The Calculation of Sensitometric Properties of 1,2,4-Triazolo[1,5-*a*]pyrimidines by Use of a Neural Network

R. Meusinger*, G. Fischer, and **R. Moros**

Leipzig, Institut für Analytische Chemie der Universität

Received August 25th, 1998, respectively April 19th, 1999

Dedicated to Prof. Dr. Manfred Mühlstädt on the Occasion of his 70th Birthday.

Keywords: Computer chemistry, Nitrogen heterocycles, NMR spectroscopy, Neural Network, Sensitometry

Abstract. Quantitative structure–property relationships were proposed by using artificial neural networks and information received from ^{13}C NMR spectra. The suitability of 1,2,4-triazolo[1,5-*a*]pyrimidines as stabilizers in photographic silver halide materials was determined from their chemical structures. For the numeric coding of the chemical structures of differently substituted 1,2,4-triazolo[1,5-*a*]pyrimidines **1**–**44** only information available from their ^{13}C NMR spectra was used. Even an assignment of the ^{13}C NMR chemical shift values to the carbons was not necessary. The best results were achieved by combination of the ^{13}C NMR chemical shifts of carbons of the basic heterocycle and the relative fog D_{rel} using a feed-forward two-layer neural network. For some compounds with a good stabilizing effect the calculated results strongly differ from experimental values giving indication of a mechanism which is not covered by the ^{13}C NMR chemical shifts.

The diminution of light caused by a developed photographic layer is designated as its optical density *D*. If *D* is plotted *versus* the logarithm of the product of light intensity *I* and exposure time *t*, a characteristic curve is received. In Figure 1, the typical shape of a characteristic curve is represented. From this graphic, the photographically relevant properties of photomaterials can be derived. In such a way, the minimum density D_{min}, the sensitivity or speed *E* and the gradient $\gamma = \tan \alpha$ (the slope of curve in the usable exposure range) can be represented. A usable blackening of the photographic material begins above the minimum density D_{min}. The unwanted blackening of unexposed grains in the photographic emulsion occurring below this threshold value is designated as 'fogging' of silver halide photographic layers. For reduction of the minimum density values, benzimidazoles and phenylmercaptotetrazoles are often added to the photographic emulsion as so-called anti-fogging agents or antifoggants. In addition, stabilizers are applied to maintain the essential sensitometric characteristics, in particular, the speed *E* and the minimum density D_{min}, of photographic materials during storage. The best-known emulsion stabilizer, the 7-hydroxy-5-methyl-1,2,4-triazolo[1,5-*a*]pyrimidine, was already discovered by Birr at AGFA Wolfen in 1935 [1]. Connections between the anti-fogging effect and the low solubility of the silver salts and the adsorption properties of photographic additives were discussed by James in 1977 [2]. In the last decade some reviews were presented concerning the characteristics of stabilizers and anti-fogging agents and the existing ideas about the mechanism [3]. However, in spite of numerous efforts, up to now there is no uniform theory generally explaining the effects of the anti-fogging agents. Therefore, all previous successes in the field of stabilization and the decrease of fog were essentially achieved empirically. Thus, no quantitative relationships could be established between the structures of triazolo[1,5-*a*]pyrimidines and their sensitometric properties, up to now.

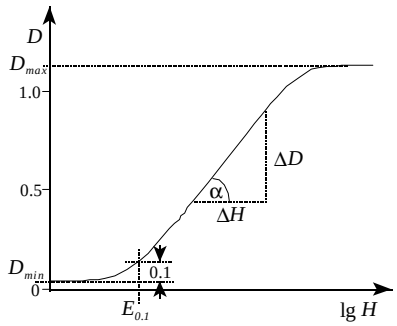


Fig. 1 Typical characteristic curve in which the optical density *D* of the individual steps of a step–wedge exposure is represented as a function of the exposure $H = I \cdot t$. The minimum density D_{min}, the speed $E_{0.1}$ at 0.1 density units over D_{min}, the slope α as well as the maximum density D_{max} are depicted.

