

Shiling Yuan  
Zhengting Cai  
Guiying Xu  
Yuansheng Jiang

## Quantitative structure–property relationships of surfactants: prediction of the critical micelle concentration of nonionic surfactants

Received: 11 September 2001  
Accepted: 8 January 2002  
Published online: 24 April 2002  
© Springer-Verlag 2002

S. Yuan · G. Xu  
Key Laboratory of Colloid & Interface  
Chemistry of the Education Ministry,  
Shandong University, Jinan 250100,  
China

S. Yuan · Z. Cai (✉) · Y. Jiang  
Institute of Theoretical Chemistry,  
Shandong University, Jinan 250100,  
China  
E-mail: shilingyuan@sdu.edu.cn  
Tel.: +86-0531-8567117  
Fax: +86-0531-8564464

Y. Jiang  
College of Chemistry and Chemical  
Engineering, Nanjing University,  
Nanjing 210093, China

**Abstract** The quantitative structure–property relationship was used to predict the critical micelle concentration (cmc) of a nonionic surfactant in aqueous solution. Several structural, electronic, spatial and thermodynamic properties were selected as descriptors to build the relationship between the macroscopic properties (such as the cmc) and microscopic structures. These descriptors include the octanol/water partition coefficient, the heat of formation, the volume of a molecule and the energy of the lowest unoccupied molecular orbital. A general four-parameter structure–property relationship between the logarithm of the cmc and the descriptors,

which was developed for a set of 37 nonionic surfactants ( $r^2 = 0.990$ ,  $F$ -test = 802.970), can be used to predict the cmc of nonionic surfactants without the experimental data. The success of this approach indicates that the prediction of the cmc for other surfactants, such as anionic, cationic and zwitterionic surfactants, as well as the prediction of other surfactant properties, can be obtained.

**Keywords** Quantitative structure–property relationship · Nonionic surfactant · Critical micelle concentration

### Introduction

The critical micelle concentration (cmc) is always considered as an important parameter of the surface activity of a surfactant because a lower cmc indicates that the concentration at which micelles form in aqueous solution is lower. This ability to adsorb at an interface and to reduce interfacial tension at a lower cmc is of great importance to many processes of technological interest, such as emulsification, foaming, solubilization, detergency and so on. Additionally, the cmc is also considered as a watershed of distinct change of surfactant properties. So, many researchers concentrated on the cmc of different systems, especially the effect of additives on the cmc of surfactant in solution [1, 2].

On the basis of a vast amount of experimental data concerning the cmc of surfactants, several simple empirical relationships between the cmc and the chemical structure of surfactants were found. In the 1960s, Becher [3] obtained the relationship between the cmc of ionic surfactants and the number of C atoms in the hydrocarbon chain for a surfactant with a single terminal headgroup. He found that the cmc is usually reduced to approximately half its previous value with the addition of each methylene group in the alkyl chain for the number of C atoms from 8 to 16, and for the nonionic surfactant, the cmc decreases to one-tenth of its previous concentration for every increase of two C atoms. He got the following equation:

$$\log(\text{cmc}) = A - Bn \quad (1)$$

where  $A$  and  $B$  are constants specific to the homologous series and  $n$  is the number of C atoms in the alkyl chain. A number of equations for surfactants have been listed in the literature [4, 5].

In the same way, the effect of the hydrophilic group on the cmc in the nonionic surfactant was discussed. The relationship between the cmc and the number of ethylene oxide (EO) groups has the same expression as Eq. (1):

$$\log(\text{cmc}) = A' + B'm \quad (2)$$

where  $A'$  and  $B'$  are constants which depend on the temperature and the type of hydrophobic group and  $m$  is the number of EO groups. Some types of nonionic surfactant have explicit equations such as  $\text{C}_{12}\text{H}_{25}(\text{EO})_n\text{-OH}$  and  $\text{C}_9\text{H}_{19}\text{-C}_6\text{H}_4(\text{EO})_n\text{-OH}$  [6, 7].

