ORIGINAL ARTICLE

QSPR modeling of stability constants of diverse 15-crown-5 ethers complexes using best multiple linear regression

Jahanbakhsh Ghasemi · Saadi Saaidpour

Abstract Quantitative structure–property relationships (QSPR) models for the stability constants of 58 complexes of 1,4,7,10,13-pentaoxacyclopentadecane ethers (15C5) were established with the Comprehensive Descriptors for Structural and Statistical Analysis (CODESSA) program. Experimental stability constants (log K) for a diverse set of 58 complexes of 15C5 structures are correlated with computed structural descriptors using CODESSA. Stability constants for complexes of 15C5 ethers with potassium cation (K$^+$) have been determined at 25 °C in methanol solution. Standard quantum chemistry packages are used for optimizing the molecular geometry and for semi-empirical quantum computations. The QSPR model for the stability constants (log K) is obtained by selecting descriptors from a wide diversity of constitutional, geometrical, topological, electrostatic, quantum chemical, and thermodynamic molecular descriptors. The best multilinear regression method (BMLR) encoded in CODESSA software was used to select significant descriptors for building multilinear QSPR model and the predictive power of model is estimated with the leave-one-out (LOO) cross-validation method. The proposed model can be used for the prediction of the stability constants of 15C5 complexes. The best QSPR model with five descriptors has $R^2 = 0.9452$, $s^2 = 0.0110$, and $F = 67.0312$.

Keywords 15-Crown-5 ethers · Stability constants · QSPR · BMLR · Prediction

J. Ghasemi (✉) · S. Saaidpour
Chemistry Department, Faculty of Sciences, Razi University, Kermanshah, Iran
e-mail: Jahan.ghasemi@gmail.com

Abbreviations

QSPR	Quantitative structure–property relationships
15C5	1,4,7,10,13-Pentaoxacyclopentadecane
CODESSA	Comprehensive Descriptors for Structural and Statistical Analysis
BMLR	Best multilinear regression
log K	Stability constant of complex formation
AM1	Austin Method 1
TMSA	Total molecular surface area
PNSA1	Partial negatively charged molecular surface area
WNSA1	Surface weighted charged partial negative charged surface area

Introduction

Crown ethers are compounds with multiple oxygen heteroatom (three or more) incorporated in a monocyclic carbon backbone. They were first synthesized by Pedersen in 1967 [1, 2]. Their generic name originates from their molecular shape, reminiscent of a royal crown. Owing to the nature of their binding sites and to the presence of a hydrophilic cavity delineated by a lipophilic envelope, crown ethers exhibit a strong affinity and high selectivity for alkali and alkaline earth metal ions. They were the first synthetic ligands for which this pronounced selectivity was identified. Crown ethers were extensively studied in parallel with natural ion-selective cyclic antibiotics such as valinomycin or enniatin for which they serve as simple models, helping to explain the transport of these biologically relevant cations and the mechanism of neurotransmission [3–6].

Crown ethers have found applications in many areas based on their ability to selectively recognize metal and ammonium ions. In analytical chemistry, their selective

metal ion binding properties are exploited in separation and transport processes for the recovery or the removal of cations, in their concentration from very dilute solutions (trace enrichment of radionuclides) and in the design of ion-selective electrodes. They have also been used bonded to the stationary phase in chromatographic techniques. Owing to their ability to dissolve salts in organic media, by reducing the cation/anion interaction (i.e., by shielding the cation and activating the anion), they have been used in many syntheses, and as catalysts in phase-transfer catalysis or enzyme mimics. They also have medical applications as diagnostic or therapeutic agents [7–9].

Crown ethers have been extensively utilized as phase-transfer catalysts in organic synthesis [10, 11], as neutral carriers in ion sensors [12], and for enhancing the separation selectivity in chromatography, electrophoresis, membrane transfer, and solvent extraction of cations [13–17]. Crown ethers have also received special interest from physical chemists, who seek to understand the origins of their complexation selectivity. A variety of efficient approaches have been utilized to elucidate the structural and thermodynamic aspects of this chemistry. X-ray crystallography has, for example, played an important role in revealing unambiguously the solid-phase structures of crown ether complexes [18]. Based on such studies, it was directly confirmed that a cation is accommodated in the cavity of a crown ring. This structure has led to the well-known size fit theory, which has been successfully applied to explain crown ether complexation. However, we must bear in mind that a size fit is just one of the factors governing the selectivity in crown ether complexation. Crown ether complexation in solution has often been discussed according to thermodynamic data obtained with calorimetric, spectroscopic, and electrochemical and separation methods. Izatt and his coworkers have made great contributions to the accumulation of thermodynamic data of crown ether complexation, mainly by means of calorimetry, and have published comprehensive reviews in which extensive discussions are given based on a number of thermodynamic and kinetic data compiled for various systems [19, 20]. Although these data have allowed us to discuss systematically crown ether complexation in solution, it is a problem that thermodynamic studies do not give direct information on solution phase structures and therefore most discussions strongly depend on the corresponding solid-phase structures revealed by X-ray crystallography. Thus, a number of researchers still show great interest in the solution phase complexation of crown ethers. Recent developments in computer science have allowed discussions of the relations between complexation and solution phase structures of crown ethers [21–24].

The stability constants of macrocyclic ligand–metal ion complexes in solutions have been measured by various experimental techniques such as potentiometry, conductometry, and calorimetry [25]. The cation–crown ether interaction depends on several factors, such as the properties of the cation, the ligand (the ratio the diameters of the cation and the crown ether cavity), and the solvent, particularly its ability to solvate cation and ligand molecules [26–32].

Crown ether moiety is one of the most popular host compounds in host–guest chemistry field, and they showed remarkable ability to form strong complexes with alkali metal, alkaline metal, and organic cations [33, 34]. When the metal cations radii exactly fit the size of the crown ether unit, it always forms a 1:1 host/guest complex. However, compounds that consist of more than one crown ether units in the same molecule, it may produce ''sandwich type'' complexes with the metal cations. When the cation size exceeds that of the crown ether cavity such as the ''butterfly crown ethers'' [35, 36], it may form a host/guest 2:1 sandwich type complex [37–40].

Crown ethers are widely used as complexing agents that can capture selectively metal cations in their cavity. Their ability to bind metal cations depends on many factors, such as the cavity size, the nature of heteroatoms (oxygen, nitrogen, or sulfur) and substituents in the macrocycle, and the solvent. Recently, crown ethers have found new application as components of supramolecular systems [41].

In our previous papers, we reported on the application of quantitative structure–property/activity relationships (QSPR/QSAR) techniques in the development of a new, simplified approach to prediction of compounds properties using different models [42–46]. The aim of the present study is to build QSPR multiple regression model, which could correlate and predict the stability constants (log K) of complexation between 58 diverse 15-crown-5 ethers and metal ion K$^+$ in methanol solution. We used CODESSA software for the prediction of stability constants (log K); we believe the modules integrated in this software package provide an optimum way to high quality results. QSPR modeling, by multilinear regression utilized in the CODESSA program, which applies up to 575 different constitutional, geometrical, topological, electrostatic, quantum chemical, and thermodynamic molecular descriptors. The predictive power of the resulting model is demonstrated by testing them on unseen data that were not used during model generation. A physicochemical explanation of the selected descriptors is also given.

