called the loadings of the variables over the principal component [18].

The PCA analysis was performed by the MASCA package [19].

Multiple Linear Regression (MLR) Analysis

Multiple Linear Regression Analysis relates one experimental variable y_k to one or several structural variables x_i by the equation [20]:

$$y_k = b_o + \sum_i b_i \cdot x_{ik} + e_k \quad (1)$$

where b represents partial regression coefficients and e the deviations and residuals. The leave-one-out (similar to the leave-n-out [21]) and leave-half-out [22] cross validation procedure was applied in order to verify the reliability of our results. In this procedure one, respectively half of all molecules are held out from the set, the correlation equation is computed for the rest of 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 and 'predictive r^2' (r^2_{CV}) was obtained. MLR calculations were performed by the SYSTAT package [23].

RESULTS AND DISCUSSION

The experimental R_f chromatographic values have been analyzed by the PCA method. The correlation matrix of the R_f values is presented in Table 2. Significant intercorrelations between these values are observed. Therefore the scores of the first principal component were used as independent variable in further regressional calculations. The dependence of the obtained score values for the first two principal components (t[2] versus t[1]) is presented in figure 1.

Table 2.

Correlation matrix of the chromatographic R_f values

	Rf(75/25)	Rf(70/30)	Rf(65/35)	Rf(60/40)	Rf(80/20)	Rf(85/15)
Rf(75/25)	1.00	0.94	0.89	0.77	0.97	0.86
Rf(70/30)		1.00	0.97	0.87	0.89	0.75
Rf(65/35)			1.00	0.91	0.86	0.77
Rf(60/40)				1.00	0.70	0.58
Rf(80/20)					1.00	0.94
Rf(85/15)						1.00

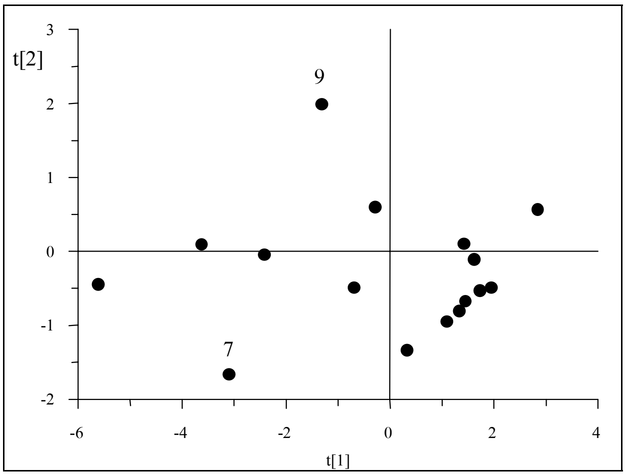


Figure 1. Dependence of the first principal component scores of the R_f matrix

Even if two direct dyes seemed to act differently in comparison to the others: dye 7 having as coupling component monoazoic compound 5-nitro-anthranilic acid→H acid and dye 9 having as coupling component 2-amino-phenol-5-sulphoneamide→resorcine, they were not found to be outliers in the R_f space.

In figures 2 the variation of the first two principal components loadings (p[2] vs. p[1]) in the R_f space is presented.

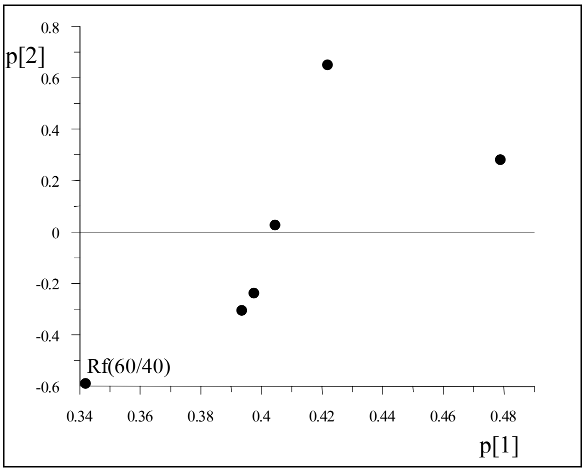


Figure 2. Dependence of the first two principal components loadings

Even if the chromatographic R_f values are interrelated, a different behavior in the case of the chromatographic mobility of the eluting solvent isopropanol:NH_3 25%= 60:40 (R_f (60/40)) could be observed from figure 2.

Principal Component Analysis was then applied to the descriptor matrix. The correlation matrix between the original descriptors and the first three principal components is presented in Table 3. Volume and bulkiness parameters produce an important contribution to the first PC. Polarity parameters contribute to the second principal component. The first PC describes, mainly, a measure of the steric parameters of the dye molecules, the second PC is a measure of hydrophobic and electronic effects.

Table 3.

