Journal of Solution Chemistry (2019) 48:875–890
https://doi.org/10.1007/s10953-019-00887-x

On the Head Group Effect on Critical Micelle Concentration of Cationic Surfactants Using Molecular Connectivity Indices and Atomic Partial Charges

Anna Mozrzymas[1]

Received: 13 October 2018 / Accepted: 9 April 2019 / Published online: 18 June 2019

Abstract
The influence of the head group structure on the critical micelle concentration of cationic surfactants, mainly quaternary ammonium bromide salts with one alkyl tail, has been studied theoretically by QSPR using molecular connectivity indices and also the atomic partial charges. Two models obtained allow to study the effect of head group structure modification on the value of the critical micelle concentration. The model, which contains molecular connectivity indices together with the charge localized on the central nitrogen atom, has a wider descriptive capability than the model which contains topological indices only. The present semi-empirical charge investigations of different quaternary ammonium salts show that the charges of the quaternary ammonium head groups are approximately the same, but one of the obtained models shows that the significant part of that charge, which is localized on the central nitrogen atom, affects the value of the critical micelle concentration.

Keywords Cationic surfactants · Head group · Critical micelle concentration · Molecular connectivity indices · Atomic partial charges

1 Introduction

Cationic surfactants have attracted the attention of chemists for a long time due to their wide application in various fields of industry and medicine. Their utility results from, among other things, their antimicrobial activities [1–3] and anticorrosion properties [4, 5]. The medical applications of cationic surfactants include their use as antimicrobial agents, as drug delivery agents and also as gene delivery agents. Studies of the complexing and transfection of DNA have been presented in recent years [6–10]. The use of cationic surfactants continues to increase and new surfactants are constantly synthesized. Recently, cationic surfactants and their critical micelle concentrations were investigated by Quantitative Structure–Property Relationship (QSPR) studies [11–17].

✉ Anna Mozrzymas
anna.mozrzymas@upwr.edu.pl

[1] Department of Physics and Biophysics, Wrocław University of Environmental and Life Sciences, ul. Norwida 25, 50-375 Wrocław, Poland

Many properties of the surfactants that make them so useful are due to their unique structure. The common feature of all the surfactants is their amphiphilic nature. They contain both hydrophobic groups called tails and hydrophilic groups called heads. The tail in many cationic surfactants is a linear alkyl chain and the head group is a quaternary ammonium group.

The architecture of the hydrophilic head group is the factor which plays an important role in regulating physical/chemical properties [18–20] and biological activities [21–23] of many cationic surfactants. It is well known that the electrostatic repulsion of the surfactant's ionic head groups counteracts micelle formation but some modifications of the structure of the head group can reduce the electrostatic repulsion of the head groups and facilitate aggregation. Experimental studies [19, 20] show that variations in the surfactants' head group structure influence the values of critical micelle concentration and the changes of the *cmc* values caused by structure modification can be even larger than one order of magnitude.

Recently, the effect of the spacer group of cationic gemini surfactants on the critical micelle concentration using the molecular connectivity indices has been reported [24]. This effect was studied using gemini surfactants with fixed alkyl chains and head groups but with various spacer groups. Continuing these studies, the author has decided to examine the effect of the head groups on the critical micelle concentration of cationic (monomeric) surfactants using, as in the previous paper [24], only the molecular connectivity indices and also a combination of the molecular connectivity indices with atomic charges.

The main aim of the present work is to find simple formulae which allow the study, in particular, the effect of the head group structure modifications on the value of the critical micelle concentration. The formulae have been derived for compounds with the same alkyl chains and the same counterion, but with various head groups. To the author's best knowledge, there was no paper presenting theoretical studies on the head group structure's effect on the critical micelle concentration using different descriptors, in particular the topological indices.

2 Data and Methods

To determine the influence the surfactants head group structure on the value of critical micelle concentration the surfactants with various head group sizes and varied hydrophilicity of the head were taken into account. The data set contains 15 compounds of the training set (Fig. 1) and 4 compounds of the test set (Fig. 2). The compounds in the training set have a hexadecyl hydrocarbon tail and bromide as the counterion.

The first series of the training set (compounds 1–4) consists of compounds with quaternary ammonium head groups with three linear alkyl chains containing one to four carbon atoms attached to the nitrogen atom [20]. In the second series (compounds 5–7) the $–CH_3$ groups of compound 1 are successively replaced by$–CH_2CH_2OH$ groups [8]. The compounds of the third series (compounds 8–10) have cyclic head groups with the quaternary ammonium groups incorporated into rings in size from five to seven atoms [20]. The head group of the compound 11 is a pyridinium ring [25]. The replacement of the –OH groups of compounds from the second series by$–OCH_3$ groups [19] gives the last series of the training set (compounds 13–15).

The test set compounds contains surfactants with primary, secondary, tertiary and quaternary ammonium head groups. To compare the effect of the variation in the head group

