



Quantitative Structure-Retention Relationship Modelling of Aroma Components in Fenjiu, China

HUA QIAO^{1,*}, LI FENG² and TIJIAN SUN¹

¹Basic Medical College, Shanxi Medical University, Taiyuan 030001, P.R. China

²Shanxi Institute of Geological Survey, Taiyuan 030001, P.R. China

*Corresponding author: Tel./Fax: +86 351 4135645; E-mail: qiaohua0410@126.com

(Received: 26 January 2012;

Accepted: 10 December 2012)

AJC-12514

Aroma compounds are very important for identification and quality control of liquors and prediction of the retention times of given aroma compounds can be a useful method for characterizing liquors. In this paper, Fenjiu was analyzed by using gas chromatography-mass spectrometry without pretreatment. Quantitative structure-retention relationship models were constructed by various molecular descriptors of identified aroma compounds. These models show high correlation coefficients and low standard errors. To test the predictive quality of the quantitative structure-retention relationship models, leave-one-out cross validation technique was employed. Results indicate that the models are of good predictive quality.

Key Words: Fenjiu, Aroma compounds, Quantitative structure-retention relationship, Gas chromatography-mass spectrometry.

INTRODUCTION

As favourite and widely consumed drinks, liquors have hundreds of varieties in China. Fenjiu is one of the oldest famous Chinese liquor. It is well-known for its unique style as fragrant liquor which is brewed by the traditional method with local high-quality sorghum as main material. Aroma compounds play an important role in the recognition and quality control of Fenjiu¹. During analysis of aroma compounds gas chromatography and gas chromatography-mass spectrometry are commonly used, whereby retention time is a vital parameter in these analyses.

Quantitative structure-retention relationships are statistically derived relationships between the chromatographic parameters of analytes and the descriptors accounting for the structural differences among the analytes tested². Gas chromatographic quantitative structure-retention relationships models have been successfully developed for a large number of compounds: alkylbenzenes³, alcohols⁴, polycyclic aromatic hydrocarbons⁵, cyclic adenosine monophosphates⁶, carbazoles⁷, halogenated environmental contaminants⁸, proteins⁹. The advantage of quantitative structure-retention relationships is that once such a relationship is established, it can be useful in the prognosis of the chromatographic behaviour of new molecules, even before they are actually synthesized^{10,11}. There are some published data on the quantitative structure-retention relationships study of aromas of Fenjiu^{12,13}, but no quantitative structure-retention relationships studies have been

performed for aroma compounds of Fenjiu sample in GC-MS system to date.

In this work, the aroma compounds in Fenjiu were directly analyzed using GC-MS without pretreatment. The retention data were used to generate quantitative structure-retention relationships models by using molecular structural descriptors. The obtained models were evaluated by leave-one-out cross validation and show good predictive ability.

EXPERIMENTAL

Fenjiu used in the experiment was kindly provided by Shanxi Xinghua Cun Fen Chiew (Group) Co. Ltd. The provided Fenjiu sample was used for direct analysis without special pretreatment. All the used chemicals were of analytical purity grade.

An Agilent 7890 A gas chromatography equipped with a dean switch device was used for the separation of the aroma compounds in Fenjiu. Detection of the substances was achieved by an Agilent 5975C detector and a flame ionization detector. A HP-FFAP MS capillary column (30 m, 0.25 mm, 0.25 μ m film thickness) was used as the primary column and an Agilent HP-5 MS capillary column (30 m, 0.25 mm, 0.25 μ m film thickness) was used as the secondary column. The primary column was directly connected to the MS detector. The carrier gas was helium. The optimized oven temperature program used in the analysis was programmed as follows: 45 °C for 4 min, then 3.5 °C/min to 230 °C for 20 min. The injector and detector temperatures both were 250 °C and 1 μ L of Fenjiu sample was

vapourized in the injector port, with a split ratio of 30:1. Switch time of dean switch device was 7.0-9.6 min, while ethanol was switched to the secondary column. A 29-400 m/z mass range was recorded in full-scan mode, with the ion source, quadrupole and transfer line temperatures maintained at 230, 150 and 280 °C, respectively. The ionization energy was 70 eV.

