Correlation between Molecular Structures and Relative Electrophoretic Mobility in Capillary Electrophoresis: Alkylpyridines

YAO, Xiao-Jun[a,b] **(姚小军)** **FAN, Bo-Tao**[*,a] **(范波涛)** **DOUCET, J. P.**[a] **PANAYE, A.**[a]
LIU, Man-Cang[b] **(刘满仓)** **ZHANG, Rui-Sheng**[b] **(张瑞生)** **HU, Zhi-De**[b] **(胡之德)**
[a] *Université Paris 7-Denis Diderot, ITODYS 1, Rue Guy de la Brosse, Paris 75005, France*
[b] *Department of Chemistry, Lanzhou University, Lanzhou, Gansu 730000, China*

The quantitative relationship between relative electrophoretic mobility in capillary electrophoresis for a series of 31 closely related alkylpyridines and their molecular structures was studied by using CODESSA. According to the *t*-test on the results, we found that the three most important descriptors affecting the mobility are the relative number of rings (*NR*), Min e-n attraction for a C—N bond (*MEN*) and average complementary information index (*ACIC*). With these structure descriptors a good three-parameter linear model was developed to correlate the mobility of these compounds with their structures. This model can not only correctly predict the migration behavior of these compounds, but also find the structural factors which are responsible for the migration behavior of these compounds, thus can help to explain the separation mechanism of these compounds. The method used in this work can also be extended to the mobility-structure relationship research of other compounds.

Keywords HPCE, mobility, QSPR, alkylpyridine

Introduction

High performance capillary electrophoresis (HPCE) has now been widely used to separate and determine analytes from small inorganic and organic ions to macromolecular species such as DNA and proteins. Over the past few years, there have been considerable interests in the developing methods using capillary electrophoresis to separate small organic molecules, as a complementary method to HPLC separations.[1-3] Electrophoretic mobility (μ) is the most important parameter governing the separation of solutes in capillary zone electrophoresis (CZE). Excellent discussions have been made as to the effect of numerous experimental variables such as pH and buffer ion strength on the apparent mobility of an ion.[4-6] These models discuss how these parameters affect the intrinsic mobility of the full charged ion (μ_{A-} or μ_{B+}), but the discussion of actual intrinsic mobility (μ_o) of an ion is generally limited to the simplistic statement that mobility is related to the charge-to-size ratio of the solute. The intrinsic mobility behavior of ions is much more complex. It is possible to have a sufficient understanding of the separation mechanism if a model correlating the migration behavior of an analyte to its structural parameters can be developed. Fu and Lucy[7,8] have developed empirical expressions for the prediction of electrophoretic mobility of mono-amines and aliphatic carboxylic acids. They have correlated the mobility of the analytes with the molecular mass, molar volume and dissociation constant by using non-linear equations. Liang and coworkers[9] have studied the correlation between the electrophoretic mobility of 13 flavonoids and topological indices. The relationship between the structures of a range of analytes and their mobility has also been examined by Kaliszan and coworkers[10] as part of a quantitative structure (electrophoretic)-retention relationship (QSRR). Jalali-Heravi and Garkani-Nejad[11] have used artificial neural networks for the prediction of electrophoretic mobility of sulfonamides in capillary zone electrophoresis based on several parameters calculated from AM1 semi-empirical methods in MOPAC. Li *et al*.[12] have constructed a model by means of a multilayer neural network by using extended delta-bar-delta (EDBD) algorithm to estimate complex property of electrophoretic mobility of aliphatic carboxylates and amines from simpler experimental properties.