A nonlinear term in the form of the product of the alkane C number and the EO number,  $nm$ , was used to improve the correlation equation [8]. This equation is

$$\log(\text{cmc}) = A + Bn + Cm + Dnm \quad (3)$$

Many investigators have developed empirical equations relating the cmc to the various structural units in different series of surfactants; however, these equations, like Eqs. (1), (2) and (3), are pure empirical relationships between the cmc and the number of C atoms or EO groups in the hydrocarbon chain. Strictly speaking, these equations cannot be considered as the relationship between the cmc and molecular chemical structure because they cannot take into account the properties of the chemical structure, such as the heat of formation, the spatial structure of molecule and the molecular orbitals. With the development of computer techniques, it is possible to research the relationship between the macroscopic property (the cmc) and the microscopic structures (molecular structure) from structural properties of molecules using quantum or molecular mechanics; however, none of these properties were considered in Eqs. (1), (2) and (3).

Although a large amount of knowledge concerning the prediction of physical properties for organic compounds has been accumulated over the last 2 decades with the help of quantitative structure–property relationships (QSPR) [9, 10], there is little information available for the prediction of solution properties of surface-active compounds [11]. Here, we discuss the relationship between the cmc and the molecular structure properties using the heat of formation,  $H_f$ , the molecular volume,  $V_m$ , the lowest unoccupied molecular orbital (LUMO) and the distributed coefficient between octanol and water ( $\log P$ ) as descriptors. From the equation for the macroscopic property and the microscopic structures, the cmc of every nonionic surfactant can be calculated. On the other hand, given only the molecular structure of the surfactant, the cmc of a nonionic surfactant can be predicted.

## Theory: the mathematical foundation of QSPR

A QSPR is a multivariate statistical correlation between a property of the material and its key geometric or chemical characteristics of a molecular system. By computing and analyzing the QSPR, the critical factors for the property can be chosen to assist the design and control of the properties of materials and chemicals. The mathematical foundation of QSPR is based on the principle of multilinearity. According to the principle, the multivariate mathematical relationship between a macroscopic property of the system being studied,  $P$ , such as the cmc, solubility, which is experimental data, and a set of physicochemical microscopic structure properties,  $X_i$ , which are the structural factors of the molecule, is assumed to be linear. The experimental property  $P$  depends mainly on the microscopic structure properties  $X_i$ , and the relationship between  $P$  and  $X_i$  can be found using the multilinear least-squares method:

$$P = P_0 + \sum_{i=1}^n a_i X_i \quad (4)$$

This indicates that the property may depend on more than one microscopic structure property if  $n$  is not 1. In addition, it is known that nonlinear functions should also be used to describe the macroscopic property. Thus, some descriptors, such as the exponential function,  $X_i = \exp(x_i)$ , the logarithmic function,  $X_i = \log x_i$ , and so on, can be used to depict the macroscopic property. When the front expressions are introduced into Eq. (4), the final relationship between  $P$  and  $X_i$  is still linear.

## Methods

All the computational work was performed using a Silicon Graphics Indigo workstation, at Shandong University, and the software programs Cerius2 and QSAR (quantitative structure–activity relationship) from MSI, San Diego. In the program package, the microscopic structure properties, like  $H_f$ ,  $V_m$  and LUMO, were calculated using the MOPAC 6.0 program. We used the visualizer model to build the molecular structure of the surfactants, then selected a force field to optimize the geometry using a molecular dynamics method in the open force field model, and finally calculated the relationship between the macroscopic property and the microscopic structure properties using the QSAR program. QSAR is a chemical structure–activity and structure–property statistical analysis. The program can generate a large number (about 60) of molecular descriptors on the basis of the thermodynamic, geometrical and electronic structure of a molecule. The statistical analysis technique includes stepwise multiple linear regression analysis, multiple simple linear analysis, multiple linear regression, and genetic function approximation analysis [12]. These analysis methods can be used to get a better relationship between the cmc and the microscopic structures.