Kleinpeter and co-workers have studied in detail in the past few years the ^{13}C NMR spectra, ^{15}N NMR spectra and quantum-chemical calculations of a large variety of substituted 1,2,4-triazolo[1,5-*a*]pyrimidines [4].

  1436-9966/99/3410X-449 $ 17.50 + .50/0

Application of Neural Network

This paper presents results of the use of a backpropagation neural network (NN) for quantitative structure–property relationships of 1,2,4-triazolo[1,5-*a*]pyrimidines. During the past several years there has been a growth of interest in use of NNs in the field of chemistry [5]. An overview of chemistry related applications was given by Burns and Whitesides [6]. Applications of NNs exist in almost all fields of analytical chemistry. They have also been applied to the analysis of spectroscopic data, as well as for mass [7], near-infrared [8], fluorescence [9] and NMR [10] spectra, to list only a few current examples.

An artificial neural network is a very simplified model of natural neural systems. Analogous to the neural cells of a living organism the artificial neurons are computer simulated. One of the unique properties of neural networks is their ability to learn by observation. An NN can by adapting itself, learn to recognize features in a data set by repetitiously examining examples of the same or similar data. So, an NN based analysis can be realized without the developer having any *a priori* assumptive knowledge of the problem's parameters [9]. They can also serve as a powerful tool to assist in knowledge acquisition, helping the chemist to gain insight into poorly understood systems. For this reason NNs are also used for finding relationships between structural information and properties of single compounds or complex systems when the mathematical description of the structure–property relationships is not known. However, only a few applications of the use of NNs for the prediction of quantitative structure–property relationships (QSPR) were described up to now. A typical example of QSPR is the estimation of boiling points from structural parameters. Balaban *et al*. used multiple linear regression (MLR) to set up a correlation between boiling points and structural parameters of 185 ethers and other compounds [11]. Lohninger observed a significant decrease of the prediction error by using a radial basis function neural network to approximate the correlation of the same parameters and properties [12]. Cherqaoui and Villemin also constructed models of relationships between the structures and boiling points of 150 alkanes by means of multilayer neural networks [13]. Devillers and co-workers used a backpropagation neural network for estimating the *n*-octanol/water partition coefficients of over 7 700 organic molecules from their structure described by means of a modified autocorrelation method [14]. An overview of the current usage of the NN in QSPR and quantitative structure-activity relationships (QSAR) studies is given also by Devillers [15]. However, in all cases the molecular structures were represented by numerical codes, containing information about topology and connectivities. Another way is the description of structural information by using the numerical values of the NMR chemical shifts. In this way, the relationships between the ^{13}C NMR chemical shift information and the boiling points of 150 *n*- and *iso*-alkanes were simulated with an NN [16].

Experimental

Fourtyfour differently substituted 1,2,4-triazolo[1,5-*a*]pyrimidines were tested as stabilizers [17]. Each compound was evenly distributed in a high-speed AgBr emulsion in a concentration of 3–9 mmol per mole silver halide except for the -SH or -SCN containing compounds (0.3–0.9 mmole) and for compound 43 (40 mmole). Then the minimum density D_{min_k} and the sensitivity $E_{0.1\_k}$ were determined as the most important sensitometric parameters. All specifications were compared to the photographic layer without any addition of stabilizer, the so-called 'control' D_{min_0} to compensate for variations due to different layer thicknesses and emulsion processing. The relative fog D_{rel} and the sensitivity $\Delta E_{0.1}$ were calculated from these values by the simple equations:

$$D_{\mathrm{rel}} = D_{\mathrm{min_k}}/D_{\mathrm{min_o}} \qquad \text{and} \qquad \Delta E_{0.1_k} - E_{0.1_o}$$

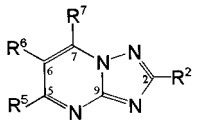
The determinations of sensitometric parameters were examined both with an emulsion prepared freshly (marked here as 'fr') and after an accelerated artificial aging by storage for ten days at a temperature of 50 °C (marked here as 'ag').

