Dyes and Pigments 2 (1981) 305–318

QUANTITATIVE RELATIONSHIPS BETWEEN CHEMICAL STRUCTURE AND TECHNICAL PROPERTIES OF ARYLAZOINDOLE SULPHONIC ACID DYES

RENATO GRECU and MARCELLO PIERONI

Dipartimento di Chimica Organica, Centro Ricerche Istituto,
G Donegani S p A , Via Fauser 4, Novara, Italy

and

ROSARINA CARPIGNANO

Istituto di Chimica Organica Industriale, Università di Torino,
C so Massimo D'Azeglio 48, Torino, Italy

SUMMARY

A series of arylazoindole sulphonic acid dyes has been synthesised and their spectrophotometric properties have been determined Their technical properties, such as lightfastness, fastness to washing and substantivity, on nylon and on wool, have been measured.

The values obtained of each of these properties have been analysed by a modified Free–Wilson technique Good correlations with substituent effects have been detained, demonstrating the validity of the additive model in the study of quantitative structure–activity relationships for this set of dyes. The quantification of substituent contributions makes it possible to point out the most active substituents as well as the most critical positions in order to improve the technical properties.

1. INTRODUCTION

One of the main purposes in dye research is to find good relationships between chemical structure and technical properties Many attempts have been made in order to identify structural features having a general influence on properties.[1–3] An interesting paper on the effect of terminal groups in 4-aminoazobenzene disperse dyes has recently appeared.[4] But generally the relationships reported in literature are of a qualitative kind. In the present paper a series of arylazoindole sulphonic acid dyes are examined.

Dyes and Pigments 0143-7208/81/0002-0305/$02 50 © Applied Science Publishers Ltd, England, 1981
Printed in Great Britain

The statistical analysis of quantitative structure–activity relationships (QSAR), used in drug design, has been applied to the most important technical properties of this series of dyes. By quantifying the effect of substituents, the analysis should be useful in indicating the most critical structural features in order to design dyes with improved technical properties

Part I of the present paper is concerned with the chemistry and tinctorial properties of the set of dyes examined. Part II deals with the QSAR analysis

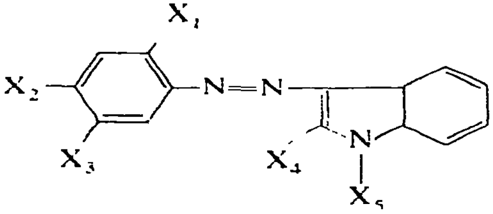
2 PART I

Azo compounds containing the indole residue as coupling component have been known since the early 1900s $^{5-8}$

The existing literature consists almost exclusively of patents

A great deal of research has been carried out in recent years particularly in the field of arylazoindole sulphonic acid dyes, as the tinctorial properties of these dyes on both natural and synthetic fibres are generally very good

We have synthesised a series of arylazoindole derivatives, many of which have not been described before, of general formula **I**



(I)

$X_1, X_3 = H, Cl, SO_3H$
$X_2 = H, SO_3H$
$X_4 = H, Me, Ph$
$X_5 = H, Me$

For these dyes we have measured the lightfastness, the fastness to washing and the substantivity, both on nylon 6 6 and on wool

2 1 *Experimental*

Dyes were synthesised according to the usual methods,[9] using a mixture of glacial acetic acid and dimethylformamide as solvent for the coupling agents

The compounds obtained were pure (i e showing only one distinct spot on TLC), except for the derivatives of indole itself, for which repeated crystallisations from ethanol–water were needed The chromatography was carried out on silica gel, using a mixture (4 4 2 1) of toluene, ethyl acetate, acetic acid and water as eluent

Ultraviolet spectra of dyes (sodium salts) were recorded in methanol solution (10^{-5}–10^{-6}M) on a Pye Unicam SP 1800 spectrophotometer

Dyeings were carried out on samples of nylon 6 6 and of wool at pH 4 5–5, at such a concentration (*ca* 1 1 %) as to obtain comparable depths (1/1)

