



Applying group functions to description of ionic liquids



Ilya V. Popov^{a,c}, Andrei L. Tchougréeff^{a,b,c,*}

^a Moscow Center for Continuous Mathematical Education, Russia

^b Chair of Solid State and Quantum Chemistry, RWTH Aachen University, Aachen, Germany

^c A.N. Frumkin Institute of Physical Chemistry and Electrochemistry of RAS, Russia

ARTICLE INFO

Article history:

Received 24 December 2016

Received in revised form 1 April 2017

Accepted 1 April 2017

Available online 6 April 2017

Keywords:

New concept of semi-empirism

Group functions

Ionic liquids

ABSTRACT

Concept of group functions (GF) in quantum chemistry explicates chemical idea of ‘chromophores’ – fragments of molecules responsible for their physical properties and chemical reactivity. This concept permits to construct quantum chemical methods which provide correct qualitative and quantitative results for complex systems being unpretentious in computer resources. This is achieved by basing the form of the trial electronic wave function on adequate ‘chromophores’ present in the system under study. If it can be said so, substantial percentage of needs in computational resources are switch on the researcher’s mind which should be able to realize what parts of the molecule are adequate ‘chromophores’ at given conditions by this giving a chance to a “new concept of semi-empirism”.

In our previous studies group functions were applied to transition metal complexes of 3d-elements (effective Hamiltonian crystal field – EHCF) and to organic molecules with single and isolated multiple bonds (antisymmetrized product of strictly localized geminals – APSLG). It was shown that these methods can be successfully applied for theoretical study of complex systems which represent considerable difficulties for standard quantum chemistry methods. For organic molecules deductive molecular mechanics which represents a simplified version of the APSLG method was also proposed.

In the present paper recent results obtained with use of the GF theory are discussed. Deductive molecular mechanics force field stemming from the APSLG method is adopted for many-particle molecular assemblies and possibility of its application for modeling liquid state is demonstrated on an example of binary mixtures of hydrophobic-hydrophilic ionic liquids with water.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

For the long history of progress in understanding nature of chemical compounds theoretical concepts which allow to explain and predict experimentally observed relations between properties and structure of the system have been developed. These concepts are based on the idea of picking out fragments of the molecule responsible for its specific properties like reactivity within a class of chemical transformations and to focus on this fragment, considering the rest to be inert at present conditions and taking its influence into account as ‘perturbation’. This approach – dividing into parts is the most general way of studying the reality which was formulated as philosophic maxima yet by Rene Descartes: ‘to divide each of the difficulties under examination into as many parts as possible, and as might be necessary for its adequate solution’ [1].

Creation of such concepts was decisive for the development of chemistry. The most convincing example is the concept of atom which allows chemists to understand nature and reactivity of matter and to provide a huge number of substances, materials and methods of their synthesis. Similar paradigm applies to physical properties of molecules as well. E.g., color of substances is explained with use of the concept of chromophore proposed by Witt [2]. To emphasize the correctness of the idea to pick out the fragment responsible for the system’s properties we generalize term ‘chromophore’ and use it for any distinguishable part of the system responsible for its properties following the suggestion of Ruedenberg [3] (see [4] for a detailed elaboration on this topic).

Unfortunately in quantum chemistry (QC) as we know it now there are no general approaches to explicate the idea of ‘chromophores’ in QC (excluding several examples which explicate it in an intuitive way: crystal field theory [5] and Hückel theory [6]). The main reason which causes such state of affairs is that most ‘chromophores’ which are observable in chemistry are considered not to exist in QC. For example, chemical bonds are considered

* Corresponding author at: Chair of Solid State and Quantum Chemistry, RWTH Aachen University, Aachen, Germany.

E-mail address: andrei.tchougreeff@ac.rwth-aachen.de (A.L. Tchougréeff).

not to be observable, because there is no operator for them [7]; atoms are not existing in most the purist *ab initio* point of view, rather nuclei and electrons. Thus it is important to notice that all entities proposed by human are objects in mind which form the foundations of some symbolic systems (languages) describing the reality. And the only criterion of entity existence is the comparison of such description and reality itself (more detailed discussions can be found in [4]). At this point of view, atoms and bonds are observable entities, thus they should be somehow expressed in terms of QC.

We abstract the idea of 'chromophores' can come back to quantum chemistry if one takes the trial wave function (WF) [4] in the McWeeny's group functions approximation (GF) [8] as proposed in [9] assuming the groups to represent (generalized) chromophores. According to the group functions approximation total electronic wave function has the general form (see Appendix A also):

$$\Psi(x_1 \dots x_N) = \hat{A} \left\{ \Phi^A(x_1 \dots x_{N_A}) \cdot \Phi^B(x_{N_A+1} \dots x_{N_B}) \cdot \dots \right\}$$

$$N = N_A + N_B + \dots$$

where A, B refer to subsystems in which the original system is divided, Φ^K are their electronic wave functions and N_K are numbers of electrons at these subsystems. $\hat{A}$ is projection operator to the antisymmetric irreducible representation of the symmetric group S_N . When total WF Ψ is expressed that way, it is called 'group function' [8]. SCF-like wave function is obviously a special case of group function which is obtained when all N_K are equal to unity.

Applying GF approach to transition metal complexes (containing 3d-elements) and organic molecules with single and isolated multiple bonds, two methods were performed in our previous studies: Effective Hamiltonian of Crystal Field (EHCF) [11] and Antisymmetrized Product of Strictly Local Geminals (APSLG) [10,12] respectively. Discussions of EHCF method is beyond the scope of this paper, further information may be found in [11,13,14]. In the present study we are going to focus on APSLG method and its descendants.

In the APSLG method the trial WF may be expressed as:

$$\Psi = \hat{A} \prod_m g_m$$

where g_m are two-electronic functions which correspond to chemical bonds and lone electronic pairs in organic molecules, so called 'geminals'. This form of trial function takes chemical bonds as (generalized) 'chromophores' – observable entities – which is a natural choice for organic molecules with single and isolated multiple bonds [15–24]. In the second quantization formalism they can be expressed as:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m (r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+)$$

where u_m, v_m – amplitudes of ionic contributions; w_m – amplitude of covalent contributions; $r_{m\sigma}^+, l_{m\sigma}^+$ are Fermi-operators creating an electron with spin projection σ in the one-electronic states (OES) $|r_m\rangle$ and $|l_m\rangle$, respectively. For the lone pairs only one contribution remains:

$$g_m^+ = r_{m\alpha}^+ r_{m\beta}^+$$

As it was discussed previously [16] significant success of MM as applied to organic molecules is based on underlying representation of molecular electronic structure: in all MM schemes chemical bonds are assumed to be local and transferable pretty much independent on their surrounding. Chemical bonds are 'chromophores' in this case and using, say, APSLG approximation for the underlying wave function is a good starting point to deduce rather than simply parameterize an MM force field. Truly, strictly localized hybrid

orbitals and transferable of one- and two electronic density matrix elements coming from APSLG permit to present the energy of a molecule in the classical form as combination of additive potentials with parameters determined from QC calculations (see Appendix B). Such alternative molecular mechanics dubbed as 'deductive molecular mechanics' (DMM) [20,21] was constructed in our previous works. In [18] it was shown that DMM is successful for organic molecules.