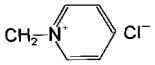
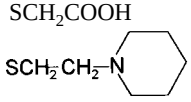
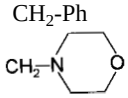
All ^{13}C NMR chemical shift values used here were taken from a study by Kleinpeter, Thomas and Fischer [4a]. Computations were carried out with personal computers. Multiple linear regressions and cluster analysis were done using the SYSTAT statistic software (STATCON). An internally developed computer program was used for the back-propagation neural network calculations [18]. Here, the number of the input units of the hidden layers and of the hidden neurons were optimized by several tests. Always one neuron was used as output neuron. The input data were separated into sets for training, monitoring, and testing by random selections. All data were normalized before their input. A sigmoidal function was always used as transfer function. Training and monitoring data sets were run simultaneously. The iterative process was stopped if the deviation for the monitoring data incresed.

Results and Discussion

The chemical structures of 44 differently substituted 1,2,4-triazolo[1,5-*a*]pyrimidines and their experimentally determined relative fog values are summarized in Table 1. We used the numerical values of the ^{13}C NMR chemical shifts to describe the structures of these 1,2,4-triazolo[1,5-*a*]pyrimidines quantitatively. In such a way, the constitutions, the equilibrium between possible tautomers, the effects induced by the neighbouring groups and the spatial arrangements of the carbon atoms could be described simultaneously by few numeric values. So, the chemical structure of every compound were encoded by a set of the ^{13}C NMR chemical shifts of the five carbon atoms C-2, C-5, C-6, C-7 and C-9 of the basic heterocycle. However, the differences between the ^{13}C NMR chem-

Table 1 Sensitometric properties of 1,2,4-Triazolo[1,5-*a*]pyrimidines



					D_{min_k}/D_{min_0}			
Nr.	R^2	R^5	R^6	R^7	fr. [a])	ag. [b])	calc. [c])	s.d. [d])
1	H	CH_3	H	OH	0.43	0.31	0.65	0.07
2	CH_2OH	CH_3	H	OH	0.45	0.29	0.55	0.05
3	$(CH_2)_6CH_3$	CH_3	H	OH	0.55	0.52	0.45	0.08
4	CH_2OCH_2COOH	CH_3	H	OH	1.00	0.65	0.55	0.04
5	CH_2SH	CH_3	H	OH	0.42	0.62	0.58	0.03
6		CH_3	H	OH	0.78	0.51	0.55	0.04
7	COOH	CH_3	H	OH	0.91	0.78	0.64	0.06
8	$COOCH_2CH_3$	CH_3	H	OH	1.00	0.70	0.58	0.07
9	$COS(CH_2)_3CH_3$	CH_3	H	OH	0.67	0.62	0.60	0.09
10	COS-*n*C_8H_{17}	CH_3	H	OH	0.67	0.40	0.36	0.04
11	$COSCH_2COOH$	CH_3	H	OH	0.15	0.13	0.57	0.01
12	COSPh	CH_3	H	OH	0.37	0.21	0.52	0.09
13	OH	CH_3	H	OH	0.85	0.82	0.74	0.06
14	SH	CH_3	H	OH	0.68	0.61	0.71	0.11
15	SCH_3	CH_3	H	OH	0.71	0.46	0.53	0.05
16	SCH_2COOH	CH_3	H	OH	0.57	0.56	0.51	0.01
17		CH_3	H	OH	0.59	0.50	0.48	0.02
18	SCH_2-Ph	CH_3	H	OH	0.41	0.33	0.39	0.04
19	-S–S- (disulfide)	CH_3	H	OH	1.00	0.67	0.71	0.12
20	SO_2CH_3	CH_3	H	OH	0.93	0.82	0.71	0.12
21	$NHNO_2$	CH_3	H	OH	0.48	0.37	0.71	0.11
22	H	CH_3	$(CH_2)_2OH$	OH	0.42	0.34	0.74	0.04
23	H	CH_3	$(CH_2)_2Cl$	OH	0.80	0.71	0.68	0.06
24	H	CH_3	CH_2-Ph	OH	0.97	0.88	0.86	0.03
25	H	CH_3		OH	1.00	0.88	0.50	0.11
26	H	CH_3	Cl	OH	1.00	0.58	0.67	0.03
27	H	CH_3	Br	OH	0.49	0.32	0.57	0.10
28	H	CH_3	I	OH	0.50	0.47	0.46	0.02
29	H	CH_3	NO_2	OH	1.00	1.00	1.05	0.06
30	SCH_3	CH_3	Br	OH	0.28	0.42	0.38	0.03
31	SCH_3	CH_3	SCN	OH	0.89	0.64	0.59	0.17
32	H	$(CH_2)_2CH_3$	H	OH	0.89	0.50	0.54	0.05
33	H	$(CH_2)_4CH_3$	H	OH	0.83	0.44	0.55	0.06
34	H	CH_2-Ph	H	OH	0.89	0.48	0.47	0.03
35	H	CH_2SH	H	OH	0.42	0.94	0.71	0.04
36	H	$CH_2S(CH_2)_3CH_3$	H	OH	0.48	0.33	0.46	0.04
37	H	SCH_3	H	OH	1.00	0.60	0.59	0.04
38	H	SO_2CH_3	H	OH	1.00	0.65	0.68	0.04
39	COOH	$(CH_2)_4CH_3$	H	OH	1.00	0.83	0.69	0.10
40	$COOCH_2CH_3$	$(CH_2)_4CH_3$	H	OH	0.74	0.44	0.54	0.05
41	SCH_3	CH_2Ph	H	OH	0.64	0.37	0.51	0.12
42	H	H	$COOCH_2CH_3$	OH	0.89	1.26	– [e])	– [e])
43	H	CH_3	H	CH_3	0.80	0.45	0.51	0.12
44	H	CH_3	H	SH	0.53	0.33	0.36	0.05
				R_{train}	0.25 [f])	0.58 [f])	0.52 [g])	
				R_{test}	0.24 [f])	0.56 [f])	0.79 [g])	