The identification of compounds was made by comparing the mass spectra and retention times with those of the standards and comparison of the mass spectra obtained with those present in the NIST 05 spectral library.

Model generation: Sixty five compounds identified by GC-MS were classified into five groups *i.e.*, alcohols, aldehydes and ketones, acids, esters, ethers and hydrocarbons based on the difference of functional group. Based on stepwise multiple linear regression, quantitative structure-retention relationships models were developed for every group by using the steric and electronic descriptors (SDEs) (Table-2).

In order to obtain good models with perfect predictability, lots of molecular descriptors representing a wide variety of molecular properties should be considered. The descriptors used in this paper included size, shape, symmetry, charge distribution and topological indices of the 3D structures, which were called SEDs descriptors¹⁴ and computed with the structures having minimized energy.

Energy minimization of the chemical structures was performed with MM2 method until the root mean square gradient value became smaller than 0.1 kcal/mol Å. Re-optimization of minimized molecules was carried out *via* AM1 method until the root mean square gradient attained a value smaller than 0.1 kcal/mol Å using MOPAC. The descriptor values for all the molecules were available in the "analyze" option of the Chem3D packing. Then, the following descriptors were computed: Balaban index (BIdx), Cluster count (ClsC), Connolly solvent-excluded volume (SEV), Connolly molecular area (MS), Connolly accessible area (SAS), Dipole length (DPLL), diameter (Diam), Electronic energy (ElcE), Exact mass (Mass), HOMO energy (Homo), LUMO energy (Lumo), Molecular topological index (TIdx), Molecular weight (MW), Ovality, Principal moment of inertia-X (PMIX), Principal moment of inertia-Y (PMIY), Principal moment of inertia-Z (PMIZ), Radius (Rad), Repulsion energy (NRE), Shape attribute (ShpA), Shape coefficient (ShpC), Sum of degrees (SDeg), Sum of valence degrees (SVDe), Total valence connectivity (TVCon), Total connectivity (TCon), Total energy (TotE), Wiener index (WIdx).

Stepwise multiple linear regression is very popular in quantitative structure-retention relationships studies^{15,16} and is implemented in almost all statistical software packages, including SPSS program. In this paper, the SPSS statistical package was used in the regression analysis to select the most relevant descriptors and produce multivariate regression models for each group of compounds. At last the leave-one-out cross validation were employed to evaluate the stability of the models.

RESULTS AND DISCUSSION

Aroma compound analysis: Without pretreatment, the aroma profile of Fenjiu was analyzed by applying GC-MS

analyzer and was shown in Fig. 1. As shown in Fig. 1, a total of 73 compounds are detected by GC-MS and 65 compounds are identified, most of which are esters while others are alcohols, aldehydes, ketones, acids, *etc.* (Table-1). Esters are identified as the most abundant class of identified compounds; they are followed by alcohols. Among the esters identified, Wang¹⁷ reported that ethyl acetate and ethyl lactate were the predominant constituents (except ethanol) in Fenjiu sample and consequently the main body flavour components. The ethyl acetate contributes towards the general perception of "fruitiness" in the wine¹⁸, while the odor of ethyl lactate when dilute is mild, buttery, creamy, with hints of fruit and coconut¹⁹ and they combined make Fenjiu taste fragrant.

