Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2987–2991

BIOORGANIC & MEDICINAL CHEMISTRY

The 3-D QSAR Study of Anticancer 1-*N*-substituted Imidazo- and Pyrrolo-quinoline-4,9-dione Derivatives by CoMFA and CoMSIA

Myung-Eun Suh,* Min-Jung Kang and So-Young Park

Division of Medicinal Chemistry, College of Pharmacy, Ewha Womans University, Seoul 120-750, South Korea

Received 7 February 2001; accepted 26 May 2001

Abstract—The 3-D QSAR analysis with new imidazo- and pyrrolo-quinolinedione derivatives was conducted by Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA). When crossvalidation value (q^2) is 0.844 at four components, the Pearson correlation coefficient (r^2) of the CoMFA is 0.964. In the CoMSIA, q^2 is 0.709 at six components and r^2 is 0.969. Unknown samples were analyzed, using QSAR analyzed results from the CoMFA and CoMSIA methods. Excellent agreement was obtained between, with an error range of 0.01–0.15 the calculated values and measured in vitro cytotoxic activities against human lung A-549 cancer cell lines.

Introduction

Computer-aided molecular modeling allows us to predict their biological activity, toxicity and the nature of the pharmacophore. Nowadays, many medicinal chemists use Quantitative Structure Activity Relationships (QSAR) methods because it hoped that the procedure will minimize the number of compounds that synthetic chemists should prepare and the time needed to discover new drug candidates. Actually, the correlation of physicochemical properties with biological activity is believed to offer a useful tool for the design of new drugs.

In the present study, new 1-*N*-substituted imidazo-[1] and pyrrolo-quinoline-4,9-dione derivatives[2,3] were synthesized as the prodrugs of anticancer agents (Table 1). The structures have the required disposition to be DNA intercalators according to Moore's theory.[4]

QSAR based on the three-dimensional (3-D) structures of the ligands alone involve two methods, Comparative Molecular Field Analysis (CoMFA),[5] and Comparative Molecular Similarity Indices Analysis (CoMSIA).[6]

Recently, more advanced techniques have attempted to model the receptor environment from the perspective of the ligand structure. QSAR studies incorporate 3-D information for ligands and provides a more detailed analysis of ligand–receptor interactions.

The CoMFA program places the drug molecules with a steric or an electrostatic probe at evenly spaced grid points.

The CoMSIA program is known as one of the new 3-D QSAR descriptors. In CoMSIA, the steric and electrostatic features, hydrogen bond donors, and acceptors and hydrophobic fields are considered.

The 3-D methods illustrate the conformation or spatial orientation of molecules. In addition, they provide indicators useful for the design of new compounds or drug candidates. The cytotoxicities of 1-*N*-substituted imidazo- and pyrrolo-quinoline-4,9-dione derivatives in vitro were compared with the values predicted by the QSAR methods—CoMFA and CoMSIA (Figs 1 and 2).

Methods

Molecular 3-D structure building

Structures of entire sets of 1-*N*-substituted imidazoquinoline-4,9-dione analogues were built using Sybyl 6.6 version[7] Molecular Modeling Software. Structural energy minimization was performed using the standard

*Corresponding author. Tel.: +822-3277-3040; fax: +822-3277-2851; e-mail: suh@mm.ewha.ac.kr

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00196-1

Table 1. Functional groups of the imidazo- and pyrrolo- quinoline-4,9-dione derivatives (Fig. 2)

Structure	Name	Fuctional group (R,X)	Structure	Name	Fuctional group (R,X)
A	**K1**	$-CH_3$	B	**K12**	–O–
A	**K2**	$-p$-$C_6H_4NO_2$	C	**P1**	$-CH_2CH_3$
A	**K3**	$-p$-C_6H_4Cl	C	**P2**	$-CH_2CH_2CH_3$
A	**K4**	$-p$-$C_6H_4OCH_3$	C	**P3**	–Cyclopropyl
A	**K5**	$-p$-$C_6H_4CH_3$	C	**P4**	$-CH_2CH_2OH$
A	**K6**	$-p$-$C_6H_4CF_3$	C	**P5**	$-CH_2C_6H_5$
A	**K7**	$-CH_2CH_3$	C	**P6**	$-C_6H_5$
A	**K8**	$-CH_2CH_2Cl$	C	**P7**	$-C_6H_4F$
A	**K9**	$-CH_2C_6H_5$	C	**P8**	$-C_6H_4Cl$
A	**K10**	$-CH_2CH_2OH$	C	**P9**	$-C_6H_4Br$
B	**K11**	$-CH_2-$			

Tripos molecular mechanics force field and Gasteiger–Hückel charge in the software, with a 0.001 kcal/mol energy gradient convergence criterion on the Silicon Graphics IRIS Octane computer system.