## Results and discussion

### Selection of descriptors

In the QSAR program, there are ten descriptors including several properties, such as conformational, electronic, spatial, structural and thermodynamic properties. The selected descriptors should capture necessary information on factors which influence the phenomena controlling the physical parameters that we wish to predict. We selected about ten descriptors in order to depict perfectly the relationship between the molecular structure and the macroscopic property. They are  $\log P$ , the octanol/water partition coefficient, which is related to a molecule tending to partition into bilayers and is like the hydrophile-lipophile balance (HLB) value in colloid chemistry,  $H_f$ , the heat of formation, which is a thermodynamic property,  $V_m$ , the volume of a molecule,

and the density of a molecule, which are spatial properties, the highest occupied molecular orbital (HOMO) and the LUMO, and the dipole moment, which is an electronic property. In addition, some other parameters were selected as descriptors in order to obtain a good expression, such as hydrogen-bond acceptors. The selected descriptors and the values of the microscopic structure properties of the surfactant molecules are shown in Table 1. The values were taken from the work of Huibers and Lobanov [11] and Rosen [13]. All the values used were those measured in purified water, at 25 °C. From Table 1, the last equation, which calculates the logarithm of the cmc using the molecular structure properties, is

$$\log(\text{cmc}) = 2.91598 - 0.696867 \times \log P + 3.4772 \times 10^{-7} \times H_f - 0.0073381 \times V_m - 0.28405 \times \text{LUMO} \quad (5)$$

**Table 1.** The values of the descriptors and the experimental and the calculated logarithm of the critical micelle concentration (cmc) for the different nonionic surfactants

