Journal of Data Science **1**(2003), 447-460

Multivariate Chemometric Study on the Interfacial Properties of Nucleic-Acid Bases

Hai-Bin Luo, Yuen-Ting Fong, and Yuen-Kit Cheng
Hong Kong Baptist University

Abstract: Systematic quantitative structure-retention relationship studies of nucleic acid bases were carried out by the combined use of multivariate analysis and experimental chromatographic technique. The results revealed a multiple linear relationship between the chromatographic retention and the molecular structural parameters yielding a regression R^2 value of 0.8113 (cross-validated $Q^2 = 0.6945$). Five molecular descriptors, viz., moment of inertia (I_x, I_y and I_z), molar volume, and polar surface area, are able to account for the retention behavior of the compounds. Principal component analysis and factor analysis results indicate that the descriptors moment of inertia and molar volume have a primary influence on the chromatographic retention. The results provide useful insights for the future experimental and theoretical studies on the medicinal research of nucleic acid-base compounds.

Key words: Chemometrics, factor analysis, HPLC, multiple linear regression, nucleic acid, principal component, QSRR, retention.

1. Introduction

Nucleic acids, absolutely essential in living organisms, are constructed from nucleotides, which in turn are made up of five purine or pyrimidine nucleic-acid bases (nucleobases). Nucleobases and their derivatives/analogs are commonly found among designer drug molecules (Coulson 1994). Our main goal here is a systematic QSRR study on pyrimidines between the

molecular descriptors (variables) and their chromatographic retention (experimental unit). We performed the analysis involving various regression approaches in relating retention to a size-specific, shape-specific, and polarity parameters. Redundant variables will be identified.

The quantitative structure-retention relationship (QSRR) (Kaliszan 1987) between experimental chromatographic retention data and molecular descriptors has been extensively studied for three main reasons: (1) explanation of the mechanism of chromatographic separations, (2) prediction of retention, and (3) characterization of solute physicochemical properties of importance for reactivity and especially for bioactivity (Kindsvater *et al.* 1974, Roland and Robert 1980). In QSRR studies, molecular descriptors are either determined from experiments or computed by molecular mechanics or even semiempirical quantum chemical techniques.

The rationale of choosing suitable molecular descriptors for a specific problem starts with the existing simple empirical theories of chromatography up to the highly sophisticated *ab initio* calculations by quantifying the relevant intermolecular interactions. In reversed-phase high-performance liquid chromatography (RP-HPLC), it is generally agreed that two types of intermolecular interactions govern solute retention: (1) polar forces resulting from permanent or induced dipole from solute, stationary phase, and mobile phase molecules, and (2) nonpolar forces resulting from dispersive interactions. The abilities of solutes to undergo dispersive interactions are generally expressed by physicochemical variables that reflect the size and shape of the molecule, such as moment of inertia and surface area. The polarity of solutes is expressed in terms of electronic parameters, such as dipole moment, polar surface area, and ionization potential.

There exists a handful of robust statistical regression methods for QSRR studies, such as Multiple Linear Regression (MLR), Principal Component Analysis and Regression (PCA and PCR), Partial Least Squares Regression (PLSR), and Factor Analysis (FA). PLSR is especially useful when there are more variables than experimental samples. A cross-validated R^2, commonly referred as Q^2, is computed analogously to the conventional R^2. Factor Analysis (FA) has been used frequently in chromatography for two main purposes (Kindsvater *et al.* 1974, Roland and Robert 1980). One is to ascertain how many factors are necessary to account for the variance of the

retention data. This useful information can be obtained without having to identify the factors. Second, it is possible to interpret the abstract factors with physically meaningful parameters.