[a]) The 'fresh' values were measured immediately after coating the photographic layer. [b]) The 'aged' values were measured after an accelerated aging (storage for ten days at 50 °C). [c]) The averaged consistent values of D_{min_k}/D_{min_0} (after aging) of five calculations with different data sets and a neural network. [d]) Standard deviations computed over five calculations. [e]) Compounds with a relative fog of >1 density units were not considered here. [f]) Regression coefficients R for the values computed for 33 training compounds and for 10 randomly selected validating compounds (test) with the multiple linear regression. [g]) Averaged regression coefficients from five different calculations with neural networks. The results of the randomly selected validating and test data sets were combined here (see also Figure 3).

ical shifts of the nitrogen–neighbouring carbon atoms C-2, C-5, C-7 and C-9 were small within the respective compounds. Only the highfield shifted signals of the C-6 carbons distinguishes clearly from the other heterocyclic–carbon atoms. Furthermore, the assignment of any NMR signals to the individual carbon atoms could not be carried out with certainty. So in compound **43** a neighbouring difference of 17.5 ppm was measured between the two signals uncertainly assigned for the C-5 or C-7 position [4a]. Therefore, the chemical shifts of the five carbons in each triazolopyrimidine were assigned to five input units of the NN in the same increasing order of values which was observed experimentally. This procedure allowed the exclusion of interferences with subjective interpretation of the experimental data. Unfortunately, the variations of the individual ^{13}C chemical shifts of the carbons between the different compounds were also negligible [4a]. Here, standard deviations were found of only 1.5 ppm for the C-9 carbons. They increase to 7.3 ppm for the C-6 carbons. For this reason, the number of the carbons in all substituents was applied auxiliary as structure information data.