TABLE-1
AROMA COMPOUNDS OF FENJIU BY GC-MS
AND THEIR RETENTION TIMES

Peak No.	Compounds	t_r (Exp) ^f	t_r (Cal) ^g
1	Acetaldehyde ^b	4.618	3.315
2	Isobutyraldehyde ^b	5.137	9.151
3	Ethyl formate ^d	5.549	8.138
4	Ethyl acetate(including acetal) ^d	6.522	6.302
5	Ethanol ^a	10.076	8.410
6	2-Butanol ^a	10.689	11.633
7	n-Propanol ^a	11.726	13.857
8	1,1-Diethoxyisopentane ^c	12.233	14.740
9	Isobutanol ^a	13.778	12.571
10	Isoamyl acetate ^d	15.015	16.202
11	Ethyl valerate ^d	16.041	19.293
12	Butanol ^a	17.384	16.752
13	Isopentanol ^a	18.304	18.090
14	Ethyl caproate ^d	19.753	19.727
15	Pentanol ^a	19.912	20.208
16	3-Methylbut-3-en-1-ol ^a	19.945	20.224
17	Unknown	–	–
18	1,1,3-Triethoxypropane ^c	21.498	19.784
19	Ethyl heptanoate ^d	21.879	25.196
20	3-Hydroxybutanone ^b	22.017	17.905
21	Anisole ^c	22.545	25.351
22	Ethyl lactate ^d	23.467	19.458
23	Hexanol ^a	24.640	23.572
24	Ethyl octanoate ^d	25.696	26.265
25	Ethyl 2-hydroxy-3-methylbutanoate ^d	26.215	22.394
26	1,2-Propanediol ^a	26.353	27.054
27	Heptanol ^a	26.913	27.629
28	Acetic acid ^c	27.664	27.561
29	Furaldehyde ^b	28.161	29.577
30	Benzaldehyde ^b	29.621	29.577
31	Propanoic acid ^c	30.340	30.497
32	(R,R)-2,3-Butanediol ^a	30.753	31.980
33	Isobutyric acid ^c	31.884	32.674
34	meso-2,3-Butanediol ^a	32.094	31.294
35	Ethyl decanoate ^d	33.755	32.814
36	Butyric acid ^c	34.042	33.527
37	gamma-Butyrolactone ^d	34.878	34.540
38	Isovaleric acid ^c	35.458	36.094
39	Diethyl succinate ^d	35.660	31.813
40	2,2-Diethoxyethylbenzene ^e	37.002	34.916
41	Unknown	–	–
42	Unknown	–	–
43	Pentanoic acid ^c	38.167	36.932
44	Ethyl phenylacetate ^d	38.674	39.840
45	1-Nitro-4-ethenylbenzene ^e	39.605	35.309
46	Phenethyl acetate ^d	40.018	40.730
47	Ethyl laurate ^d	40.737	39.369
48	Hexanoic acid ^c	41.149	40.948
49	Unknown	–	–