Structure	$\log P$	$H_f$ (kcalmol <sup>-1</sup> )	$V_m$ (Å <sup>3</sup> )	Density	HOMO (eV)	LUMO (eV)	Dipole	$\log(\text{cmc})$ (obs.)	$\log(\text{cmc})$ (cal.)
C <sub>8</sub> Ph(EO) <sub>1</sub>	5.580	-96.539	269.351	0.93	-11.6276	3.9354	0.851	-4.305	-4.067
C <sub>8</sub> Ph(EO) <sub>2</sub>	5.060	-131.103	312.148	0.94	-11.5503	3.9401	0.808	-4.116	-4.020
C <sub>8</sub> Ph(EO) <sub>3</sub>	4.540	-174.875	355.281	0.95	-11.6173	3.9418	1.489	-4.013	-3.975
C <sub>8</sub> Ph(EO) <sub>4</sub>	4.020	-209.585	398.666	0.96	-11.6534	3.8485	1.010	-3.886	-3.904
C <sub>8</sub> Ph(EO) <sub>5</sub>	3.500	-240.146	441.163	0.97	-11.6347	3.8594	1.310	-3.824	-3.857
C <sub>8</sub> Ph(EO) <sub>6</sub>	2.980	-280.732	483.690	0.97	-11.6415	3.8629	1.192	-3.678	-3.807
C <sub>8</sub> Ph(EO) <sub>7</sub>	2.460	-312.584	527.061	0.98	-11.7478	3.7389	1.127	-3.602	-3.728
C <sub>8</sub> Ph(EO) <sub>8</sub>	1.940	-361.907	569.310	0.98	-11.9742	3.5181	1.080	-3.553	-3.613
C <sub>8</sub> Ph(EO) <sub>9</sub>	1.420	-399.631	612.770	0.98	-12.0868	3.3646	1.006	-3.523	-3.526
C <sub>8</sub> Ph(EO) <sub>10</sub>	0.900	-410.660	653.937	0.99	-11.9472	3.5008	1.154	-3.481	-3.504
C <sub>6</sub> (EO) <sub>3</sub>	0.640	-192.382	250.054	0.94	-13.6132	6.7091	1.541	-1.000	-1.271
C <sub>6</sub> (EO) <sub>6</sub>	-0.920	-300.115	379.072	0.97	-13.2593	6.4800	1.116	-1.164	-1.065
C <sub>8</sub> (EO) <sub>1</sub>	2.740	-120.566	198.170	0.98	-13.6641	6.7060	1.408	-2.310	-2.353
C <sub>8</sub> (EO) <sub>3</sub>	1.700	-195.995	284.089	0.92	-13.6037	6.6981	2.037	-2.125	-2.256
C <sub>8</sub> (EO) <sub>6</sub>	0.140	-296.476	412.096	0.96	-13.3282	6.5479	3.433	-2.004	-2.066
C <sub>8</sub> (EO) <sub>9</sub>	-1.420	-411.708	540.375	0.97	-13.3326	6.5923	3.268	-1.886	-1.932
C <sub>10</sub> (EO) <sub>3</sub>	2.760	-205.070	317.860	0.91	-13.5000	6.6484	1.494	-3.222	3.243
C <sub>10</sub> (EO) <sub>4</sub>	2.240	-237.988	360.064	0.93	-13.4671	6.7157	1.778	-3.167	-3.195
C <sub>10</sub> (EO) <sub>6</sub>	1.200	-299.373	446.827	0.95	-13.2145	6.4970	2.120	-3.046	-3.045
C <sub>10</sub> (EO) <sub>8</sub>	0.160	-379.360	531.669	0.96	-13.2275	6.5203	2.023	-3.000	-2.949
C <sub>10</sub> (EO) <sub>9</sub>	-0.360	-417.709	574.231	0.97	-13.3494	6.4091	2.550	-2.886	-2.868
C <sub>12</sub> (EO) <sub>2</sub>	4.340	-176.543	308.897	0.89	-13.5931	6.6975	1.541	-4.481	-4.278
C <sub>12</sub> (EO) <sub>3</sub>	3.820	-203.469	352.214	0.90	-13.2291	6.7143	1.627	-4.284	-4.238
C <sub>12</sub> (EO) <sub>4</sub>	3.300	-243.270	393.975	0.92	-13.5161	6.5814	2.229	-4.194	-4.144
C <sub>12</sub> (EO) <sub>5</sub>	2.780	-280.188	437.391	0.93	-13.5023	6.5650	1.853	-4.194	-4.096
C <sub>12</sub> (EO) <sub>6</sub>	2.260	-323.076	480.667	0.94	-13.1609	6.5449	2.531	-4.060	-4.045
C <sub>12</sub> (EO) <sub>7</sub>	1.740	-356.462	522.220	0.95	-13.3205	6.4872	1.585	-4.080	-3.971
C <sub>12</sub> (EO) <sub>8</sub>	1.220	-390.530	565.767	0.95	-13.3195	6.4552	2.048	-4.000	-3.920
C <sub>12</sub> (EO) <sub>9</sub>	0.700	-432.056	608.166	0.96	-13.4868	6.4421	3.046	-4.000	-3.865
C <sub>12</sub> (EO) <sub>12</sub>	-0.860	-545.792	735.744	0.97	-13.5890	6.2505	3.355	-3.854	-3.659
C <sub>12</sub> (EO) <sub>14</sub>	-0.900	-615.635	823.203	0.98	-13.3455	5.9328	4.177	-4.260	-4.260
C <sub>14</sub> (EO) <sub>6</sub>	3.320	-344.770	513.912	0.93	-13.5300	6.5744	2.189	-5.000	-5.036
C <sub>14</sub> (EO) <sub>9</sub>	2.280	-405.073	598.680	0.95	-13.3958	6.5505	1.506	-5.046	-4.927
C <sub>16</sub> (EO) <sub>6</sub>	4.380	-325.199	547.731	0.93	-13.4152	6.4583	2.115	-5.780	-5.990
C <sub>16</sub> (EO) <sub>7</sub>	3.860	-379.583	590.693	0.93	-13.3368	6.6236	1.313	-5.770	-5.990
C <sub>16</sub> (EO) <sub>9</sub>	2.820	-451.982	677.258	0.94	-13.3025	6.4264	1.479	-5.678	-5.845
C <sub>16</sub> (EO) <sub>12</sub>	1.260	-560.155	804.633	0.96	-13.4785	6.1130	3.081	-5.638	-5.603

and is obtained through the process described in the following section.