Material and methods

Dataset

The chemical structures and experimental values for the stability constants of 58 15C5 ether derivatives taken from

the literature [47] are presented in Tables 1 and 4, respectively. The dataset was split into a training set and a testing set. The training set of 46 complexes was used to adjust the parameters of the models, and the test set of 12 complexes was used to evaluate its prediction ability. Since the temperature and solvent also affect the stability constants, we used only data obtained at standard temperature (25 °C) and just in methanol solution.

Molecular modeling and descriptor calculation

The structures were drawn in HyperChem 7.5 [48] and exported in a file format suitable for AMPAC 8.16 [49]. The geometry optimization was performed on a HP Pavilion dv6000 Laptop computer with AMD Turion64 X2 Mobile Technology CPU with windows XP operating system with the semi-empirical quantum method Austin Method 1 (AM1) [50] incorporated in the AMPAC 8.16 program. The gradient norm criterion 0.01 kcal/Å was applied in the geometry optimization for all structures. The HyperChem mol files and the AMPAC output files were used by the CODESSA program [51] to compute more than 575 structural descriptors for the 58 crown ethers. CODESSA computes six classes of structural descriptors: constitutional (number of various types of atoms and bonds, number of rings, molecular weight, etc.); topological (Wiener index, Randic indices, Kier–Hall shape indices, Balaban index, etc.); geometrical (moments of inertia, molecular volume, molecular surface area, etc.); electrostatic (minimum and maximum partial charges, polarity parameter, charged partial surface area descriptors, etc.); and quantum chemical (reactivity indices, dipole moment, HOMO and LUMO energies, etc.) and thermodynamic descriptors (vibrational enthalpy of the molecule, translational enthalpy of the molecule, vibrational entropy of the molecule, rotational entropy of the molecule, etc.). The generation of the descriptors is carried out without taking into account of the solvation of the ligands molecules. It means that the generated descriptors are carried out using the gas-phase geometry calculation of AM1. We did not use any extra descriptor to consider the lariat effect and also the changing in the conformation of ligands was not considered.

Results and discussion

Multilinear regression modeling

An important stage in the multilinear regression QSPR methodology is the search for the best multilinear equation among a given pool of descriptors. The best multilinear regression (BMLR) procedure [52] available in the

framework of the CODESSA was used to find the best correlation models from selected non-collinear descriptors. Equation 1 gives the mathematical representation of the equation that should correlate the best log K with a certain number (n) of molecular descriptors (d_i) weighted by the regression coefficients b_i:

$$\log K = b_0 + \sum_{i=1}^{n} b_i d_i \tag{1}$$

The best multilinear regression method (BMLR) encoded in CODESSA Pro software was used to select significant descriptors for building multilinear QSPR models [53]. The treatment started with the reduction of the number of molecular descriptors. If two descriptors were highly correlated, then only one descriptor was selected; the descriptors with insignificant variance were also rejected. This helps to speed up the descriptor selection and reduces the probability of including unrelated descriptors by chance. The strategy used to develop physically meaningful multilinear QSPR equations from the very large pool of descriptors is a combination of the multilinear regression and forward selection procedures. This strategy involved the following steps:

(1) Detection of all orthogonal pairs of descriptors i and j from the given descriptor space. Pairs of descriptors with a correlation coefficient $R_{ij}^2 > 0.6$ were considered inter-correlated and such pairs were eliminated at this stage.

(2) Search for superior multi-parameter regression equations: for each descriptor pair, retained in the previous step, additional non-collinear descriptor vectors were successively added, and the appropriate ($n + 1$)-parameter regression treatment was carried out. When the Fisher criterion F (or cross-validation coefficient R_{cv}) obtained for any of these correlations was lower than for the best correlation of the previous rank (n), the latter was designed as the final result and the search was terminated. Otherwise, the descriptor sets with the highest coefficient of determinations were stored and the current step was repeated with the number of parameters (descriptors) increased by one ($n + 2$). The final result had therefore the maximum value of the Fisher criterion and the highest cross-validated coefficient of determination. A major decision in developing successive QSPR is when to stop adding descriptors to the model during the stepwise regression procedure.

A simple rule (''breaking point'' rule) was used to decide the optimum number of descriptors by considering the improvement of the R^2 by addition of a further descriptor to the model. If the difference between the models with n and $n + 1$ descriptors is improved by a value of less than 0.02

Table 1 Chemical structures of 58 15-crown-5 ethers

No.	Structure	No.	
1		12	
2		13	
3		14	
4		15	
5		16	
6		17	
7		18	
8		19	
9		20	
10		21	
11		22	