Recently, quantitative structure property relationship (QSPR) method has been widely used for the prediction of various physico-chemical properties of organic compounds. Software CODESSA[13,14] developed by Katritzky group has been successfully used in various QSPR research. But there is no reported research work on the application of CODESSA to the correlation between electrophoretic mobility and molecular structures. The primary goal of the present work was to develop a uniform empirical model for the prediction of electrophoretic mobility of alkylpyridines with known experimental mobility and to determine the underlying factors governing the mobility of these compounds. Investigation of the relationship between the structures of a range of analytes and their mobility is of impor-

* E-mail: fan@paris7.jussieu.fr
Received April 4, 2003; revised and accepted July 10, 2003.
Project supported by PRA SI 00-05 of Association Franco-Chinoise pour la Recherche Scientifique & Technique (AFCRST).

tance for electrophoresis, because such study can reveal the manner in which the mobility changes depending upon the variation in structure and can find the main factors affecting the migration and thus discloses the separation mechanism. This research work is also useful for the optimization of the separation conditions.

Method

The structures and electrophoretic mobility of 31 alkyl pyridines were taken from previous work of Andrew *et al*..[15] In their work, Andrew and coworkers predicted the relative electrophoretic mobility of a series of closely related alkyl pyridines by proposing that they experience a preferred orientation under the influence of the applied electrical field. Several good three-parameter models were developed for each sub-class of compounds, such as *n*-alkyl pyridines, unsaturated and branched alkylpyridines.[15] Electrophoretic mobility of analytes was determined by using a lithium phosphate buffer under pH = 2.5. The analytes were separated at 25 ℃ using a voltage of 15 kV. Under these conditions, the compounds, whose pK_a values range from 5.2 to 7.4, were in a full positive charge. So the holistic charge of molecules has no effect on the difference of mobility of these compounds and the main determinant in the mobility are mass term such as size, shape, branching and the charge distribution of these compounds. The experimental mobility and names of compounds investigated are all listed in Table 1.

To develop a QSPR model, the main steps involved are: structure entry, optimization, descriptor generation, feature selection, model development and finally model evaluation. All molecules were drawn into Hyperchem[16] and pre-optimized by using MM + molecular mechanics force field. A more precise optimization was done with semi-empirical AM1 method in MOPAC.[17,18] All calculations were carried out at restricted Hartree Fock level with no configuration interaction. The molecular structures were

Table 1 Electrophoretic Mobility of alkylpyridine ($cm^2 \cdot V^{-1} \cdot s^{-1}$) ($\times 10^4$)

Number	Name	Mobility	Predicted value	Residual
1	Pyridine	4.176	4.2045	0.0285
2	2-Methylpyridine	3.581	3.6016	0.0206
3	3-Methylpyridine	3.721	3.7426	0.0216
4	4-Methylpyridine	3.722	3.7381	0.0161
5	2-Ethylpyridine	3.222	3.2187	-0.0033
6	3-Ethylpyridine	3.366	3.3547	-0.0113
7	4-Ethylpyridine	3.397	3.3511	-0.0459
8	2,3-Dimethylpyridine	3.236	3.2244	-0.0116
9	2,4-Dimethylpyridine	3.196	3.2206	0.0246
10	2,5-Dimethylpyridine	3.236	3.2258	-0.0102
11	2,6-Dimethylpyridine	3.168	3.177	0.009
12	3,4-Dimethylpyridine	3.349	3.3309	-0.0181
13	3,5-Dimethylpyridine	3.285	3.3361	0.0511
14	2-Propylpyridine	2.923	2.941	0.018
15	4-Propylpyridine	3.097	3.0695	-0.0275
16	4-Isopropylpyridine	3.059	3.0573	-0.0017
17	3-Ethyl-4-methylpyridine	3.071	3.0673	-0.0037
18	5-Ethyl-2-methylpyridine	2.976	2.9568	-0.0192
19	6-Ethyl-2-methylpyridine	2.904	2.9198	0.0158
20	2,4,6-Trimethylpyridine	2.849	2.8872	0.0382
21	3-Butylpyridine	2.848	2.8456	-0.0024
22	2-*Tert*-butylpyridine	2.748	2.6842	-0.0638
23	4-*Tert*-butylpyridine	2.828	2.8306	0.0026
24	2-Pentylpyridine	2.534	2.5298	-0.0042
25	2-(1-Ethylpropyl)pyridine	2.521	2.5268	0.0058
26	2-Hexylpyridine	2.391	2.3782	-0.0128
27	2,4,6-Tri-*tert*-butylpyridine	1.809	1.81	0.001
28	2-Vinylpyridine	3.388	3.328	-0.06
29	4-Vinylpyridine	3.597	3.5374	-0.0596
30	(*Z*)-2-(3-Pentenyl)pyridine	2.64	2.671	0.031
31	(*E*)-2-(3-Pentenyl)pyridine	2.602	2.6731	0.0711