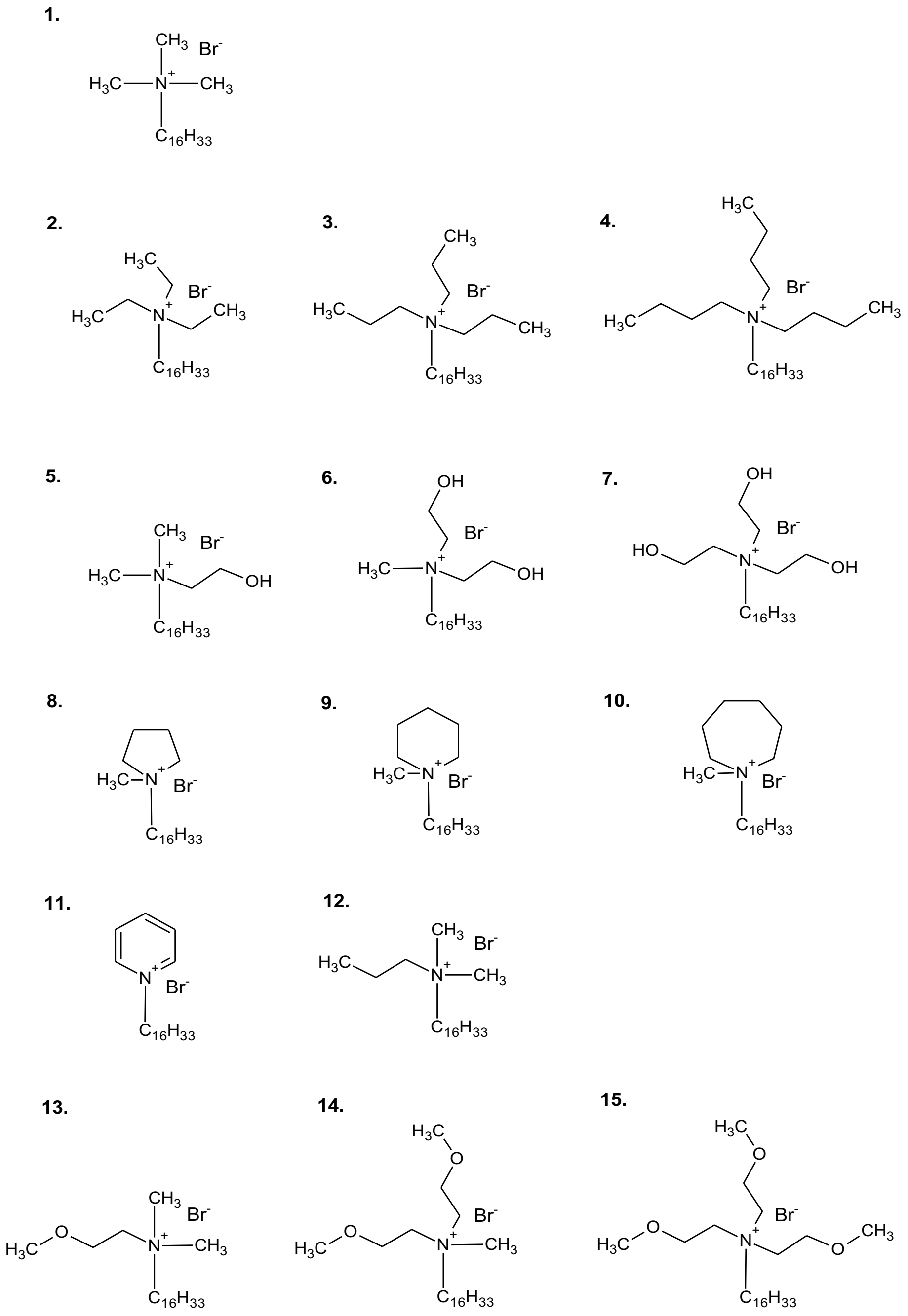

Fig. 1 Structures of compounds from the training set

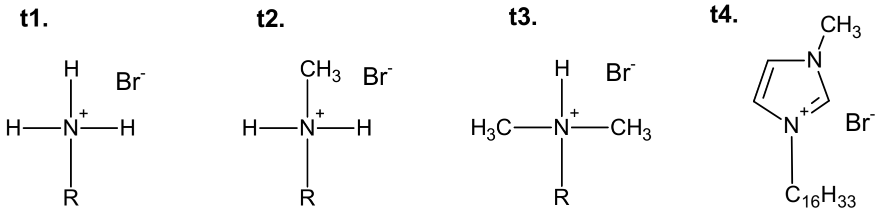

Fig. 2 Structures of compounds from the test set

structure on the values of critical micelle concentration, the experimental values of *cmc*, available in the literature, have been taken also (except compound t4) for compounds with a dodecyl hydrocarbon tail ($R = C_{12}H_{25}$) and bromide as counterion.

The chemical structures of the surfactants and the experimental values of *cmc* were taken from the literature [8, 19, 20, 25–27]. The experimental values of *cmc* of compounds from the training set are given in Table 1. The *cmc* values were measured in pure water [8, 19, 25] or in aqueous buffer of pH = 7.0 [20]. All *cmc* values were measured at 25.00 °C.

Some *cmc* values used in the calculations, which were measured in aqueous buffer solution (0.03 mmol·L^{-1} Tris buffer) of pH = 7.0 [20], were compared with values obtained in pure water [28] and found to be good agreement. The values taken from Ref. [28] are: 1.01 (mmol·L^{-1}), 0.799 (mmol·L^{-1}) and 0.585 (mmol·L^{-1}) for compounds 1, 2 and 3, respectively.

In the figure below (Fig. 3) are presented the structures of additional surfactants, proposed by the author, which will be needed in the discussion.

The Kier and Hall [29] molecular connectivity indices were calculated based on the graphic structural formulae of the molecules using the expression shown in Appendix A. The atomic partial charges were calculated using the semi-empirical molecular orbital package MOPAC 7 included in the VEGA program [30], employing the semi-empirical Hamiltonian PM3. The statistical calculations were performed using the program STATISTICA 12 [31].

3 Results

The values of all descriptors for compounds from the training set are contained in Table 2.

As follows from Table 2, ten topological descriptors: five molecular connectivity indices from zeroth to fourth order, and five valence molecular connectivity indices also from zeroth to fourth order, were used in the search for the best model. In this study other descriptors namely the charge localized on the central nitrogen atom (Q_{N^+}) and the total head group charge (Q_{tot}), were also taken into account. The total head group means the head group that contains α-methylene group, i.e. the first $-CH_2-$ group of the alkyl tail attached to the nitrogen atom.

Each formula expressing the relationship between the $\log_{10}$ *cmc* and the descriptors was generated using the least-squares method. The final equation was obtained using the stepwise method. The quality of the derived formulae was tested using three statistical parameters: the correlation coefficient (r), the Fisher ratio (F) and the standard deviation (s). The best model was selected according to the following statistical principles: highest values of r and F, lowest value of s and smallest possible number of significant descriptors in the model. Thus, the search for the best equation was stopped on three descriptors.