After removal of the dyed fabrics, an undyed pattern was entered into each bath,

TABLE 1

PHYSICAL PROPERTIES, ANALYSIS AND ABSORPTION MAXIMA OF DYES OF GENERAL FORMULA **I**

Dye no	Structure					Mp or dec °C	Molecular formula	Found		Calc		λ_{max} (nm)	log ε	Ref
	X_1	X_2	X_3	X_4	X_5			C	H	C	H			
1	SO_3H	H	H	H	H	206-8	$C_{14}H_{11}N_3O_3S$	/	/	56 33	3 64	415	/	8
2	SO_3H	H	H	Me	H	263 (dec)	$C_{15}H_{13}N_3O_3S$	57 13	4 17	57 13	4 15	393	4 32	5,8
3	SO_3H	H	H	Ph	H	290 (dec)	$C_{20}H_{15}N_3O_3S$	63 71	3 86	63 68	4 01	408	4 34	—
4	SO_3H	H	H	Ph	Me	265 (dec)	$C_{21}H_{17}N_3O_3S$	64 42	4 27	64 44	4 38	399	4 33	—
5	H	H	SO_3H	H	H	>300	$C_{14}H_{11}N_3O_3S$	56 31	3 64	56 33	3 68	375	4 32	8
6	H	H	SO_3H	Me	H	272 (dec)	$C_{15}H_{13}N_3O_3S$	57 07	4 26	57 13	4 15	381	4 38	5,8
7	H	H	SO_3H	Ph	H	260 (dec)	$C_{20}H_{15}N_3O_3S$	63 70	3 94	63 68	4 01	398	4 35	—
8	H	H	SO_3H	Ph	Me	>300	$C_{21}H_{17}N_3O_3S$	64 39	4 47	64 44	4 38	390	4 35	—
9	H	SO_3H	H	H	H	201-3	$C_{14}H_{11}N_3O_3S$	/	/	56 33	3 64	375	/	8
10	H	SO_3H	H	Me	H	272 (dec)	$C_{15}H_{13}N_3O_3S$	57 06	4 29	57 13	4 15	390	4 34	5,8
11	H	SO_3H	H	Ph	H	234 (dec)	$C_{20}H_{15}N_3O_3S$	63 80	4 02	63 68	4 01	404	4 35	—
12	H	SO_3H	H	Ph	Me	221-3	$C_{21}H_{17}N_3O_3S$	64 47	4 35	64 44	4 38	395	4 26	—
13	Cl	H	SO_3H	Ph	H	296-8	$C_{20}H_{14}ClN_3O_3S$	58 37	3 44	58 34	3 43	412	4 35	—
14	Cl	H	SO_3H	Ph	Me	>300	$C_{21}H_{16}ClN_3O_3S$	59 21	3 83	59 22	3 79	405	4 37	—
15	Cl	SO_3H	Cl	Ph	H	275 (dec)	$C_{20}H_{13}Cl_2N_3O_3S$	53 82	3 12	53 82	2 94	425	4 36	—
16	Cl	SO_3H	Cl	Ph	Me	235 (dec)	$C_{21}H_{15}Cl_2N_3O_3S$	54 70	3 40	54 79	3 28	417	4 36	—
17	Cl	H	SO_3H	Me	H	233-5	$C_{15}H_{12}ClN_3O_3S$	51 52	3 44	51 50	3 46	395	4 34	—
18	Cl	SO_3H	Cl	Me	H	274-6	$C_{15}H_{11}Cl_2N_3O_3S$	46 90	2 65	46 77	2 88	406	4 35	—
19	Cl	H	SO_3H	H	H	275 (dec)	$C_{14}H_{10}ClN_3O_3S$	50 02	3 13	50 08	3 00	388	4 36	—
20	Cl	SO_3H	Cl	H	H	210 (dec)	$C_{14}H_9Cl_2N_3O_3S$	/	/	45 42	2 45	400	/	—

[/] Purity was not considered sufficient for characterisation and dyeing was continued until exhaustion was complete. By comparing the depth of the latter dyed specimen with standard depths,[10] substantivity was rated on a 1–5 scale corresponding to depths of 1/3, 1/6, 1/12, 1/25, and undyed specimen respectively

The evaluations of lightfastness (xenon lamp) and fastness to washing (60 °C) were made according to the ISO Standard Methods [11]

2.2 *Dyes and their characterisation*

In Table 1 the synthesised dyes of general formula **I**, with their characterisation data, are listed

The effect of substituents on the principal absorption band of this class of dyes may be explained considering their structure as a donor-acceptor complex chromogen,[12] in which the indole nucleus acts as a donor