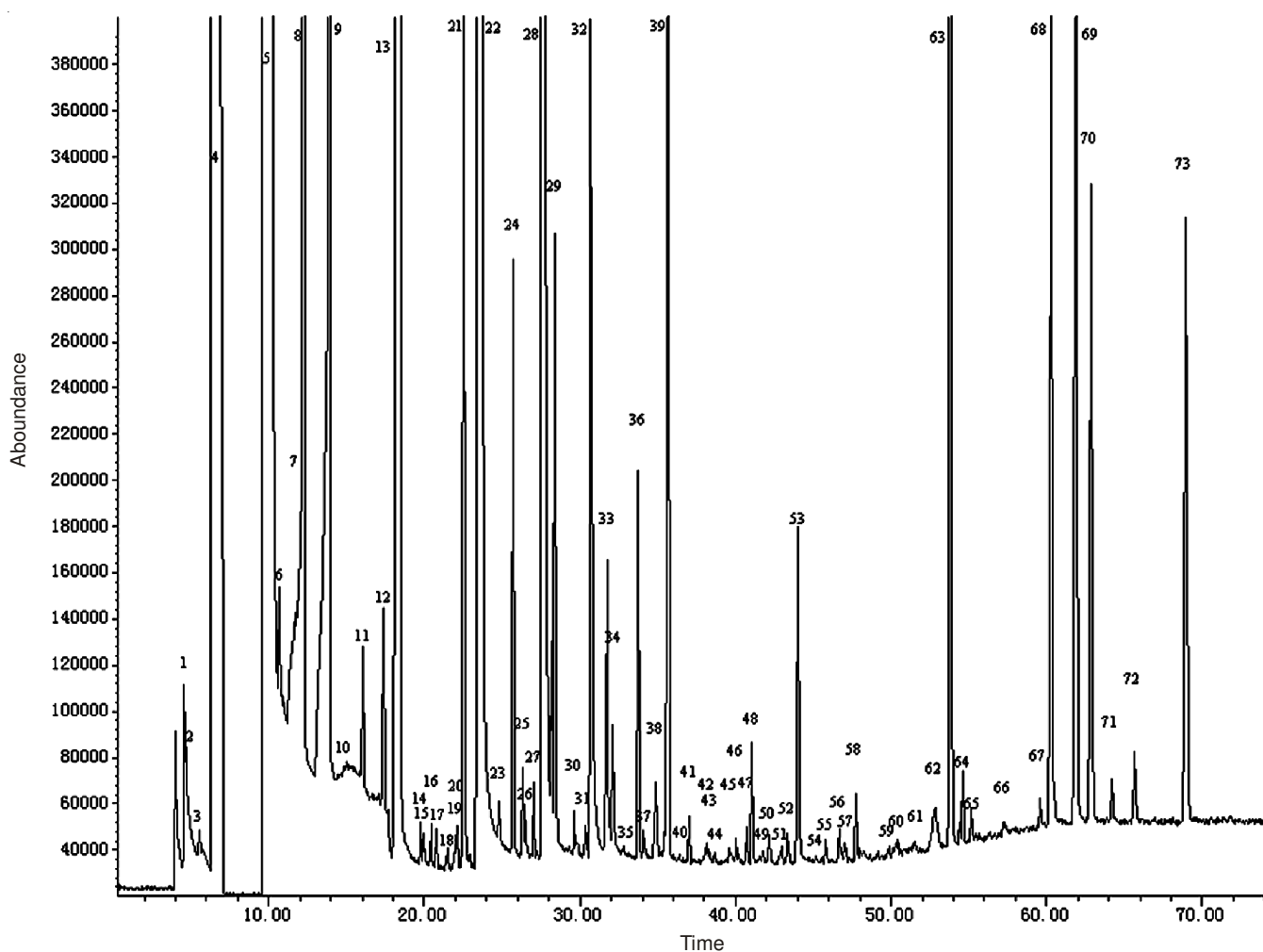


Fig. 1. GC-MS chromatogram of Fenjiu. The peak numbers in Fig. 1 correspond to the compounds in Table-1

Peak No.	Compounds	t_R (Exp) ^f	t_R (Cal) ^g
50	Unknown	—	—
51	Benzoic acid ^c	42.968	43.420
52	Diisobutyl succinate ^d	43.360	44.859
53	Phenylethyl alcohol ^a	44.025	43.698
54	Unknown	—	—
55	Heptanoic acid ^c	45.771	45.775
56	Unknown	—	—
57	Ethyl myristate ^d	47.009	45.926
58	4-Pentyl- γ -butyrolactone ^d	47.728	46.309
59	2-Methylbiphenyl ^e	49.896	50.441
60	Diisobutyl adipate ^d	50.372	54.560
61	Ethyl pentadecanoate ^d	51.535	50.422
62	4,4'-Dimethoxy-diphenylmethane ^c	52.878	55.092
63	Ethyl palmitate ^d	53.904	52.486
64	Ethyl 11-hexadecenoate ^d	54.392	54.424
65	Ethyl 9-hexadecenoate ^d	55.142	54.416
66	Propane-1,2,3-triol ^b	57.299	56.902
67	Ethyl stearate ^d	59.546	59.047
68	Ethyl oleate ^d	60.257	61.420
69	Ethyl linoleate ^d	61.872	62.419
70	Diisobutyl phthalate ^d	62.848	65.323
71	Ethyl α -linolenate ^d	64.209	64.145
72	Unknown	—	—
73	Dibutyl phthalate ^d	68.994	66.203