Correlation matrix between the original descriptors and the first three principal components

	PC1	**PC2**	**PC3**
A_{COSMO}	0.98	-0.11	-0.13
V_{COSMO}	0.98	-0.03	-0.15
SAS	0.97	-0.09	-0.11
MS	0.98	-0.11	-0.12
SEV	0.99	-0.02	-0.09
MW	0.98	0.04	-0.14
OVAL	0.88	-0.38	-0.22
E_{HOMO}	0.57	0.54	0.38
E_{LUMO}	-0.68	-0.50	0.34
μ	0.02	0.78	-0.22
QMAX	0.80	-0.25	0.44
QMIN	-0.79	0.26	-0.44
Q+	0.38	0.23	0.65
ClogP	-0.60	-0.59	-0.15
MR	0.99	-0.06	-0.10
Eigenvalues	10.09	1.88	1.28
Eigenvalues cumulative, %	67.24	79.75	88.30

The dependence of the principal component scores of the descriptor matrix is presented in figure 3. Two direct dyes, having as coupling component resorcine (for dye 3), respectively salicylic acid (for dye 10) have a different behaviour in the descriptor space in comparison to the other 14 dyes. This can be explained by the presence of only one phenylic group in each of the coupling components.

Starting from the set of descriptors described above, MLR approach has been applied to model the scores of the chromatographic R_f values. Starting from the entire data set of 16 structural descriptors, intercorrelations between these descriptors have been inspected for the set of 16 compounds. Variable selection by a stepwise regression procedure based on the Fischer test was performed. All the statistical tests were tested at a significance level of 5 % or less.

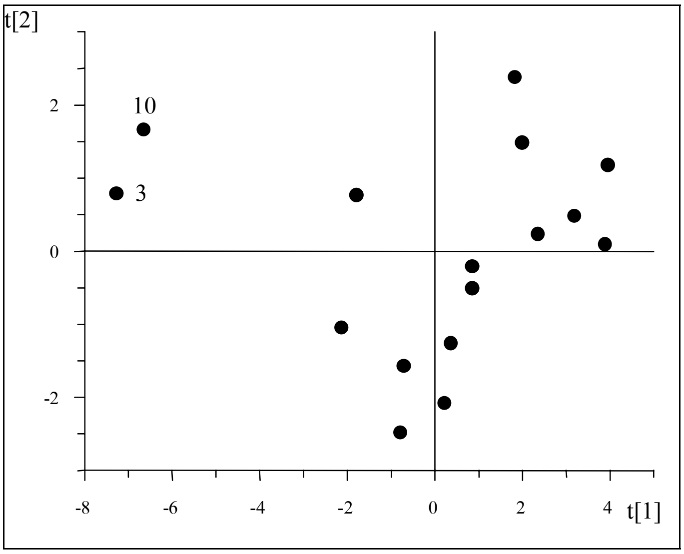


Figure 3. Dependence of the first principal scores of the descriptor matrix

Following MLR equation (correlation coefficient r, standard error s, F-test, and crossvalidated correlation coefficients $r^2_{CV_{LHO}}$ and $r^2_{CV_{LOO}}$ resulting from the leave-half-out and leave-one-out procedures, respectively) was obtained:

$$\begin{gathered} t[1] = -2.92(\pm 0.748) + 0.66(\pm 0.15) C\log P \\ r^2 = 0.572 \quad s = 1.625 \quad F = 18.69 \quad r^2_{CV_{LHO}} = 0.372 \quad r^2_{CV_{LOO}} = 0.488 \end{gathered} \tag{2}$$

Equation (2) indicates that dye hydrophobicity influences significantly the lipophilic dye behaviour.

CONCLUSIONS

This paper presents an experimental and theoretical evaluation of chromatographic mobilities i.e. R_f values of a series of 16 direct dyes containing 4,4'-diaminobenzanilide as a middle component. Molecular dye structures were energy minimized using the MM2 force field and by the semiempirical continuum-solvation model COSMO-AM1. Several molecular descriptors were derived from the energetically minimized molecules. Principal Component Analysis was then applied to the descriptor matrix. Volume and bulkiness parameters produced an important contribution to the first PC. Polarity parameters contributed to the second principal component. Two direct dyes, having only one phenylic group in each of the coupling component (resorcine, respectively salicylic acid) have a different behaviour in the descriptor space in comparison to the other 14 dyes. The R_f values were obtained by reverse-phase thin-layer chromatography and then were quantitatively analysed by MLR (Multiple Linear Regression) analysis. Good correlation with the scores of the R_f values obtained by principal component analysis and predictability were noticed. Hydrophobicity of the direct dyes influences significantly their lipophilic behaviour.