Table 1 continued
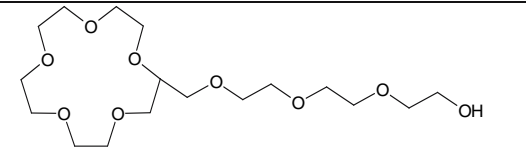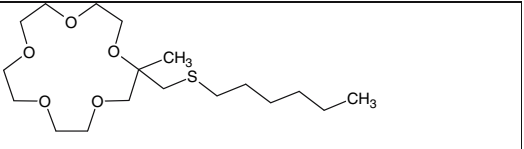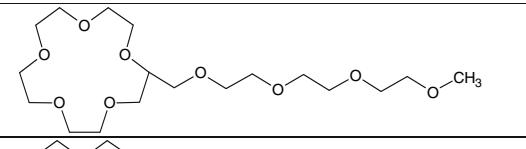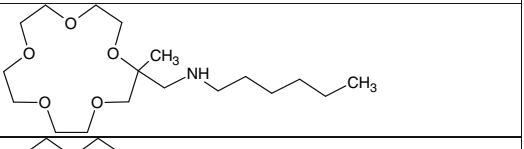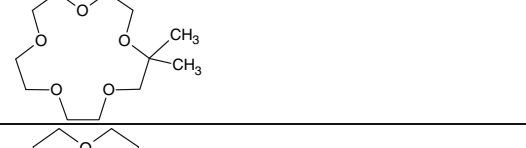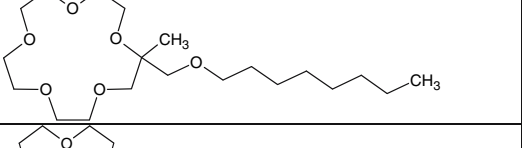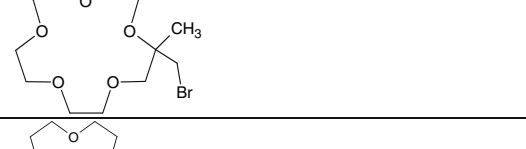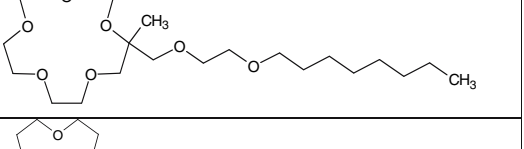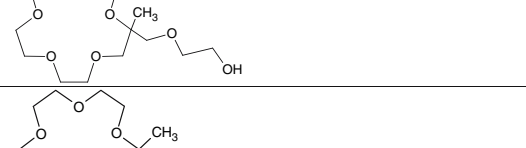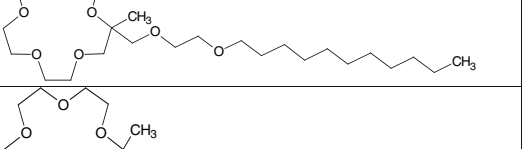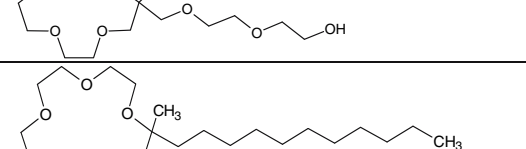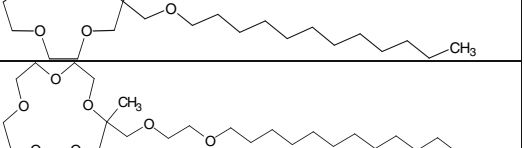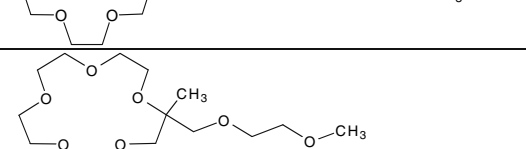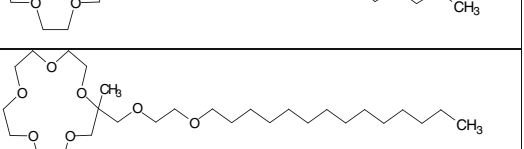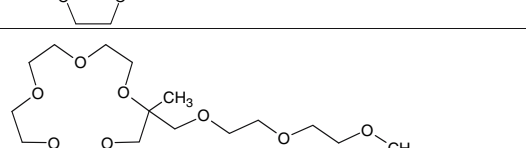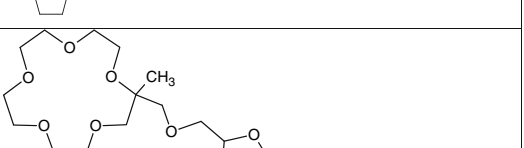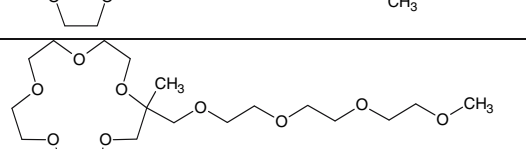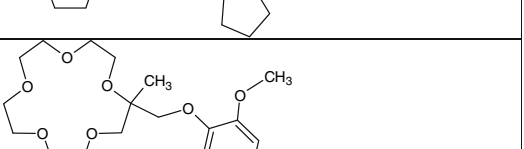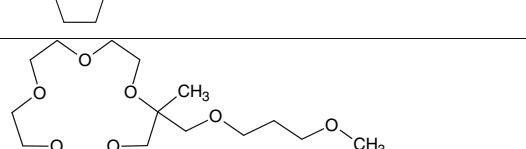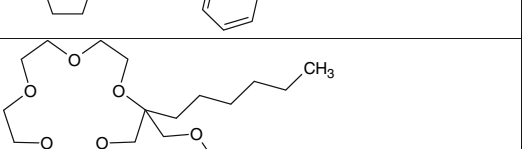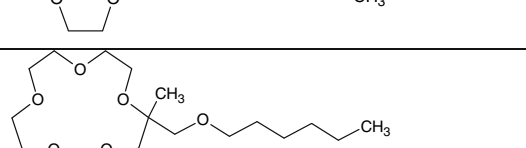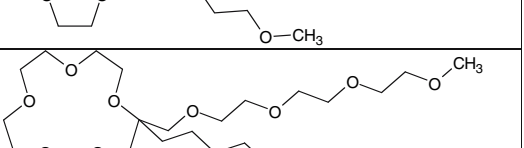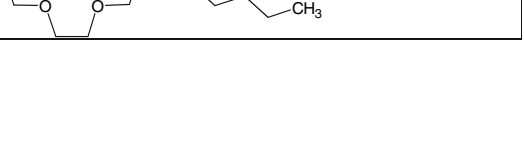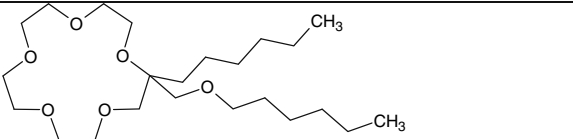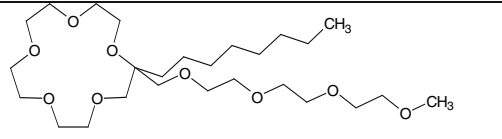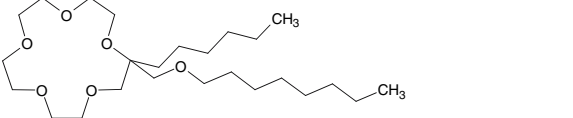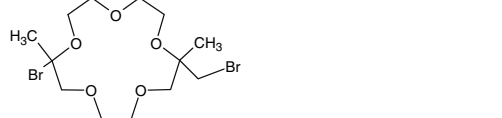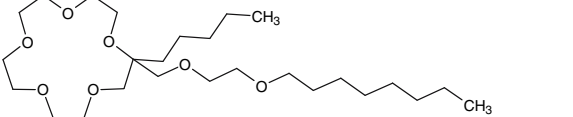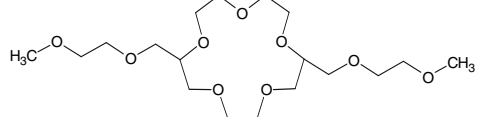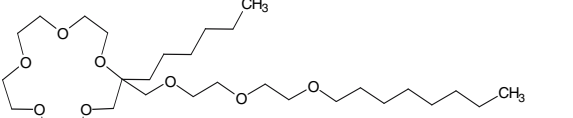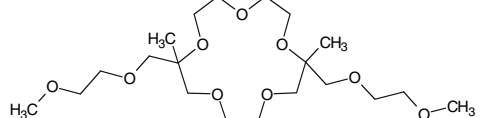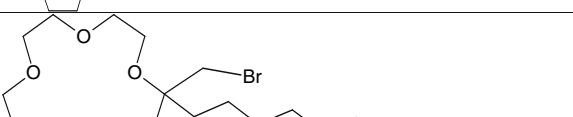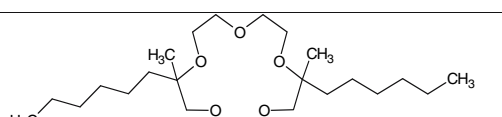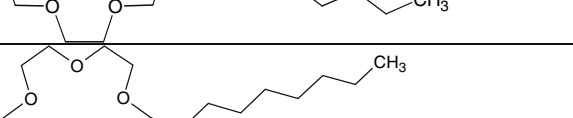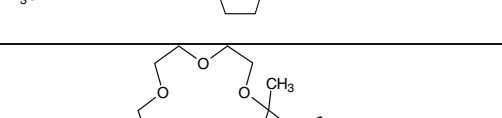
23		35	
24		36	
25		37	
26		38	
27		39	
28		40	
29		41	
30		42	
31		43	
32		44	
33		45	
34		46	