In present work we investigate a possibility of using DMM for molecular modeling of more complex system: hydrophobic-hydrophilic ionic liquids (HHILs) and their binary mixtures with water. HHILs are the new class of room temperature ionic liquids [25]. Binary biphasic systems containing HHILs and water are perspective for the heavy metal extraction because of HHILs' unusual property – they have high content of water being almost insoluble in it at the same time. Specifically, we address the HHIL *tetra*-octylammonium lauroylsarcosinate (TOALS) [25] in the framework of Monte-Carlo approach using the APSLG based force field with parameters having clear QC source. It is shown that such an approach gives acceptable results in molecular modeling of such complex system as a ionic liquid and its mixture with water. All presented results were obtained with use of our library CARTESIUS FORT implementing the GF concept and containing tools for modeling complex system of different composition and structure.

2. Is there nanostructural organization in ionic liquids? – APSLG based study

2.1. Force field

Force fields (FF) based on the assumption of transferability and additivity of various intra- and intermolecular interactions are widely used for modeling of liquids. Such an assumption allows one to represent the potential energy surface (PES) as a combination of some additive potentials with empirical parameters:

$$E = E^{\text{bond}} + E^{\text{bend}} + E^{\text{torsion}} + E^{\text{Coulomb}} + E^{\text{vdW}}$$

First three terms correspond to intramolecular interactions of bonded atoms (elongation of valence bonds, deformation of valence and torsion angles, respectively); the last two correspond to interactions of nonbonded atoms (electrostatic and van der Waals interactions). Each energy contribution depends on atomic coordinates and may have different functional forms. The quality of molecular simulations substantially depends on the functional form of potentials and values of parameters in force field.

Being successfully used for modeling of huge number of molecular systems (specifically, for ionic liquids [26–35]), these molecular mechanics FFs do not have their own charge distributions rather use interpolated ones obtained through training against QC calculations. Substantial advantage of DMM force field used in present study is that it has its own charge distribution and at the same time requires much less computational resources than any QC method.

It should be noted that additional terms in the total energy appear because of the mutual influence of the atomic charge distributions (known as van der Waals interactions). In principle they can be obtained in APSLG by using perturbation theory [15], but it is beyond the scope of this paper. Thus we take the van der Waals interactions into account by applying Lennard-Jones potential in its classical (6–12) form with parameters given in [27]. In conclusion, the form of the force field used in present work is:

$$E = E_{\text{total}}^{\text{SLG}} + E^{\text{VDW}}$$

(for complete expression for $E_{\text{total}}^{\text{SLG}}$ see Appendix B) with MINDO/3 [17] parametrization used for APSLG part. The single molecule

was calculated by APSLG method to obtain explicit values of the electronic structure parameters and to optimize geometry. Then energy structure parameters were taken fixed while MC calculations was going, so only changes of atomic coordinates were applied in the thermodynamics calculation. Monte-Carlo simulations in isobaric-isothermal (NPT) ensemble were used to compute physical properties of pure TOALS. The Gibbs ensemble Monte Carlo (GEMC) method [36] was used for molecular modeling of TOALS/water binary system and calculation of mutual solubilities.

2.2. Simulation details

2.2.1. Simulations of pure TOALS

Monte-Carlo simulations in isobaric-isothermal (NPT) ensemble with Metropolis acceptance rules [37] were used to compute physical properties of pure TOALS. Simulations in NPT ensemble with periodic boundary conditions were performed for liquid TOALS at 1 atm and 25 °C. Several models containing 16, 32 and 64 ionic pairs have been studied. Initial value of density was assumed to be equal to 1 g/cm³ for all systems. Initial coordinates and orientations of ions were taken from the preliminary MC simulations of the box without boundary conditions and volume changes. Initial positions for such preliminary simulations were chosen randomly. Ewald algorithm [38] have been used to compute long-range charge-charge part of the electrostatic potential. Cutoff distances of 15 Å were applied for all non-hydrogen atoms. Volume changes were tried with the probability of 1% and their maximum changes as well as the maximum changes for the particle translations and rotations were adjusted to yield acceptance rates for the new configurations of approximately 35–50%. The system was assumed to be equilibrated when average values of energy and volume were fluctuating around constant value (no further monotonous change). The length of equilibration period was approximately 5–15 millions of steps and producing period consisted of 10 millions of steps. Density ρ was calculated by:

$$\rho = \frac{MN}{0.6022\langle V \rangle}$$

where M – molecular weight of TOALS, N – number of ionic pairs in the MC box, $\langle V \rangle$ – evaluated volume of the MC box.

It was shown in [39] that vapor of ionic liquids consists of ionic pairs, thus in order to obtain the enthalpy of vaporization we performed gas-phase MC simulations for the ionic pair of TOALS as well. Enthalpy of vaporization were obtained then as follows:

$$\Delta H_{vap} = -\frac{\langle E \rangle_{liquid}}{N} + RT + \langle E \rangle_{gas}$$

where $\langle E \rangle$ – liquid-phase energy obtained from MC calculations and $\langle E \rangle_{gas}$ – gas-phase energy of the ionic pair.

2.2.2. GEMC simulations of TOALS/water binary system

GEMC method is usually employed to simulate phase co-existence properties for thermodynamically equilibrated bulky phases [36]. In the present study NPT-GEMC method was used to obtain mutual solubilities for binary system TOALS/water. GEMC calculations were performed for two regions in each phase with their own periodic boundary conditions (that is away from the interphase boundary) in parallel way at 25 °C and 1 atm. Initial system contains two cubic MC boxes with pure TOALS (16 ionic pairs) and water (512 molecules) respectively. To provide more efficient convergency initial coordinates and orientations of particles were taken equal to those computed by the above NPT MC calculations for pure liquids. Systems with several initial values of densities (when density was not equal to equilibrated value, initial

coordinates were linear scaled) were under study to demonstrate that obtained results are independent on those values.

To provide the equilibrium of two phases three types of MC steps were implemented: particles displacements and rotations in each subsystem (temperature equilibration); volume changes for each subsystem (pressure equilibration); transfer of particles between two subsystems (chemical potential equilibration). The ratio of different steps was as follows: 57% probability for particles' displacements and rotations, 3% probability for volume changes and 40% probability for particles transfer. Receiving and supplying subsystems were chosen randomly with 50% probability. The amplitudes for coordinates, angles and volume changes were adjusted for each subsystem to provide 35–50% rate of new configuration acceptance. The particle for the transfer move was selected randomly with some special conditions for ions. It was allowed to transfer only ionic pair (cation + anion), but not a single ion. It means that when ion was randomly chosen to be transferred, cation from the same phase is chosen to be transferred too. There is no total guarantee that this assumption about concert transfer of ionic pair is reliable. But it looks reasonable to maintain electroneutrality of bulky phases and also allows a clear interpretation of solubility. For transfers of water molecules multiple trial insertions technique proposed in [40] was used to increase efficiency of calculations. Unfortunately, this technique was not fairly efficient for transfers of the TOALS ionic pairs because of their structure. TOALS ions contains long alkyl chains, thus it is almost impossible to find acceptable “voids” for them in the water phase acting “randomly”. To solve this problem the following moves have been performed. Ionic pair was transferred into random positions of the box first. Then the positions of the ions were equilibrated by MC simulations while the positions of other molecules were fixed to find “low-energy” positions of ionic pair (in analogous with “ghost” particle method). These “low-energy” positions were then used for possible insertions in the multiple insertion technique. As in the case of pure components the system was assumed to be equilibrated when the average energy and number of molecules in both subsystems were fluctuating around a constant value (no monotonous variation). An equilibration period required 15–20 millions MC steps and a production period usually consist of 15 millions cycles.

