

3D Molecular Theory of Solvation for Nanochemistry in Solution

Andriy Kovalenko^{*,**}

^{*}National Institute for Nanotechnology

11421 Saskatchewan Dr., Edmonton, AB, T6G 2M9, Canada, andriy.kovalenko@nrc.ca

^{**}Department of Mechanical Engineering, University of Alberta, Edmonton, AB, T6G 2G8, Canada

ABSTRACT

Statistical-mechanical, molecular theory of solvation (a.k.a. 3D-RISM-KH) predicts from the first principles the solvation structure and thermodynamics of nanosystems and properly accounts for chemical functionalities by representing both electrostatic and non-polar features of solvation structure such as hydrogen bonding and solvophobicity, salt bridges, structural solvent, associative and electrochemical effects. We coupled 3D-RISM-KH with *ab initio* methods in a self-consistent field description (including analytical gradients) of electronic structure, optimized geometry, and chemical reactions in solution, and have extensively validated the KS-DFT/3D-RISM-KH multiscale method against experimental data for solvation thermochemistry, conformational equilibria, tautomerization energies and activation barriers in different solvents. The method explains the electronic and solvation structure of ionic liquids; mechanisms of self-assembly, conformational stability and solvent-driven supramolecular chirality of synthetic organic rosette nanotubes; and mechanisms of supercapacitance in nanoporous carbon electrodes.

Keywords: 3D molecular theory of solvation, quantum chemistry, solution, ionic liquids, organic rosette nanotubes, nanoporous carbon, supercapacitor

1 INTRODUCTION

Nanoscale objects and phenomena are different from the properties of the atomic constituents as well as from those described by the macroscopic laws of continuous media. The behavior of nanostructures can change in a wide range, which constitutes the promise of nanotechnology on control over materials. Predictive modeling of nanosystems should operate at length scale from one to hundreds nanometers and time scale to microseconds and more, and yet derive their properties from the chemical functionalities of the constituents, thus requiring multiscale treatment. A number of important applications involve processes in solution. Their molecular modeling means long-time description of millions molecules, which is by far not feasible with *ab initio* methods, challenging for molecular simulations, and on the other hand, problematic for continuum theories which are phenomenological and thus non-transferable.

A genuine challenge of multiscale modeling is theoretical coupling of methods at different scales, so that “observables” at lower-level scales are analytically linked

to “force fields” of more coarse-grained models at higher-level scales. Statistical mechanics itself is in a sense an example of such theoretical coupling between microscopic variables and thermodynamic, macroscopic properties. The statistical-mechanical, molecular theory of solvation, a.k.a. three-dimensional reference interaction site model, starts from an explicit solvent model with all-atom force field but operates with 3D correlation functions of species in a statistical ensemble rather than with individual molecular trajectories and predicts the solvation structure and thermodynamics from the first principles of statistical mechanics [1]. Complemented with the Kovalenko-Hirata (KH) closure approximation [1], the 3D-RISM-KH theory properly accounts for chemical functionalities of both solute and solvent species by representing both electrostatic and non-polar features of the solvation structure, such as hydrogen bonding, hydrophobicity, salt bridges, structural solvent, associative and electrochemical effects, etc [2-12]. Moreover, the 3D-RISM-KH theory analytically yields the solvation thermodynamics in terms of the 3D correlation functions, including the solvation free energy, its energetic and entropic decomposition, and partial molar volume effects [1-12]. It also readily allows decomposition of any solvation thermodynamic property into partial contributions of solute sites, providing a basis for spatial decomposition analysis (SDA) of association effects in solution [13].

We coupled the 3D-RISM-KH theory with *ab initio* methods (KS-DFT and CASSCF) in a self-consistent field description (including analytical gradients) of electronic structure, optimized geometry, chemical reactions and nanocatalysis in solution [1,2]. We have implemented the KS-DFT/3D-RISM-KH multiscale method in the ADF computational chemistry package [2] and extensively validated it against experimental data for solvation thermochemistry, conformational equilibria, and activation barriers for various nanosystems in different solvents [2-5].