Methods of QSAR analysis

Low energy conformation was searched for using a systematic and grid conformational search. All of the structures generated were aligned in a 3-D lattice box using imidazoquinoline-4,9-dione as a common fitting structure.

In this work, we measured r^2 and q^2 values. The r^2 value being the Pearson correlation coefficient, in our case, this refers the correlation between the calculated and observed biological activity; q^2 is the predicted value based on a leave-one-out (LOO) cross-validation study.[8]

The two QSAR methods are ligand-based QSAR techniques. In this study, the CoMFA and CoMSIA modules of Sybyl (version 6.6, Tripos Inc.) were used.

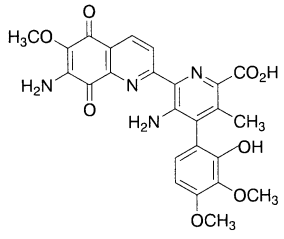

Figure 1. Structure of the streptonigrin.

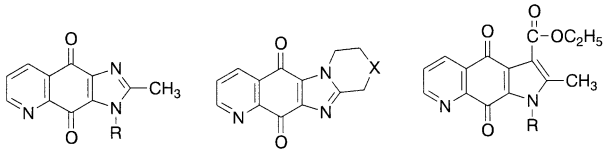

Figure 2. Structures of the imidazo- and pyrrolo-quinoline-4,9-dione derivatives.

$$q^2 = 1 - \frac{\sum_{i=1}^{N}(y_{i\,\text{observed}} - y_{i\,\text{predicted}})^2}{\sum_{i=1}^{N}(y_{i\,\text{observed}} - y_{i\,\text{observed}})^2}$$

In this study, 21 compounds were analyzed, but compounds **K2, K9** and **K10** were omitted from the CoMFA and CoMSIA. As a result of factor analysis, the compounds having the bad factors influenced inaccurately on QSAR methods were omitted.

Comparative molecular field analysis (CoMFA). CoMFA is one of the more well known 3-D QSAR methods. It incoporates steric and electrostatic values as well as C logP values and ClogP accounts for the hydrophobic ligand parameters.

In CoMFA analysis, ligands are placed in a 3-D lattice and their steric and electrostatic fields are calculated at each lattice grid point. The resulting field matrix is analyzed by Partial Least Squares (PLS) for the compounds.

The 3-D lattice set up was $22\times16\times19$ Å^3 with a 1 Å grid spacing for both the steric and electrostatic fields, the default cutoff used was 30 kcal/mol.

Comparative molecular similarity indices analysis (CoMSIA). CoMSIA is a new descriptor. Gerhard Klebe, Ute Abraham, and Thomas Mietzner developed CoMSIA while at BASF Ludwigshafen, Germany. This technique is most commonly used to determine the common features that are important in the binding of a drug to the biologically relevant receptor.

In CoMSIA, steric and electrostatic feature, hydrogen bond donors, and acceptors and hydrophobic fields are considered.

The equation used to calculate the similarity indices is as follows:[9]

$$A^{q}_{F,K,(j)} = -\Sigma W_{probe,k} W_{ik}\, e^{-\alpha\gamma 2iq}$$

A is the similarity index at grid point q, summed over all atoms 'i' of the molecule 'j'. $W_{probe,k}$ is the probe atom of radius 1 Å charge +1, hydrophobicity index +1, hydrogen bond donating +1, and hydrogen bond accepting +1. W_{ik} is the value of the physicochemical property 'k' of atom 'i'. The distance between the probe atom at grid point 'q' and atom 'i' of the test molecule is represented by r_{iq}, and α is the attenuation factor, larger values of which result in steeper Gaussian function curves and a strong attenuation of the distance-dependent effects of molecular similarity.