Mutual solubilities in GEMC have been calculated as follows:

$$x_1 = \frac{\langle N_1 \rangle_2}{\langle N_1 \rangle_2 + \langle N_2 \rangle_2}$$

$$x_2 = \frac{\langle N_2 \rangle_1}{\langle N_1 \rangle_1 + \langle N_2 \rangle_1}$$

where $\langle N_i \rangle_j$ is the average number of molecules of the i -th component in the j -th phase.

2.3. Results and discussion

2.3.1. Energy structure of TOALS molecule according to SLG calculations

Parameters of energy structure: geminal amplitudes and HO coefficients were obtained for single molecule of TOALS. Optimized geometry of TOALS is presented in Fig. 1.

2.3.2. Properties of pure TOALS

To validate force field and simulation package, standard NPT MC calculations of pure TOALS have been performed. In the ideal situation a series of physical properties which could be used to assess the simulations' quality (for example density, enthalpy of vaporization and heat capacity) should be obtained from these calculations and compared with experimental values. However, in case



Fig. 1. Geometries of TOALS anion and cation.

of HHILs most of the mentioned experimental quantities are not available for pure HHILs. The reason is that these species do not exist in the pure form and usually contain a large amount of water [25] (hydrophilic!). In the original work [25] only the densities of pure HHILs could be experimentally measured. On the other hand in the literature on molecular modeling of liquids density is considered to be a reliable quantity to validate force field. Calculated physical properties and available experimental values are presented in Table 1. Obtained value of density is in a good agreement with experiment.

Calculated value of enthalpy of vaporization for TOALS is much higher than those experimental values for imidazolium based ILs (30–40 kcal/mol). In Ref. [41], enthalpies of vaporization of tetraalkylphosphonium borates with long alkyl chains were obtained. It was shown that for ILs with such bulky cations enthalpies of vaporization should be substantially higher (50–70 kcal/mol) than the respective values for imidazolium ILs. According to such trend, TOALS should have even higher vaporization enthalpy than ILs [41]. It should be mentioned, however, that the partial charges in the FF used in Ref. [41] have no clear physical meaning: carbon atoms in alkyl chains have counterintuitive positive and hydrogen atoms negative charges. In our FF partial charges of carbons in alkyl chains are all negative, and this alone may assure the correct enthalpy trend as compared to Ref. [41].

2.3.3. Mutual solubilities

Results of GEMC calculations are presented in Table 2.

Analyze of results shows that GEMC-APSLG calculations allow to obtain phase coexisting properties in quite good agreement with experimental data, so the quality of APSLG based force field is good enough.

2.3.4. Nanostructural organization in pure TOALS and in water rich TOALS phase

Pure TOALS. Recently quite few papers have been devoted to structure of pure liquid ILs. The hypothesis of some kind of nanostructural organization of ILs to be responsible for unusual properties of ILs [42–47] was around. To analyze structure of pure TOALS we calculated several radial distribution functions (RDFs – see Fig. 2). The most interesting feature of these RDFs is that they have nearly discrete appearance: set of tight peaks formed by several closely disposed discrete bands. Such view of RDFs is the result of bulky structure and topology of ions: both anions and cations if ILs under study are substantially asymmetric which causes the formation of quite tough 'network' where small displacements

of ions are almost not allowed because they could greatly increase total energy of the system (repulsion of alkyl chains). In all works devoted to structural analysis of ionic liquids at least one of ions (cation or anion) is small and more symmetric particle, thus TOALS demonstrate such unusual properties in comparison with other ILs. Anyway several usual (for all ILs) important features may be observed in presented RDFs: first solvation shell (at 7 Å for cation-cation, 4.5 Å for anion-anion and 3.8 Å for cation-anion) and second solvation shell (at 12 Å for cation-cation, 11.5 Å for anion-anion).

Water in TOALS phase. Recently, the ILs/water binary mixtures and their unusual properties become a topic of lively discussions not only because of their practical importance, but also because of quite unusual behavior of those mixtures. Hypothesis about water nano-clustering in those mixtures has been proposed to explain their observed properties. Also numerical studies devoted to structure of liquid ILs/water mixtures based on molecular modeling appeared [47–49]. All of them use various radial distribution functions for deriving conclusions about formation of such clusters. For this end a geometrical criterion of 'connectivity' is used [47]. This approach is valid only when water molecules form clearly distinguishable clusters with significant intercluster separations well above the characteristic sizes of the clusters themselves and there is no 'free' molecules (which do not belong to any cluster) in the phase. When those conditions are not satisfied, some 'wrong' clusters may appear. In order to cope with this uncertainty we applied well established methods of 'cluster analysis' [50] which proposes rigorous conditions of clustering to our molecular assemblies.

Nevertheless, we analyze first some basic RDFs produced for water rich TOALS phase (presented in Fig. 3). These RDFs look quite similar (water-water and cation-water especially) to those reported in Ref. [47]. Characteristic maximum at 3 Å of water-water RDF is an argument in favor of clustering, but it does not guarantee it. Unfortunately, cation-water and anion-water RDFs are much less informative: they do not disprove clustering and do not prove it. All we can say is that water molecules prefer to locate closer to anions than cations, which can be explained by formation of hydrogen bond between hydrogen atom of water and oxygen carbonyl atom of the anion (see Fig. 4).

For further investigation of water clustering we applied cluster analysis. It allows to sort a set of points (positions of water molecule in TOALS phase) into groups (clusters) which ideally are located more closely to each other than to other points. Technically, the modified k-means algorithm [51] was used to separate water molecules in MC box into clusters. Within this algorithm, first, the optimal number of clusters had been obtained by finding clusters distributed normally. After that standard k-means algorithm was used. Such analysis gives an opportunity to find optimal number of clusters, average number of molecules per cluster, average size of clusters. The 'degree' of clustering is quantified by the following parameter:

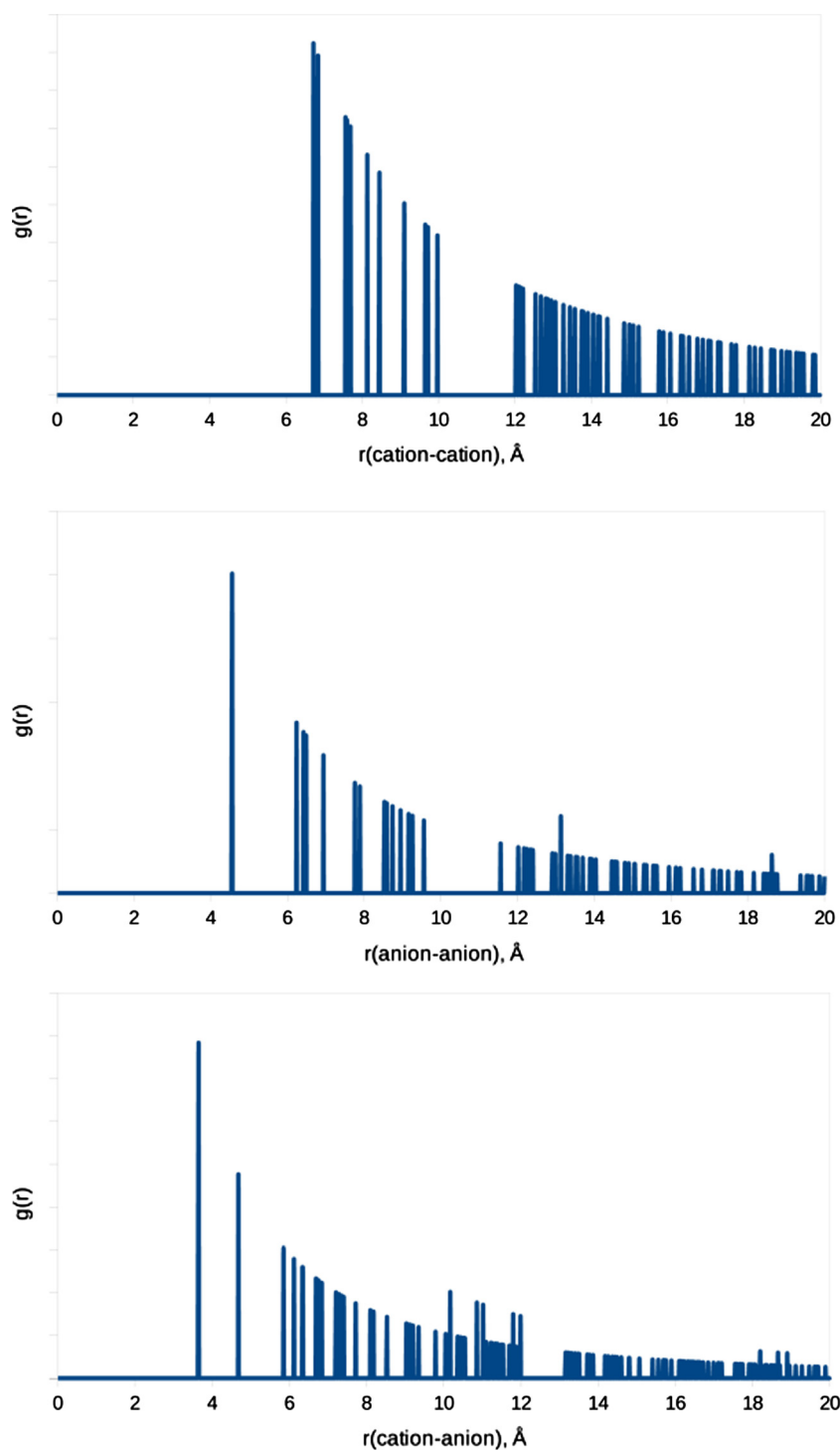
Table 1
Physical properties of pure TOALS: calculated and experimental values.

	ΔH_{vap} , kcal/mol	ρ , g/cm ³
Calc.	104.62 ± 0.02	0.941 ± 0.002
Exp. [25]	–	0.945

Table 2

Densities of TOALS and water phases and mutual solubilities obtained from GEMC calculations.

	$\rho_{\text{water phase}}, \text{g/cm}^3$	$\rho_{\text{TOALS phase}}, \text{g/cm}^3$	Molar fraction of TOALS in water phase, %	Molar fraction of water in TOALS phase, %
Calc.	0.992 ± 0.002	0.958 ± 0.002	$(2.3 \pm 1.0) \cdot 10^{-5}$	0.81 ± 0.03
Exp. [25]	–	–	$(1.1 \pm 0.5) \cdot 10^{-5}$	0.86

**Fig. 2.** Radial distribution functions obtained for pure TOALS. The first plot corresponds to the distance between quaternary nitrogen atoms of cations, the second one – to the distance between quaternary nitrogen atoms of anions and the last one – between those nitrogen atoms in cation and anion.

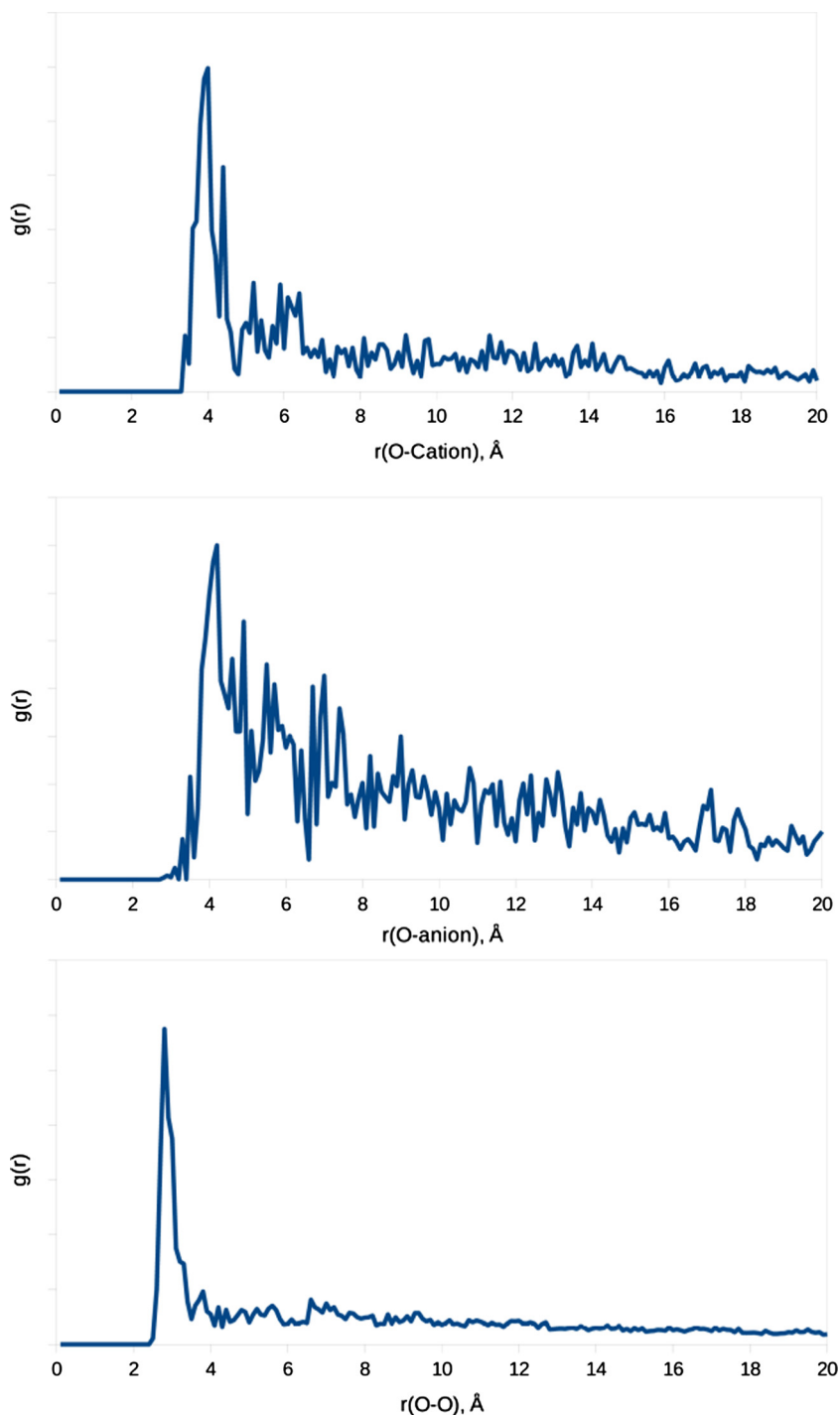


Fig. 3. Radial distribution functions obtained for water rich TOALS phase. The first plot corresponds to the distance between quaternary nitrogen atom of cation and oxygen atom of water, the second one – to the distance between quaternary nitrogen atom of anion and oxygen atom of water, the last one – between oxygen atoms of water molecules.

$$\zeta = \sum_k \sum_{i \in k} (x_i - x_k)^2$$

where x_k – coordinates of clusters' centers of mass, x_i – particle's coordinates (in our case of water oxygen atoms), which must be smaller in a clustered situation.