Table 1 Values of *cmc* of training set compounds

Compound	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
cmc/mmol·L^{-1}	0.91	0.81	0.48	0.36	0.202	0.153	0.036	0.83	0.76	0.67	0.64	0.612	0.203	0.122	0.073
[Refs.]	[20]	[20]	[20]	[20]	[8]	[8]	[8]	[20]	[20]	[20]	[25]	[19]	[19]	[19]	[19]

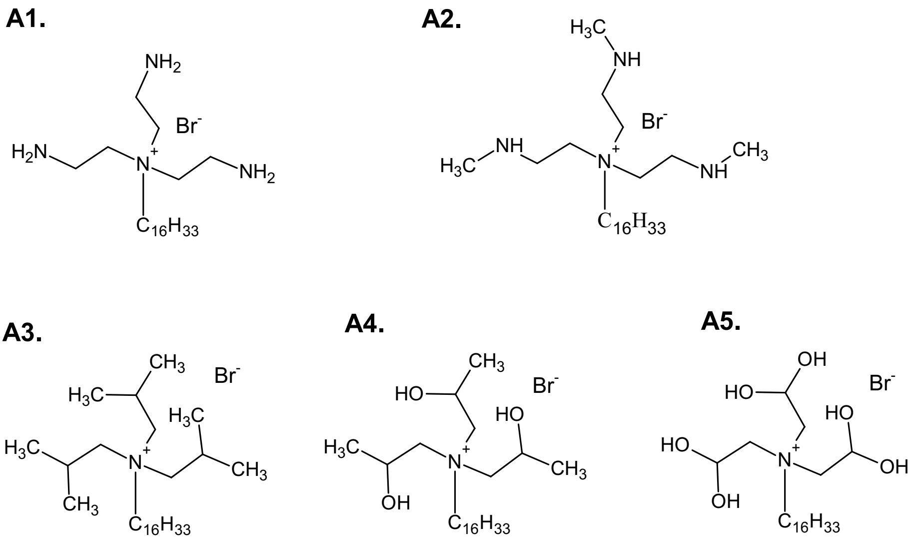

Fig. 3 Structures of additional compounds

3.1 Model 1, Using Molecular Connectivity Indices

The correlation formulae which contain the molecular connectivity indices only were derived basing on data contained in Tables 1 and 2. In the case of Model 1 the search for the best equation consisted of three steps. The results of all correlations for Model 1 are presented in Table 3.

As follows from Table 3, the highest values of the correlation coefficients in the third step are for the relationships containing the ${}^{0}\chi$, ${}^{0}\chi^{v}$, and ${}^{3}\chi_{c}$ indices and also for ${}^{0}\chi$, ${}^{0}\chi^{v}$, and ${}^{3}\chi_{c}^{v}$. Because the ${}^{3}\chi_{c}$, ${}^{3}\chi_{c}^{v}$ indices highly correlate ($r=0.997$), and also the ${}^{3}\chi_{c}$ index encodes only information about the number of branches and their environment while the ${}^{3}\chi_{c}^{v}$ index additionally encodes the information about heteroatoms, i.e., the relationship which contains the ${}^{3}\chi_{c}^{v}$ index is richer in structural information [29], thus the relationship which contains the ${}^{3}\chi_{c}$ index will be ignored in further considerations.

Thus, Model 1 is defined by the following molecular connectivity indices: the zeroth order molecular connectivity index (${}^{0}\chi$), the zeroth order valence molecular connectivity index (${}^{0}\chi^{v}$) and the cluster valence molecular connectivity index of third order (${}^{3}\chi_{c}^{v}$). The final formula for Model 1 is the following:

$$\begin{aligned}\log_{10} cmc = &-(1.002 \pm 0.325) - (0.777 \pm 0.060)\cdot {}^{0}\chi \\ &+ (0.674 \pm 0.062)\cdot {}^{0}\chi^{v} - (0.436 \pm 0.107)\cdot {}^{3}\chi_{c}^{v}\end{aligned} \tag{1}$$

The plot of the experimental values of $\log_{10} cmc$ versus the values of $\log_{10} cmc$ calculated using Eq. 1, along with the plot of residuals versus the experimental values of $\log_{10} cmc$, are presented in Fig. 4.

Table 2 Values of the descriptors

	$^{0}\chi$	$^{1}\chi$	$^{2}\chi$	$^{3}\chi_{c}$	$^{4}\chi_{pc}$	$^{0}\chi^{v}$	$^{1}\chi^{v}$	$^{2}\chi^{v}$	$^{3}\chi_{c}^{v}$	$^{4}\chi_{pc}^{v}$	$Q_{\mathrm{N}^{+}}$	Q_{tot}
1[a]	15.10660	9.56066	7.90685	1.56066	0.75	15.05381	9.36497	7.61012	1.39590	0.67082	0.6052	0.894
2	17.22792	11.24264	7.90685	0.70711	1.96599	17.17514	11.09334	7.61012	0.63246	1.75843	0.5542	0.893
3	19.34924	12.74264	9.09619	0.70711	1.5	19.29646	12.59334	8.83226	0.63246	1.34164	0.5687	0.901
4	21.47056	14.24264	10.1568	0.70711	1.5	21.41778	14.09334	9.89292	0.63246	1.34164	0.5774	0.921
5	16.52081	10.62132	8.26041	1.20711	1.20711	15.91524	10.05022	7.70275	1.07967	1.07967	0.6028	0.894
6	17.93503	11.68198	8.65685	0.92678	1.43566	16.77667	10.73546	7.83373	0.82893	1.28409	0.5944	0.893
7	19.34924	12.74264	9.09619	0.70711	1.5	17.63810	11.42070	8.00308	0.63246	1.34164	0.6	0.900
8	15.93503	10.7678	8.36396	0.92678	1.43566	15.88224	10.60300	8.09362	0.82893	1.28409	0.5666	0.898
9	16.64214	11.2678	8.71751	0.92678	1.43566	16.58935	11.10300	8.44718	0.82893	1.28409	0.5722	0.899
10	17.34924	11.7678	9.07107	0.92678	1.43566	17.29646	11.60300	8.80073	0.82893	1.28409	0.5743	0.898
11[a]	15.71949	10.93185	7.88890	0.20412	0.43301	14.94057	9.87307	6.70951	0.10541	0.19625	0.494	0.902
12	16.52081	10.62132	8.26041	1.20711	1.20711	16.46803	10.44110	7.97914	1.07967	1.07967	0.5908	0.898
13	17.22792	11.12132	8.61396	1.20711	1.20711	16.87628	10.43091	7.97194	1.07967	1.07967	0.607	0.888
14	19.34924	12.68198	9.36396	0.92678	1.43566	18.69874	11.49685	8.37212	0.82893	1.28409	0.5926	0.897
15	21.47056	14.24264	10.1568	0.70711	1.5	20.52120	12.56279	8.81066	0.63246	1.34164	0.5878	0.894