optimized by using the Polak-Ribiere algorithm until the root mean square gradient was 0.01. The resulted geometry was transferred into software CODESSA that can calculate constitutional, topological, electrostatic and quantum chemical descriptors. Constitutional descriptors are related to the number of atoms and bonds in each molecule. Topological descriptors include valence and non-valence molecular connectivity indices calculated from the hydrogen-suppressed formula of the molecule, encoding information about the size, composition and the degree of branching of a molecule. The quantum chemical descriptors include information about binding and formation energies, partial atom charge, dipole moment and molecular orbital energy levels.

Once molecular descriptors are generated, CODESSA uses the heuristic or the best multilinear regression method (BMLR) to search for the set of best multilinear correlations: pre-selection of descriptors by eliminating the descriptors that are not available for structure or those having a small variation in magnitude or the descriptors having a F-test value less than 1 or t-values less than 1.5. In this work the best multilinear regression method is used to build a model in the form of an equation which gives mobility in terms of calculated descriptors.

Results and discussion

The mobility of an analyte was affected by its structure and many experimental conditions, such as electrolyte concentration, ionic strength, buffer pH. According to experiment of previous work, altering the separation conditions, such as electrolyte concentration, temperature, had no effect on the separation of these compounds. They could be fully resolved by changing the buffer pH which influences the partial charge of the analytes. At the separation pH of 2.5, all the analytes bear a full positive charge. Since they have the same holistic charge, the main structural factors affecting the migration of an analyte are the size, shape, symmetry and charge distribution of the analytes. Various molecular descriptors were generated to describe the structures of the analytes. After using the best linear regression routine in CODESSA, the best linear model was obtained. It contains three generated molecular descriptors. Among them, there are one constitutional descriptor, one topological descriptor and one quantum chemistry descriptor. The values of the generated descriptors are listed in Table 2. The best three-descriptor correlation model is shown in detail in Table 3. This model produces a root mean square error of 0.0303 mobility units and a squared correlation coefficient of 0.9956. The squared correlation coefficient and root mean square error for leave-one-out cross-validation is 0.9938 and 0.0360 mobility units, respectively. Fig. 1 shows the plot of predicted vs. experimental mobility. As can be seen from Table 1 and Fig. 1, the predicted mobility is very consistent with that of experiment. From the results obtained above, we can see that the models proposed can depict the migration behavior of these compounds successfully and the descriptors used have high ability in describing the difference in migration behavior of the observed compounds.

By interpreting of the descriptors in the regression model, it is possible to gain some insight into factors that are likely to govern the migration behavior of these compounds under given electronic field. According to the t-test, the most important descriptor affecting the mobility is a constitutional descriptor, the relative number of rings (*NR*). It is defined as the number of rings divided by the number of atoms in a molecule and is related to the size and shape of the molecule. For the molecules with the same number of rings and similar structures, the smaller its value, the larger the size of the molecules. Since it receives a positive coefficient in the regression, it indicates that this descriptor has a positive effect on the mobility of molecule. However, this descriptor varies little among isomers, and thus the addition of another topological descriptor, namely average complementary information index (*ACIC*), can account for the differences in the shape of these molecules. The average complementary information index[19,20] is defined on the basis of the Shannon's information theory and gives information on how many atoms with similar connectivity pattern in the molecule and reflects the relative position and size of alkyl groups connected to the pyridine ring. The combination of these two descriptors, comprising size and shape information about molecules, adequately represents mass term effects on the mobility of a molecule.