### Correlation analysis

The squared correlation coefficient (or coefficient of multiple determination),  $r^2$ , is a measure of the fit of the regression model. Correspondingly, it represents the part of the variation in the observed (experimental) data that is explained by the model. Correlation coefficient values closer to 1.0 represent the better fit of the model. The  $F$ -test reflects the ratio of the variance explained by the model and the variance owing to the error in the model (i.e. the variance not explained by the model). High values of the  $F$ -test indicate that the model is statistically significant [11]. The stepwise multiple linear regression analysis in the QSAR program is used to look for the best equation with the biggest squared correlation coefficient and  $F$ -test, which can be used to depict the relationship between the cmc and the microscopic properties. We can see from Table 2 that the  $\log P$  descriptor is always found in the relationship formula in spite of the different types of nonionic surfactant; however, for the octyl poly(EO), octylphenyl poly(EO) and hexadecyl poly(EO) surfactants, two descriptors, the density, a spatial property, and  $\log P$ , a thermodynamic property, are found together in the formula with the biggest squared correlation coefficient. It is easily understood that the density is depends on the numbers of C atoms and  $\log P$  depends on the HLB of the surfactants. So these equations are similar to the Eqs. (1) and (2). For dodecyl poly(EO) surfactants, the dipole, an electronic property, is added to the equation, but  $r^2$  is still low, and maybe some other selected descriptors should be added to the relationship. Since the  $\log P$  descriptor always exists in the equations for different surfactants, we build the relationship for nonionic surfactant beginning from the  $\log P$  descriptor. Thus, the  $\log P$  descriptor, which is considered as the main effect on the cmc, is selected first, then the bigger  $r^2$  is used to distinguish which formula fits to depict the cmc; next another descriptor is added to the formula until the biggest squared correlation coefficient is obtained. The process of selecting the biggest  $r^2$ , i.e. the process of selecting the

best relationship, is shown in Table 3 and Figs. 1, 2, 3 and 4.

So, we get the last equation, Eq. (5)

Table 3 shows which descriptors are important in the selecting the descriptors using  $r^2$  and the  $F$ -test. For the 37 nonionic surfactants, each descriptor cannot solely depict the relationship between the cmc and the descriptor. At the beginning, the equations with the  $\log P$  and dipole descriptors have the same biggest  $r^2$ , which is 0.263. For the equation using  $\log P$  as the descriptor results in a bigger  $F$ -test, and from the previous discussion,  $\log P$  is selected first. When the descriptor  $V_m$  is added to depict the relationship, a bigger  $r^2$  is obtained. After another descriptor besides  $\log P$  and  $V_m$  is selected in the equation, all the  $r^2$  for different descriptors are more than 0.900; this means that a better equation between the macroscopic property and the microscopic structure descriptors can be obtained if more than two descriptors are selected. When the third descriptor LUMO, an electronic property, is added to the equation, three-parameter structure properties including electronic, spatial and thermodynamic properties have been selected. After the fourth descriptor is added to the expression, the increase in  $r^2$  is not large. Additional descriptors cannot increase  $r^2$  after Eq. (5) is obtained. So Eq. (5) is the last one which can calculate the cmc of nonionic surfactants. Equation (5) and Eqs. (6), (7) and (8) in Table 3 indicate the process of selecting descriptors.