Thus electron donor groups (Me) or substituents enlarging conjugation (Ph) on the indole ring cause a bathochromic shift of the visible absorption band (6–7 nm for methyl group, 23–25 nm for phenyl ring) Replacement of hydrogen substitution by a methyl group on indole nitrogen produces a general hypsochromic effect (7–9 nm) This may be attributed to increased steric hindrance, which reduces the *p*-orbital overlap between phenyl ring and indole system

TABLE 2
TECHNICAL PROPERTIES OF DYES OF GENERAL FORMULA **I**

Dye no	*Colour on fibre*	*Lightfastness*		*Fastness to washing at 60 °C*[a]						*Substantivity*	
		Nylon	*Wool*	*Nylon*			*Wool*			*Nylon*	*Wool*
1	Brownish-orange	1–2	1–2	3	3	3–4	1–2	2	3–4	4	2
2	Brownish-yellow	3	3	2	3	3	1–2	3	3–4	3	2
3	Reddish-yellow	4-5	4	2–3	3	3	1–2	2–3	4	5	4
4	Yellow	6	5	2–3	4	3	1–2	2–3	4	4–5	3-4
5	Greenish-yellow	2–3	2–3	2	4	3	1–2	3–4	3–4	3	1–2
6	Greenish-yellow	5	4–5	1–2	3	3	1	2–3	3–4	3	2
7	Yellow	6	5	2–3	3	3	1–2	2–3	3–4	5	4
8	Greenish-yellow	6	5	2–3	3	3	1–2	2–3	3–4	4–5	3–4
9	Brownish-yellow	1–2	2	3	3	3	1–2	2–3	3–4	4	2–3
10	Brownish-yellow	5	4–5	2–3	3	4	1	3	3–4	3	2–3
11	Reddish-yellow	6	5	3	3	3	1–2	2–3	4	5	4
12	Yellow	6	5	3	3	3	1–2	2	3–4	4–5	3–4
13	Reddish-yellow	6–7	6	4	3	2–3	2–3	2	2	5	4–5
14	Yellow	7	6	4	3–4	3	2–3	2–3	3	5	4–5
15	Reddish-yellow	7	6	4	3	3–4	3	2–3	3–4	5	4–5
16	Reddish-yellow	7	6–7	4	3–4	3	3–4	2	3	5	4–5
17	Greenish-yellow	5–6	5–6	3–4	3	3	2	2	3	4	4
18	Yellow	6	6	3–4	3	2–3	2	2	2–3	4–5	4

[a] Values in the first column refer to the change of colour, in the second column to the staining of the same kind of fibre as the specimen, and in the third column to the staining of wool for nylon and of cotton for wool

Examining the acceptor moiety, we may note that:

(a) the *meta*-isomer is the least bathochromic among the aminobenzenesulphonic acids,
(b) a chlorine atom *ortho* to the azo group causes a bathochromic shift (14–15 nm).

In Table 2 the technical properties of the dyes on nylon 6.6 and wool are listed.

3 PART II

3.1 *Background*

The study of quantitative structure–activity relationships (QSAR) has gained a large and increasing practical importance in drug design.[13]

The development of QSAR in biochemistry and medicinal chemistry has been recently surveyed by C. Hansch.[14] There are two main kinds of approaches in QSAR. In the Hansch approach[14] the biological activity of a set of congeners is correlated with some physicochemical properties (hydrophobic constants, electronic and steric parameters, etc.). The Free–Wilson method of analysis[15] is a statistical approach independent of physicochemical properties. The basis of the Free–Wilson model is the additive contribution to activity of each substituent, independently on other substituents in the molecule.

Thus for a congeneric series of n compounds, the activity A for each member i of the set, under equivalent conditions of assay, can be expressed as

$$A_i = \Sigma_p \Sigma_s C_{i\,ps} + \mu \qquad (i = 1, 2, \ldots, n) \tag{1}$$

where C is the activity contribution of the substituent s, dependent on the position p; μ represents the activity of a constant portion (parent structure) of the series. If the activities of a number of congeneric molecules have been measured, the values of the individual group contributions or *de novo* substituent constants can be calculated by regression analysis.

The interpretation of the parameter μ will depend upon arbitrary choices to be specified. In the classical form of the Free–Wilson model based on symmetry equations, μ is the overall average of the biological activities and is interpreted as the activity value of a hypothetical 'naked' compound (all substituents absent).