Table 1 continued

($\Delta R^2 < 0.02$), then the optimum model is taken to have n descriptors. The plot of the number of descriptors involved in the obtained models versus square correlation coefficient (R^2) and the cross-validated square correlation coefficient (R_{cv}^2) values corresponding to those models and selection of the optimum number of descriptors is shown in Fig. 1. Consequently, the model corresponding to the breaking point is considered the best/optimum model. In addition, the Fisher criterion was also monitored for a significant improvement in the correlation coefficient value with
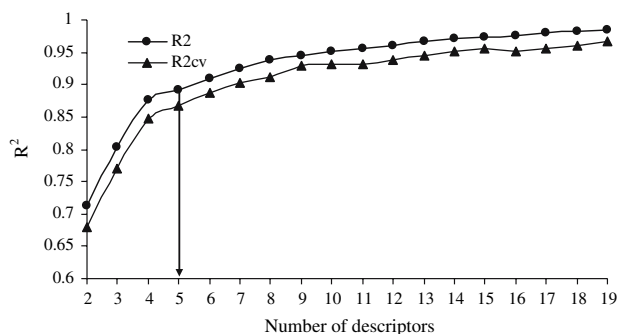
Fig. 1 Correlation and cross-validated correlation coefficients (R^2 and R_{CV}^2) versus number of descriptors for log K values of 58 crown ethers

respect to the number of descriptors. The final QSPR model selected for the crown ethers is shown in Table 2. In Table 2, X, ΔX, t-test, R^2, and s^2 are the regression coefficients, standard errors of the regression coefficients, t significance, coefficient of determination, and squared standard deviation, respectively.

The log K is assumed to be highly dependent upon the maximum net atomic charge (most positive) for hydrogen atoms, d_1; average information content (order 2), d_2; maximum nuclear–nuclear repulsion for C–H bonds, d_3; maximum nuclear–nuclear repulsion for C–C bonds, d_4; and weighted partial negative charged surface area (WNSA1), d_5. The values of five descriptors for 58 15C5 are listed in Table 3. The QSPR model was generated using a training set of 46 crown ethers. The test set of 12 crown ethers (Table 4) with regularly distributed log K values was used to assess the predictive ability of the QSPR model produced in the regression.

BMLR analysis

The best multilinear regression (BMLR) procedure available in the framework of the CODESSA PRO was used to

Table 2 The multilinear QSPR model obtained for 46 crown ethers for log K ($R^2 = 0.9452$, $F = 67.0312$, and $s^2 = 0.0110$)

No.	X	ΔX	t-Test	Descriptor
0	−7.1531	1.4498	−4.9340	Intercept
1	−1.6602	0.3534	−4.6973	Max net atomic charge for a H atom, d_1
2	−0.5744	0.0748	−7.6827	Average information content (order 2), d_2
3	0.3352	0.0488	6.8685	Max n–n repulsion for a C–H bond, d_3
4	0.0166	0.0071	2.3326	Max n–n repulsion for a C–C bond, d_4
5	0.0101	0.0008	13.2083	WNSA1 Weighted PNSA (PNSA1*TMSA/1000), d_5

Table 3 The descriptors values for 58 compounds

No.	d_1	d_2	d_3	d_4	d_5
1	0.2029	1.3788	28.0880	114.2850	26.8451
2	0.2403	2.5816	29.1660	122.6700	30.5213
3	0.2280	2.9017	29.0920	122.7320	46.3491
4	0.2191	2.8873	29.0880	122.7290	42.9662
5	0.2105	2.8451	29.0930	122.7270	51.3847
6	0.2360	2.7913	29.0940	122.7270	58.1360
7	0.2407	2.2485	28.1330	115.2020	46.2836
8	0.2325	2.4115	28.5170	114.3940	46.0634
9	0.1928	2.9377	29.0830	122.7280	65.7881
10	0.2487	2.7047	28.0350	122.1930	76.7420
11	0.2107	3.1394	28.2040	123.4380	85.0576
12	0.2642	3.1771	28.1360	123.5210	90.7570
13	0.2282	2.9608	28.1220	123.4580	122.3442
14	0.2286	3.2620	28.5080	123.5100	114.4621
15	0.2616	2.5558	28.1980	114.8220	77.7206
16	0.2280	2.7599	28.7980	114.2670	75.1756
17	0.2293	3.1283	29.1070	122.7370	88.1416
18	0.2074	2.3163	28.5150	114.2690	64.1022
19	0.1935	2.6563	29.0380	122.6710	64.2308
20	0.2349	2.8619	29.0790	122.7280	73.3991
21	0.2084	2.2379	28.4850	114.2690	77.3927
22	0.2131	2.7908	29.0820	122.7270	87.4824
23	0.2539	2.0776	28.2320	114.2990	92.4574
24	0.1912	2.1724	28.4860	114.3030	98.1164
25	0.2158	2.3193	29.1270	116.1800	32.4641
26	0.3860	2.5816	28.8740	115.9930	34.1826
27	0.2496	2.4120	29.3220	117.0080	65.0123
28	0.2488	2.3266	29.1740	116.3760	79.1355
29	0.2530	2.2549	29.3540	117.3120	96.5612
30	0.2217	2.5009	29.2390	116.0560	59.3033
31	0.2272	2.4141	29.1770	116.8020	82.1789
32	0.1951	2.3404	29.1140	116.7120	98.5604
33	0.2254	2.8512	29.0220	117.0840	62.7907
34	0.2234	2.9531	29.2020	122.7310	51.0221
35	0.2591	3.2974	29.0600	124.0850	55.8527
36	0.2588	3.3640	29.0850	122.7280	52.6296
37	0.2086	2.9801	29.0860	122.7280	54.7154

Table 3 continued

No.	d_1	d_2	d_3	d_4	d_5
38	0.2317	2.9094	29.3000	122.7260	67.8314
39	0.2066	2.8418	29.1450	122.7270	93.3363
40	0.2184	2.9307	29.0890	122.7270	61.8624
41	0.2259	2.8931	29.2890	122.7270	85.0030
42	0.2224	2.8626	29.2780	122.7280	86.3221
43	0.2153	2.9430	29.1990	122.1350	64.3698
44	0.2410	3.2726	29.1170	123.5000	92.5329
45	0.2605	3.0161	29.1350	122.7240	63.9585
46	0.2424	2.8530	29.1060	122.7280	100.7149
47	0.2689	3.0035	29.1150	122.7300	61.0331
48	0.2141	2.9697	29.1160	122.7340	64.4290
49	0.2359	2.9343	29.1070	122.7280	74.8760
50	0.2540	2.8946	29.0790	122.7330	105.9792
51	0.4945	3.0715	29.0810	122.7290	48.9480
52	0.2663	3.0215	29.1170	122.7250	64.3869
53	0.2080	2.8869	29.1060	122.7270	120.5140
54	0.3799	3.0062	29.0620	117.0410	31.9018
55	0.2596	2.5972	28.5210	114.2700	78.2697
56	0.2225	2.7976	29.1890	117.7410	87.9364
57	0.2668	2.7976	29.3450	117.0640	101.3521
58	0.2448	3.1079	29.2510	122.7040	40.8798