As expected, with conventional linear methods of data analysis no significant relations were found between these ^{13}C NMR data and the sensitometric properties of the 1,2,4-triazolo[1,5-*a*]pyrimidines. That is exemplified by the small regression coefficients which resulted from multiple linear regression (*R* values in Table 1). By use of a hierarchical cluster analysis considering all sensitometric and spectroscopic information which were available the best similarity was observed between the ^{13}C NMR chemical shift values and the relative fog in the photographic layer after an accelerated aging (D_{min_k}/D_{min_0} 'ag'). However, the results of the sensitivity measurements $\Delta E_{0.1}$ both in the fresh material and in the aged one showed no similarity with the NMR data. Therefore, the following computations were limited to the values of the relative fog in the aged material.

For the determination of the completely unknown connections between this sensitometric property and the ^{13}C NMR chemical shifts a multi-layered backpropagation neural network was used. Always six input units were required whereas the output layer was formed by a single individual neuron providing the calculated value of the relative fog D_{rel}. This is schematically shown in Figure 2. Three data sets were formed from the available experimental values, a larger one for training the neural network and two smaller ones for validating and testing the NN. During training, the inputs are first multiplied by random weights, summed and used as input for the selected transfer function in each hidden neuron. The outputs of the hidden neurons are also multiplied by random weights and summed to provide the output neuron. Finally, the output neurons are compared to the known values. The resulting differences are used to modify the connection weights by a backward pass during which the weight changes are propagated back to the network. So the network produces more and more correct outputs. These forward and backward processes are continued until the outputs converge on the desired values.

If a network is allowed to train too long, it will lose its ability to generalize. The NN will start to fit the noise in the training set, and the result is a so-called 'overtrained' net. By using the trained net to predict the output of a second monitoring data set simultaneously, it is possible to determine the optimum number of iterations for the training set [8]. During

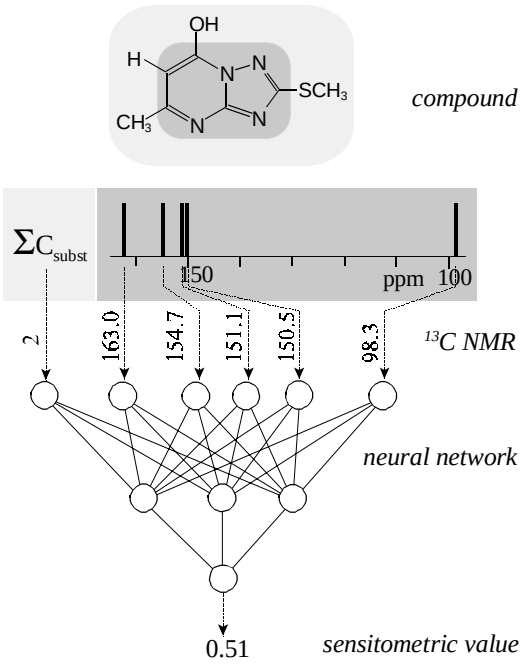


Fig. 2 Schematic presentation of the method used here for the computation of quantitative structure–property-relationships. Here, the chemical structure of compound **15** is numeric coded by ^{13}C NMR spectral information. The chemical shift values of the five carbons in the basic heterocycle (dark-gray area) are used as input values in the order of increasing chemical shifts. Also the number of all carbons in the substituents (light-gray area) is used as input for the neural network. Circles describe the artificial neurons. The trained network is able to compute the sensitometric value which leaves the single output neuron.

the learning phase the mean deviation between the computed and the correct output values of the training set decreased constantly. The iterative process was stopped when the deviation for the monitoring set increases. Several network architectures with a different number of layers and hidden neurons were tested. The models were trained with a set of 33 randomly selected samples, validated with five selected samples, and finally tested with five randomly selected samples which the network never 'saw' before.