^aalcohol group; ^baldehyde and ketone group; ^cacid group; ^dester group; ^eether and hydrocarbon group; ^fexperimental retention times; ^gretention times calculated from equation 1-5

For identification of alcohols, *n*-propanol (with fruity sensory)¹⁸, isobutanol (with nail polish sensory)²⁰, isopentanol (with rancid and nail polish sensory)²⁰ are the most common alcohols next to ethanol in wine. The presence of these alcohol components defines the liquor that has good balance. Besides, phenylethyl alcohol, an important aroma compound in liquor with pleasing rose scent¹, is also identified in Fenjiu. The acids identified in Fenjiu are acetic acid, hexanoic acid, *etc.* The content of acetic acid is the highest in acids, while the flavour of hexanoic acid is the strongest in acids. These acid components make Fenjiu have the harmonious flavour. At the same time, important aroma components furfural (belongs to aldehyde class) with the odor of almond and sweet²⁰, 1,1,3-triethoxypropane and 1,1-diethoxyisopentane (belong to hydrocarbon class) with fruit type fragrance¹, are all identified in Fenjiu. In addition, those compounds which are not discussed here are indispensable that complements the other flavours making for a successful overall taste balance.

All the 65 compounds identified are mainly formed during alcoholic fermentation and demonstrate the main constituents of Fenjiu. By comparing the liquor' profile, retention time and peak area to that of the standard sample, its recognition and quality control could be made. Hence, the method provides a much simpler approach for the analysis of liquors without sample pretreatment.

TABLE-2
EQUATIONS BUILT FOR FIVE GROUPS OF COMPOUNDS

Equation No.	Group name	Equation	N ^a	R ^b	R _{CV} ^c	F ^d	S ^e
1	Alcohol	$t_R = 197.830 + 6.843SVDe - 8.506SDeg + 7.125Diam + 16.787Lumo + 25.853Homo$	15	0.997	0.990	285.849	1.284
2	Aldehyde and ketone	$t_R = -11.275 + 1.459SVDe$	5	0.969	0.932	45.925	3.499
3	Acid	$t_R = 32.820 + 0.017Tindx - 11.377TCon$	9	0.995	0.991	286.291	0.716
4	Ester	$t_R = -12.095 + 7.583SDeg - 13.847ClSC - 11.100ShpC + 11.231Lumo + 0.981SVDe + 2.256DPLL$	29	0.993	0.985	271.253	2.363
5	Ether and hydrocarbon	$t_R = 26.398 + 2.522SDeg - 62.098ShpC$	7	0.984	0.944	60.518	3.358

^aThe sample number; ^bThe multiple correlation coefficient; ^cThe coefficient of leave-one-out cross validation correlation; ^dFisher criterion; ^eThe standard deviation

Quantitative structure-retention relationships study of aroma compounds

Model generation: It is generally recognized that R values close to 1 indicate good model fit. The built models and their statistics are listed in Table-2. As it can be seen from Table-2, the models' correlation coefficients are ≥ 0.969 , indicating the high reliabilities of the quantitative structure-retention relationships models obtained. Meanwhile the Fisher criterion can be interpreted as a ratio of the variance explained by the model and the variance not explained by the model. Thus, high F values indicate that the models are statistically credible. Meanwhile, the calculated retention times were compared with the corresponding experimental values in Table-1. It is observed that the t_{RS} (Cal) for most compounds are close to their t_{RS} (Exp), showing that the derived variables have significant relevance to the chromatographic retention times of aroma compounds.

Model validation: In principle, cross validation is an important and practical method to verify a model¹⁶. Among the common types of cross validation, leave-one-out procedure is more suited to a small amount of samples^{21,22}. Thereupon, to validate the models obtained, the leave-one-out method was used to do the cross validation. Coefficient of leave-one-out correlation (R_{LOO}) is a measure of the predictability of the model on one side and the quality of the database on the other. A more stable model has a R_{LOO} coefficient closer to R . In this paper, cross validation results of built models are listed in Table-2. As can be seen from Table-2, each model's R_{LOO} is close to its R , which suggests that the built model is stable. As shown in Fig. 2, the retention times provided by the leave-one-out cross validation model are in good agreement with the chromatographic experimental data. These results indicate that the models derived from employed SEDs parameters are of good quality.