2. Method

2.1 Experimental Details: Equipments and Chemicals

The RP-HPLC columns used was Xterra^{TM} MS C_{18}-modified polymeric silica-based column (150 mm long $\times$ 4.6 mm internal diameter with 5 μm particles). The nucleobases and their derivatives were purchased from Aldrich: 2-chloro-4-(trifluoromethyl) pyrimidine 99% (2), 2,4,6-trimethoxypyrimidine 99% (3), 2-amino-4-chloro-6-methoxypyrimidine 95% (4), 2-amino-4-chloro-6-methylpyrimidine 97% (5), 1,3-dimethyl-5-fluorouracil 99% (8), 4,6-dichloro-2,5-diphenylpyrimidine 99% (9), 5-ethyluracil 98% (10), ethyl(6-amino-1,2,3,4-tetrahydro-1,3-dimethyl-2,4-dioxo-5-pyrimidinyl)-carbamate 99% (12), 6-(chloromethyl)uracil 98% (13), thymine 99% (15), N4-acetylcytosine 99% (16), sulfadiazine 99% (17), 5-nitrouracil 98% (18), 2,6-diamino-4-chloropyrimidine 98% (20), 5-carbethoxy-2-thioluracil 98% (21), uracil 99% (22), sulfisomidine 98% (23), 2-thioluracil 97% (24), 2,4,6-triaminopyrimidine 97% (27), cytosine 97% (28), 4,5-diaminopyrimidine 95% (29), 4-amino-5-chloro-2,6-dimethylpyrimidine 96% (30), 2-amino-4,6-dimethylpyrimidine 97% (31), and 4,6-diamino-2-pyrimidinethiol 99% (32); and from Sigma: 4-amino-6-chloro-2-methylthiopyrimidine 97% (1), 2-amino-5-bromo-6-methyl-4-pyrimidinol 99% (6), 2- isopropyl-6-methyl-4-pyrimidine 99% (7), 5-iodouracil 99% (11), dithiouracil 95% (19), and 5-fluorouracil 99% (25), and 2-mercaptopyrimidine 98% (26). Numbers in the parentheses refer to the pyrimdines in Figure 1. The percentage value for each compound indicates the chemical purity specified by the supplier.

2.2 Experimental Details: Experiments

The nucleobase standards were prepared in 40% (v/v) methanol/water and chromatographed using a HP 1050 liquid chromatograph. The HP liquid chromatograph is equipped with a UV diode-array detector set to