### Interpretation of descriptors

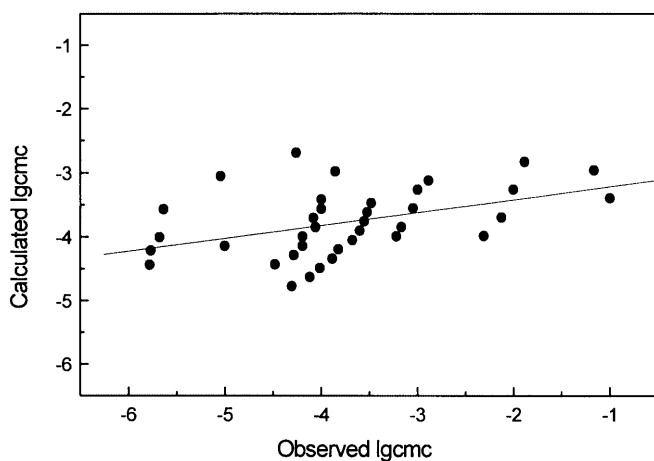
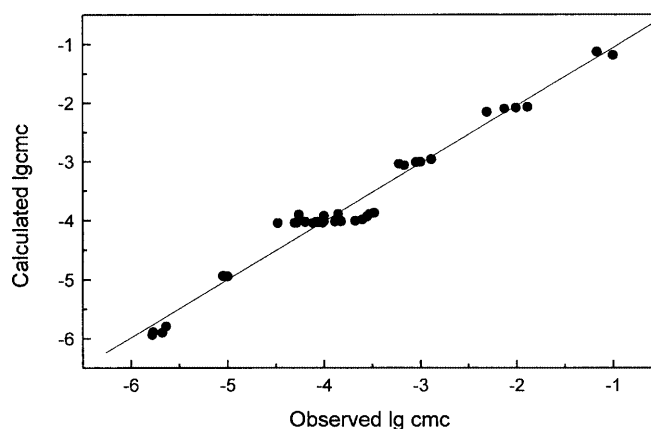
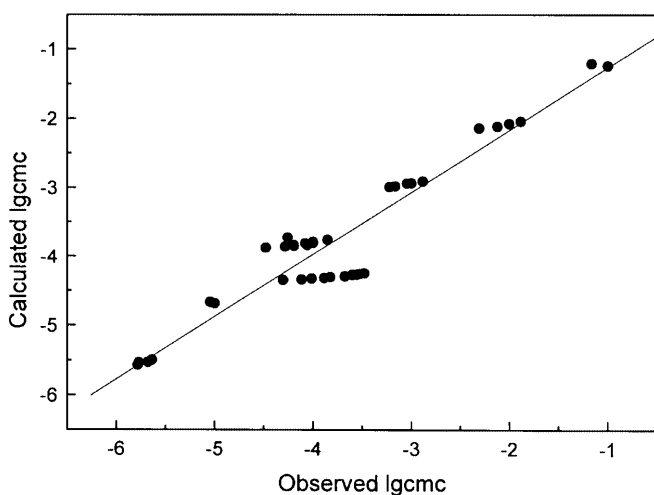
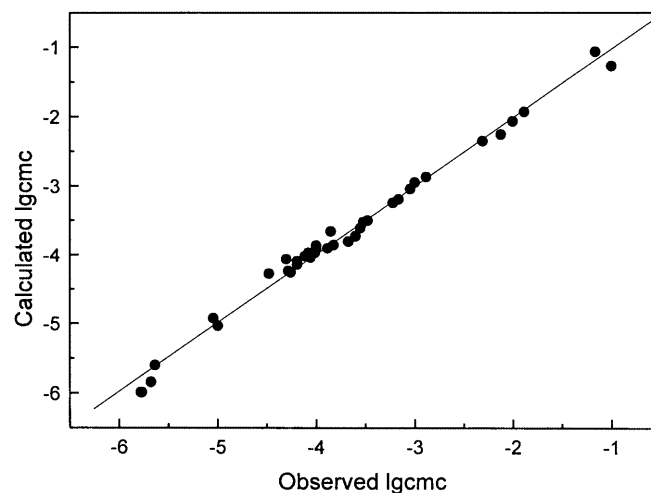
In the best parameter model (Eq. 5), the four descriptors selected capture the nature of the micellization. The  $\log P$  descriptor is widely used in extractive chemistry. It is well known that the cmc depends on the HLB value of the surfactant. For the surfactant molecule,  $\log P$  is similar to HLB in colloid chemistry. For the micellization of a surfactant in aqueous solution, the formation of a micelle leads to the creation of a hydrophobic microenvironment in the core of the micelle. The micellization potentials of amphiphiles depend essentially on the water volume displaced by the hydrophobic fragment [8]. So, the energy change is very important for the micellization. The HOMO and the LUMO are the

**Table 2.** The relationship formula between the cmc and descriptors using the stepwise method when the squared correlation coefficient,  $r^2$ , is the biggest

Structure	Number of sample	Relationships between $\log(\text{cmc})$ and descriptors	$r^2$
$C_6(\text{EO})_n$	2		
$C_8(\text{EO})_n$	4	$\lg(\text{cmc}) = -4.3091 + 2.41138 \times \text{density} - 0.040781 \times \log P$	0.991
$C_{10}(\text{EO})_n$	5	$\lg(\text{cmc}) = -6.03933 - 0.237349 \times \text{HOMO} - 0.108148 \times \log P$	0.998
$C_{12}(\text{EO})_n$	10	$\lg(\text{cmc}) = -116.815 + 0.069075 \times \text{dipole} + 0.136401 \times V_m + 22.4549 \times \text{density} + 11.5071 \times \log P$	0.946
$C_{14}(\text{EO})_n$	2		
$C_{16}(\text{EO})_n$	4	$\lg(\text{cmc}) = -12.1806 + 6.77588 \times \text{density} - 0.02281 \times \log P$	0.976
$C_8\text{Ph}(\text{EO})_n$	10	$\lg(\text{cmc}) = -14.0991 + 10.7706 \times \text{density} - 0.030538 \times \log P$	0.994