Several hundred runs were carried out with different network architectures. The best results could be achieved by using a one-hidden-layered neural network whose hidden layer consisted of three neurons. In Figure **3a** the results of a run with this network construction are shown. The comparison of calculated and experimentally determined relative fog values are depicted for the training, monitoring and test data sets. Here, the training process was stopped after about 200.000 iterations. The quality of the computed results were reflected by the regression coefficients *R* for the comparison of calculated with experimental values. However, with randomly selected data sets, slightly different results were always obtained for the individual compounds during different runs. So, in Figure **3b** as also in Table 1 the calculated relative fog values averaged over five different runs are shown. In addition to the comparison between averaged computed and experimen-

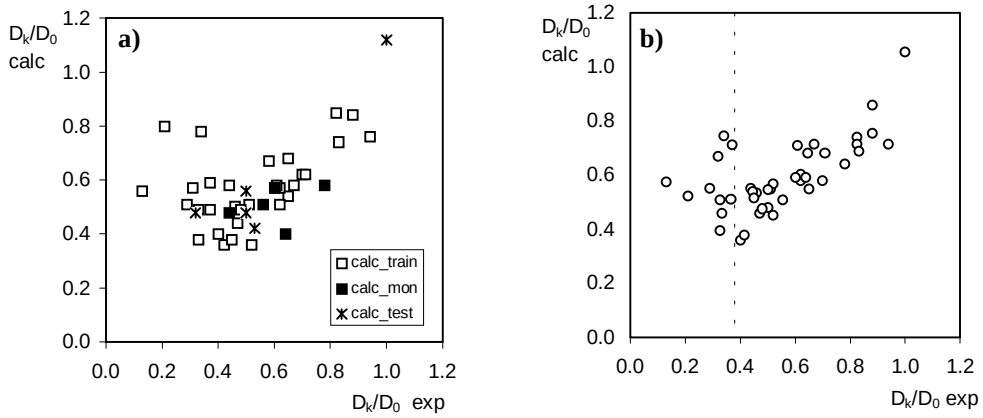


Fig. 3 Correlation between computed and experimentally determined values for the relative fog (D_{min_k}/D_{min_0}) of 43 differently substituted 1,2,4-triazolo[1,5-*a*]pyrimidines. In Figure **3a** the results of a single run with a trained neural network are shown. The data were randomly separated into sets for training (33 compounds), monitoring, and testing (5 compounds always). The regression coefficients and the standard deviations (values in parentheses) are R_{train} = 0.458 (0.183), R_{mon} = 0.360 (0.13) and R_{test} = 0.927 (0.1). The regression coefficient for the training data corrects clearly itself if values only larger than 0.4 were considered (R_{train} = 0.870 (0.1)). In Figure **3b** the averaged results from five different runs with randomly selected data sets always are shown in the same manner. Here the regression coefficients are R_{train} = 0.518 (0.172) and R_{test} = 0.788 (0.12) and after disregard of the small relative fog values R_{train} = 0.868 (0.1).

tal values, also the standard deviations from the five calculations are represented in Table 1. It is clearly shown, that these values are independent of the calculated relative fog values. The stabilizing action and the underlying physicochemical effects of triazolopyrimidines are influenced by their electronic structures which also express themselves in the NMR spectra. Therefore different substituents influence both photographic activities and NMR chemical shift values. This becomes clear with the correlation achieved between the computed and the experimental values, represented in Figure **3a**. However, it is shown also that for the compounds with a good stabilizing effect the calculated results strongly differ from experimental values. That becomes clear on the left side of the dashed line in the Figure **3b**. On the other hand, the standard deviations are nearly of the same size across the entire range of values (Table 1). This is an indication of a mechanism which is not covered by the ^{13}C NMR chemical shifts used here for the numeric coding of the chemical structures. This becomes especially clear on closer inspection of the results obtained for the 7-hydroxy-5-methyl-1,2,4-triazolo[1,5-*a*]pyrimidin **1**. For this compound, already used for a long time as emulsion stabilizer and antifoggant in film production, an averaged relative fog value of 0.65 density units was computed which is very clearly over the experimental value of 0.31 density units. Chambers proposed that the photographic properties of hydroxytriazolopyrimidines are those of the corresponding silver-complexed, mesomeric anions [19]. However, the outstanding photographic suitability of triazolopyrimidines must be co-determined by the properties of the bicyclic parent system. Therefore even the non-acidic derivative **43**, exhibiting in higher concentrations a small but distinct stabilizing effect [20, 21], fits in with the structure–property relationship.