The second most important descriptor in this regression is a quantum chemical descriptor Min e-n attraction for a C—N bond (*MEN*). This descriptor relates to the strength of intramolecular interactions and the charge distribution of molecules. The inclusion of this descriptor reflects that of the charge distribution on the migration behavior of these compounds. From the result it is suggested that for molecules with similar structure and same holistic charge, the main factors affecting the behavior are the size, shape, branching and charge distribution of these compounds. These factors are responsible for the differences in the effective hydrodynamic radii and the preferred orientation of the compounds in the given electrical field and can thus affect the mobility of these compounds.

Conclusion

A successful three-parameter linear model correlating molecular structure with its mobility has been proposed. This model can not only correctly model the migration behavior of these compounds, but also find the structural factors that are responsible for the migration behavior of these compounds, thus can help us to explain the separation mechanism of these compounds. The results obtained indicate that for molecules with similar structures and holistic charge, its size, shape, and charge distribution play an important role in governing its migration behavior in HPCE. The method used in this work can also be extended

Table 2 The molecular descriptors and their corresponding mobility of alkylpyridine

Number	Name	*NR*	*MEN*	*ACIC*	Mobility
1	Pyridine	0.0909	329.8445	1.6695	4.176
2	2-Methylpyridine	0.0714	326.5162	1.2507	3.581
3	3-Methylpyridine	0.0714	330.0198	1.1968	3.721
4	4-Methylpyridine	0.0714	329.9001	1.1968	3.722
5	2-Ethylpyridine	0.0588	326.5594	1.4332	3.222
6	3-Ethylpyridine	0.0588	329.972	1.3888	3.366
7	4-Ethylpyridine	0.0588	329.8763	1.3888	3.397
8	2,3-Dimethylpyridine	0.0588	326.6829	1.4273	3.236
9	2,4-Dimethylpyridine	0.0588	326.5839	1.4273	3.196
10	2,5-Dimethylpyridine	0.0588	326.7214	1.4273	3.236
11	2,6-Dimethylpyridine	0.0588	326.6611	1.707	3.168
12	3,4-Dimethylpyridine	0.0588	330.0295	1.545	3.349
13	3,5-Dimethylpyridine	0.0588	330.1678	1.545	3.285
14	2-Propylpyridine	0.05	326.5938	1.6203	2.923
15	4-Propylpyridine	0.05	329.8374	1.5826	3.097
16	4-Isopropylpyridine	0.05	329.9547	1.6826	3.059
17	3-Ethyl-4-methylpyridine	0.05	330.0222	1.6377	3.071
18	5-Ethyl-2-methylpyridine	0.05	326.6483	1.5377	2.976
19	6-Ethyl-2-methylpyridine	0.05	326.7172	1.7755	2.904
20	2,4,6-Trimethylpyridine	0.05	326.6871	1.9642	2.849
21	3-Butylpyridine	0.0435	329.9358	1.8491	2.848
22	2-*Tert*-butylpyridine	0.0435	326.3262	2.0017	2.748
23	4-*Tert*-butylpyridine	0.0435	330.0668	1.9689	2.828
24	2-Pentylpyridine	0.0385	326.5663	2.1371	2.534
25	2-(1-Ethylpropyl)pyridine	0.0385	326.3592	2.108	2.521
26	2-Hexylpyridine	0.0345	326.559	2.3745	2.391
27	2,4,6-Tri-*tert*-butylpyridine	0.0213	326.5985	3.5674	1.809
28	2-Vinylpyridine	0.0667	324.4703	1.6271	3.388
29	4-Vinylpyridine	0.0667	329.8032	1.5768	3.597
30	(*Z*)-2-(3-Pentenyl)pyridine	0.0417	326.5659	1.8298	2.64
31	(*E*)-2-(3-Pentenyl)pyridine	0.0417	326.6233	1.8298	2.602

Table 3 The linear model between structure and mobility ($R^2 = 0.9956$, $F = 2026.88$, $s = 0.0303$)