[a] For these compounds the values of some molecular connectivity indices were published in Ref. [13]

Table 3 Values of correlation coefficients in each step of Model 1

Connectivity index	${}^0\chi$	${}^1\chi$	${}^2\chi$	${}^3\chi_c$	${}^4\chi_{pc}$	${}^0\chi^v$	${}^1\chi^v$	${}^2\chi^v$	${}^3\chi_c^v$	${}^4\chi_{pc}^v$
Step 1	0.621	0.563	0.536	0.130	0.210	0.415	0.260	0.132	0.111	0.215
Step 2	–	0.660	0.631	0.630	0.633	0.936	0.920	0.791	0.629	0.630
Step 3	–	0.961	0.939	0.974	0.938	–	0.961	0.939	0.974	0.940

3.2 Model 2 Using Molecular Connectivity Indices and Atomic Partial Charges

The correlation formulae which contain the molecular connectivity indices and the charges were derived based on data contained in Tables 1 and 2. The search for the best equation of Model 2 consists of three steps. The results of all correlations for Model 2 are presented in Table 4.

As shown in Table 4, two molecular connectivity indices define Model 2: the zeroth order molecular connectivity index (${}^0\chi$) and the zeroth order valence molecular connectivity index (${}^0\chi^v$). The third descriptor which defines Model 2 is the charge (Q_{N^+}) localized on the central nitrogen atom. The final formula for Model 2 is the following:

$$\begin{aligned}\log_{10} cmc = {} & (0.684 \pm 0.584) - (0.675 \pm 0.053) \cdot {}^0\chi \\ & + (0.600 \pm 0.056) \cdot {}^0\chi^v - (4.451 \pm 0.989) \cdot Q_{\mathrm{N}^+}\end{aligned} \tag{2}$$

The plot of the experimental values of $\log_{10} cmc$ versus the values of $\log_{10} cmc$ calculated using Eq. 2, along with the plot of residuals versus the experimental values of $\log_{10} cmc$, are presented in Fig. 5.

The values of $\log_{10} cmc$ calculated using Models 1 and 2 and the experimental ones are given in Table 5 for the compounds of the training set and in Table 6 for the test set compounds.

Models 1 and 2 were derived based on the training set compounds so all calculated values of $\log_{10} cmc$ are for compounds with the hexadecyl hydrocarbon tail and

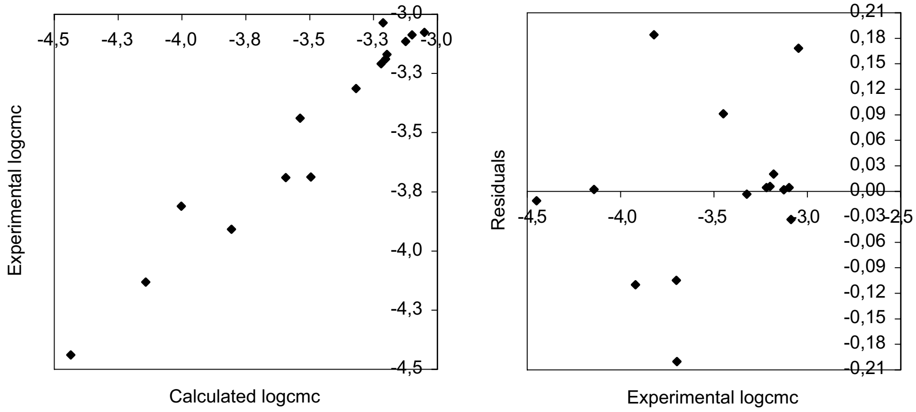

Fig. 4 Plot of the experimental $\log_{10} cmc$ values versus the calculated using Eq. 1 for the training set ($r = 0.974$, $F = 242.104$, $s = 0.102$), and plot of residuals versus the experimental $\log_{10} cmc$ values

Table 4 Values of correlation coefficients in each step of Model 2

Descriptor	$^0\chi$	$^1\chi$	$^2\chi$	$^3\chi_c$	$^4\chi_{pc}$	$^0\chi^v$	$^1\chi^v$	$^2\chi^v$	$^3\chi_c^v$	$^4\chi_{pc}^v$	$Q_{\mathrm{N^+}}$	Q_{tot}
Step 1	0.621	0.563	0.536	0.130	0.210	0.415	0.260	0.132	0.111	0.215	0.450	0.118
Step 2	–	0.660	0.631	0.630	0.633	0.936	0.920	0.791	0.629	0.630	0.704	0.755
Step 3	–	0.961	0.939	0.974	0.938	–	0.961	0.939	0.974	0.940	0.978	0.957

bromide as counterion. To compare the effect of the variation in the head group structure on the *cmc* value, the experimental $\log_{10}$ *cmc* values (those available in literature) contained in Table 6 have been taken (except for compound t4) for compounds with a dodecyl hydrocarbon tail ($R = C_{12}H_{25}$), but the calculated values using Models 1 and 2 are for compounds with the hexadecyl hydrocarbon tail ($R = C_{16}H_{33}$). Table 6 contains also the $\log_{10}$ *cmc* value of compound 1 from the training set which will be helpful in the discussion.