A generalization of molecular theory of solvation to disordered nanoporous materials, the so-called replica RISM-KH theory, enables predictive molecular modeling of electrolyte solutions in (functionalized) nanoporous materials, and reveals the mechanisms of sorption and supercapacitance in nanoporous carbon electrodes[1,14-16].

Below we briefly summarize the 3D molecular theory of solvation coupled with quantum chemistry and illustrate how it predicts the electronic and solvation structure of a carbon nanotube [3] and ionic liquid [5], and reveals the mechanisms of self-assembly, conformational stability and possible functions of organic rosette nanotubes [6-9].

2 THREE-DIMENSIONAL MOLECULAR THEORY OF SOLVATION

The solvation structure is represented by the probability density of finding site γ of solvent molecules at 3D space position \mathbf{r} around the solute (supra)molecule, $\rho_\gamma(\mathbf{r}) = \rho_\gamma g_\gamma(\mathbf{r})$, which is determined by the average number density ρ_γ in the solution bulk times the 3D distribution function (normalized density distribution) $g_\gamma(\mathbf{r})$ of solvent site γ . The latter indicates site density enhancement when $g_\gamma(\mathbf{r}) > 1$ or depletion when $g_\gamma(\mathbf{r}) < 1$ relative to the average density at a distance from the solute in the solution bulk where $g_\gamma \rightarrow 1$.

The 3D distribution functions of solvent interaction sites are obtained from the 3D-RISM integral equation [1]

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}' c_\alpha(\mathbf{r} - \mathbf{r}') \chi_{\alpha\gamma}(\mathbf{r}'), \quad (1)$$

where $h_\gamma(\mathbf{r})$ is the 3D total correlation function of solvent site γ related to the 3D site distribution function by $g_\gamma(\mathbf{r}) = h_\gamma(\mathbf{r}) + 1$, and $c_\gamma(\mathbf{r})$ is the 3D direct correlation function which has the asymptotics of the solute-solvent site interaction potential, $c_\gamma(\mathbf{r}) \sim -u_\gamma(\mathbf{r}) / (k_B T)$; the site-site susceptibility of pure solvent $\chi_{\alpha\gamma}(r)$ is an input to the 3D-RISM theory; and indices α and γ enumerate all sites on all sorts of solvent species. Another relation between the 3D total and direct correlation functions, called a closure, is necessary to complement the 3D-RISM integral equation (1). The exact closure can be formally expressed as a series in terms of multiple integrals of the combinations of the total correlation functions. However, it is extremely cumbersome, and in practice is replaced with amenable approximations. We use the 3D-KH closure approximation, proven to be appropriate to describe various association effects in complex liquids and electrolyte solutions [1-5], in supramolecular synthetic organic [6-9,13] and biomolecular [10-12] systems in solution,

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp(d_\gamma(\mathbf{r})) & \text{for } d_\gamma(\mathbf{r}) \leq 0 \\ 1 + d_\gamma(\mathbf{r}) & \text{for } d_\gamma(\mathbf{r}) > 0 \end{cases}, \quad (2)$$

$$d_\gamma(\mathbf{r}) = -\frac{u_\gamma(\mathbf{r})}{k_B T} + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}),$$

where $u_\gamma(\mathbf{r})$ is the 3D interaction potential between the whole solute and solvent site γ specified by the molecular force field, and $k_B T$ is the Boltzmann constant times the solution temperature. The closure (3) couples in a nontrivial way the so-called mean spherical approximation (MSA) and hypernetted chain (HNC) approximation, the former being applied to the spatial regions of solvent density enrichment ($g_\gamma(\mathbf{r}) > 1$) such as association peaks, and the latter to those of solvent density depletion ($g_\gamma(\mathbf{r}) < 1$) including the repulsive core [1].