**Table 3.** The process of selecting the equation

	$\log P$	$V_m$	LUMO	$H_f$	HOMO	Density	Dipole
$r^2$	0.263	0.208	0.004	0.006	0.003	0.017	0.263
$F$ -test	12.478	9.211	0.151	0.205	0.119	0.596	0.070
$\log(\text{cmc}) = -3.06163 - 0.324149 \times \log P$ (Eq. 6)							
$r^2$	—	0.902	0.271	0.342	0.289	0.291	0.377
$F$ -test	—	156.303	6.331	8.843	6.915	6.962	10.275
$\log(\text{cmc}) = 0.882468 - 0.599766 \times \log P - 0.0069813 \times V_m$ (Eq. 7)							
$r^2$	—	—	0.980	0.909	0.975	0.976	0.935
$F$ -test	—	—	553.06	110.214	422.991	453.841	214.324
$\log(\text{cmc}) = 2.93586 - 0.679331 \times \log P - 0.0075495 \times V_m - 0.279726 \times \text{LUMO}$ (Eq. 8)							
$r^2$	—	—	—	0.990	0.981	0.984	0.981
$F$ -test	—	—	—	802.970	402.926	477.293	405.369
$\log(\text{cmc}) = 2.91598 - 0.696867 \times \log P + 3.4772 \times 10^{-7} \times H_f - 0.0073381 \times V_m - 0.28405 \times \text{LUMO}$							

**Fig. 1.** Scatter plot of the calculated  $\log(\text{cmc})$  versus the observed  $\log(\text{cmc})$  using  $\log P$  as a descriptor**Fig. 3.** Scatter plot of the calculated  $\log(\text{cmc})$  versus the observed  $\log(\text{cmc})$  using  $\log P$ ,  $V_m$  and the lowest unoccupied molecular orbital (LUMO) as descriptors**Fig. 2.** Scatter plot of the calculated  $\log(\text{cmc})$  versus the observed  $\log(\text{cmc})$  using  $\log P$  and  $V_m$  as descriptors**Fig. 4.** Scatter plot of the calculated  $\log(\text{cmc})$  versus the observed  $\log(\text{cmc})$  using  $\log P$ ,  $V_m$ , the LUMO and  $H_f$  as descriptors

**Table 4.** The calculated cmc of some nonionic surfactants compared with the observed cmc

Structure	$\log P$	$V_m$	LUMO	$H_f$	$\log(\text{cmc})$ (obs.)	$\log(\text{cmc})$ (cal.)
C <sub>8</sub> (EO) <sub>2</sub>	2.240	241.609	6.7021	-155.699	-2.237 [14]	-2.322
C <sub>9</sub> Ph(EO) <sub>10</sub>	1.430	689.655	3.6604	-414.822	-4.125 [15]	-4.181
C <sub>11</sub> (EO) <sub>8</sub>	0.690	547.829	6.5955	-382.866	-3.523	-3.485
C <sub>15</sub> (EO) <sub>8</sub>	2.810	615.623	6.5637	-397.913	-5.456	-5.424
C <sub>13</sub> (EO) <sub>8</sub>	1.750	582.336	6.5740	-393.489	-4.569	-4.444
C <sub>12</sub> (EO) <sub>3</sub> (PO) <sub>6</sub>	3.880	707.008	6.6243	-436.032	-5.849 [16]	-6.858
C <sub>12</sub> (EO) <sub>4</sub> (PO) <sub>5</sub>	3.350	690.863	6.4799	-426.472	-5.202 [16]	-6.329
C <sub>12</sub> (EO) <sub>5</sub> (PO) <sub>4</sub>	2.820	673.408	6.5426	-417.572	-4.978 [16]	-5.849