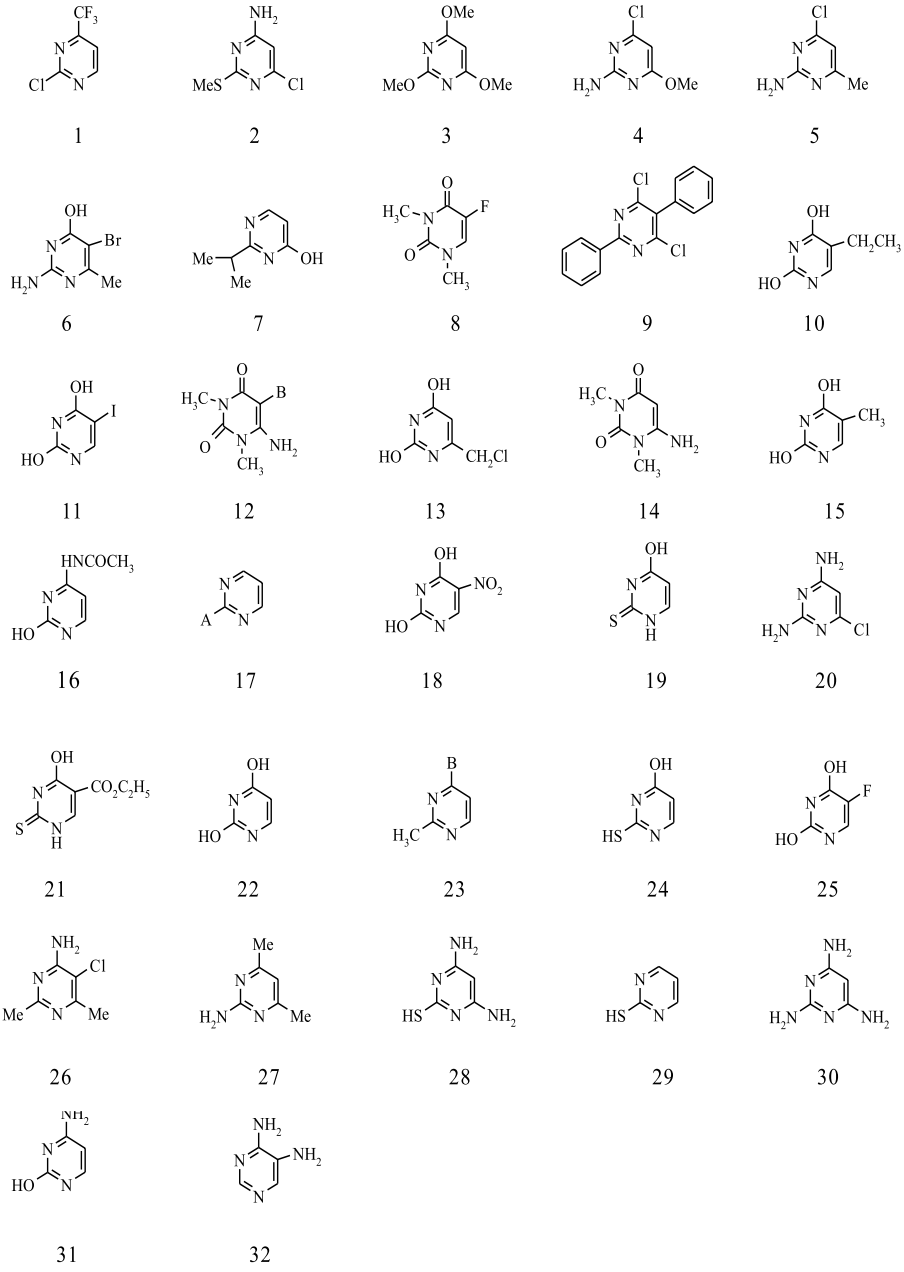

Figure 1. The structures of the pyrimdines used in this study.

254 nm. The column temperature and the flow rate were kept at 35^{o}C and 0.5 ml/min, respectively, for the experiments. The retention of the nucleobases was expressed by their capacity factors log k, which were determined under various organic mobile phases by RP-HPLC: acetonitrile (ACN), methanol (MeOH), and tetrahydrofuran (THF).

2.3 Modeling

Initial geometry optimization for the thirty two pyrimidine compounds (Figure 1) were done with the molecular mechanics MMX force field using the program PCMODEL, version 7.5 (Schlecht 1998). The dipole moment (DM), polar surface area (PSA), molar volume (MV: molecular volume times the Avogadro constant), and molecular moment of inertia along the three principal axes ($I_x \leq I_y \leq I_z$) of the compounds were also calculated using the same software (Table 1).

2.4 Data analysis

Standard multivariate analysis (Johnson 1998, Johnson and Wichern 2002) was used to probe the correlation between the retention parameter $\log(k)$ and the molecular descriptors using the publicly available QSAR routine (Fedders and Ponder 1996) and data analysis software (Lohninger 1999, Malinowski 2002). As an acceptable practice in QSRR studies, the criterion $R^2 \geq 0.81$ for MLR, PCR, and PLSR is employed to decide whether a model is internally self-consistent, and a cross-validated $Q^2 \geq 0.5$ for the robustness and absence of over-fitting in a model by the equation $Q^2 = 1 - \sum_1^n (y_{pred} - y_{obs})^2 / \sum_1^n (y_{pred} - \overline{y_{obs}})^2$. The summation term in the numerator is often referred as"PRESS" in the statistics literature. The leave-one-out procedure was adopted which means that each y_{pred} term in the summation is predicted from the remaining $(n-1)$ experimental units (y_{obs}'s), and this is repeated n times until each of the experimental unit has been left out once for the summation. The Q^2 value for a model with good predictive performance will be close to 1. FA was performed by SAS (2001).