Table 4 Experimental and predicted stability constants of 15-crown-5 ethers complexes

No.	Log K for K$^+$ in MeOH at 25 °C			
	Experimental	Predicted	Residual	Relative error (%)
1	3.33[a]	3.31	0.02	−0.63
2	3.19[a]	3.09	0.10	−3.01
3	3.00[a]	3.07	−0.07	2.23
4	3.04[a]	3.06	−0.02	0.49
5	3.15[b]	3.18	−0.03	0.95
6	3.29[a]	3.24	0.05	−1.61
7	3.09[b]	2.97	0.12	−3.75
8	3.18[a]	3.01	0.17	−5.44
9	3.22[a]	3.30	−0.08	2.45
10	3.30[b]	3.09	0.21	−6.33
11	3.11[a]	3.07	0.04	−1.41
12	2.86[a]	2.99	−0.13	4.62
13	3.53[b]	3.49	0.04	−1.13
14	3.44[a]	3.37	0.07	−2.12
15	3.13[a]	3.10	0.03	−1.05
16	3.19[a]	3.20	−0.01	0.34
17	3.24[a]	3.36	−0.12	3.83
18	3.34[a]	3.28	0.06	−1.71
19	3.37[a]	3.43	−0.06	1.72
20	3.24[a]	3.35	−0.11	3.33

Table 4 continued

No.	Log K for K^+ in MeOH at 25 °C			
	Experimental	Predicted	Residual	Relative error (%)
21	3.50[a]	3.45	0.05	−1.40
22	3.38[a]	3.57	−0.19	5.59
23	3.45[a]	3.54	−0.09	2.49
24	3.62[a]	3.73	−0.11	2.98
25	2.95[a]	3.18	−0.23	7.93
26	2.70[b]	2.68	0.02	−0.74
27	3.36[a]	3.48	−0.12	3.69
28	3.72[a]	3.62	0.10	−2.77
29	3.89[b]	3.90	−0.01	0.36
30	3.42[a]	3.38	0.04	−1.26
31	3.68[a]	3.64	0.04	−1.03
32	4.00[a]	3.88	0.12	−2.98
33	3.14[a]	3.15	−0.01	0.32
34	3.25[a]	3.13	0.12	−3.72
35	2.98[a]	2.90	0.08	−2.82
36	2.94[b]	2.81	0.13	−4.35
37	3.15[a]	3.14	0.01	−0.41
38	3.47[b]	3.34	0.13	−3.63
39	3.79[a]	3.63	0.16	−4.20
40	3.29[a]	3.22	0.07	−2.04
41	3.52[a]	3.53	−0.01	0.37
42	3.68[a]	3.57	0.11	−3.10
43	3.49[b]	3.27	0.22	−6.22
44	3.35[a]	3.32	0.03	−0.84
45	3.29[b]	3.14	0.15	−4.56
46	3.72[a]	3.63	0.09	−2.50
47	2.98[a]	3.10	−0.12	3.93
48	3.07[a]	3.24	−0.17	5.64
49	3.25[b]	3.33	−0.08	2.43
50	3.56[a]	3.63	−0.07	1.91
51	2.61[a]	2.55	0.06	−2.30
52	3.17[a]	3.13	0.04	−1.39
53	3.79[b]	3.87	−0.08	1.98
54	2.38[a]	2.50	−0.12	5.21
55	3.23[a]	3.18	0.05	−1.52
56	3.46[a]	3.51	−0.05	1.36
57	3.61[a]	3.61	0.00	0.03
58	3.07[a]	2.92	0.15	−4.95

[a] Training

[b] Testing

find the best correlation models from selected non-collinear descriptors. The BMLR selects the best two-parameter regression equations, the best three-parameter regression equations, etc., on the basis of the highest R^2 value in the stepwise regression procedure. During the BMLR procedure, the descriptor scales are normalized and centered automatically, and the final result is given in natural scales. The result obtained by BMLR is the best representation of the property in the given descriptors' pool.

Consequently, the model corresponding to the break point shows the optimum number of descriptors to be used in modeling that property (log K).

The BMLR technique was performed on the 15C5 of the training set shown in Table 4. After regression analysis, a few suitable models were obtained among which the best model was selected and presented in Eq. 2. A small number of molecular descriptors (five descriptors) proposed were used to establish a QSPR model. BMLR analysis provided a useful equation that can be used to predict the log K of 15C5 complexes based upon these parameters. The best equation obtained for the stability constants of the 15C5 complexes is:

$$\log K = -7.1531 - 1.6602d_1 - 0.5744d_2 + 0.3352d_3 \\ + 0.0166d_4 + 0.0101d_5$$

$$(2)$$

Positive values in the regression coefficients indicate that the indicated descriptor contributes positively to the value of log K, whereas negative values indicate that the greater the value of the descriptor the lower the value of log K. In other words, increasing the d_1 and d_2 will decrease log K and increasing the d_3, d_4 and d_5 increases extent of log K of the 15C5 complexes.

Model validation techniques are needed in order to distinguish between true and random correlations and to estimate the predictive power of the model. The real predictive ability of any QSPR model cannot be judged solely by using internal validation, it has to be validated on the basis of predictions for log K of complexes not included in the training set. For evaluation of the predictive power of the generated BMLR, the optimized model was applied for prediction of log K values of 12 complexes in the prediction set, which were not used in the optimization procedure. The predicted values of log K, residuals and the percent relative errors (RE%) of prediction obtained by the BMLR method are presented in Table 4. The plots of predicted log K versus experimental log K and the residuals (experimental log K − predicted log K) versus experimental log K value, obtained by the BMLR modeling, and the random distribution of residuals about zero mean are shown in Figs. 2 and 3, respectively. The stability and validity of model was tested by prediction of the response values for the prediction set. The average relative errors ($\overline{RE}$%) of prediction and square of correlation coefficient (R^2) are −2.08% and 0.9235 for BMLR model, respectively.