Fujita and Ban[16] modified the Free–Wilson approach: the activity contribution of each substituent is relative to hydrogen whose contribution is defined as equal to zero ($C_H = 0.00$), μ, obtained by the least-squares method, is the theoretically predicted activity value for an all-hydrogen substituted compound.

Cammarata and Yau[17] used a similar, but statistically different, modified Free–Wilson model, where the observed activity value of the unsubstituted compound is arbitrarily taken as the constant term μ.

Hansch's linear multiple regression model and the modified form of the Free–Wilson additive model have been shown to be theoretically interrelated[18] and numerically equivalent,[19] due to the fact that the individual groups' contribution can be interpreted as a weighted sum of several physicochemical constants of the substituents.

A mixed approach has been developed, which is a combination of both models and takes advantage of each: the *de novo* constants, which can be regarded as discrete variables (the term indicator variable is more appropriately used), can be combined with the continuous variables of Hansch's analysis to increase the potential of QSAR. Interest in such mixed analysis is increasing.[14]

In the present paper the statistical approach of QSAR analysis is applied to the technical properties listed in Table 2 of the series of dyes of general formula **I**.

3.2 *Method*

The technique used is the Fujita–Ban[16] modification of the Free–Wilson approach (eqn 1).

The 18 congeneric structures examined were encoded in the matrix of Table 3, where the number one in any position indicates the presence of the substituent.

Regression analyses were performed on lightfastness, fastness to washing and substantivity, for nylon and wool. The stepwise procedure of the SPSS programs[20] was used. Calculations were carried out on a IBM 370/158 computer. The results are given in Tables 4–10.

TABLE 3

MATRIX USED FOR CALCULATION OF GROUP CONTRIBUTIONS (EQNS 2–7)

Dye no	X_1		X_2	X_3		X_4		X_5
	SO_3H	*Cl*	*SO_3H*	*SO_3H*	*Cl*	*Me*	*Ph*	*Me*
1	1							
2	1					1		
3	1						1	
4	1						1	1
5				1				
6				1		1		
7				1			1	
8				1			1	1
9			1					
10			1			1		
11			1				1	
12			1				1	1
13		1		1			1	
14		1		1			1	1
15		1	1		1		1	
16		1	1		1		1	1
17		1		1		1		
18		1	1		1	1		

3.3 *Results and discussion*

The results of the statistical analysis of data relating to each property are discussed separately.

The substitution pattern of the set of congeners under examination impose the following limitations on the analysis:

(1) the effect of Cl at the X_3-position can be assessed only in the presence of Cl at X_1,
(2) the effect of Cl at the X_1-position can be assessed only in the presence of a substituent (Ph or Me) at X_4;
(3) SO_3H at X_1 without any other substituent at X_4 occurs in one compound only, therefore the net effect of X_1—SO_3H cannot be assessed with sufficient accuracy.

Lightfastness Equations (eqns 2 and 3) with good correlation (of the form 1) could be formulated relating the lightfastness, on nylon and wool respectively, to the most significant variables, as judged from the statistical tests:

Eqn	*Fibre*	*n*[a]	*k*[b]	*R*[c]	*s*[d]	R^2[e]	*F*[f]	
2	Nylon 6 6	18	7	0·979	0·472	95·9	33·9	(+ + +)
3	Wool	18	7	0·984	0·336	96·9	45·0	(+ + +)

[a] Number of compounds used for the regression analysis
[b] Number of explanatory variables entered in the regression equation, the explanatory variables were allowed to enter the regression if they exceeded a *F* value of 0·1
[c] Multiple correlation coefficient
[d] Standard deviation of the estimate
[e] Percentage of the variance of the dependent variable explained by the regression equation
[f] *F*-ratio and significance level (+ + +) $P < 0{\cdot}001$, (+ +) $P < 0{\cdot}01$, (+) $P < 0{\cdot}05$

Tables 4 and 5 list the calculated contributions of the unsubstituted compound (μ) and of the substituents included in the correlation equations. The lightfastness values calculated using the contributions in Tables 4 and 5 and the differences between experimental and calculated values (i.e the residuals) are listed in Table 10