Major contributing molecular descriptors: It is generally accepted that when all chromatographic parameters (column, temperature, etc.) are constant, the molecular structure of the compound determines its retention time. So it is crucial to describe the structure of a compound appropriately. In this paper, the selected descriptors, SVDe, Tindx and SDeg express the main molecular structure information of different kinds of aroma compounds.

The retention times of alcohols, aldehydes and ketones are directly related to SVDe, which is the sum of all δ^v values in a molecule as what is defined by Kier and Hall^{23,24}, containing

detailed information of the molecular size and configuration of all non-hydrogen atoms²⁵. The coefficients of correlation between the retention times of alcohols, aldehydes and ketones and SVDe are 0.887 and 0.969, respectively. According to equations 1 and 2 (Table-2), the SVDe increases the retention times of alcohols, aldehydes and ketones showing that the molecular size and configuration are two significant factors affecting their retention times.

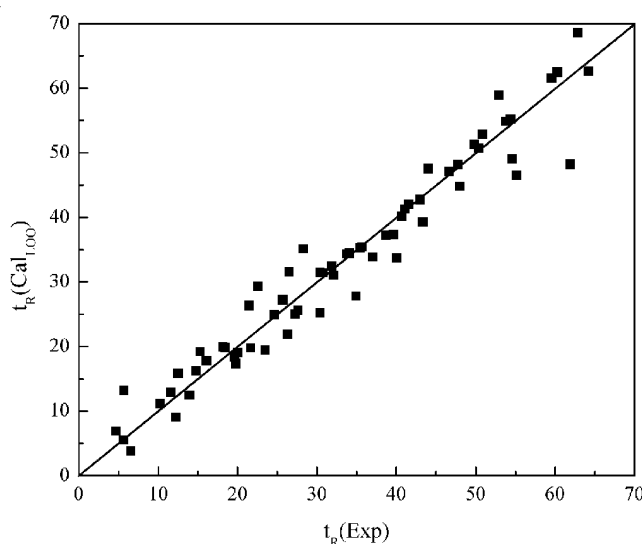


Fig. 2. Correlation of $t_R(\text{Cal}_{100})$ to $t_R(\text{Exp})$

The result in equation 3 (Table-2) indicates that retention times of acids are strongly dependent on the Tindx, with the correlation coefficient being 0.986. The Tindx provides a relative intricacy (branched) value for a molecule²⁶ and increases as the carbon number of acids increases. The positive contribution of Tindex suggests that steric-branching is a vital factor influencing acid retention times.

The retention times of esters, ethers and hydrocarbons are directly related to SDeg, the sum of the degrees of every atom. The correlation coefficients between the retention times of esters, ethers and hydrocarbons and SDeg are 0.937 and 0.816, respectively. The models 4 and 5 (Table-2) show that retention times of esters, ethers and hydrocarbons increase with raise in the magnitude of SDeg, implying that molecular size or bulkiness is an important factor affecting retention times of these compounds.

Conclusion

Without pretreatment, Fenjiu was analyzed by GC-MS, with 73 aroma compounds detected and 65 of them identified. And then quantitative structure-retention relationships models based on stepwise multiple linear regressions were developed using different molecular descriptors. Results indicate that the models have high correlation coefficients and show good predictive ability, thus these models can be used to estimate the retention times of unknown aroma compounds. Furthermore, such retention predictions may be helpful to improve the confidence in identification of aroma compounds and can serve as a supplement in characterizing liquors. The limitation of this work is that all identified compounds can not be modeled in one equation because of the large diversity in compounds' structures. Further studies on whether these 65 compounds can be modeled in an equation are in progress.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. Zhiliang Li for technical assistances. The work was supported by the foundation of 331 plan of Basic Medical College, Shanxi Medical University.