Table 1: The experimental retention parameters (log k's) and the computed values of the five molecular descriptors for the 24 pyrimidines (outliers 8, 9, 21, 30, 31, and 32 are excluded).

No.	log k	PSA (Å^2)	MV (cm^3)	I_x (g·Å^2/mol)	I_y (g·Å^2/mol)	I_z (g·Å^2/mol)
3	1.3244	56.246	124.370	81.129	104.230	183.650
4	1.0752	65.661	101.390	57.017	92.074	148.520
5	0.6823	54.303	93.536	40.370	74.746	114.570
6	0.3865	75.969	101.810	47.193	126.970	173.610
7	0.3472	42.516	107.060	33.543	81.020	111.340
10	0.2740	69.962	96.548	34.585	82.649	116.140
11	0.2717	70.239	89.357	28.527	168.250	196.780
12	0.2600	126.570	164.070	114.910	225.990	267.840
13	0.2379	70.160	93.535	43.787	114.830	148.700
14	0.2340	87.235	109.060	49.845	74.824	123.540
15	0.2123	70.911	81.732	26.706	56.853	83.013
16	0.2107	87.697	101.560	27.986	129.560	157.000
17	0.2070	111.660	160.980	62.329	421.460	434.070
18	0.1951	115.430	82.299	37.762	92.765	126.530
19	0.1919	88.884	73.500	25.723	57.115	82.838
20	0.1916	84.112	85.275	39.439	74.129	113.570
22	0.1297	71.860	64.148	21.454	39.839	61.293
23	0.0938	107.940	175.270	71.950	486.860	505.650
24	0.0590	47.206	77.686	24.149	60.939	85.087
25	0.0573	71.717	66.066	27.199	57.155	84.353
26	-0.1233	23.233	75.964	13.449	47.838	61.288
27	-0.1488	118.080	81.923	38.753	44.869	83.622
28	-0.1528	80.531	71.398	20.813	40.504	61.317
29	-0.1546	84.095	75.025	25.554	32.928	58.482

See Figure 1 for the structures of the molecules.

Table 2. The substituion pattern and the RP-HPLC retention value (log k) in two mobile phases of the pyrimidines. The core framework and atom numbers of the pyrimidine core is