A further analysis was run to prove the equivalence of X_2- and X_3-positions for the SO_3H substituent Correlation equations which are almost equally successful to eqns 2 and 3 respectively could be formulated ($R = 0{\cdot}978$, $s = 0{\cdot}462$, $R^2 = 95{\cdot}7$, $P < 0{\cdot}001$ and $R = 0{\cdot}984$, $s = 0{\cdot}325$, $R^2 = 96{\cdot}8$, $P < 0{\cdot}001$). The results of the statistical analysis indicate that in the set of dyes examined the substituents contribute to the lightfastness almost independently and additively. From the data in Tables 4 and 5 it may be noted:

(a) the low contribution of the parent structure and the important contributions of the substituents, specially for nylon;

(b) substituent positions X_4 and X_1, *ortho* to the azo group, appear to be the most significant contributors to the regression equation: position X_4 alone accounts for nearly 80% of the total variance (nylon), positions X_1 and X_4 account for nearly 50% and 44% of the total variance, respectively (wool),

(c) the most positive substituents are in the order Ph, Me at X_4, Cl at X_1 (not always statistically significant),

(d) X_1 is the most negative position for the SO_3H substituent, particularly when there is a substituent at X_4, X_2- and X_3-positions may be regarded as equivalent.

The lightfastness of dyes of the series not tested, composed of the substituents analysed, may be calculated from the group contributions of Tables 4 and 5. The dyes with the greatest predicted fastness were included in the analysis (nos. 16, 14, 15).

Examination of the pattern of the residuals[21] indicates no abnormality.

From Δ values in Table 10, dye no. 4 shows the greatest positive deviation and dye no. 2 the greatest negative deviation, both for nylon and for wool. Further investigations are required to ascertain the meaning of these anomalies.

Fastness to washing. The analysis is generally unsatisfactory if staining data are used, while colour change data give good correlations. These data were therefore used in the analysis. All dyes of the series give higher fastness values on nylon 6.6 than on wool, in both cases the variance of the data is rather small (2.5 units).

Equations 4 and 5, corresponding to the matrix in Table 3, give good correlation, as judged from the statistical tests.

Eqn	*Fibre*	*n*	*k*	*R*	*s*	R^2	*F*	
4	Nylon 6.6	18	5	0·979	0·182	96·0	57·5	(+++)
5	Wool	18	5	0·956	0·232	91·4	25·6	(+++)

Tables 6 and 7 list the contributions of the parent structure and of the most significant substituents. The calculated values of fastness of each congener and the residuals are listed in Table 10.

Data in Tables 6 and 7 show some differences in the behaviour of dyeings on nylon 6.6 and on wool. The SO_3H groups account for about 8% of the total variance for nylon, but only for about 0.5% of the total variance for wool. If SO_3H substituents are dropped from the analysis, the correlation becomes poorer for nylon ($R = 0·920$, $s = 0·319$, $R^2 = 84·7$, $P < 0·001$) but is almost equally successful for wool ($R = 0·953$, $s = 0·239$, $R^2 = 90·9$, $P < 0·001$).

In this set of dyes the SO_3H appears therefore to contribute to the fastness to washing on nylon, with a positive effect in the order $X_2 > X_1 > X_3$. The three

TABLE 4

CALCULATED GROUP CONTRIBUTIONS TO LIGHTFASTNESS ON NYLON, BASED ON $C_H = 0·00$ (EQN 2)

Position	*Substituent*	C^g	s^h	F^f	R^2 *change*[i]
X_1	SO_3H	−0·87	0·33	6·8(+)	10·1
	Cl	0·60	0·37	2·6	4·5
X_3	SO_3H	0·25	0·33	0·56	0·2
	Cl	0·58	0·51	1·3	0·3
X_4	Me	2·58	0·36	51·6(+ + +)	31·9
	Ph	3·68	0·36	104·9(+ + +)	48·3
X_5	Me	0·40	0·30	1·8	0·7
Parent structure (μ)		2·04			

[f] F-ratio and significance level (+ + +) $P < 0·001$, (+ +) $P < 0·01$, (+) $P < 0·05$

[g] Calculated group contribution, based on $C_H = 0·00$

[h] Standard deviation for each explanatory variable

[i] Percentage of the total variance explained by each explanatory variable

TABLE 5

CALCULATED GROUP CONTRIBUTIONS TO LIGHTFASTNESS ON WOOL, BASED ON $C_H = 0·00$ (EQN 3)