REFERENCES

1. Y.L. Ding, Master thesis, Studies on Characteristic Aroma Compounds in Fen-liquor, Jiangnan University (2008).
2. K. Héberger, *J. Chromatogr. A*, **1158**, 273 (2007).
3. K. Bouharis, M.L. Souici and D. Messadi, *Asian J. Chem.*, **23**, 1044 (2011).
4. Y.Z. Song, J.F. Zhou, S.J. Zi, J.M. Xie and Y. Ye, *Bioorg. Med. Chem.*, **13**, 3169 (2005).
5. H.Y. Xu, J.W. Zou, Y.J. Jiang, G.X. Hu and Q.S. Yu, *J. Chromatogr. A*, **1198-1199**, 202 (2008).
6. S. Muresan, C. Bologa, M. Mracec, A. Chiriac, B. Jastorff, Z. Simon and G. Náray-Szabó, *J. Mol. Struct. (Theochem.)*, **342**, 161 (1995).
7. N. Nakajima, C. Lay, H.B. Du and Z. Ring, *Energy Fuels*, **20**, 1111 (2006).
8. S. Rayne and M.G. Ikononou, *Anal. Chem.*, **75**, 1049 (2003).
9. C.B. Mazza, N. Sukumar, C.M. Breneman and S.M. Cramer, *Anal. Chem.*, **73**, 5457 (2001).
10. P. Bruneau, *J. Chem. Inf. Comput. Sci.*, **41**, 1605 (2001).
11. J. Ghasemi, S. Saaidpour and S.D. Brown, *J. Mol. Struct. (Theochem.)*, **805**, 27 (2007).
12. C.Q. Liu, *Food Mach.*, **1**, 28 (2002).
13. J.R. Kou, H. Qiao, Y.G. Hu, M.P. Li and S.W. Zhang, *Liquor-making Sci. Tech.*, **5**, 58 (2005).
14. X.H. Liu, Z.S. Gao, L.S. Wang and H.X. Yu, *Chinese J. Environ. Chem.*, **26**, 184 (2007).
15. R. Hemlatha, L.K. Soni, A.K. Gupta and S.G. Kaskhedikar, *Indian. J. Chem.*, **45B**, 1242 (2006).
16. L. Xu and X.G. Shao, *Methods of Chemometrics*, Scientific Press of China, Beijing, edn 2 (2004).
17. Y.T. Wang, *Liquor-making Sci. Tech.*, **3**, 27 (2004).
18. Y. Fang and M.C. Qian, *Flavour Fragr. J.*, **20**, 22 (2005).
19. R.A. Peinado, J. Moreno, J.E. Bueno, J.A. Moreno and J.C. Mauricio, *Food Chem.*, **84**, 585 (2004).
20. W.L. Fan and M.C. Qian, *Flavour Fragr. J.*, **21**, 333 (2006).
21. L.J. Xie, X.Q. Ye, D.H. Liu, Y.B. Ying and X.Q. Rao, *Food Res. Int.*, **44**, 2198 (2011).
22. R. Karoui, A. Mouazen, É. Dufour, L. Pillonel, E. Schaller, D. Picque, J. De Baerdemaeker and J.O. Bosset, *Eur. Food Res. Technol.*, **223**, 44 (2006).
23. A. Afantitis, G. Melagraki, H. Sarimveis, P.A. Koutentis, J. Markopoulos and O. Igglessi-Markopoulou, *Bioorg. Med. Chem.*, **14**, 6686 (2006).
24. L.B. Kier and L.B. Hall, *Molecular Connectivity in Structure Activity Analysis*, Wiley, New York (1986).
25. J. Xu, B. Chen, Q.J. Zhang and B. Guo, *Polymer*, **45**, 8651 (2004).
26. H.P. Schultz, *J. Chem. Inf. Comput. Sci.*, **29**, 227 (1989).