The values of calculated $\log_{10}$ *cmc* for compounds of the additional set are given in Table 7.

4 Discussion

As follows from Table 5, the calculated values of $\log_{10}$ *cmc* using Models 1 and 2 are very close to the experimental ones. Table 6 shows that Model 1 better predicts the $\log_{10}$ *cmc* value of test compound t4 than Model 2, that means it may have better predictive ability. Perhaps this is due to the fact that the error of the additive parameter in Model 2 (Eq. 2) is only slightly lower than the value of the additive parameter itself, what makes that equation weaker in respect of the estimation capability. But the main aim of this work is to find the formula that allows to study the effect of the head group structure modifications on the *cmc* value, i.e. the formula which will show how the *cmc* values of cationic (monomeric)

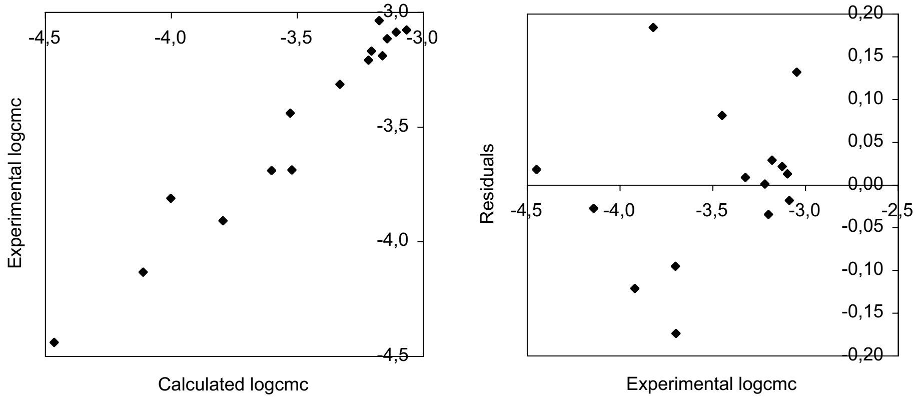

Fig. 5 Plot of the experimental $\log_{10}$ *cmc* values versus the calculated using Eq. 2 for the training set ($r = 0.978$, $F = 284.876$, $s = 0.09$), and plot of residuals versus the experimental $\log_{10}$ *cmc* values

Table 5 Calculated and literature values of $\log_{10}$ *cmc* of compounds from the training set

Compound	Experimental $\log_{10}$ *cmc*	Calculated $\log_{10}$ *cmc* Model 1	Residual	Calculated $\log_{10}$ *cmc* Model 2	Residual
1	−3.041 [20]	−3.208	0.167	−3.172	0.131
2	−3.091 [20]	−3.095	0.003	−3.103	0.012
3	−3.319 [20]	−3.314	−0.005	−3.326	0.007
4	−3.444 [20]	−3.533	0.090	−3.524	0.080
5	−3.695 [8]	−3.589	−0.106	−3.598	−0.096
6	−3.815 [8]	−3.998	0.183	-3.998	0.183
7	−4.444 [8]	−4.431	−0.012	−4.461	0.017
8	−3.081 [20]	−3.046	−0.034	−3.062	−0.019
9	−3.119 [20]	−3.120	0.001	−3.139	0.02
10	−3.174 [20]	−3.193	0.019	−3.202	0.028
11	−3.194 [25]	−3.198	0.004	−3.158	−0.036
12	−3.213 [19]	−3.216	0.003	−3.213	−0.0001
13	−3.692 [19]	−3.491	−0.202	−3.517	−0.175
14	−3.914 [19]	−3.802	−0.111	−3.791	−0.122
15	−4.137 [19]	−4.137	0.001	−4.108	−0.029

Table 6 Calculated and literature values $\log_{10}$ *cmc* of compounds t1–t4 from the test set and compound 1 from the training set

Compound	Experimental $\log_{10}$ *cmc*	Calculated $\log_{10}$ *cmc*	
		Model 1	Model 2
t4	−3.100 [26]	−3.193	−2.687
t1	−1.951[a] [27]	−2.498	−4.130
t2	−1.959[a] [27]	−2.461	−3.716
t3	−1.873[a] [27]	−2.669	−3.393
1	−1.836[a] [27]	−3.208	−3.172

[a] The *cmc* values were determined for compounds with the dodecyl hydrocarbon tail

surfactants change with change in the head group structure, which means that the additive parameter is not so significant in these investigations and therefore Model 2 will be also taken into account.

An inspection of data contained in Table 5 reveals that, in agreement with the experiments, as the head group size increases the critical micelle concentration decreases. The slight decrease in the *cmc* value is then when with the increase of the head group size the hydrophobicity of the head group chains also increases (compounds 2–4 and 8–10). The exceptions are compounds 1 and 2 for which the calculated values of the *cmc* increase a little when the $–CH_3$ groups are replaced by $–CH_2CH_3$ groups. A larger decrease in the *cmc* is then when the size of the head group increases by replacing all of the $–CH_3$ groups by $–CH_2CH_2OH$ groups. A comparison of the head groups which contain hydrocarbon

Table 7 Calculated values $\log_{10}$ *cmc* of compounds from the additional set

Compound	Calculated $\log_{10}$ *cmc*	
	Model 1	Model 2
A1	-4.161	-4.225
A2	-3.944	-3.870
A3	-4.109	-3.545
A4	-4.932	-4.805
A5	-5.917	-5.793

chains and those which contain heteroatoms shows that replacement of the $–CH_2–$ group by –O– in the head group (compounds 4 and 15) causes a decrease the experimental and also the calculated using the Models 1 and 2 values of the *cmc*. Also, the experimental *cmc* values and those calculated using Models 1 and 2 are lower as the $–CH_3$ group is replaced by an –OH group (compounds 3 and 7). As follows from Tables 5 and 7, the replacement the $–CH_2–$ group by a –NH– or $–CH_3$ group by $–NH_2$ (compounds 3 and A1, 4 and A2) causes a decrease the theoretical *cmc* values too. From the comparison of compounds 7 with A1 and compounds 15 with A2, it follows that the replacement the –OH group by $–NH_2$, or –O– by the –NH– group causes an increase the *cmc* values calculated using Models 1 and 2.