Position	*Substituent*	*C*†	*s*†	*F*†	R^2 *change*†
X_1	SO_3H	−0·75	0·24	9·9(+)	4·8
	Cl	0·94	0·26	12·9(+ +)	47·2
X_3	SO_3H	0·12	0·24	0·27	0·1
	Cl	0·46	0·36	1·6	0·4
X_4	Me	2·12	0·26	68·9(+ + +)	21·2
	Ph	2·62	0·26	105·2(+ + +)	22·6
X_5	Me	0·30	0·21	2·0	0·6
Parent structure (μ)		2·21			

† See Table 4 for definition

TABLE 6

CALCULATED GROUP CONTRIBUTIONS TO FASTNESS TO WASHING ON NYLON BASED ON $C_H = 0·00$ (EQN 4)

Position	*Substituent*	*C*†	*s*†	*F*†	R^2 *change*†
X_1	SO_3H	0·37	0·13	8·5(+)	2·8
	Cl	1·76	0·14	159·2(+ + +)	71·5
X_2	SO_3H	0·75	0·13	33·9(+ + +)	5·4
X_3	Cl	−0·75	0·20	14·5(+ +)	3·0
X_4	Me	−0·60	0·10	39·4(+ + +)	13·2
Parent structure (μ)		2·28			

† See Table 4 for definition

TABLE 7

CALCULATED GROUP CONTRIBUTIONS TO FASTNESS TO WASHING ON WOOL BASED ON $C_H = 0\ 00$ (EQN 5)

Position	*Substituent*	*C*†	*s*†	*F*†	*R^2 change*†
X_1	SO_3H	0 12	0 14	0 77	0 5
	Cl	0 99	0 16	39 0(+ + +)	71 9
X_3	Cl	0 50	0 19	6 9(+)	4 9
X_4	Me	−0 49	0 13	13 5(+ +)	13 5
X_5	Me	0 11	0 13	0 65	0 5
Parent structure (μ)		1 47			

† See Table 4 for definition

positions of the above substituent may be considered equivalent for wool The most active substituent, both on nylon and on wool, is Cl This substituent at X_1 accounts for *ca* 71 % of the total variance, a second Cl at X_3 explains 3 % and 5 % of the total variance for nylon and wool respectively, and makes a negative contribution for nylon but a positive one for wool

Me at X_4 is a negative substituent and accounts for *ca* 13 % of the total variance Neither Ph at X_4 nor Me at X_5 has any statistical significance

The fastness of dyes in the series not reported, may be calculated from the values in Tables 6 and 7 The dyes with the greatest calculated fastness on nylon are the ones with $X_1 = Cl$, $X_2 = SO_3H$, $X_3 = H$, $X_4 = H$ or Ph, $X_5 = Me$ or H, none of them was contained within the series The most promising dyes for wool probably are those with $X_1 = Cl$, $X_2 = SO_3H$, $X_3 = Cl$, $X_4 = H$ or Ph, $X_5 = Me$ or H, two of them were already included in the analysis (nos 16, 15)

Substantivity The small variance of substantivity data (2–2 5 units) of the series of dyes under study and the low accuracy of the measurements limit the success of the analysis

Using the matrix in Table 3, eqns 6 and 7 with the variables listed in Tables 8 and 9

TABLE 8

CALCULATED GROUP CONTRIBUTIONS TO SUBSTANTIVITY ON NYLON BASED ON $C_H = 0\ 00$ (EQN 6)

Position	*Substituent*	*C*†	*s*†	*F*†	*R^2 change*†
X_1	Cl	0 62	0 19	10 5(+ +)	9 4
X_3	SO_3H	−0 22	0 18	1 5	1 9
X_4	Me	−0 40	0 27	2 1	2 6
	Ph	1 10	0 27	16 3(+ +)	69 4
X_5	Me	−0 30	0 23	1 7	2 1
Parent structure (μ)		3 74			

† See Table 4 for definition

were formulated for nylon and wool respectively. The correlations are fairly good, as judged from the statistical tests:

Eqn	*Fibre*	*n*	*k*	*R*	*s*	R^2	*F*	
6	Nylon	18	5	0·924	0·358	85·5	14·1	(+ + +)
7	Wool	18	7	0·958	0·383	91·8	15·9	(+ + +)

The calculated values of substantivity and the residuals are listed in Table 10. Data in Tables 8 and 9 show that SO_3H groups account for 1·1 % and 1·9 % of the total variance and are not statistically significant, it may be assumed therefore that their contributions are not significantly different at positions X_1, X_2 and X_3. Almost all dyes of the set show a greater substantivity for nylon than for wool: the difference may be attributed to the parent structure ($\mu = 3·74$ and 1·83 respectively), while the substituent contributions are greater for wool than for nylon.