Descriptor	Coefficient	*t*-test value
Intercept	−10.5144	−9.117
Relative number of rings (*NR*)	28.1071	41.2558
Min e-n attraction for a C—N bond (*MEN*)	0.0377	10.7351
Average Complementary Information content (order 1) (*ACIC*)	−0.1671	−8.0566

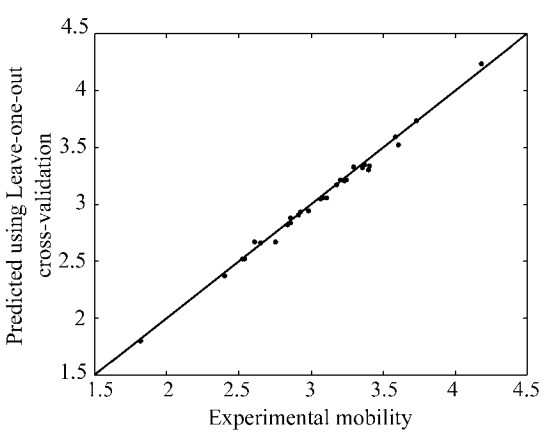

Fig. 1 The predicted vs. experimental mobility.

to the mobility-structure relationship research of other compounds.

Acknowledgements

The authors thank the Association Franco-Chinoise pour la Recherche Scientifique & Technique (AFCRST) for supporting this study (Programme PRA SI 00-05).

References

1 Zhao, Y. K.; Cao, Q. E.; Xiang, Y. Q.; Hu, Z. D. *J. Chromatogr., A* **1999**, *849*, 277.

2 Hu, Z. D.; Jia, L.; Zhang, Z. P.; Shi, Y. P.; Wang, K. T. *J. Liq. Chromatogr. Relat. Technol.* **1997**, *20*, 1211.

3 Jia, L.; Zhang, H. X.; Kou, X. L.; Hu, Z. D. *Chromatographia* **1995**, *41*, 605.
4 Smith, S. C.; Khaledi, M. G. *Anal. Chem.* **1993**, *65*, 193.
5 Rawjec, Y. Y.; Williams, R. L.; Vigh, G. *J. Chromatogr., A* **1993**, *625*, 233.
6 Wang, K. T.; Liu, H. T.; Chen, X. G.; Zhao, Y. K.; Hu, Z. D. *Talanta* **2001**, *54*, 753.
7 Fu, S.; Lucy, C. A. *Anal. Chem.* **1998**, *70*, 173.
8 Fu, S.; Lucy, C. A. *Analyst* **1998**, *123*, 1487.
9 Kaliszan, R.; Turowski, M.; Bucinsjki, A.; Hartwick, R. A. *Quant. Struct.-Act. Relat.* **1995**, *14*, 356.
10 Liang, H.; Vuorela, H.; Riekkola, M.; Hiltunen, R. J. *J. Chromatogr., A* **1998**, *798*, 233.
11 Jalali-Heravi, M.; Garkani-Nejad, Z. *J. Chromatogr., A* **2001**, *927*, 211.
12 Li, Q. F.; Dong, L. J.; Jia, R. P.; Chen, X. G.; Hu, Z. D.; Fan, B. T. *Comput. Chem.* **2002**, *26*, 245.
13 Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA: Training Manual*, University of Florida, Gainesville, **1995**.
14 Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *CODESSA: Reference Manual*, University of Florida, Gainesville, **1994**.
15 Andrew, G. M.; Smith, R. M.; Rowe, R. C.; Wren, S. A. C. *Anal. Chem.* **1999**, *71*, 497.
16 *HyperChem* 4.0, Hypercube, Inc, Gainesville, USA, **1994**.
17 *MOPAC 6.0*, *Quantum Chemistry Program Exchange*, Program 455, Indiana University, Bloomington, IN, USA.
18 Stewart, J. P. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1244.
19 Kier, L. B. *J. Pharm. Sci.* **1980**, *69*, 807.
20 Basak, S. C.; Niemi, G. J.; Veith, G. D. *J. Math. Chem.* **1990**, *4*, 185.

(E0304042 ZHAO, X. J.)