To the best of our knowledge it was successful for the first time to compute quantitative relationships between physicochemical properties and ^{13}C NMR chemical shifts for a group of compounds with the aid of a neural network. It could be shown that a neural network is able to register the complex connections between structure-relevant NMR data and material qualities. Unfortunately, the networks offer only small possibilities for interpretation of the discovered structure–property relationships. However, the similarity between computed and experimental relative fog values shows that ^{13}C NMR chemical shifts are basically suitable for a simple numeric coding of chemical structures and also quantitative structure descriptions.

One of the authors (G. F.) thanks Chem. Ing. Sigrid Wischek for performing the photographic tests.

References

[1] E. J. Birr, Stabilization of photographic silver halide emulsions, Focal Press, London, New York 1974, p. 140
[2] T. H. James, The theory of the photographic process, 4th ed., New York, London 1977, p. 396
[3] a) G. Fischer, J. Inf. Rec. Mater. **1988**, *16*, 91; b) G. Fischer, J. Inf. Rec. Mater. **1993**, *21*, **29**
[4] a) E. Kleinpeter, S. Thomas, G. Fischer, J. Mol. Struct. **1995**, *355*, 273; b) E. Kleinpeter, A. Koch, G. Fischer, C.-P. Askolin, J. Mol. Struct. **1997**, *435*, **65**
[5] J. Zupan, J. Gasteiger, Neural Networks for Chemists – An Introduction, VCH Verlagsgesellschaft, Weinheim, New York, Basel, Cambridge, Tokyo 1993
[6] J. A. Burns, G. M. Whitesides, Chem. Rev. **1993**, *93*, 2583
[7] E. Anklam, M. R. Bassani, T. Eiberger, S. Kriebel, M. Lipp, Fresenius J. Anal. Chem. **1997**, *357*, 981
[8] C. W. Brown, S. C. Lo, Anal. Chem. **1998**, *70*, 2983
[9] J. M. Andrews, St. H. Liebermann, Anal. Chim. Acta **1994**, *285*, 237
[10] P. J. G. Lisboa, S. P. J. Kirby, A. Vellido, Y. Y. B. Lee, W. El-Deredy, NMR Biomed. **1998**, *11*, 225
[11] A. T. Balaban, L. B. Kier, N. Joshi, J. Chem. Inf. Comput. Sci. **1992**, *32*, **237**

[12] H. Lohninger, J. Chem. Inf. Comput. Sci. **1993**, *33*, 736
[13] Cherqaoui, D Villemin, J. Chem. Soc., Faraday Trans. **1994**, *90*, 97
[14] J. Devillers, D. Domine, C. Guillon, S. Bintein, W. Karcher, SAR and QSAR in Environmetal Research, Overseas Publishers Association, Amsterdam 1997, p. 151
[15] J. Devillers, Neural Networks in QSAR and Drug Design, Academic Press, London 1996
[16] J. Meiler, R. Meusinger in J. Gasteiger (Ed.), Software-Development in Chemistry 10, Gesellschaft Deutscher Chemiker, Frankfurt a. Main 1996, p. 259
[17] a) G. Fischer, Z. Chem. **1990**, *30*, 305; b) G. Fischer, Adv. Heterocycl. Chem. **1994**, *57*, 81
[18] J. Meiler, Thesis, Department of Chemistry and Mineralogy, University of Leipzig, 1998
[19] V. C. Chambers, Photogr. Sci. Eng. **1959**, *3*, 268
[20] E. J. Birr, Z. Wiss. Photogr. Photophys. Photochem. **1952**, *47*, 2
[21] K. Murofushi, Y. Kuwabara, Sh. Baba, K. Aoki, Bull. Soc. Sci. Phot. Japan **1955** (4/5), 23

Address for correspondence:
Dr. R. Meusinger
Institut für Analytische Chemie
Universität Leipzig
Linnéstr. 3-5
D-04103 Leipzig
Fax: Internat. code (0) 341-9736115
e-Mail: rmeusi@rz.uni-leipzig.de