When applied to molecular assemblies, this analysis should be performed for each configuration within the producing period of the Monte-Carlo procedure. Then all parameters (ζ , size of clusters, number of molecules per cluster) should be averaged. Described

cluster analysis was performed for pure water and water in TOALS phase to compare parameters of clustering. Results are presented in Table 3. First of all we conclude that dividing into clusters gives much higher value of the ζ -parameter for pure water, than for water molecules in TOALS which indicates an explicit tendency for clustering in the latter case. Moreover, the average size (diameter) of clusters in pure water is larger than the analogous value for water in TOALS although the clusters in pure water contain less molecules than in TOALS/water mixture. Fig. 5 showing oxygen atoms of clusters in pure water and in TOALS/water mixture (for

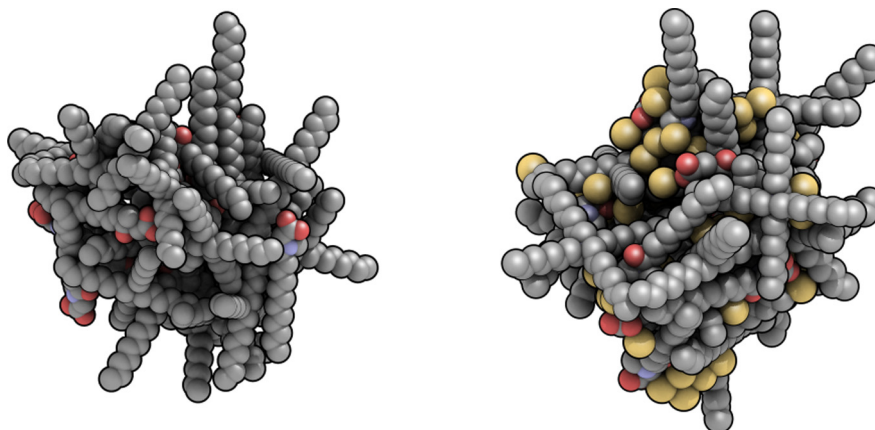


Fig. 4. Snapshots of pure TOALS (on the left) and water rich TOALS phase (on the right). Color coding: gray – C; blue – N; red – O (in TOALS); yellow – O (in water). Hydrogen atoms are not shown. Pictures were produced by QuteMol [52]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

Averaged parameters of clusters for pure water and water in TOALS obtained from k-means analysis.

	Pure water	Water in TOALS/water mixture
ξ	3.42	1.64
Number of molecules in cluster	5–6	9–10
Size of cluster, nm	1.15	1.01

more clear representation the ions of TOALS are not presented there) give quite a spectacular presentation of this result. The presence of well-distinguishable clusters of water molecules in TOALS phase is clear. At the same time in pure water there are no such clusters: the k-means algorithm tries to divide them as much as possible, but not successfully (this fact is reflected in a larger value of ξ), as one can see: the obtained clusters in pure water rather strongly overlap. It is particularly easy to see in 3D *.mol files presenting the results of our cluster analysis in Supporting Information (see Fig. 6).

Another important result is that there are two types of water molecules in TOALS/water mixtures: free water (one to three molecules per cluster – 15–18%) and well clustered water (82–85%). Free water was not taken into account while average size of clusters and number of molecules per cluster were calculated. In pure

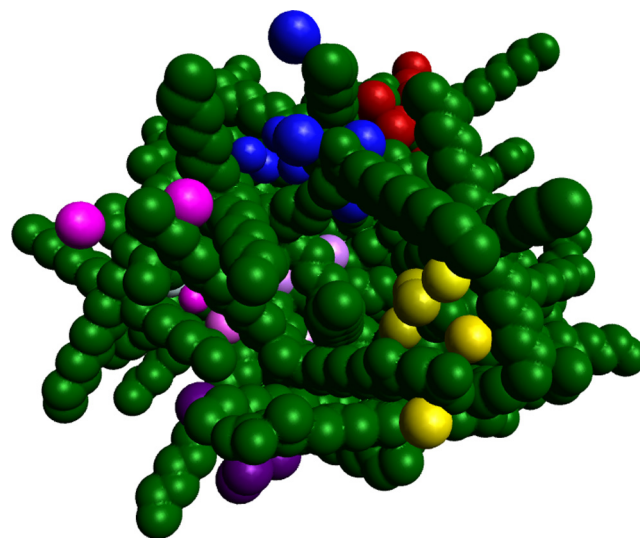


Fig. 6. The water clusters in water rich TOALS phase are marked by different colors. Green color corresponds to the atoms in the TOALS ions. One can see why removing TOALS completely was profitable for visually establishing the water clusterization: major part of water molecules is screened from view by the atoms belonging to TOALS. On the other hand it also can be seen that different water clusters are effectively separated from each other by the alkyl “rods”. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

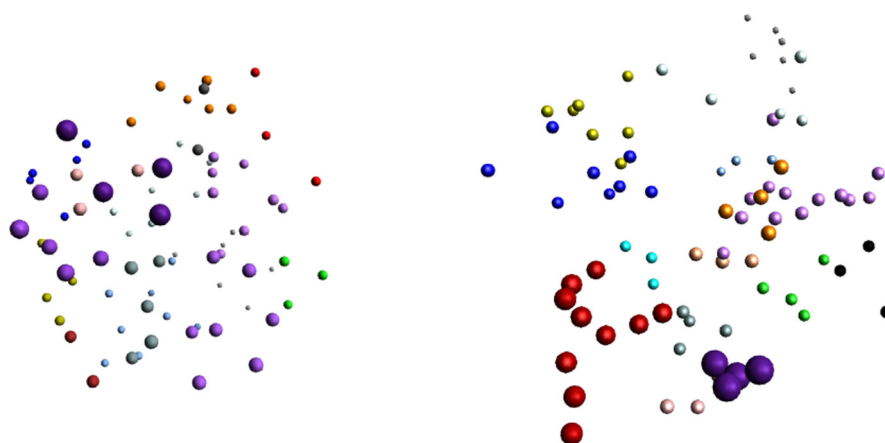


Fig. 5. Clusters in pure water (on the left) and in TOALS phase (on the right) as derived by k-means analysis from our MC data. Different colors mark the different clusters. For TOALS phase ions are not presented.

water molecules distributed evenly on clusters and there are no such 'free' molecules. Regarding the average size of the clusters existing (to the validity of the used criterium) of 1,01 nm and their maximal size of 1,6 nm we can conclude that these water clusters in TOALS occupy an intermediate position between molecular-particles and nano-particles.

3. Conclusions

In the present paper we describe recent applications of semiempirical group functions theory. Principal possibility of using APSLG based force field for modeling complex binary systems like TOALS/water is demonstrated. Deductive molecular mechanics force field is validated against several physical properties of pure TOALS calculated by NPT MC simulations with use of the said FF. Obtained value of density is in a good agreement with experiment. Theoretical explanation of abnormally high solubility of water in TOALS based on structural analysis is given. It is shown that in pure TOALS phase bulky alkyl chains form structural voids. The sizes of such voids are large enough to accommodate nano-clusters of water molecules when TOALS are mixed with water.

Acknowledgments

The authors thank Moscow Supercomputer Center RAS for the use of supercomputer facilities and Prof. Dr. Igor V. Pletnev (Division of Analytical Chemistry, Chemistry Department, MSU) for discussions and valuable comments. Mr. T.S. Kushnir is gratefully acknowledged for his kind assistance with the parallelizing the FORTRAN code. The authors are greatly thankful to the Referee for his instructive and benevolent comments which helped to significantly improve the manuscript. The financial support by the Russian Science Foundation (RSF) through the grant No. 17-73-30038 is acknowledged.