Interpretation of the selected descriptors

It is well-known that the process of cation–macrocycle association depends on several factors including the relative cation and ligand cavity sizes, the number and arrangement of the ligand binding sites, the substitution on
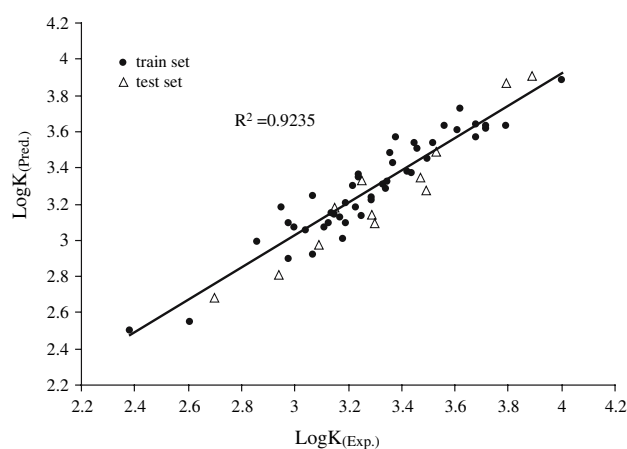

Fig. 2 Scatter plot of the predicted versus experimental log K values for train and test sets of 15-crown-5 ethers complexes
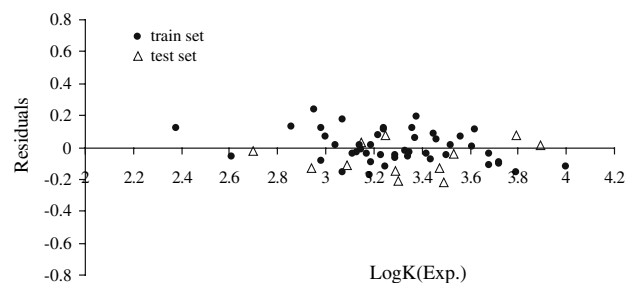
Fig. 3 Scatter plot of the experimental log K values versus residuals

the macrocyclic ring and, especially, the nature of the solvent and, etc. Here to eliminate the effect of solvent and temperature on the stability constants, we used only data at 25 °C and in methanol solution.

The ring sizes of crown ethers are not known precisely, but on the basis of molecular models, the diameter of ring sizes of 15C5 ethers are 1.7–2.2 Å. The ionic diameter of the potassium cation is 2.66 Å [54]. In these cases, the complexation takes place outside the circular bidimensional cavity and the cation completes its coordination sphere with a second ligand, leading to a sandwich complex. The stoichiometry of the binding of K^+(M) with 15C5 ethers (L) studied in the present work is 1:2 (ML_2). It is noted that the size of the crown cavity, charge density of ions (i.e., Coulombic interactions) as well as hydrophobic interaction plays a major role in governing the occurrence and stability of the complexed species.

The descriptors involved in the QSPR model are: (i) maximum net atomic charge for H atoms, d_1; (ii) average information content (order 2), d_2; (iii) maximum nuclear–nuclear repulsion for C–H bonds, d_3; (iv) maximum nuclear–nuclear repulsion for C–C bonds, d_4; (v) surface weighted charged partial negative charged surface area (WNSA1), d_5.

In this QSPR modeling according to selected descriptors, complexation phenomenon is mainly related to: (i) cation/ligand electrostatic interactions (d_1 and d_5); (ii) steric deformation of the ligand and conformational changes of the ligand accompanying the complexation (d_2); and (iii) repulsion between neighbors C–H and C–C bonds (d_3 and d_4). As the stability constant is a mean parameter which reflect the whole properties of the ligand in complexation reaction (i.e., cavity size, lariat effect, geometry of donating centers, solvation of the ligand and metal ion, etc.) to have a better correlation between stability constants and structural descriptors it is better to enter some extra theoretical or experimental descriptor to include the effect of solvation of the ligand and other important effect conformational variation due complexation and also lariat effect.

The first selected significant descriptor (d_1) involved in the Eq. 2 is maximum net atomic charge for H atoms [55], defined as in Eq. 3

$$Q_{max} = \max(Q_H^+),\tag{3}$$

where Q^+ is positive atomic partial charges. The higher values of the most positive net atomic charges on hydrogen atoms (Q_H^+) of 15C5 could lead to the increase repulsion with alkali metal cation (K^+) and the decrease of log K values. Therefore, with increasing the maximum (most positive) atomic partial charges on hydrogen atoms, decreases the binding strength of K^+ to the crown ether and complex is less stable.

The second significant descriptor (d_2) involved in the Eq. 2 is average information content (order 2) [56], defined as in Eq. 4

$$^k IC = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n},\tag{4}$$

where n_i is the number of atoms in the ith class, n is the total number of atoms in the molecule, and k denotes the number of atomic layers in the coordination sphere around a given atom. The diffusion of molecules in and between different media is another essential factor that probably influences the property of compounds. In general, it depends both on the shape and on the symmetry of the crown ether molecules. The topological descriptor, d_2, describes the connectivity and branching in a molecule and relates to molecular shape and symmetry. In addition, d_2 larger values lead to decrease stability constants (log K) of crown ether complexes. The effects of d_1 and d_2 lead to less stable complexes and decrease of stability constants. It is well-known that the topological features of the macrocyclic crown ethers involving the cavity size, the nature and number of donor atoms in the ring, and the type of substituents on the ligand play important roles in both the

selectivity and stability of the resulting macrocycle–metal ion complexes.

The third and fourth descriptor in the QSPR model is maximum nuclear–nuclear repulsion for C–H bonds (d_3) and maximum nuclear–nuclear repulsion for C–C bonds (d_4), respectively, [57].

Maximum nuclear repulsion energy between given two atomic species in the molecule, calculated as follows:

$$E_{nn}(C-H) = \frac{Z_C Z_H}{R_{CH}}\tag{5}$$

and

$$E_{nn}(C-C) = \frac{Z_C Z_C}{R_{CC}},\tag{6}$$

where Z_H and Z_C are the nuclear (core) charges of atoms H and C, respectively, and R is the distance between them. This energy describes the nuclear repulsion driven processes in the molecule and may be related to the conformational (rotational, inversional) changes or atomic reactivity in the molecule. The descriptors of d_3 and d_4 have similar effects on the QSPR model.

The maximum nuclear–nuclear repulsion between the neighboring bonds of C–H and C–C lead to steric repulsion and conformational changes of the ligand as well as the size of the rings formed upon complexation may be changed. These descriptors reflect characteristics of the charge distribution of the molecule. With increasing d_3 and d_4 leads to wide cavity of crowns, lower the steric effect with cation, strong complexation and the increase the stability constants (log K).

The final descriptor is WNSA1 [58, 59], d_5, defined as in Eq. 7

$$WNSA1 = \frac{PNSA1 \times TMSA}{1000},\tag{7}$$

where WNSA1 is surface weighted charged partial negative charged surface area, PNSA1 is partial negatively charged molecular surface area, and TMSA is total molecular surface area.

The most important property of crown ethers is their selective complexing ability. They bound the metal cations in their cavity via ion–dipole interactions between the metal cations and negatively charged donor atoms of the polyether ring. Crown ethers are characterized by a negatively charged internal cavity able to host metal cations with suitable size. WNSA1 can be related to the concerning charge distribution, which induces electron redistribution in metal cations and thereby established ion–dipole attractions. WNSA1 weighted PNSA1 describes the electrostatic component of the host–guest interaction energy and account for formation of complexes. In principle, stronger electrostatic attraction between the donor atoms

(partially negative charge) and positive cation (K^+) should increase the highly stable complexes.