The site-site susceptibility of solvent breaks up into the intra- and intermolecular terms,

$$\chi_{\alpha\gamma}(r) = \omega_{\alpha\gamma}(r) + \rho_\alpha h_{\alpha\gamma}(r), \quad (3)$$

where the intramolecular correlation function

$$\omega_{\alpha\gamma}(r) = \delta_{\alpha\gamma} \delta(r) + (1 - \delta_{\alpha\gamma}) \delta(r - l_{\alpha\gamma}) / (4\pi l_{\alpha\gamma}^2)$$

represents the geometry of solvent molecules with site-site separations $l_{\alpha\gamma}$ specified by the molecular force field (z-matrix in quantum chemistry), and $h_{\alpha\gamma}(r)$ is the radial total correlation function between sites α and γ enumerating all sites on all sorts of molecules in bulk solvent. In advance to the 3D-RISM-KH calculation, $h_{\alpha\gamma}(r)$ are obtained from the dielectrically consistent RISM theory (see references in [1]) coupled with the KH closure (DRISM-KH), applied to the bulk solvent with counter ions, co-solvent, and ligands at a given concentration. The susceptibility (3) of bulk solvent is then input into the 3D-RISM integral equation (1).

The solvation free energy of the solute supramolecule in multicomponent solvent following from the 3D-RISM-KH integral equations (1) and (2) is given by the closed analytical expression [1]

$$\mu_{\text{sol}} = k_B T \sum_\gamma \rho_\gamma \int d\mathbf{r} \left[\frac{1}{2} h_\gamma^2(\mathbf{r}) \Theta(-h_\gamma(\mathbf{r})) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r}) c_\gamma(\mathbf{r}) \right], \quad (4)$$

where $\Theta(x)$ is the Heaviside step function.

The potential of mean force $W_\gamma(\mathbf{r})$ acting on solvent site γ near the biomolecule is defined in terms of the 3D site distribution function as $W_\gamma(\mathbf{r}) = -k_B T \ln g_\gamma(\mathbf{r})$.

For a bound solvent molecule, the binding strength is determined as the difference between the potential of mean force of solvent site “s” in the effective potential well at the solute macromolecule and that in the first peak of bulk solvent. The binding strength is thus expressed in terms of the solvent peaks at the macromolecule and in bulk solvent,

$$A_s^{\text{binding}} = -k_B T \ln [g_s(\mathbf{r}^{\text{max at solute}}) / g_s(\mathbf{r}^{\text{max in bulk}})]. \quad (5)$$

The solvation free energy (4) and binding energy (5) obtained by 3D-RISM-KH can be decomposed into partial contributions of solute sites, providing a basis for spatial decomposition analysis (SDA) of association effects [13].

To properly treat electrostatic forces in electrolyte solution with polar molecular solvent and ionic species when evaluating the convolution in the 3D-RISM-KH integral equations (1)-(2), the radial correlations (3) of bulk solvent, and in the thermodynamics expressions (4) and (5), the electrostatic asymptotics of all the correlation functions (both the 3D and radial ones) are treated analytically [1,4]. The non-periodic electrostatic asymptotics are separated out in the direct and reciprocal space and the remaining short-range terms of the correlation functions are discretized on a rectangular grid in a non-periodic uniform box large enough to ensure decay of the short-range terms at the box boundaries [4]. The convolution of the short-range terms in

the integral equation (1) is calculated using 3D fast Fourier transform (3D-FFT) in double size box to prevent aliasing. Accordingly, the electrostatic asymptotics terms in the thermodynamics integral (4) are handled analytically, reducing to one-dimensional integrals easy to compute [4]. This analytical treatment of the electrostatics eliminates the periodicity artifacts arising when using the Ewald summation and 3D-FFT supercell for the electrostatic terms, in particular, those in the values of $g_\gamma(\mathbf{k}=0)$ coming from the cancellation or renormalization of the Coulomb singularities of $c_\gamma(\mathbf{k} \rightarrow 0)$ [1,4].