TABLE 9

CALCULATED GROUP CONTRIBUTIONS TO SUBSTANTIVITY ON WOOL, BASED ON $C_H = 0·00$ (EQN 7)

Position	*Substituent*	*C*†	*s*†	*F*†	R^2 *change*†
X_1	SO_3H	0·12	0·27	0·21	0·2
	Cl	1·28	0·30	18·2(+ +)	31·5
X_2	SO_3H	0·37	0·27	1·9	0·9
X_3	Cl	−0·37	0·41	0·82	0·5
X_4	Me	0·46	0·29	2·4	2·0
	Ph	1·76	0·29	36·4(+ +)	55·3
X_5	Me	−0·30	0·24	1·5	1·3
Parent structure (μ)		1·83			

† See Table 4 for definition

The most critical substituent with regard to the substantivity is Ph at X_4, which makes a strong positive contribution and accounts for 69 % and 55 % of the total variance on nylon and wool, respectively. The second active substituent is Cl at X_1 which accounts for 9·4 % and 31·5 % of the total variance on the two fibres respectively. The substantivity of dyes in the series, not reported, may be calculated from the values in Tables 8 and 9. The greatest predicted substantivity for both fibres would correspond to the dye with $X_1 = Cl$, $X_2 = SO_3H$, $X_3 = H$, $X_4 = Ph$, $X_5 = H$, not included in the series, or, for nylon, to the dye with the same structure except for $X_3 = Cl$, already analysed (no. 15).

TABLE 10

CALCULATED VALUES OF PROPERTIES AND RESIDUALS

Dye no	Fastness to light				Fastness to washing				Substantivity			
	Nylon		Wool		Nylon		Wool		Nylon		Wool	
	Calc′	Δ″	Calc′	Δ″	Calc′	Δ″	Calc′	Δ″	Calc′	Δ″	Calc′	Δ″
1	1·17	0·33	1·46	0·04	2·67	0·33	1·58	−0·08	3·75	0·25	1·96	0·04
2	3·74	−0·74	3·58	−0·58	2·04	−0·04	1·11	0·39	3·35	−0·35	2·41	−0·41
3	4·84	−0·34	4·08	−0·08	2·64	−0·14	1·61	−0·11	4·85	0·15	3·71	0·29
4	5·24	0·76	4·38	0·62	2·64	−0·14	1·71	−0·21	4·55	−0·05	3·41	0·09
5	2·29	0·21	2·33	0·17	2·29	−0·29	1·46	0·04	3·50	−0·50	1·83	−0·33
6	4·87	0·13	4·46	0·04	1·67	−0·17	0·98	0·02	3·10	−0·10	2·29	−0·29
7	5·97	0·03	4·96	0·04	2·27	0·23	1·48	0·02	4·60	0·40	3·59	0·41
8	6·37	−0·37	5·26	−0·26	2·27	0·23	1·58	−0·08	4·30	0·20	3·29	0·21
9	2·04	−0·54	2·21	−0·21	3·04	−0·04	1·46	0·04	3·75	0·25	2·21	0·29
10	4·62	0·38	4·33	0·17	2·42	0·08	0·98	0·02	3·35	−0·35	2·66	−0·16
11	5·72	0·28	4·83	0·17	3·02	−0·02	1·48	0·02	4·85	0·15	3·96	0·04
12	6·12	−0·12	5·13	−0·13	3·02	−0·02	1·58	−0·08	4·55	−0·05	3·66	−0·16
13	6·57	−0·07	5·90	0·10	4·03	−0·03	2·47	0·03	5·27	−0·27	4·87	−0·37
14	6·97	0·03	6·20	−0·20	4·03	−0·03	2·57	−0·07	4·97	0·03	4·57	−0·07
15	6·90	0·10	6·23	−0·23	4·03	−0·03	2·97	0·03	5·43	−0·43	4·87	−0·37
16	7·30	−0·30	6·53	−0·03	4·03	−0·03	3·07	0·43	5·13	−0·13	4·57	−0·07
17	5·47	0·03	5·40	0·10	3·43	0·07	1·97	0·03	3·77	0·23	3·57	0·43
18	5·80	0·20	5·73	0·27	3·43	0·07	2·47	−0·47	3·93	0·57	3·57	0·43