Results and Discussion

A trial set of 21 compounds was synthesized. CoMFA, CoMSIA, and C logP were used as descriptors and the predicted activities were compared to the cytotoxicity

against human lung cancer cell lines (A-549) as a dependent column.

In the studies, the q^2 value of CoMFA was 0.844 at four compounds and r^2 value was 0.964. In CoMSIA, q^2 value was 0.709 at six compounds and r^2 was 0.969 (Table 3). The r^2 would have been comparatively accurated itself if q^2 had over 0.5. The q^2 of all compounds was very available, so r^2 value was correct.

The Pearson correlation coefficient, r^2, showed the level to which the predicted activity approximated to the biological activity in vitro. The respecteive measured r^2 values for CoMFA and CoMSIA were 0.964 and 0.969, which means that the analyzed results have a 96.4 and 96.9% fittness compared to the biological in vitro test results. Therefore, CoMFA and CoMSIA seem to be very reliable predictors of the antitumor activity of the 1-*N*-substituted imidazo- and pyrroloquinoline-4,9-dione derivatives.

1. CoMFA

Table 4 shows the relative contributions by CoMFA analysis. The optimum value of cross-validated r^2 for 10 components was 0.844 at four compounds. In this analysis, the standard error of estimation was 0.142, r^2 was 0.964 and the F values was 88.144 ($n1 = 4$, $n2 = 13$).

2. CoMSIA

Table 5 shows the relative contributions by CoMSIA. The optimum value of cross-validated r^2 for 10 compounds was 0.709 at six compounds. By analysis, the standard error of estimation was 0.140, r^2 was 0.969 and the F values was 61.641 ($n1 = 6$, $n2 = 12$).

The results of CoMFA and CoMSIA are summarized in Tables 6 and 7, which show a comparison of the predicted activities and the actual biological activities.

Table 2. Functional groups of the unknown samples. (Fig. 2)

Structure	Name	Functional group	Structure	Name	Functional group
A	**KT1**	–*p*-C_6H_4Br	B	**KT6**	–S–
A	**KT2**	–C_3H_7	C	**PT1**	–CH_3
A	**KT3**	–C_4H_9	C	**PT2**	–$CH_2CH_2OCH_3$
A	**KT4**	–Isopropyl	C	**PT3**	–Furfuryl
A	**KT5**	–C_6H_5	C	**PT4**	–C_6H_4I

Table 3. Summary output of CoMFA and CoMSIA

	CoMFA	CoMSIA
Optimum number of components	4	6
Cross-validation, q^2	0.844	0.709
Cross-validated, r^2	0.964	0.969
Standard error of estimate	0.142	0.140
F value	($n1 = 4$, $n2 = 13$) 88.144	($n1 = 6$, $n2 = 12$) 61.641

Using the results of the analysis, we designed new compounds using QSAR optimization and synthesized 10 test compounds (Table 2). The cytotoxicities of the test compounds were compared with the corresponding CoMFA and CoMSIA predicted values by QSAR analysis (Tables 8 and 9). The results showed that the range of their errors were comparatively small, and varied between 0.03 and 0.24 in CoMFA and 0.03 and 0.3 in CoMSIA.

In the steric CoMFA map (Fig. 3), the large green colored area around the substituted group of the template molecule indicates that a sterically bulky group in this position could result in the enhancement of cytotoxicity. In the electrostatic CoMFA map (Fig. 3), the

Table 4. Relative contributions of CoMFA

Relative contributions	
CoMFA (steric)	0.520
CoMFA (electrostatic)	0.254
ClogP	0.226

Table 5. Relative contributions of the CoMSIA

Relative contributions	
CoMSIA (steric)	0.141
CoMSIA (electrostatic)	0.185
CoMSIA (hydrophobic)	0.535
CoMSIA (acceptor)	0.139

Table 6. The results of CoMFA and the comparative error values versus actual cytotoxicities