The equations (1)-(2) are converged to a root mean square tolerance (typically 10^{-5}) ensuring accurate structure and thermodynamics, using the modified direct inversion in the iterative subspace (MDIIS) accelerated numerical solver of integral equations of liquid state theory [1].

3 KS-DFT/3D-RISM-KH METHOD

In the self-consistent field KS-DFT/3D-RISM-KH multiscale method [1,2] that we implemented in the ADF computational chemistry package [2], the total Helmholtz free energy of the solute and solvent is defined as

$$A[n_e(\mathbf{r}), \{\rho_\gamma(\mathbf{r})\}] = E[n_e(\mathbf{r})] + \mu_{\text{solv}}[n_e(\mathbf{r}), \{\rho_\gamma(\mathbf{r})\}] \quad (6)$$

where E is the solute electronic energy comprising the standard components, and $n_e(\mathbf{r})$ is the 3D density of solute electrons. Variation of the functional (6) with the solvation free energy (4) with respect to the 3D solvent site density $\rho_\gamma(\mathbf{r})$ yields the 3D-RISM-KH integral equations (1)-(2). Variation of (6) with respect to $n_e(\mathbf{r})$ subject to the normalization condition for N_e solute valence electrons gives the Kohn-Sham equation for the wavefunction ψ_j and energy ε_j of solute valence electron j , modified due to the presence of solvent,

$$\left[-\frac{1}{2}\Delta + v_i(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) + v_{\text{solv}}(\mathbf{r})\right]\psi_j(\mathbf{r}) = \varepsilon_j\psi_j(\mathbf{r}) \quad (7)$$

where v_i , v_H and v_{xc} are the ion core, Hartree and exchange-correlation potentials, and v_{solv} is the effective potential of solvent acting of solute valence electrons. The latter as well as the classical potential of the solute acting on solvent site γ are obtained analytically by variational differentiation of the functional (6) with respect to the 3D density of solute valence electrons $n_e(\mathbf{r})$ and the 3D solvent site density $\rho_\gamma(\mathbf{r})$,

$$v_{\text{solv}}(\mathbf{r}) = \frac{\delta\mu_{\text{solv}}}{\delta n_e(\mathbf{r})} = \sum_\gamma \int d\mathbf{r}' \rho_\gamma(\mathbf{r}-\mathbf{r}') v_\gamma^{(\text{ps})}(\mathbf{r}'), \quad (8)$$

$$u_\gamma(\mathbf{r}) = \frac{\delta\mu_{\text{solv}}}{\delta \rho_\gamma(\mathbf{r})} = \int d\mathbf{r}' n_e(\mathbf{r}-\mathbf{r}') v_\gamma^{(\text{ps})}(\mathbf{r}'), \quad (9)$$

where the solvent site pseudopotentials $v_\gamma^{(\text{ps})}(\mathbf{r})$ can be represented as a sum of the core repulsion, dispersion, and electrostatic terms, which results in subdivision of $u_\gamma^{(\text{ps})}(\mathbf{r})$ into the Lennard-Jones and Coulomb parts [1,2].

4 NANOSTRUCTURES IN SOLUTION PREDICTED BY KS-DFT/3D-RISM-KH

The 3D-RISM-KH equations (1)-(2) were solved on a uniform 3D grid with 0.25 to 0.5 Å resolution in a cubic box of size 32 to 128 Å, large enough to accommodate the macromolecule together with sufficient solvation space. Further refinement of the grid did not affect the results much. The classical solute-solvent interaction potentials $u_{ij}(\mathbf{r}-\mathbf{R}_i)$ comprised the Coulomb and Lennard-Jones terms with the parameters from the OPLS-AA force field.

Figure 1 exhibits the predictions of the KS-DFT/3D-RISM-KH theory for the 3D density distributions of water in the first hydration shell inside as well as outside the (6,6) single-wall carbon nanotube (SWNT) [3]. The theory yields the 3D hydration map, including the orientations of water molecules, both around the nanotube and in its inner part, in excellent agreement with the molecular simulation results. Water molecules in contact with the outer surface of the nanotube are located mostly at the on-top and bridge-site positions of the surface lattice. Their hydrogens are oriented outward from the nanotube, with the HOH molecular plane at small angles to the nanotube axis. This alignment of water dipoles outward from the surface is caused by the strong polarization of the nanotube.