Comparison of the compounds with straight and branched hydrocarbon chains and the same number of atoms in the head group (compounds 4 and A3) shows that the branches cause a decrease the calculated *cmc* values (Tables 5 and 7). In the case of calculated *cmc* values using Model 1, this decrease is significant: 0.293 mmol·L^{-1} and 0.078 mmol·L^{-1} for compounds 4 and A3, respectively. But in the case of *cmc* values calculated using Model 2, this decrease is slight: 0.299 mmol·L^{-1} and 0.285 mmol·L^{-1} for compounds 4 and A3, respectively. From Tables 5 and 7 it also follows that the comparison of compounds with straight and branched chains containing heteroatoms and the same number of atoms in the head group (compounds 15 and A4) shows that the presence of heteroatoms on branches, in comparison with those in straight chains, cause the decrease the calculated *cmc* values too. The comparison of the compounds A3–A5 (Table 7) shows that, similarly to compounds with straight chains in the head group, the replacement of the hydrophobic branch by a hydrophilic one, namely the $–CH_3$ group by the –OH group, results in lower calculated using Models 1 and 2 values of the critical micelle concentration and, in the case of Model 2, the changes of *cmc* value are at least one order of magnitude.

An inspection of data contained in Table 6 shows also that the experimental values of critical micelle concentration of bromide dodecyl chain surfactants with primary, secondary, tertiary and quaternary ammonium head groups (compounds t1–t3, 1) increase with the increase of the head group order. For dodecyl chain surfactants with primary, secondary, tertiary and quaternary ammonium head groups the experimental *cmc* values are: 11.2 mmol·L^{-1}, 11 mmol·L^{-1}, 13.4 mmol·L^{-1} and 14.6 mmol·L^{-1} [27]. As follows from Table 6, the *cmc* values of the hexadecyl chain surfactants (compounds t1–t3, 1) calculated using Model 1 at first increase a little and next decrease with further increase of the head group order, but by using Model 2 the changes in the calculated *cmc* values for these compounds are in good agreement with the experimental ones.

Model 1 (Eq. 1) contains the zeroth order molecular connectivity index (${}^0\chi$), the zeroth order valence molecular connectivity index (${}^0\chi^v$) and the cluster valence molecular connectivity index of the third order (${}^3\chi_c^v$), whereas Model 2 (Eq. 2) contains also the zeroth order molecular connectivity index (${}^0\chi$) and the zeroth order valence molecular connectivity index (${}^0\chi^v$) but along with the charge (Q_{N^+}) localized on the central nitrogen atom. All the molecular connectivity indices were calculated based on the hydrogen-suppressed graphic structural formula of the molecule. The values of the ${}^0\chi$ and ${}^0\chi^v$ indices increase when the number of atoms in the head group increases but the values of the ${}^3\chi_c^v$ index increase with the increase number of branches in the head group. The analysis of the variation of the values of nitrogen charges (Q_{N^+}) in Table 2 shows that the charge localized on the quaternary nitrogen atom increases with the increase of size caused by elongation of the hydrocarbon chains (compounds 2–4), or it decreases with the increase of the number of $-CH_2CH_2OCH_3$ groups in the head group (compounds 13–15). For the compound with a pyridinium ring (compound 11) the value of the nitrogen charge (Q_{N^+}) differs significantly from the others. In the remaining cases (compounds 5–7, 8–10) the increase in size of the head group does not cause any significant changes in the value of Q_{N^+}.

The cluster valence molecular connectivity index of the third order (${}^3\chi_c^v$) appearing in Model 1 represents three-bond cluster terms within the molecule and it differentiates heteroatoms [29]. Thus the ${}^3\chi_c^v$ index is calculated from that part of the molecules of the training set which includes the quaternary nitrogen atom, and only the changes in the nearest environment of that atom influence the value of the ${}^3\chi_c^v$ index for most investigated compounds. Whereas the charge on the central nitrogen atom depends on all atoms in the head group. Maybe therefore the inclusion of the the nitrogen atom charge (Q_{N^+}) in the correlation analysis gives better description capability. In general, Model 1 is very good in describing mainly compounds with quaternary ammonium head groups, namely those compounds which were used in the correlation analysis and probably the inclusion to the correlation analysis of the compounds from the test set could give a model a better descriptive capability.

The results obtained for cationic monomeric surfactants, the experimental and the presented theoretical are consistent with those obtained for gemini surfactants. For example, when the head group size of the gemini surfactants increases with successive replacement the methyl groups by hydroxyethyl groups, the *cmc* values decrease [32]. Also, when the hydrocarbon parts of the head group of gemini surfactants increase from methyl to butyl, the *cmc* values decrease too [33].

It is worth to note that for the training set compounds, i.e. the compounds with a quaternary ammonium head group, the charges calculated on the total head group are approximately equal to 0.9 (Table 2). As was described in the "Results" section, the total head group means the head group which contains an α-methylene group. Recent calculations of the charge on the α-methylene group for different ionic surfactants show that the charge on that group is significant [34, 35] and often this group is hydrated [36]; this suggests that the α-methylene group should be considered as a part of the head group [34]. Thus, assuming that the α-methylene group is included in the head group, the charges of the head groups are about 0.9 regardless of whether that group is more or less hydrophilic, smaller or larger. Similar values (0.885 and 0.891) were obtained by Huibers [34] for trimethylammonium and pyridinium surfactants and also by Zhao et al. [35] for trimethylammonium and tripropylammonium surfactants (0.893, 0.899), using also PM3 calculations. This means that the changes of the head group structure do not affect significantly the value of the quaternary ammonium head group charge, while the critical micelle concentration in some cases changes significantly. The values of the total head group charges (Q_{tot}) included in Table 2 show that only for compounds 2–4 there is a slight decrease of the head group charge along with the increase the length of the hydrophobic head chains.