Name	A549	C logP	TRIP_Std	Log A549	CoMFA	Error
K1	0.13	−0.26	82.00	0.89	0.91	0.02
K2	4.97	1.40	114.00	−0.70	—	—
K3	0.15	2.34	110.00	0.82	0.58	0.24
K4	0.37	2.24	124.00	0.43	0.55	0.12
K5	0.17	2.12	114.00	0.77	0.72	0.05
K6	0.62	2.51	116.00	0.21	0.25	0.04
K7	0.14	0.27	88.00	0.85	0.83	0.02
K8	4.83	0.34	92.00	−0.68	—	—
K9	0.08	1.31	120.00	1.10	—	—
K10	5.34	−1.09	92.00	−0.73	—	—
K11	0.05	0.32	96.00	1.30	1.33	0.03
K12	0.03	−1.27	94.00	1.52	1.54	0.02
P1	0.93	2.17	124.00	0.03	0.05	0.02
P2	0.34	2.70	130.00	0.47	0.58	0.11
P3	0.26	2.00	128.00	0.59	0.59	0.00
P4	0.33	0.82	126.00	0.48	0.36	0.12
P5	0.31	3.21	156.00	0.51	0.38	0.13
P6	0.68	3.52	134.00	0.17	0.06	0.11
P7	4.57	3.67	138.00	−0.66	−0.32	0.34
P8	5.09	4.24	130.00	−0.71	−0.76	0.05
P9	6.15	4.39	136.00	−0.79	−0.85	0.06

logA549, Log value of activity against human lung cancer cell lines (A549); C logP, hydrophobic parameter; TRIP_Std, in the Tripos SYBYL program, steric and electrostatic field were considered as basic values; CoMFA, Activity values in the CoMFA; Error, the difference between logA549 and CoMFA.

Table 7. CoMSIA results and comparative cytotoxic error values

Name	A549	LogA549	Steric	ES	HP	H-acceptor	CoMSIA	Error
K1	0.13	0.89	6.44	1.86	3.18	5.48	0.88	0.01
K2	4.97	−0.70	7.65	1.90	12.83	6.71	—	—
K3	0.15	0.82	7.41	1.87	6.58	5.19	0.72	0.10
K4	0.37	0.43	8.05	2.00	5.40	5.19	0.40	0.03
K5	0.17	0.77	7.70	1.88	6.05	5.19	0.75	0.02
K6	0.62	0.21	7.74	2.19	6.99	5.18	0.62	0.41
K7	0.14	0.85	6.86	1.85	3.29	5.50	0.70	0.15
K8	4.83	−0.68	6.87	1.92	4.39	5.49	—	—
K9	0.08	1.10	7.75	1.97	4.94	5.49	1.08	0.02
K10	5.34	−0.73	6.97	1.94	3.82	6.32	—	—
K11	0.05	1.30	7.11	1.86	3.67	5.49	1.26	0.04
K12	0.03	1.52	6.82	1.98	3.73	6.25	1.56	0.04
P1	0.93	0.03	7.96	2.10	3.80	5.38	−0.01	0.04
P2	0.34	0.47	8.26	2.46	4.19	5.37	0.39	0.08
P3	0.26	0.59	8.11	2.09	3.94	5.36	0.50	0.09
P4	0.33	0.48	8.09	2.21	4.27	6.15	0.60	0.12
P5	0.31	0.51	8.75	2.22	5.22	5.37	0.54	0.03
P6	0.68	0.17	8.42	2.17	5.65	5.37	0.03	0.14
P7	4.57	−0.66	8.42	2.17	5.84	5.36	−0.39	0.27
P8	5.09	−0.71	8.44	2.17	6.59	5.36	−0.69	0.02
P9	6.15	−0.79	8.43	2.18	7.24	5.38	−0.91	0.12

logA549, log value of activity against human lung cancer cell lines (A549); Steric, Each value represents a steric feature of the CoMSIA; ES, the electrostatic compotent of CoMSIA; HP, Hydrophobic parameter; H-acceptor; Hydrogen bonding acceptor; CoMSIA, The predicted activity values by CoMSIA; Error, The difference between logA549 and CoMSIA.