energy responses for electron transfer in a molecule; the larger the difference between the HOMO and the LUMO in a molecule, the more difficult the electron transfer from the occupied orbital to the exited unoccupied orbital. The formation of a micelle can be considered as the result of the energy change from the monomer to the aggregate of the surfactant molecule. So, it is very easy to understand that the LUMO is selected as a descriptor. Additionally, for linear molecules, the increase in the volume of the hydrophobic chain is simply related to the alkane C number. At the same time, the heat of formation is also related to the number of atoms in molecules.  $V_m$  and  $H_f$  of surfactant molecules increase with increasing number of non-hydrogen atoms in the hydrophobic fragment. The combination of these descriptors in connection with the micellization indicates the micellization depends on the spatial and energy properties of surfactant molecules, so the equation obtained by the method described earlier can depict the common macroscopic property, i.e. the cmc.

#### Estimating the equation

In order to check whether the equation can depict the cmc from the molecular structure properties or not, some nonionic surfactants were used to estimate the equation using the calculated and experimental results. The results are shown in Table 4.

In this table, one conclusion is shown that the QSPR between the logarithm of the cmc and the descriptors can be used to predict the cmc of nonionic surfactants. Since the QSAR mainly comes from nonionic surfactants with alkyl poly(EO), it may be difficult to predict the nonionic surfactants with poly(propylene oxide). However the difference between the observed and calculated logarithm of the cmc is not too big, so the equation may also be considered as

a qualitative description for the nonionic surfactants. If the cmc of these poly(propylene oxide) surfactants needs to be calculated or predicted, some known cmc of poly(propylene oxide) surfactants must be added to the process of searching for the best equation again. However, for alkyl poly(EO) surfactants, Eq. (5) can be used to calculate and predict the cmc of the nonionic surfactants.

#### Conclusions

The QSPR between the logarithm of the cmc and the descriptors, which include the octanol/water partition coefficient, the heat of formation, the volume of a molecule and the energy of the LUMO can be used to predict the cmc of nonionic surfactants. Each descriptor cannot be used singly to depict the relationship between the macroscopic property and the microscopic structure properties owing to the small correlation coefficient. If two spatial and thermodynamic properties,  $\log P$  and  $V_m$ , are selected as descriptors, the correlation coefficients are more than 0.900 when another descriptor is added to the equation. However, adding other descriptors to the last equation cannot change the correlation coefficient. This means that Eq. (5) is the best relationship between the cmc and the microscopic structure properties.

Owing to the success of this approach for predicting the cmc of nonionic surfactants using the QSAR program, we can say that the prediction of the cmc for other surfactants, such as anionic, cationic and zwitterionic surfactants, as well as the prediction of other surfactant properties, would be achieved.

**Acknowledgement** The authors acknowledge financial support from the National Natural Science Foundation of China (no. 29973023).

---

## References

1. Guerin CBE, Szbeifer I (1999) *Langmuir* 15:7901
2. Brackman JC, Engberts JBFN (1991) *Langmuir* 7:46
3. Becher P (1967) In: Schick MJ (ed) *Nonionic surfactants*. Dekker, New York, Chap 15
4. Kreshek G (1975) In: Frank F (ed) *Water, a comprehensive treatise*, vol 4. Plenum, New York, Chap 2
5. Meguro K, Takasawa Y (1981) *J Colloid Interface Sci* 83:50
6. Schick MJ (1962) *J Colloid Sci* 17:801
7. Schick MJ, Atlas SM, Eirich FR (1962) *J Phys Chem* 66:1326
8. Ravey JC, Gherbi A (1988) *Prog Colloid Polym Sci* 76:234
9. Kier LB, Hall LH (1976) *Molecular connectivity in chemistry and drug research*. Academic, New York
10. Stanton DT, Jurs PC (1990) *Anal Chem* 62:2323
11. Huibers PDT, Lobanov VS (1996) *Langmuir* 12:1451
12. Molecular Simulations Inc. (October 1998) *QSAR users guide*. Molecular Simulations, San Diego
13. Rosen MJ (1989) *Surfactant and interfacial phenomena*, 2nd edn. Wiley, New York
14. Shinoda K (1963) *Colloidal surfactants*. Academic, New York, Chap 1
15. Schick MJ, Gilbert AH (1965) *J Colloid Sci* 20:464
16. Li Y (1999) PhD thesis. Shandong University