′ Values calculated from group contributions in Tables 4–9 respectively

″ Δ = observed − calculated value. Observed values are reported in Table 2. Intermediate values of the scale have been expressed by adding 0·5 to the lowest value (e.g. 1–2 = 1·5)

4 CONCLUSIONS

The results of the statistical analysis of substituent effects on technical properties of a series of arylazoindole sulphonic acid dyes demonstrate the validity of the Free–Wilson additive model in the study of QSAR for this set of dyes The quantification of substituent effects makes it possible to point out the most active substituents as well as the most critical positions in terms of improving the technical properties The present analysis showed that the following features are most important

(1) the presence of Ph or Me at the X_4-position and of Cl at X_1 for the lightfastness both on nylon 6 6 and on wool (these substituents together account for *ca* 85% and 91% of the total variance for the two fibres respectively),
(2) the presence of Cl at X_1 for the fastness to washing (*ca* 71% of the total variance explained),
(3) the presence of Ph at X_4 and of Cl at X_1 for the substantivity (these substituents together account for *ca* 79% and 87% of the total variance on nylon 6 6 and wool respectively)

The extension of this type of statistical analysis to other series of dyes should give further data for the substituents studied and for others not included in the present paper Only when a sufficient set of data is available, can more general conclusions be drawn concerning the quantitative structure–property relationships of dyes

The results of the Free–Wilson approach can be further analysed in terms of physicochemical substituent parameters in order to gain insight into their mechanism of action

ACKNOWLEDGEMENTS

We thank Mr C Paffoni (Dip Chimica Organica, Ist G Donegani, Novara) for carrying out the application of dyes and for evaluation of their technical properties

REFERENCES

1 K Venkataraman, *The chemistry of synthetic dyes*, vol 1 New York, Academic Press (1952)
2 N A Evans and I W Stapleton, in *The chemistry of synthetic dyes*, vol 8, ed K Venkataraman, p 221 London, Academic Press (1978)
3 C H Giles, D G Duff and R S Sinclair, in *The chemistry of synthetic dyes*, vol 8, ed K Venkataraman, p 279 London, Academic Press (1978)
4 G Hallas, *J Soc Dyers Col*, 95, 285 (1979)
5 Farbenfabriken, Germ Patent 141,354 (1902)
6 Farbenfabriken, Germ Patent 160,674 (1904)

7 Farbenfabriken, Germ Patent 163,141 (1904)
8 H Pauly and K Gundermann, *Chem Ber* **41**, 4004 (1908) (footnote 4)
9 H E Fierz-David and L Bangley, *Grundlegende operationen der farbenchemie*, 8th edn Wien, Springer-Verlag (1952)
10 *Standard depths glossy* London British Standards Institution (1961)
11 *Standard methods for the determination of the colour fastness of textile and leather* Bradford, Society of Dyers and Colourists (1978)
12 J Griffiths, *Colour and constitution of organic molecules*, p. 172 London, Academic Press (1976)
13 W P Purcell, G E Bass and J M Clayton, *Strategy of drug design a guide to biological activity* London, J Wiley (1973)
14 C H Hansch, in *Correlation analysis in chemistry—Recent advances*, ed N B Chapman and J Shorter, p 397 New York, Plenum Press (1978)
15 S M Free, Jr, and J W Wilson, *J Med Chem* 7, 395 (1964)
16 T Fujita and T Ban, *J Med Chem*, **14**, 148 (1971)
17 A Cammarata and S J Yau *J Med Chem* 13, 93 (1970)
18 A Cammarata, *J Med Chem*, **15**, 573 (1972)
19 H Kubinyi and O H Kehrhahn, *J Med Chem*, **19**, 578 (1976)
20 N H Nie C H Hull, J G Jenkins K Steinbrenner and D H Bent, *SPSS-Statistical package for the social sciences* 2nd edn New York, McGraw-Hill (1975)
21 N Draper and H Smith, *Applied regression analysis* New York, J Wiley (1966)