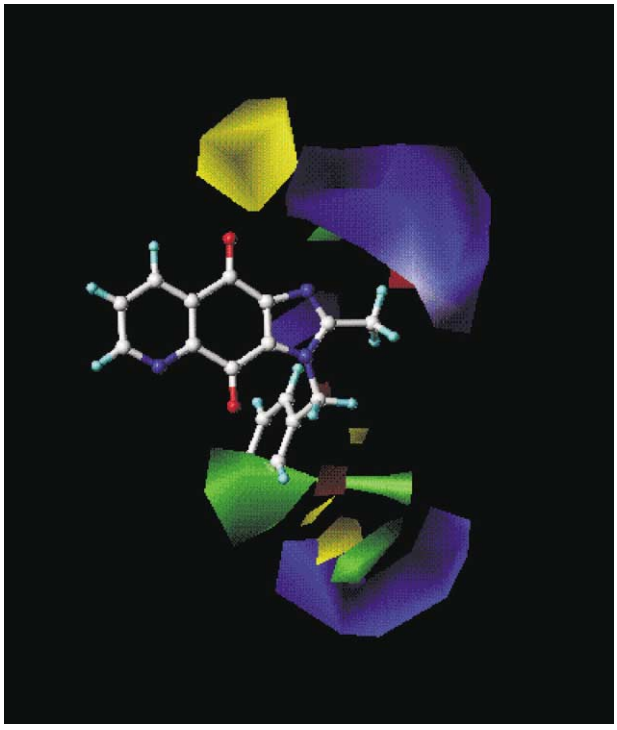

Figure 3. The steric and electrostatic CoMFA map. Red color is a negative charge region, blue a positive charge region, green a positive sterically active region, and yellow a negative sterically active region.

Table 8. The testing results of the CoMFA with unknown samples

Name	A549	TRIP_Std	log A549	CoMFA	Error
KT1	0.13	112.00	0.89	0.56	0.24
KT2	4.97	96.00	0.82	0.79	0.03
KT3	0.15	106.00	0.85	0.70	0.15
KT4	0.37	100.00	0.21	0.99	0.78
KT5	0.17	106.00	0.89	0.74	0.15
KT6	0.62	94.00	1.52	1.38	0.04
PT1	0.14	116.00	−0.08	0.12	0.20
PT2	4.83	140.00	0.43	0.58	0.15
PT3	0.08	134.00	−0.05	0.07	0.12
PT4	5.34	136.00	−0.66	−0.88	0.22

Table 9. The testing results of the CoMSIA with the unknown samples

Name	A549	logA549	Steric	ES	HP	H-acceptor	CoMSIA	Error
KT1	0.13	0.89	7.41	1.87	7.25	5.19	0.83	0.03
KT2	4.97	0.82	7.22	1.85	3.74	5.50	0.75	0.07
KT3	0.15	0.85	7.55	1.85	4.12	5.50	0.80	0.05
KT4	0.37	0.21	7.26	1.86	3.57	5.54	0.47	0.26
KT5	0.17	0.89	7.39	1.87	5.66	5.19	0.50	0.39
KT6	0.62	1.52	6.81	1.89	3.54	5.57	1.35	0.17
PT1	0.14	−0.08	7.60	2.10	3.70	5.37	0.20	0.28
PT2	4.83	0.43	8.32	2.26	4.27	6.43	0.87	0.44
PT3	0.08	−0.05	8.43	2.20	4.29	5.38	0.34	0.39
PT4	5.34	−0.66	8.44	2.17	8.53	5.36	−0.92	0.28

logA549, log value of activity against human lung cancer cell lines (A549); Steric, Each value represents a steric feature of the CoMSIA; ES, The electrostatic component of the CoMSIA; HP, Hydrophobic parameter; H-acceptor, Hydrogen bonding acceptor; CoMSIA, The predicted activity values by CoMSIA; Error: The difference between logA549 and CoMSIA.