Appendix A. General ideas of group functions approximation

According to the group functions approximation total electronic wave function (WF) in the problem of separation of variables in quantum chemistry has the general form:

$$\Psi(x_1 \dots x_N) = \hat{A} \left\{ \Phi^A(x_1 \dots x_{N_A}) \cdot \Phi^B(x_{N_A+1} \dots x_{N_B}) \cdot \dots \right\}$$

$$N = N_A + N_B + \dots$$

where $A, B, \dots$ are indices of subsystems which the original system is divided into, Φ^K are their electronic wave functions and N_K are numbers of electrons at these subsystems. $\hat{A}$ is projection operator to some antisymmetric irreducible representation of the symmetric group S_N .

Hamiltonian operator for original system can be written as:

$$\hat{H} = \hat{H}^0 + \hat{V} = \sum_K \hat{H}_K + \sum_{K>L} \sum \hat{V}_{KL}$$

where $\hat{H}_K$ is hamiltonian operator of subsystem K and $\hat{V}_{KL}$ is operator of interaction between subsystems K and L . Obviously, the GF Ψ is the eigenvector of hamiltonian $\hat{H}^0$ with eigenvalue E :

$$E = E_A + E_B + \dots$$

If we introduce according to the second quantization formalism number of electron operators $\hat{N}_K$ for subsystem K , group function Ψ will be also its eigenvector with eigenvalue N_K . It means that

$$\forall K [\hat{H}^0, \hat{N}_K] = 0$$

We can now formulate the conditions when group function Ψ approximate real WF in quite good way:

$$\text{if } \exists \{\hat{N}_K\} \forall K [\hat{H}, \hat{N}_K] \rightarrow 0$$

which can be formulated as: GF is a good approximation for real WF when the numbers of electrons in each subsystem are 'good' quantum numbers. If the condition of subsystems' WFs strong orthogonality is imposed, it can be said that the total space of one electronic states (OES) corresponding to the original system is dividing into mutually orthogonal subspaces with fixed constant dimensions corresponding to subsystems. It means that each function Φ^K can be constructed from OES corresponding *only* to the K -th subspace.

More detailed and complete information about GF concept can be found in [9,15].

Appendix B. Strictly localized geminals and deductive molecular mechanics

In this section we are going to discuss some general ideas of SLG based molecular mechanics with mainly focusing on nonbonded interactions.

Total electronic WF of organic molecules can be written in terms of geminals (g_m) – electronic groups corresponding to singlet state of two electrons in the molecule (chemical bond or lone electron pair) as follows:

$$|\Psi\rangle = \prod_m g_m^+ |0\rangle$$

where g_m^+ is the creation operator of two electrons in geminal g_m corresponding to the m -th bond in the molecule. It is reasonable to represent such defined geminals in valence bond view:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m (r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+)$$

where u_m, v_m – amplitudes of ionic contributions; w_m – amplitude of covalent contributions; $r_{m\sigma}^+, l_{m\sigma}^+$ are Fermi-operators of creation electron in the one electronic states (OES) $|r_m\rangle$ and $|l_m\rangle$ with spin projection σ . Geminals are normalized:

$$\langle 0 | g_m g_m^+ | 0 \rangle = u_m^2 + v_m^2 + 2w_m^2 = 1$$

For the lone electron pair only one contribution remains:

$$g_m^+ = r_{m\alpha}^+ r_{m\beta}^+$$

One electronic states $|r_m\rangle$ and $|l_m\rangle$ are strictly local and reside on right and left atoms of m -th bond, respectively. They are the hybrid orbitals (HOs) oriented along this bond. System of such hybrid orbitals on each non-hydrogen atom can be obtained from the basis set of four atomic orbitals (A) (for second row elements) by using orthogonal linear transformation $\hat{h}$ corresponding to SO(4) Lee group:

$$|t_{m\sigma}\rangle = \sum_i h_{mi}(A) |a_{i\sigma}\rangle$$

$h_{ij}(A)$ – 4×4 matrix representing operator $\hat{h}$ in linear space with atomic orbitals basis set A . This space is isomorphic to four-dimensional space of quaternions $\mathbb{H}$, so each HO corresponds to normalized quaternion ($s_m, \vec{v}_m$). Scalar part of quaternion gives the contribution of atomic s-orbital and vector part is contributions of atomic p-orbitals to the hybrid orbital [20].

The independent amplitudes (u_m, v_m) and matrix elements (h_{ij}) form the full set of electronic structure parameters. Solving the 3×3 problems with effective hamiltonian for all geminals self-consistently allows to obtain explicit values of these parameters (see more details in [16]).

It is important that geminal amplitudes enter expressions of any observables only in some combinations – matrix elements of one- and two-electronic densities, defined as [18]:

$$P_m^{rr} = u_m^2 + w_m^2, P_m^{ll} = v_m^2 + w_m^2$$

$$P_m^{rl} = P_m^{lr} = (u_m + v_m)w_m$$

for one-electronic density in m -th geminal.

$$\Gamma_m^{rr} = u_m^2; \Gamma_m^{ll} = v_m^2; \Gamma_m^{rl} = \Gamma_m^{lr} = w_m^2$$

for two-electronic density in m -th geminal. Matrix elements of cross geminal two-electronic density are the products of corresponding one-electronic densities matrix elements [18].

Total energy of the system consisting of N identical molecules can be written in terms of density matrix elements in form, which is fairly similar to MM force field, because bonded and nonbonded atoms are treated in different way [18]:

$$E_{total}^{SLG} = \sum_{M=1}^N \left[\sum_{A \in M} E_{bond}^A + \sum_{A, B \in M} E_{bond}^{AB} \right] + \sum_{M=1}^N \sum_{K=M+1}^N \sum_{A \in M} \sum_{B \in K} E_{nonbond}^{AB}$$

$$E_{nonbond}^{AB} = E_{el}^{AB} + E_{exch}^{AB}$$

where indexes M and K are numbers corresponding to molecules and indexes A and B are numbers corresponding to atoms.

The exchange interactions between electrons in the groups sharing the same pair of atoms (this happens when multiple bonds are present) E_{exch}^{AB} can be written as follows [24]:

$$E_{exch}^{AB} = -2 \sum_{m < n} \left(p \in \{m, A\} \right) \left(p' \in \{n, A\} \right) \left(q \in \{m, B\} \right) \left(q' \in \{n, B\} \right) (pp'|qq')^{AB} \sum_{\sigma} P_{pq}^{\sigma} P_{p'q'}^{\sigma}$$

where m and n are numbers of bonds; p, q are atomic orbitals and $(pp'|qq')$ is two electronic two centered integral.