No.	1	2	3	4	5	6	20%ACN RP-HPLC	10% THF RP-HPLC
(1) 2-Chloro-4- (trifluoromethyl)pyrimidine		–Cl		$-CF_3$	–H	–H	∞	∞
(2) 4-Amino-6-chloro-2-methyl thiopyrimidine		–SMe		$-NH_2$	–H	–Cl	∞	∞
(3) 2,4,6-Trimethoxypyrimidine		–OMe		–OMe	–H	–OMe	1.3244	1.1980
(4) 2-Amino-4-chloro-6-methoxypyrimidine		$-NH_2$		–Cl	–H	–OMe	1.0752	1.2984
(5) 2-Amino-4-chloro-6-methylpyrimidine		$-NH_2$		–Cl	–H	–Me	0.6823	0.7249
(6) 2-Amino-5-bromo-6-methyl-4-pyrimidinol		$-NH_2$		–OH	–Br	–Me	0.3865	0.3138
(7) 2-Isopropyl-6-methyl-4-pyrimidinol		$-CH(CH_3)_2$		–H	–H	–OH	0.3472	0.2321
(8) 1,3-Dimethyl-5-fluorouracil	–Me	=O	–Me	=O	–F	–H	0.3471	0.1530
(9) 4,6-Dichloro-2,5-diphenylpyrimidine				–Cl		–Cl	0.2934	∞
(10) 5-Ethyluracil		–OH		–H	$-CH_2CH_3$	–H	0.2740	0.2250
(11) 5-Iodouracil		–OH		–OH	–I	–H	0.2717	0.3300
(12) Ethyl(6-amino-1,2,3,4-tetrahydro-1,3-dimethyl-2,4-dioxo-5-pyrimidinyl) -carbamate	–Me	=O	–Me	=O	-B	$-NH_2$	0.2600	0.1435
(13) 6-(Chloromethyl)uracil		–OH		–OH	–H	$-CH_2Cl$	0.2379	0.2488
(14) 6-Amino-1,3-dimethyluracil	–Me	=O	–Me	=O	–H	$-NH_2$	0.2340	0.1479
(15) Thymine		–OH		–OH	–Me	–H	0.2123	0.0432
(16) N4-Acetylcytosine		–OH		$-NHCOCH_3$	–H	–H	0.2107	0.0022
(17) Sulfadiazine		–B		–H	–H	–H	0.2070	0.5578
(18) 5-Nitrouracil		–OH		–OH	$-NO_2$	–H	0.1951	0.1264
(19) Dithiouracil	–H	=S		–OH	–H	–H	0.1919	1.0802
(20) 2,6-Diamino-4-chloropyrimidine		$-NH_2$		$-NH_2$	–H	–Cl	0.1916	0.1966
(21) 5-Carbethoxy-2thioluracil	–H	=S		–OH	$-CO_2C_2H_5$	–H	0.1347	0.5289
(22) Uracil		–OH		–OH	–H	–H	0.1297	-0.0224
(23) Sulfisomidine		–Me		–A	–H	–H	0.0938	0.2974
(24) 2-Thiouracil		–SH		–OH	–H	–H	0.0590	0.2651
(25) 5-Fluorouracil		–OH		–OH	–F	–H	0.0573	0.1326
(26) 4-Amino-5-chloro-2,6-dimethylpyrimidine		–Me		$-NH_2$	–Cl	–Me	-0.0749	-0.1083
(27) 2-Amino-4,6-dimethylpyrimidine		$-NH_2$		–Me	–H	–Me	-0.0926	-0.1196
(28) 4,6-Diamino-2-pyrimidinethiol		–SH		$-NH_2$	–H	$-NH_2$	-0.0984	-0.0515
(29) 2-Mercaptopyrimidine		–SH		–H	–H	–H	-0.1233	0.1313
(30) 2,4,6-Triaminopyrimidine		$-NH_2$		$-NH_2$	–H	$-NH_2$	-0.1488	-0.1335
(31) Cytosine		–OH		$-NH_2$	–H	–H	-0.1528	-0.1653
(32) 4,5-Diaminopyrimidine		–H		$-NH_2$	$-NH_2$	–H	-0.1546	-0.1635

–A: $-NHSO_2C_6H_5NH_2$ –B: $-NHCO_2CH_2CH_3$ ∞: Peaks cannot be detected in 60 minutes.

3. Results and Discussion

3.1 Six-variate linear regression

The structures of the thirty two pyrimidine nucleobases are shown in Figure 1 with their selected retention parameters (log k) (Table 2). QSRR analysis was performed for these pyrimidines using various regression approaches in relating retention to a variety of size- and shape-specific variables (molar volume and moment of inertia), and polarity variables (polar surface area and dipole moment).

In a 20% ACN in water (v/v) mobile phase, compounds 1 and 2 are out of the experimental retention range. Attempts to correlate log k with all the six molecular descriptors for the remaining 30 compounds resulted in a poor linear regression (MLR, $R^2 = 0.3898$). Further removal of six outliers (8, 9, 21, 30, 31, and 32) from the regression analysis rendered marked improvement (Table 3). The resulting six-variate MLR model (in 20% ACN in water mobile phase) is

$$\begin{aligned} \log k &= 0.0523 - 0.0164DM - 0.0057PSA - 0.0037MV \\ &\quad +0.0015I_x - 0.0196I_y + 0.0216I_z \end{aligned} \tag{1}$$

$$n = 24,\ R = 0.9013,\ R^2 = 0.8124,\ Q^2 = 0.6516,\ s = 0.1759,\ F = 12.27$$

Table 3: The six- and five-variate MLR and PLSR regression coefficients and the cross validation Q^2 values of the RP-HPLC retention models.