ACKNOWLEDGEMENTS

Parts of the results presented in this paper were supported from the GAR contract of the Romanian Academy no. 29/2001.

REFERENCES

1. W. Biedermann, Migrating cationic dyes for acrylic fibers. *Rev. Prog. Col.,* **1979**, *10*, 1-10.
2. K. Greider, Determination of Acid-dye Standard Affinities for Nylon Fibres from Sorption Isotherms. *J. Soc. Dyers Colour.,* **1976**, *92,* 8-13.
3. D.G. Duff, D.J. Kirkwood, D.M. Stevenson, Behaviour of Dyes in Aqueous Solutions. I. The Influence of Chemical Structure on Dye Aggregation – a Polarographic Study. *J. Soc. Dyers Colour.,* **1977**, *93*, 303-306.
4. H. Zollinger, Dye and substrate. Role of hydrophobic bonding in dyeing process. *J. Soc. Dyers Colour.,* **1965**, *81*, 345-350.
5. T. Imae, C. Mori, A.. Ikeda, Adsorption and micellization of a surface-active dye in aqueous methanol solutions. *J. Chem. Soc., Faraday Trans. 2,* **1982**, *78*, 1359-1368.
6. W. Biedermann, A.. Datyner, The interaction of nonionic dye with sodium dodecyl sulfate and its correlation with lipophilic parameters, *J. Colloid and Interface Sci.,* **1981**, *82*, 276-285.
7. D.M. Stevenson, D.G. Duff, D.J. Kirkwood, The Behaviour of Dyes in Aqueous Solutions. Part II – Anionic Dye – Nonionic Surfactant Interactions. *J. Soc. Dyers Colour.,* **1981**, *97*, 13-17.
8. H. Van de Waterbeemd, B. Testa, In: *"Advances in Drug Research"*, (B. Testa, *Ed.*), Academic, London, 1987, Vol.16, p. 85-225.
9. D.G. Duff, R.W. Horobin, G.B. Proctor, Estimating the Hydrophobic Character of Dyestuff: A Comparison of Partition, Reverse-Phase Thin-Layer Chromatography and Calculations. *Dyes Pigm.,* **1985**, *6*, 61-73.
10. P. Mukerjee, A.K. Ghosh, Isoextraction method and the study of the self-association of methylene blue in aqueous solutions. *J. Am. Chem. Soc.,* **1970**, *92*, 6403-6407.
11. E. Tomlinson, Chromatographic hydrophobic parameters in correlation analysis of structure-activity relations. *J. Chromatogr.,* **1975**, *113*, 1-45.
12. G. Simu, L. Kurunczi, S. Timofei, E. Dulău, M. Grad, C. Bologa, Application of the multimodal adsorption by cotton of some benzidine-free direct dyes. *Annals of West University of Timișoara, ser. chem.*, **1999**, *8*, 31-36.

13. **ChemOffice 6.0**, CambridgeSoft.Com, Cambridge, MA, U.S.A.

14. A. Klamt, G. Schüürmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin. Trans.* 1993, *2*, 799-815.

15. G. Schüürmann, Assessment of semiempirical quantum chemical continuum-solvation models to estimate pKa of organic compounds. In: *Quantitative Structure-Activity Relationships in Environmental Sciences*. F. Chen, G. Schüürmann, Eds.; SETAC Press, (FL) USA, Volume VII, 1997, p. 225-242.

16. M. L. Connolly, The Molecular Surface Package. *J. Mol. Graphics* **1993,** *11,* 139-141.

17. R. Franke, *Theoretical Drug Design Methods*, Akademie-Verlag, Berlin, 1984

18. D. J. Livingstone, *Methods in Enzimology*, Vol. 203, Academic Press, Inc., 1991, p. 613-638.

19. P. P. Mager, H. Rothe, H. Mager, H. Werner, Advances in Multivariate QSAR on IBM-PC/XT/AT under MS-DOS. In: *QSAR in Design of Bioactive Compounds,* J.R. Prous Science Publishers, 1992, pp.131-182.

20. S. Wold, W.J. Dunn III, Multivariate Quantitative Structure-Activity Relationship (QSAR): Conditions for Their Applicability. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 6-13.

21. R. Franke, *Theoretical Drug Design Methods*, Akademie-Verlag, Berlin, 1984, pp. 227

22. D. Ciubotariu, E. Deretey, T. I. Oprea, T. Sulea, Z. Simon, L. Kurunczi, A. Chiriac, Multiconformational Minimal Steric Difference. Structure-Acetylcholinesterase Hydrolysis Rates Relations for Acetic Acid Esters. *Quant. Struct.-Act. Relat.*, **1993**, *12*, 367-372.

23. **SYSTAT** package, from Systat Inc., Evenston, IL, SUA