Metal ions form complexes with crown ethers through ion–dipole electrostatic interactions, the oxygen lone pairs being attracted to the cation positive charge. The ability of crown ethers to complex with cations is well-known. Changes in the ring size and the polarizability of the ring oxygen's can be correlated with the complexing power of crown ether ring molecules. The metal–ligand binding energy also depends on the number and partially negative charge of oxygen heteroatoms present in the macrocyclic structure. This factor determines not only the size of the cavity, but also the bond energies with the cation. Therefore, with increasing WNSA1, stability constants of complexes increased.

Conclusions

A successful application of the CODESSA software system was presented in this study for the prediction of the stability constants of 58 complexes of 15C5 using theoretical descriptors derived from the molecular structure. We have shown that BMLR can predict the stability constants of 15C5 complexes based on chemical structures encoded in their chemical names. The BMLR computing thus provides a model of host–guest interaction, which can be used for predictive analysis. This study has shown that the stability constants of 15C5 complexes (log K) can be modeled in terms of structure-based descriptors solely calculated from the structure of crown ethers. The validation and cross-validation of the QSPR model suggest that the model can be used to make predictions for compounds not in the original dataset. The structural information encoded in the descriptors in the discussed model indicated the significant and specific structure information that may be useful for new compound design. The present work clearly demonstrates that analogous QSPR equation can be developed for the interpretation of complexation processes. It is useful to emphasis that to have a comprehensive QSPR model one should consider conformational changes upon metal binding, solvation of the coordinated ligand molecule and side chain, or lariat effect. The relatively high value of residuals or relative errors may due to ignoring the effect like these.

References

1. Pedersen, C.J.: Cyclic polyethers and their complexes with metal salts. J. Am. Chem. Soc. **89**(10), 2495–2496 (1967)
2. Pedersen, C.J.: Cyclic polyethers and their complexes with metal salts. J. Am. Chem. Soc. **89**(26), 7017–7036 (1967)
3. Lamb, J.D., Izatt, R.M., Christensen, J.J., Eatough, D.J.: In: Melson, G.A. (ed.) Coordination Chemistry of Macrocyclic Compounds, pp. 145–214. Plenum, New York (1979)
4. Popov, A.I., Lehn, J.M.: In: Melson, G.A. (ed.) Coordination Chemistry of Macrocyclic Compounds, pp. 537–602. Plenum, New York (1979)
5. Lifson, S., Felder, C.E., Shanzer, A., Libman, J.: In: Izatt, R.M., Christensen, J.J. (eds.) Synthesis of Macrocycles. The Design of Selective Complexing Agents, pp. 241–307. Wiley-Interscience, New York (1987)
6. Dietrich, B., Viout, P., Lehn, J.M.: Aspects de la Chimie des Composés Macrocycliques, Inter Editions, Paris (1991)
7. Potvin, P.G., Lehn, J.M.: In: Izatt, R.M., Christensen, J.J. (eds.) Synthesis of Macrocycles. The Design of Selective Complexing Agents, pp. 167–239. Wiley-Interscience, New York (1987)
8. Lindoy, L.F.: The Chemistry of Macrocyclic Ligand Complexes. Cambridge University Press, Cambridge (1989)
9. Gokel, G.: Crown Ethers & Cryptands. Royal Society of Chemistry, Cambridge (1994)
10. Liotta, C.L., Burgess, E.M., Ray, C.C., Black, E.D., Fair, B.E.: Omega-phase catalysis: a nonclassical phase transfer system. Preprints Am. Chem. Soc. Div. Petro. Chem. **30**, 367–373 (1985)
11. Chandler, K., Culp, C.W., Lamb, D.R., Liotta, C.L., Eckert, C.A.: Phase-transfer catalysis in supercritical carbon dioxide: kinetic and mechanistic investigations of cyanide displacement on benzyl chloride. Ind. Eng. Chem. Res. **37**, 3252–3259 (1998)
12. Osaki, S., Kawai, Y., Yajima, S., Kimura, K.: Neutral-carrier-type potassium ion selective electrodes based on polymer-supported liquid-crystal membranes for practical use. Anal. Sci. **20**, 1165–1169 (2004)
13. Richens, D.A., Simpson, D., Peterson, S., McGinn, A., Lamb, J.D.: Use of mobile phase 18-crown-6 to improve peak resolution between mono- and divalent metal and amine cations in ion chromatography. J. Chromatogr. A **1016**, 155–164 (2003)
14. Kudo, Y., Usami, J., Katsuta, S., Takeda, Y.: Solvent extraction of silver picrate by 3m-crown-m ethers (m = 5, 6) and its mono-benzo-derivative from water into benzene or chloroform: elucidation of an extraction equilibrium using component equilibrium constants. Talanta **62**, 701–706 (2004)
15. Takayanage, T.: Analysis of ion-association reaction in aqueous solution and its utilization by capillary zone electrophoresis. Anal. Sci. **20**, 255–265 (2004)
16. Saari-Nordhaus, R., Anderson, J.A., James, M.: Alternative approach to enhancing cation selectivity in ion chromatography. J. Chromatogr. A **1039**, 123–127 (2004)
17. Pavithran, R., Reddy, M.L.P.: Crown ethers as synergists in the extraction of trivalent lanthanoids with 3-phenyl-4-(4-fluorobenzoyl)-5-isoxazolone. Radiochim. Acta **92**, 31–38 (2004)
18. Dalley, N.K.: In: Izatt, R.M., Christensen, J.J. (eds.), Synthetic Multidentate Macrocyclic Compounds. Academic Press, New York (1978)
19. Izatt, R.M., Bradshaw, J.S., Nielsen, S.A., Lamb, J.D., Christensen, J.J.: Thermodynamic and kinetic data for cation–macrocycle interaction. Chem. Rev. **85**, 271–339 (1985)
20. An, H., Bradshaw, J.S., Izatt, R.M., Yan, Z.: Bis- and oligo(benzocrown ether)s. Chem. Rev. **94**, 939–991 (1994)
21. Dang, L.X., Kollman, P.A.: Free energy of association of the 18-crown-6: K+ complex in water: a molecular dynamics simulation. J. Am. Chem. Soc. **112**, 5716–5720 (1990)
22. Dang, L.X., Kollman, P.A.: Free energy of association of the K+:18-crown-6 complex in water: a new molecular dynamics study. J. Phys. Chem. **99**, 55–58 (1995)
23. Dang, L.X.: Free energies for association of Cs+ to 18-crown-6 in water. A molecular dynamics study including counter ions. Chem. Phys. Lett. **227**, 211–214 (1994)
24. Vayssiere, P., Chaumont, A., Wipff, G.: Cation extraction by 18-crown-6 to a room-temperature ionic liquid: the effect of solvent humidity investigated by molecular dynamics simulations. Phys. Chem. Chem. Phys. **7**, 124–135 (2005)

25. Cox, B.G., Schneider, H.: Coordination and Transport Properties of Macrocyclic Compounds in Solution. Elsevier, Amsterdam (1992), Chap. 2.3