	MLR[a]		PLSR[a]		MLR[b]		PLSR[b]	
Mobile Phase	R^2	Q^2	R^2	Q^2	R^2	Q^2	R^2	Q^2
10% ACN	.5559	.2588	.5559	.1844	.4130	-.0006	.4130	-.0041
20% ACN	.8124	.6516	.8124	.6496	.8113	.6945	.8113	.6915
10% THF	.7545	.3073	.7545	.2937	.7421	.4035	.7421	.4074
10% MeOH	.4733	-.5518	.4733	-.7232	.4193	-.3362	.4193	-.3451
20% MeOH	.5268	.2077	.5268	.1305	.4200	-.1118	.4200	-.1017

[a] six-variate regression.

[b] five-variate regression without the descriptor DM.

3.2 Five-variate linear regression

By applying PCR analysis, the principal component was reduced to 5 with a similar correlation ($R^2 = 0.8113$), which suggests redundancy in the model. Descriptor was removed one-at-a-time and the correlations were compared. Regression quality is similar to the six-variate one only if the variable dipole moment was eliminated. The results of the five-variate linear regression based on the removal of the dipole moment are compiled as in Table 3 and the model is

$$\begin{aligned} \log k &= 0.0528 - 0.0057PSA - 0.0033MV \\ &\quad +0.0017I_x - 0.0196I_y + 0.0216I_z \end{aligned} \tag{2}$$

$$n = 24,\ R = 0.9007,\ R^2 = 0.8113,\ Q^2 = 0.6945,\ s = 0.1714,\ F = 15.48$$

In comparing equations (3.1) and (3.2), it is clear that the five-variate and six-variate models are almost identical apart from the descriptor dipole moment. Thus, corroborated with the larger F-Statistic coefficient of the five-variate model, the descriptor dipole moment has a negligible effect on the retention with respect to these models.

Figure 2 gives the regression relationship of the retention observed experimentally against that calculated theoretically by equation (3.2). The line drawn in Figure 2 is an expected 45^o line passing through the origin for an optimum correlation.

Statistical analysis using the PLSR method was also carried out (Table 3). While the mobile phase is 20% ACN in water, the Q^2 value for the five- and six-variate models are 0.6915 and 0.6496, respectively. This confirmed the validity of the PLSR model without over-fitting in predicting the retention values using the five molecular descriptors.

Even with some outliers present in the data, it is necessary to mention that the prediction of the retention of the 24 pyrimidines is feasible and the model is able to show how the pyrimidines retain on the RP-HPLC systems. Further design on pyrimidine-based biological entities with desirable cellular interfacial properties can make use of the insights obtained herein.

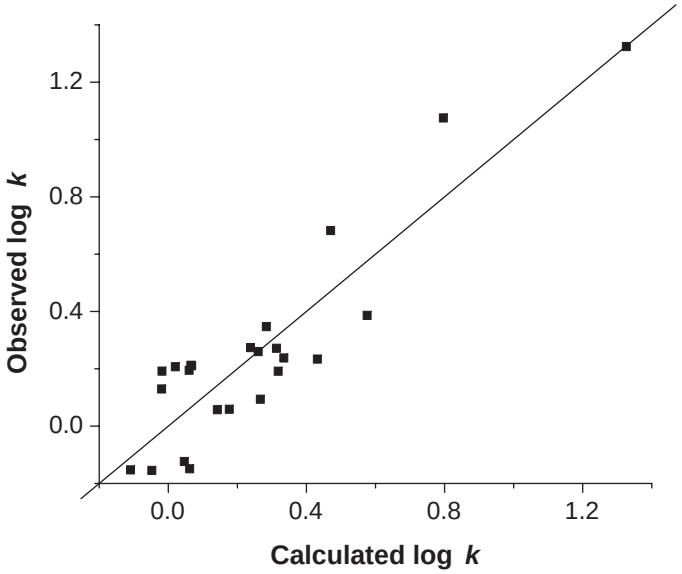

Figure 2. The plot of the relationship of the experimental versus MLR predicted RP-HPLC retention log k (equation (3.2)).