The paper [35] reported a study of the ammonium salts with methyl to butyl head chains ($C_{12}NMe_3^+$, $C_{12}NEt_3^+$, $C_{12}NPr_3^+$, $C_{12}NBu_3^+$) and also different alkyl tail lengths ($C_{10}NMe_3^{3+}$, $C_{12}NMe_3^+$). The authors of this paper noted that the *cmc* values change though these cationics surfactants have nearly the same charge on the alkyl tail. Although the authors of paper [35] do not write directly about the head groups, this observation is consistent with results of the present work obtained for a larger number of compounds and for compounds with more diverse head groups. Also, in paper [35] it was suggested that, together with the α-methylene also the second to fifth $–CH_2–$ groups of the alkyl tail also may be considered as a part of the head group. In this case the values of the head group charges are of course larger but still they are approximately the same. Thus, also in this case, the changes in the head group structure do not affect significantly the charge of the quaternary ammonium head group.

5 Conclusions

The present theoretical QSPR study confirms the experimental results that the structure and nature of the head group influence the *cmc* value.

The presence of heteroatoms in the head group of cationic surfactants causes a decrease of the critical micelle concentration. Also, an increase in the number of branches in the head group causes a decrease of the *cmc* values. The –OH and –O– groups reduce the *cmc* value more than the $–NH_2$ and –NH– groups and the –OH group more than the –O– group. Considering the effect of the head group on the value of the critical micelle concentration, the results obtained suggest that the most active surfactants would be those that contain many –OH groups in the head group, for example the A5 compound proposed in the Sect. 2.

The present investigations of the atomic partial charges of the head groups show that, in general, the head group charges of different quaternary ammonium salts are approximately the same. The structures of the surfactants are different and the *cmc* values also differ but the total head group charges (Q_{tot}) are approximately equal to 0.9. This means that the changes in the head group structure do not significantly change the charge value of the quaternary ammonium head group while the critical micelle concentration changes and in some cases even considerably.

The models obtained will be helpful in designing novel cationic surfactants that are more active in micelle formation.

Acknowledgements The statistical calculations were performed using the program Statistica 12 provided by the Wrocław University of Environmental and Life Sciences.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A

The expression for *m*-th order molecular connectivity index is the following [29]:

$$^{m}\chi_{k}=\sum_{j=1}^{n_m}\prod_{i=1}^{m+1}\left(\delta_i\right)_j^{-0.5} \tag{3}$$

where δ_i is the connectivity degree, i.e. the number of non-hydrogen atoms to which the *i*-th non-hydrogen atom is bonded, *m* is the order of the connectivity index, *k* denotes the type of the fragment of the molecule: path (p), cluster (c) and path cluster (pc), n_m is the number of fragments of type *k* and order *m*.

The replacement of δ_i by δ_i^v defined by

$$\delta_i^v=\frac{Z_i^v-h_i}{Z_i-Z_i^v-1} \tag{4}$$

where Z_i^v is the number of valence electrons in the *i*-th atom, h_i is the number of hydrogen atoms connected to the *i*-th atom and Z_i is the number of all electrons in the *i*-th atom, defines the valence molecular connectivity index $^{m}\chi_k^v$.

References

1. Zhi, L., Li, Q., Li, Y., Sun, Y.: Self-aggregation and antimicrobial activity of saccharide–cationic surfactants. Colloids Surf. A Physicochem. Eng. Asp. **456**, 231–237 (2014)
2. Brycki, B., Kowalczyk, I., Kozirog, A.: Synthesis, molecular structure, spectral properties and antifungal activity of polymethylene-α,ω-bis(*N*,*N*-dimethyl-*N*-dodecyloammonium bromides). Molecules **16**, 319–335 (2011)
3. Woźniak, E., Mozrzymas, A., Czarny, A., Kocięba, M., Różycka-Roszak, B., Dega-Szafran, Z., Dulewicz, E., Petryna, M.: Antimicrobial activity of *N*-alkoxycarbonylmethyl-*N*-alkyl-piperidinium chlorides. Z. Naturforsch. **59c**, 782–786 (2004)
4. Kaczerewska, O., Leiva-Garcia, R., Akid, R., Brycki, B.: Efficiency of cationic gemini surfactants with 3-azamethylpentamethylene spacer as corrosion inhibitors for stainless steel in hydrochloric acid. J. Mol. Liq. **247**, 6–13 (2017)
5. Malik, M.A., Hashim, M.A., Nabi, F., Al-Thabaiti, S.A., Khan, Z.: Anti-corrosion ability of surfactants: a review. Int. J. Electrochem. Sci. **6**, 1927–1948 (2011)
6. Chauhan, V., Singh, S., Mishra, R., Kaur, G.: Synthesis and bio-physicochemical properties of amide-functionalized *N*-methylpiperazinium surfactants. J. Colloid Interface Sci. **436**, 122–131 (2014)
7. Misiak, P., Wilk, K.A., Kral, T., Woźniak, E., Pruchnik, H., Frąckowiak, R., Hof, M., Różycka-Roszak, B.: New gluconamide-type cationic surfactants: interactions with DNA and lipid membranes. Biophys. Chem. **180–181**, 44–54 (2013)
8. Dasgupta, A., Das, P.K., Dias, R.S., Miguel, M.G., Lindman, B., Jadhav, V.M., Gnanamani, M., Maiti, S.: Effect of headgroup on DNA–cationic surfactant interactions. J. Phys. Chem. B **111**, 8502–8508 (2007)
9. Lléres, D., Clamme, J.-P., Dauty, E., Blessing, T., Krishnamoorthy, G., Duportail, G., Mély, Y.: Investigation of the stability of dimeric cationic surfactants/DNA complexes and their interaction with model membrane systems. Langmuir **18**, 10340–10347 (2002)
10. Fisicaro, E., Compari, C., Bacciottini, F., Contardi, L., Pongiluppi, E., Barbero, N., Viscardi, G., Quagliotto, P., Donofrio, G., Krafft, M.P.: Nonviral gene-delivery by highly fluorinated gemini bispyridinium surfactant-based DNA nanoparticles. J. Colloid Interface Sci. **487**, 182–191 (2017)
11. Jalali-Heravi, M., Konouz, E.: Multiple linear regression modelling of the critical micelle concentration of alkyltrimethylammonium and alkylpyridinium salts. J. Surfact. Deterg. **6**, 25–30 (2003)