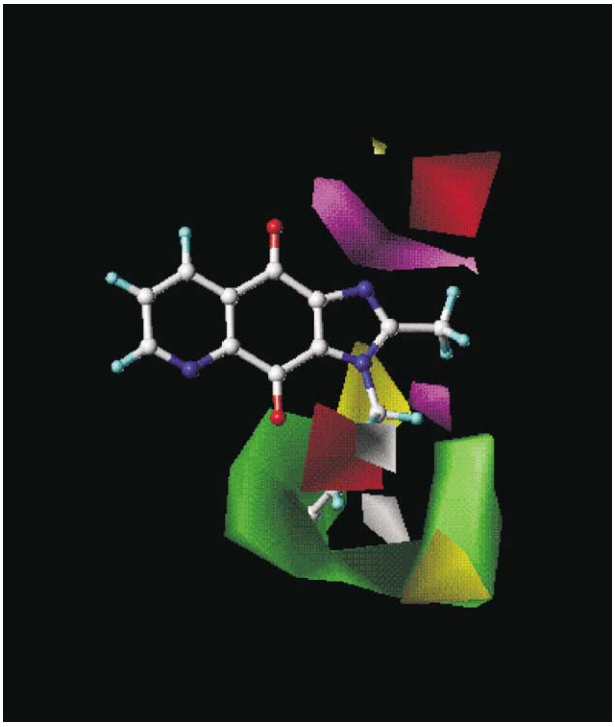

Figure 4. The steric and electrostatic CoMSIA map. Red color represents the negative charge region, blue is the positive charge region, green is the more bulky region, yellow the less bulky region, gray the hydrophilic region and purple the hydrogen acceptor region.

Table 10. Summary of the error values of CoMFA, and CoMSIA

	E_CoMFA	E_CoMSIA
Mean	0.08	0.09
Standard deviations	0.09	0.10
High	0.34	0.41
Low	0.00	0.01

red color shows better activity and conversely blue shows worse activity, that is meaning that if more electronegative groups are introduced into the red region, better activities are expected. Conversely, the blue color indicates the opposite case, less electronegative groups located in the blue region induce lower activities.

In the electrostatic and steric CoMSIA map (Fig. 4), the red color shows that introducing more electronegative groups increases antitumor activities. The green area locating more bulky and steric group indicates more powerful antitumor activities. In the hydrophobic and hydrogen bonding CoMSIA map (Fig. 4), the hydrophobic region is colored yellow, the hydrophilic region gray and the purple color indicates the region in which hydrogen bonding acceptors confer better activities.

The CoMFA and CoMSIA error indicates the extent to which CoMFA, and CoMSIA values approximate the biological values obtained by testing with human lung cells. In terms of the CoMFA error, compounds **K1**, **K7**, **K12** and **P1** proved to be lower than the other values. The difference is 0.02. In terms of the CoMFA error, the **P3** values for analyzed and the tested cytotoxicity values obtained with human lung cancer cell were identical. In addition, compounds, **K1**, **K5**, **K9**, **P5** and **P8** calculated values were also very similar to actual by CoMSIA. We believe that CoMSIA predicts the biological activity of unknown 1*N*-subsituted imidazo- and pyrrolo-quinoline-4,9-dione derivatives with a high level of accuracy.

When mean errors of CoMFA and CoMSIA were calculated, the means for both lay in the range 0.08–0.09 (Table 10). In the Table 10, the standard deviations of errors of CoMFA and CoMSIA are very similar. In conclusion, when 1*N*-subsituted imidazo- and pyrrolo-quinoline-4,9-dione derivatives are synthesized, CoMFA and CoMSIA proved very useful for predicting their biological activities, though it is true that the biological activities are not absolutely correct. However, it is very important that their biological activities can be predicted with reasonable accuracy using the QSAR methods and this was achieved despite the fact that the exact mechanism and effects of these species are unknown in the human body.

Conclusion

The 3-D QSAR analysis, CoMFA and CoMSIA were available for the imidazo- and pyrrolo-quinolinedione derivatives to predict their biological activity. The biological activities of the unknown samples, with imidazo- and pyrrolo-quinolinedione derivatives are easily predicted with 3-D QSAR analysis. It is also useful to make a plan to synthesize new compounds with good biological activities.

Acknowledgements

The authors wish to thank Korea Science and Engineering Foundation for the financial support.

References and Notes

1. Suh, M. E.; Kang, M. J.; Yoo, H. W.; Park, S. Y.; Lee, J. O. *Bioorg. Med. Chem.* **2000**, *8*, 2079.
2. Suh, M. E.; Park, H. K. *Yakhakhoeji* **1996**, *40*, 19.
3. Suh, M. E.; Park, H. K. *Yakhakhoeji* **1997**, *41*, 582.
4. Moore, M. H.; Hunter, W. H.; Kennard, O. *J. Mol. Biol.* **1987**, *206*, 693.
5. Cramer, I. I. I. R.D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
6. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 24.
7. *Sybyl version 6.5*; Tripos Associates: St. Louis, 1998.
8. So, S. S.; Karplus, M. *J. Med. Chem.* **1997**, *40*, 4360.
9. Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemisty*; Wiley: New York, 1980.