26. Christensen, J.J., Eatough, D.J.E., Izatt, R.M.: The synthesis and ion bindings of synthetic multidentate macrocyclic compounds. Chem. Rev. **74**(3), 351–384 (1974)

27. Izatt, R.M., Bradshaw, J.S., Nielsen, S.A., Lamb, J.D., Christensen, J.J.: Thermodynamic and kinetic data for cation–macrocycle interaction. Chem. Rev. **85**(4), 271–339 (1985)

28. Izatt, R.M., Pawlak, K., Bradshaw, J.S.: Thermodynamic and kinetic data for macrocycle interaction with cations, anions, and neutral Molecules. Chem. Rev. **95**(7), 2529–2586 (1995)

29. Wagner-Czauderna, E., Koczorowska, A., Kalinowski, M.K.: Stability of 15-crown-5 and benzo-15-crown-5 complexes with alkali metal cations in nonaqueous media. J. Coord. Chem. **46**, 265–276 (1999)

30. Buschmann, H.J., Schollmeyer, E.: The complexation reaction of 18-crown-6 with Ag+ in different solvents studied by potentiometric and calorimetric methods. Inorganica Chim. Acta **298**, 120–122 (2000)

31. Buschmann, H.J., Schollmeyer, E.: The interpretation of thermodynamic data for the complex formation of cations with crown ethers and cryptands. Part I: the reaction entropy. J. Incl. Phenom. Macrocycl. Chem. **38**, 85–97 (2000)

32. Filipek, S., Rudzinski, P., Kalinowski, M.K.: The influence of structural effects on the complexing ability of crown ethers. J. Coord. Chem. **50**, 131–140 (2000)

33. Mandl, C.P., Konig, B.: Luminescent crown ether amino acids: selective binding to N-terminal lysine in peptides. J. Org. Chem. **70**(2), 670–674 (2005)

34. Loeb, S.J., Tiburcio, J., Vella, S.J.: [2] Pseudorotaxane formation with N-benzylanilinium axles and 24-crown-8 ether wheels. Org. Lett. **7**, 4923–4926 (2005)

35. Shinkai, S., Nakaji, T., Ogawa, T., Shigematsu, K., Manabe, O.: Photoresponsive crown ethers. 2. Photocontrol of ion extraction and ion transport by a bis(crown ether) with a butterfly-like motion. J. Am. Chem. Soc. **103**(1), 111–115 (1981)

36. Shinkai, S., Minami, T., Kusano, Y., Manabe, O.: Photoresponsive crown ethers. 4. Influence of alkali metal cations on photoisomerization and thermal isomerization of azobis(benzocrown ethers). J. Am. Chem. Soc. **104**(7), 1960–1967 (1982)

37. Pedersen, C.J., Frensdorff, H.K.: Macrocyclic polyethers and their complexes. Angew. Chem. Int. Ed. Engl. **11**, 16–25 (1972)

38. Dietrich, B.: Coordination chemistry of alkali and alkaline-earth cations with macrocyclic ligands. J. Chem. Educ. **62**, 954–964 (1985)

39. An, H., Bradshaw, J.S., Izatt, R.M., Yan, Z.: Bis- and Oligo(benzocrown ether)s. Chem. Rev. **94**(4), 939–991 (1994)

40. Christensen, J.J., Hill, J.O., Izatt, R.M.: Ion binding by synthetic macrocyclic compounds. Science **174**, 459–467 (1971)

41. Balzani, V., Credi, A., Venturi, M.: Photochemistry and photophysics of coordination compounds: an extended view. Coord. Chem. Rev. **171**, 3–16 (1998)

42. Ghasemi, J., Saaidpour, S., Brown, S.D.: QSPR study for estimation of acidity constants of some aromatic acids derivatives using multiple linear regression (MLR) analysis. J. Mol. Struct. (Theochem) **805**, 27–32 (2007)

43. Ghasemi, J., Saaidpour, S.: QSPR prediction of aqueous solubility of drug-like organic compounds. Chem. Pharm. Bull. **55**, 669–674 (2007)

44. Ghasemi, J., Shahmirani, S., Farahani, E.V.: Development of a model to predict partition coefficient of organic pollutants in cloud point extraction process. Ann. di Chim. **96**, 327–337 (2006)

45. Ghasemi, J., Ahmadi, Sh.: Combination of genetic algorithm and partial least squares for cloud point prediction of nonionic surfactants from molecular structures. Ann. di Chim. **97**, 69–83 (2007)

46. Ghasemi, J., Asadpour, S., Abdolmaleki, A.: Prediction of gas chromatography/electron capture detector retention times of chlorinated pesticides, herbicides, and organohalides by multivariate chemometrics methods. Anal. Chim. Acta **588**, 200–206 (2007)

47. Izatt, R.M., Pawlak, K., Bradshaw, J.S.: Thermodynamic and kinetic data for macrocycle interaction with cations and anions. Chem. Rev. **91**, 1721–2085 (1991)

48. Hyperchem, v.7.5: Hypercube Inc. http://www.hyper.com (2002)

49. AMPAC 8.16: Semichem, Inc. http://www.semichem.com (2004)

50. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., Stewart, J.J.P.: AM1: a new general purpose quantum mechanical molecular model. J. Am. Chem. Soc. **107**, 3902–3909 (1985)

51. CODESSA ver. 2.7.5: Semichem Inc. http://www.semichem.com (2004)

52. Katritzky, A.R., Mu, L., Lobanov, V.S., Karelson, M.: Correlation of boiling points with molecular structure. 1. A training set of 298 diverse organics and a test set of 9 simple inorganics. J. Phys. Chem. **100**(24), 10400–10407 (1996)

53. Karelson, M.: Molecular Descriptors in QSAR/QSPR. Wiley-Interscience, New York (2000)

54. Lamb, J.D., Izatt, R.M., Swain, C.S., Christensen, J.J.: A systematic study of the effect of macrocycle ring size and donor atom type on the log K, .DELTA.H, and T.DELTA.S of reactions at 25 °C in methanol of mono- and divalent cations with crown ethers. J. Am. Chem. Soc. **102**(2), 475–479 (1980)

55. Csizmadia, I.G.: Theory and Practice of MO Calculations on Organic Molecules. Elsevier, Amsterdam (1976)

56. Kier, L.B.: Use of molecular negentropy to encode structure governing biological activity. J. Pharm. Sci. **69**, 807–810 (1980)

57. Clementi, E.: Computational Aspects of Large Chemical Systems. Springer Verlag, New York (1980)

58. Stanton, D.T., Jurs, P.C.: Development and use of charged partial surface area structural descriptors in computer-assisted quantitative structure–property relationship studies. Anal. Chem. **62**, 2323–2329 (1990)

59. Stanton, D.T., Egolf, L.M., Jurs, P.C., Hicks, M.G.: Computer-assisted prediction of normal boiling points of pyrans and pyrroles. J. Chem. Inf. Comput. Sci. **32**(4), 306–316 (1992)