Using multipole expansion the electrostatic interaction E_{el}^{AB} between charge distributions corresponding to nonbonded atoms A and B can be written as [24]:

$$E_{el}^{AB} = Z^A Z^B (ss|ss)^{AB} + \sum_{pp' \in A} \sum_{qq' \in B} P_{pp'}^A P_{qq'}^B (pp'|qq')^{AB} - Z^A \sum_{qq' \in B} P_{qq'}^B (ss|qq')^{AB} - Z^B \sum_{pp' \in A} P_{pp'}^A (pp'|ss)^{AB}$$

Only the intragroup one-electronic density elements are nonvanishing, so this expression can be rewritten:

$$E_{el}^{AB} = [-Z^A (ss|A + \sum_{pp' \in A} P_{pp'}^A (pp'|A) \bullet [-Z^B |ss\rangle + \sum_{qq' \in B} P_{qq'}^B |qq'\rangle^B]$$

The strictly local character of hybrid atomic orbitals allowed to define atomic multipoles. Monopole (atomic charge) can be obtained as follows:

$$q^A = \sum_{p \in A} P_{pp}^A - Z^A$$

Expressions for dipole moment for an sp -distribution ($\vec{\mu}^A$) and quadrupole moment ($\widehat{S}^A$) for an pp -distribution on atom A :

$$\vec{\mu}^A = D_1^A \sum_{pq \in A} P_{pq}^A (s_p^A \vec{v}_q^A + s_q^A \vec{v}_p^A)$$

$$\widehat{S}^A = (D_2^A)^2 \sum_{pq \in A} P_{pq}^A \vec{v}_p^A \otimes \vec{v}_q^A$$

where

$$D_1 = \frac{2n+1}{\sqrt{3}} \frac{(4\zeta_{ns}\zeta_{np})^{n+\frac{1}{2}}}{(\zeta_{ns} + \zeta_{np})^{2n+2}}$$

$$D_2 = \sqrt{\frac{(2n+1)(2n+2)}{20}} \zeta_{np}^{-1}$$

In terms of such defined multipole moments, electrostatic interaction can be rewritten as:

$$E_{el}^{AB} = q^A G_{00} q^B + q^A G_{01}^{\alpha} \vec{\mu}_{\alpha}^B - \vec{\mu}_{\alpha}^A G_{10} q^B - \vec{\mu}_{\alpha}^A G_{11}^{\alpha\beta} \vec{\mu}_{\beta}^B + q^A G_{02}^{\alpha\beta} \widehat{S}_{\alpha\beta}^B + \widehat{S}_{\alpha\beta}^A G_{20}^{\alpha\beta} q^B - \vec{\mu}_{\alpha}^A G_{12}^{\alpha\beta\gamma} \vec{\mu}_{\beta\gamma}^B + \widehat{S}_{\alpha\beta}^A G_{21}^{\alpha\beta\gamma} \vec{\mu}_{\gamma}^B + \widehat{S}_{\alpha\beta}^A G_{22}^{\alpha\beta\gamma\delta} \widehat{S}_{\gamma\delta}^B$$

$$G_{l_1 l_2}^{\alpha\beta\gamma\dots v} = \nabla_{\alpha} \nabla_{\beta} \dots \nabla_v f_{l_1 l_2}(R^{AB})$$

where $f_{l_1 l_2}(R^{AB})$ is the functional form of the potential; R^{AB} is the distance between atoms A and B .

Expressions for bonded interactions are quite cumbersome to present it here and can be found in [18].

Although MM like functional form of energy is obtained, it cannot be used for molecular modeling, because density matrix elements should change in some nontrivial way each time atomic coordinates are changed. Thus the most important consequence from the SLG description which allows to construct deductive MM scheme is the transferability of density matrix elements [21]. It was shown that matrix elements can be written as follows:

$$P_m^{tt'} = \frac{1}{2} + \delta P_m^{tt'} \Gamma_m^{tt'} = \frac{1}{4} + \delta \Gamma_m^{tt'}$$

Small corrections $\delta P_m^{tt'}$ and $\delta \Gamma_m^{tt'}$ can be expressed in terms of small (for typical organic bonds) parameters: bond correlation ($\zeta_m \gg 1$) and bond polarization ($\mu_m \ll 1$) [21]:

$$\delta P_m^{tt} = \frac{\mu_m}{2} \delta P_m^{rl} = -\frac{\mu_m^2}{4} - \frac{1}{4\zeta_m^2}$$

$$\delta \Gamma_m^{tt} = \frac{\mu_m}{2} - \frac{1}{4\zeta_m} \quad \delta \Gamma_m^{rl} = -\frac{\mu_m^2}{2} + \frac{1}{4\zeta_m}$$

Thus the transferability of electron density matrix elements allows not to solve self-consistent problem every time when the form of HOs or atomic coordinates are changed. It means that the MM scheme with additive form of energy dependence on HOs and atomic coordinates may be implemented.

The expressions of bonded δE^{bond} and nonbonded $\delta E^{nonbond}$ energy changes in terms of small changes of HOs parameters and atomic coordinates can be easily obtained if the transferability of densities matrix elements is taken into account [21].

As a result we have the force field which is expressed in molecular mechanics way and contains parameters determined from theoretical QM calculations.

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.comptc.2017.04.001>.

References

- [1] R. Descartes, Discours de la Méthode, J'ai Lu, 2004 (original publication: Imprimerie Ian Meyre a Leyde, 1637).
- [2] O.N. Witt, Zur Kenntnis des Baues und der Bildung färbender Kohlenstoffverbindungen, Ber. Deut. Chem. Ges. 9 (1876) 522.
- [3] K. Ruedenberg, Rev. Mod. Phys. 34 (1962) 326.
- [4] A.L. Tchougréeff, Hybrid Methods of Molecular Modeling, first ed., Springer, 2008.
- [5] H.A. Bethe, Term aufspaltung in Krystallen, Ann. Physik. 3 (1929) 133–206.
- [6] E. Hückel, Quantentheoretische Beiträge zum Benzolproblem, Zeitschrift für Physik 70 (1931) 204–286.