3.3 Principal component analysis

Intuitively, $\log k$ should increase with the molecular/molar volume (MV) variable. A quick inspection on the MLR models obtained above (equations (3.1) and (3.2)) leads to apparent contradiction. This strongly implies that certain extent of dependence exists among the selected variables if the models are robust. In order to enhance the interpretability, we have further performed PCA and FA on the data. The data has been formulated as a matrix using the five physicochemical variables as columns and the 24 pyrimidine compounds as rows. We further assumed that the samples behave random 5-variate normal.

The PCA results are shown in Table 4. The first three PCs (normalized with the sample correlation matrix), collectively, explain 99.20% of the total sample variance. Consequently, sample variation is essentially reflected by the first three PCs and a reduction in the data variables from 5 to 3 is

reasonable. Here that the variable 2, with the coefficient 0.4962, receives the greatest weight in the first PC (PC1). It also has the largest correlation (absolute value 0.9612) with PC1. The weights of the variables 3, 4, and 5 with PC1 (0.4263, 0.4785, and 0.4924) are almost as large as that for variable 2, indicating that the variables are about equally important to PC1. For PC2, however, variable 1 has the largest correlation (absolute value 0.7726), while the other four variables have small or negligible contribution.

It is clear that variables 2, 3, 4, and 5 (MV, I_x, I_y, and I_z, respectively) are intrinsically inter-dependent and have a primary influence on the chromatographic retention. The PCA results also indicate that variable 1 (PSA) reflecting the polarity of the molecules, has no negligible contribution to the retention.

Table 4: Principal Component Analysis of RP-HPLC retention model for the 24 pyrimidines.

	PC loadings				
Variable / PC	$\hat{e_1}(r_{y1,X_k})$	$\hat{e_2}(r_{y2,X_k})$	$\hat{e_3}$	$\hat{e_4}$	$\hat{e_5}$
1. PSA	.3173 (.6147)	.9161 (.7726)	-.2243	-.0955	-.0247
2. MV	.4962 (.9612)	-.2084 (-.1757)	.2037	-.8163	-.0504
3. I_x	.4263 (.8259)	.0918 (.0774)	.7845	.4240	.1210
4. I_y	.4785 (.9270)	-.2356 (-.1987)	-.4456	.1970	.6915
5. I_z	.4924 (.9539)	-.2309 (-.1948)	-.3070	.3255	-.7100
Eigenvalue	3.7531	0.7111	0.4959		
(a)	75.06	89.28	99.20		

(a) of the total sample variance explained (%). Numbers in the parentheses are the correlation coefficients.

3.4 Factor analysis

Before applying factor analysis (three-factor principal component solution with rotation by varimax (Johnson 1998, Johnson and Wichern 2002) to the data set, we have the same population distribution assumptions as in PCA above. The rotated estimated factor loadings and communalities are as shown in Table 5. The first three principal components (normalized with the sample correlation matrix) explain 99.20% of the total population

variance as in PCA. It is clear that the variables 2, 4, and 5, define factor 1, F_1^*, with high loadings 0.7331, 0.9429, and 0.9134, respectively; but with small or negligible loadings on F_2^* and F_3^*. On the other hand, variable 3 defines F_2^* while variable 1 defines F_3^*. As a whole, the last four size- and shape-specific four variables (MV and moment of inertia) define F_1^* and F_2^* and collectively have a primary influence on the chromatographic retention for the 24 pyrimidines, while the polarity variable 1 (PSA) is subordinate.