12. Kardanpour, Z., Hemmateenejad, B., Khayamian, T.: Wavelet neutral network-based QSPR for prediction of critical micelle concentration of gemini surfactants. Anal. Chim. Acta **531**, 285–291 (2005)
13. Mozrzymas, A., Różycka-Roszak, B.: Prediction of critical micelle concentration of cationic surfactants using connectivity indices. J. Math. Chem. **49**, 276–289 (2011)
14. Katritzky, A.R., Pacureanu, L.M., Slavov, S.H., Dobchev, D.A., Shah, D.O., Karelson, M.: QSPR study of the first and second critical micelle concentrations of cationic surfactants. Comp. Chem. Eng. **33**, 321–332 (2009)
15. Roy, K., Kabir, H.: QSPR with extended topochemical atom (ETA) indices, 3: modeling of critical micelle concentration of cationic surfactants. Chem. Eng. Sci. **81**, 169–178 (2012)
16. Mozrzymas, A.: Modelling of the critical micelle concentration of cationic gemini surfactants using molecular connectivity indices. J. Solution Chem. **42**, 2187–2199 (2013)
17. Mozrzymas, A.: Molecular connectivity indices for modeling of the critical micelle concentration of cationic (chloride) gemini surfactants. Colloid Polym. Sci. **295**, 75–87 (2017)
18. Buckingham, S.A., Garvey, C.J., Warr, G.G.: Effect of head-group size on micellization and phase behaviour in quaternary ammonium surfactant systems. J. Phys. Chem. **97**, 10236–10244 (1993)
19. Das, D., Roy, S., Mitra, R.N., Dasgupta, A., Das, P.K.: Head-group size or hydrophilicity of surfactants: the major regulator of lipase activity in cationic water-in-oil microemulsions. Chem. Eur. J. **11**, 4881–4889 (2005)
20. Schnee, V.P., Palmer, C.P.: Cationic surfactants for micellar electrokinetic chromatography: 1. Characterization of selectivity using the linear solvation energy relationship model. Electrophoresis **29**, 767–776 (2008)
21. Viscardi, G., Quagliotto, P., Barolo, C., Savarino, P., Barni, E., Fisicaro, E.: Synthesis and surface and antimicrobial properties of novel cationic surfactants. J. Org. Chem. **65**, 8197–8203 (2000)
22. Pernak, J., Branicka, M.: The properties of 1-alkoxymethyl-3-hydroxypyridinium and 1-alkoxymethyl-3-dimethylaminopyridinium chlorides. J. Surfact. Deterg. **6**, 119–123 (2003)
23. Haldar, J., Kondaiah, P., Bhattacharya, S.: Synthesis and antibacterial properties of novel hydrolysable cationic amphiphiles. Incorporation of multiple head groups leads to impressive antibacterial activity. J. Med. Chem. **48**, 3823–3831 (2005)
24. Mozrzymas, A.: On the spacer group effect on critical micelle concentration of cationic gemini surfactants using molecular connectivity indices. Combinatorial Chem. High Throughput Screen. **19**, 481–488 (2016)
25. Skerjanc, J., Kogej, K., Cerar, J.: Equilibrium and transport properties of alkylpyridinium bromides. Langmuir **15**, 5023–5028 (1999)
26. Baltazar, Q.Q., Chandawalla, J., Sawyer, K., Anderson, J.L.: Interfacial and micellar properties of imidazolium-based monocationic and dicationic ionic liquids. Colloids Surf. A Physicochem. Eng. Asp. **302**, 150–156 (2007)
27. Nishimura, S., Scales, P.J., Biggs, S., Healy, T.W.: An electrokinetic study of the adsorption of dodecyl ammonium amine surfactants at the muscovite mica–water interface. Langmuir **16**, 690–694 (2000)
28. Jaber, R., Wasbrough, M.J., Haldaway, J.A., Edler, K.J.: Interactions between quaternary ammonium surfactants and polyethylenimine at high pH in film forming systems. J. Colloid Interface Sci. **449**, 286–296 (2015)
29. Kier, L.B., Hall, L.H.: Molecular connectivity in structure-activity analysis. Research Studies Press Ltd, Letchworth (1986)
30. Pedretti, A., Villa, L., Vistoli, G.: VEGA—an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. J. Comput. Aided Mol. Des. **18**, 167–173 (2004)
31. STATISTICA (data analysis software system), version 12, StatSoft, Inc.; software. http://www.statsoft.com (2014)
32. Borse, M., Sharma, V., Aswal, V.K., Goyal, P.S., Devi, S.: Effect of head group polarity and spacer chain length on the aggregation properties of gemini surfactants in an aquatic environment. J. Colloid Interface Sci. **284**, 282–288 (2005)
33. Zhang, Q., Gao, Z., Xu, F., Tai, S.: Effect of hydrocarbon structure of the head group on the thermodynamic properties of micellization of cationic gemini surfactants: an electrical conductivity study. J. Colloid Interface Sci. **371**, 73–81 (2012)
34. Huibers, P.D.T.: Quantum-chemical calculations of the charge distribution in ionic surfactants. Langmuir **15**, 7546–7550 (1999)
35. Zhao, G.-X., Zhu, B.-Y., Dou, Z.-P., Yan, P., Xiao, J.-X.: Effect of charge distribution along surfactant molecules on physico-chemical properties of surfactant systems. Colloids Surf. A Physicochem. Eng. Asp. **327**, 122–126 (2008)

36. Heindl, A., Kohler, H.-H.: Rod formation of ionic surfactants: a thermodynamic model. Langmuir **12**, 2464–2477 (1996)