- [7] R. Daudel, Quantum mechanical facets of chemical bond, in: B. Pullmann, R. Parr (Eds.), *A New World of Quantum Chemistry*, D. Reidel Publ. Comp., Dordrecht, 1976, pp. 33–56.
- [8] R. McWeeny, The valence bond theory of molecular structure. I. Orbital theories and the valence-bond method, *Proc. Roy. Soc. Lon. Ser. A* 253 (1954) 63–79.
- [9] A.L. Tchougréeff, Group functions, Löwdin partition, and hybrid QC/MM methods for large molecular systems, *Phys. Chem. Chem. Phys.* 1 (1999) 1051–1060.
- [10] A.L. Tchougréeff, A.M. Tokmachev, Group functions approach based on the combination of strictly local geminals and molecular orbitals, *Int. J. Quant. Chem.* 106 (2006) 571–587.
- [11] A.V. Soudackov, A.L. Tchougréeff, I.A. Misurkin, Crystal field splittings and optical spectra of transition metal – mixed ligand complexes by effective Hamiltonian method, *Int. J. Quant. Chem.* 57 (1996) 663–671.
- [12] A.M. Tokmachev, A.L. Tchougréeff, Semiempirical implementation of strictly localized geminals for analysis of molecular electronic structure, *J. Comput. Chem.* 22 (2001) 752–764.
- [13] A.V. Sinitsky, M.B. Darkhovskii, A.L. Tchougréeff, I.A. Misurkin, Effective crystal field for trivalent first transition row ions, *Int. J. Quant. Chem.* 88 (2002) 370–379.
- [14] A.L. Tchougréeff, A.V. Soudackov, J. van Leusen, P. Kögerler, K.-D. Becker, R. Dronskowski, Effective Hamiltonian crystal field: present status and applications to iron compounds, *Int. J. Quant. Chem.* 116 (2016) 282–294.
- [15] A.L. Tchougréeff, New generation of semiempirical methods based on group function approximation, *J. Struct. Chem.* 48 (2007) 39–62.
- [16] A.M. Tokmachev, A.L. Tchougréeff, Semiempirical implementation of strictly localized geminals for analysis of molecular electronic structure, *J. Comput. Chem.* 22 (2001) 752–764.
- [17] A.M. Tokmachev, A.L. Tchougréeff, Potential energy surfaces in hybrid QM/MM methods, *Int. J. Quant. Chem.* 84 (2001) 39–47.
- [18] A.M. Tokmachev, A.L. Tchougréeff, Generic molecular mechanics as based on local quantum description of molecular electronic structure, *Int. J. Quant. Chem.* 88 (2002) 403–413.
- [19] A.M. Tokmachev, A.L. Tchougréeff, Fast NDDO method for molecular structure calculations based on strictly localized geminals, *J. Phys. Chem. A* 107 (2003) 358–365.
- [20] A.L. Tchougréeff, SO(4) group and deductive molecular mechanics, *J. Mol. Struct. Theochem.* 630 (2003) 243–263.
- [21] A.M. Tokmachev, A.L. Tchougréeff, Transferability of parameters of strictly local geminals' wave function and possibility of sequential derivation of molecular mechanics, *J. Comput. Chem.* 26 (2005) 491–505.
- [22] A.L. Tchougréeff, Deriving a mechanistic model for potential energy surface of coordination compounds of nontransition elements, *Int. J. Quant. Chem.* 107 (2007) 2519–2538.
- [23] A.L. Tchougréeff, A.M. Tokmachev, I. Mayer, Towards a possible ab initio molecular mechanics. Transferability of density matrix elements, *Int. J. Quant. Chem.* 107 (2007) 2539–2555.
- [24] A.L. Tchougréeff, A.M. Tokmachev, R. Dronskowski, A multipole model for the electron group functions method, *J. Phys. Chem. A* 113 (2009) 11406–11415.
- [25] S.V. Smirnova, T.O. Samarina, I.V. Pletnev, Hydrophobic-hydrophilic ionic liquids for the extraction and determination of metal ions with water-soluble reagents, *Anal. Methods* 7 (2015) 9629–9635.
- [26] J.K. Shah, J.F. Brennecke, E. Maginn, Thermodynamic properties of the ionic liquid 1-*n*-butyl-3-methylimidazolium hexafluorophosphate from Monte Carlo simulations, *J. Green Chem.* 4 (2002) 112–118.
- [27] Z. Liu, S. Huang, W. Wang, A refined force field for molecular simulation of imidazolium-based ionic liquids, *J. Phys. Chem. B* 34 (2004) 12978–12989.
- [28] J.K. Shah, E. Maginn, Monte Carlo Simulations of gas solubility in the ionic liquid 1-*n*-butyl-3-methylimidazolium hexafluorophosphate, *J. Phys. Chem. B* 109 (2005) 10395–10405.
- [29] T.I. Morrow, E. Maginn, Molecular dynamics study of the ionic liquid 1-*n*-butyl-3-methylimidazolium hexafluorophosphate, *J. Phys. Chem. B* 106 (2002) 12807–12813.
- [30] S.V. Sambasivarao, O. Acevedo, Development of OPLS-AA force field parameters for 68 unique ionic liquids, *J. Chem. Theory Comput.* 5 (2009) 1038–1050.
- [31] M. Ramdin, Solubility of the precombustion gases CO₂, CH₄, CO, H₂, N₂, and H₂S in the Ionic Liquid [bmim][Tf₂N] from Monte Carlo Simulations, *J. Phys. Chem. C* 41 (2014) 23599–23604.
- [32] E.J. Maginn, Atomistic simulation of the thermodynamic and transport properties of ionic liquids, *Acc. Chem. Res.* 11 (2007) 1200–1207.
- [33] J. de Andrade, E.S. Boes, H. Stassen, Computational study of room temperature molten salts composed by 1-alkyl-3-methylimidazolium CationsForce-field proposal and validation, *J. Phys. Chem. B* 106 (2002) 13344–13351.
- [34] A.F. Ghobadi, V. Taghikhani, R.J. Elliott, Investigation on the solubility of SO₂ and CO₂ in imidazolium-based ionic liquids using NPT Monte Carlo Simulation, *J. Phys. Chem. B* 23 (2011) 13599–13607.
- [35] A. Chaumont, R. Schurhammer, G. Wipff, Aqueous interfaces with hydrophobic room-temperature ionic liquids: a molecular dynamics study, *J. Phys. Chem.* 40 (2005) 18964–18973.
- [36] A.Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.* 61 (1987) 813–826.
- [37] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equations of state calculations by fast computing machines, *J. Chem. Phys.* 21 (1953) 1087–1092.
- [38] P.P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, *Ann. Phys.* 369 (1921) 253–287.
- [39] J.P. Armstrong, C. Hurst, R.G. Jones, P. Licence, K.R.J. Lovelock, C.J. Satterley, I.J. Villar-Garcia, Vapourisation of ionic liquids, *Phys. Chem. Chem. Phys.* 9 (2007) 982–990.
- [40] J. Vorholz, V.I. Harismiadis, B. Rumpf, A.Z. Panagiotopoulos, G. Maurer, Vapor + liquid equilibrium of water, carbon dioxide, and the binary system, water + carbon dioxide, from molecular simulation, *Fluid Phase Equilib.* 170 (2000) 203–234.
- [41] Y.-L. Wang, F.U. Shah, S. Glavatskih, O.N. Antzutkin, A. Laaksonen, Atomistic insight into orthoborate-based ionic liquids: force field development and evaluation, *J. Phys. Chem. B* (118) (2014) 8711–8723.
- [42] N.A. José, C. Lopes, A.A.H. Pádua, Molecular force field for ionic liquids III: imidazolium, pyridinium, and phosphonium cations; chloride, bromide, and dicyanamide anions, *J. Phys. Chem. B* 110 (2006) 3330–3335.
- [43] S. Zahn, G. Bruns, J. Thar, B. Kirchner, What keeps ionic liquids in flow?, *Phys. Chem. Chem. Phys.* 10 (2008) 6921–6924.
- [44] S.M. Urahata, M.C.C. Ribeiro, Structure of ionic liquids of 1-alkyl-3-methylimidazolium cations: a systematic computer simulation study, *J. Chem. Phys.* 120 (2004) 1855–1863.
- [45] Y. Wang, G.A. Voth, Unique spatial heterogeneity in ionic liquids, *J. Am. Chem. Soc.* 127 (2005) 12192–12193.
- [46] J.J. Hettige, J.C. Araque, H.K. Kashyap, C.J. Margulis, Nanoscale structure of tetradecyltriethylphosphonium based ionic liquids, *J. Chem. Phys.* 144 (2016) 121102–121108.
- [47] Y.-L. Wang, S. Sarman, S. Glavatskih, O.N. Antzutkin, M.W. Rutland, A. Laaksonen, Atomistic insight into tetraalkylphosphonium-bis(oxalato)borate ionic liquid/water mixtures. I. Local microscopic structure, *J. Phys. Chem. B* 119 (2015) 5251–5264.
- [48] C. Hanke, R. Lynden-Bell, A simulation study of water-dialkylimidazolium ionic liquid mixtures, *J. Phys. Chem. B* (107) (2003) 10873–10878.
- [49] X. Zhong, Z. Fan, Z. Liu, D. Cao, Local structure evolution and its connection to thermodynamic and transport properties of 1-butyl-3-methylimidazolium tetrafluoroborate and water mixtures by molecular dynamics simulations, *J. Phys. Chem. B* (116) (2012) 3249–3263.
- [50] R.C. Tryon, *Cluster Analysis: Correlation Profile and Orthometric (factor) Analysis for the Isolation of Unities in Mind and Personality*, Edwards Brothers, 1939.
- [51] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: *Proc. 5th Berkeley Symp. on Math. Statistics and Probability*, 1967, pp. 281–297.
- [52] M. Tarini, P. Cignoni, C. Montani, Ambient occlusion and edge cueing to enhance real time molecular visualization, *IEEE Trans. Visual Comput. Graphics* 12 (2006) 1–8.