Table 5. Factor Analysis of the RP-HPLC retention model for the 24 pyrimidines.

	Factor loadings*			Communalities
Variable/Factor	F_1^*	F_2^*	F_3^*	$\hat{h_i^2}$
1. Polar Surface Area PSA	0.2173	0.2004	0.9551	0.9996
2. Molar Volume MV	0.7331	0.6443	0.1508	0.9753
3. Moment of Inertia I_x	0.3211	0.9130	0.2382	0.9934
4. Moment of Inertia I_y	0.9429	0.2422	0.2224	0.9972
5. Moment of Inertia I_z	0.9134	0.3388	0.2133	0.9946
Eigenvalue	3.7531	0.7111	0.4959	
Cumulative proportion of total sample variance explained (%)	75.06	89.28	99.20	

* : Rotated by the varimax procedure

4. Conclusion

Chemometric analysis has revealed a multiple linear relationship between the physicochemical molecular descriptors and the experimental retention parameters for the twenty four pyrimidine compounds. The excellent predictive power of the QSRR models render possible the estimation of retention indices of homologous compounds whose retention values are experimentally unavailable. Subsequent PCA and FA corroborate that the four size- and shape-specific descriptors are adequate in explaining most of the RP-HPLC retention behavior, while the polarity descriptor has only a secondary influence. The analyses also indicate that the four size- and shape-specific descriptors selected for this work are inter-dependent although individually they have physical meanings highly relevant in the interpretation of the

chromatographic experiments. The convoluted effects of the moment of inertia from the data suggest that the spherical symmetry/asymmetry of the molecules is essential in the chromatographic retention. As a preliminary attempt, we have tried combining I_x, I_y and I_z or simply I_y and I_z as single variables as suggested by the results, no improvement was observed for the regression. Further investigation in this direction for the optimal variable set selection is highly desirable.

Acknowledgements

We thank the Faculty Research Grant (00-01/II72) and the statistical software support from the Faculty of Science of HKBU making this work possible.

References

Coulson, C. J. (1994). *Molecular Mechanisms of Drug Action.* Taylor and Francis, London.

Fedders, M. and Ponder, J. W. (1996). *QSAR Program.* "ftp://dasher.wustl.edu/pub/qsar/". Last visited on 27^{th} May 2003.

Johnson, D. E. (1998). *Applied Multivariate Methods for Data Analysts.* Duxbury, Pacific Grove.

Johnson, R. A. and Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis.* Prentice Hall, Upper Saddle River.

Kaliszan, R. (1987). *Quantitative Structure-Chromatographic Retention Relationships.* Wiley, New York.

Kindsvater, J. H., Weiner, P. H. and Klingen, T. J. (1974). Correlation of retention volumes of substituted carboranes with molecular properties in high pressure liquid chromatography using factor analysis, *Analytical Chemistry*, **46**, 982-989.

Lohninger, H. (1999). *Teach/Me Data Analysis. Single User Edition.* Springer, Berlin.

Malinowski, E. R. (2002). *Factor Analysis in Chemistry.* Wiley, New York.

Roland, F. H. and Robert, J. G. (1980). Factor analysis of trends in selectivity in gas-solid chromatography on cation-exchange resins, *Analytical Chemistry*, **52**, 723-728.

Schlecht, M. F. (1998). *Molecular Modeling on the PC.* Wiley, New York.

Received May 20, 2003; accepted August 2, 2003

Hai-Bin Luo
Department of Chemistry
Hong Kong Baptist University
Kowloon Tong, Hong Kong, China

Yuen-Ting Fong
Department of Chemistry
Hong Kong Baptist University
Kowloon Tong, Hong Kong, China

Yuen-Kit Cheng
Department of Chemistry
Hong Kong Baptist University
Kowloon Tong, Hong Kong, China
ykcheng@hkbu.edu.hk