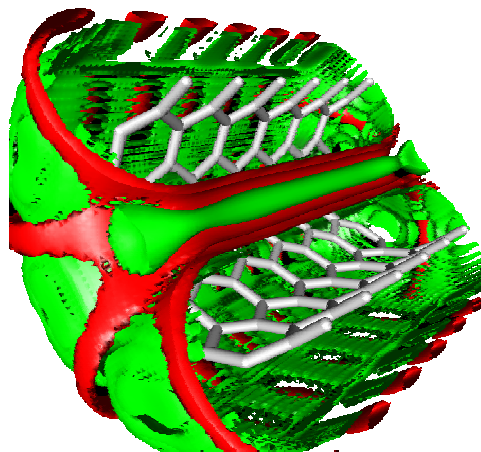


Figure 1: 3D hydration shell of the (6,6) carbon nanotube (half-cut), predicted by the KS-DFT/3D-RISM-KH multiscale theory: water oxygen (red) and hydrogen (green) [3].

The structure of the [mmim][Cl] bulk ionic liquid (IL) we obtained by the KS-DFT/3D-RISM-KH multiscale theory [5] is illustrated in Figure 2. The 3D distributions of the chloride anion around the imidazolium cation in the IL are in good agreement with the previous conclusions from classical MD or quantum-mechanical (SIESTA or CPMD) simulations. Cl⁻ is located mainly around the ring hydrogens. No Cl⁻ appears to be coordinated to the methyl groups in the axial direction. The Cl⁻ distribution function assumes larger values above the unique ring hydrogen, compared to the twin hydrogens on the opposite side, which is in agreement with literature. The 3D distribution of

$[\text{mmim}]^+$ is represented by its nitrogen which is extremely strongly localized in the regions parallel to the ring plane, in agreement with recent *ab initio* simulations showing that cations in the first coordination shell around the $[\text{mmim}]^+$ are situated mainly in parallel to the neighboring rings. Our results show the expected alternating charge layering of the cation-cation distribution, with the stacking interaction evidenced from the first maxima for both the CR and N interaction sites of $[\text{mmim}]^+$ at the same distance.

In Figure 3, we sketch the pathway of self-assembly of the synthetic organic supramolecular nanoarchitecture, rosette nanotube (RNT) decorated with the crown ether functionalities, and show the solvation structure of this RNT predicted by the 3D-RISM-KH molecular theory of solvation and its relation to the stability and possible RNTs

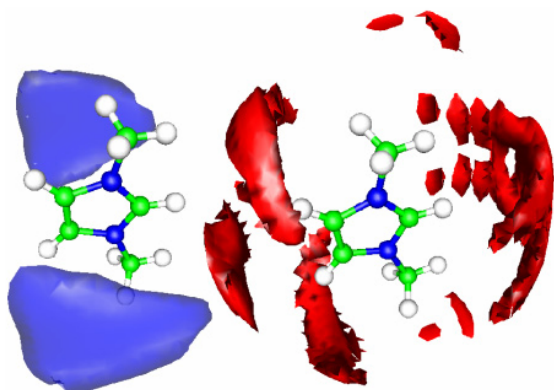


Figure 2: 3D solvation structure of Me-Me imidazolium ion in bulk liquid of $[\text{mmim}][\text{Cl}]$ at $T=400$ K, predicted by the KS-DFT/3D-RISM-KH multiscale theory [5]. Isosurfaces of the nitrogen of $[\text{mmim}]^+$ cations at $g_N(\mathbf{r}) > 2$ (blue) and of Cl^- anions at $g_{\text{Cl}}(\mathbf{r}) > 5$ (red).

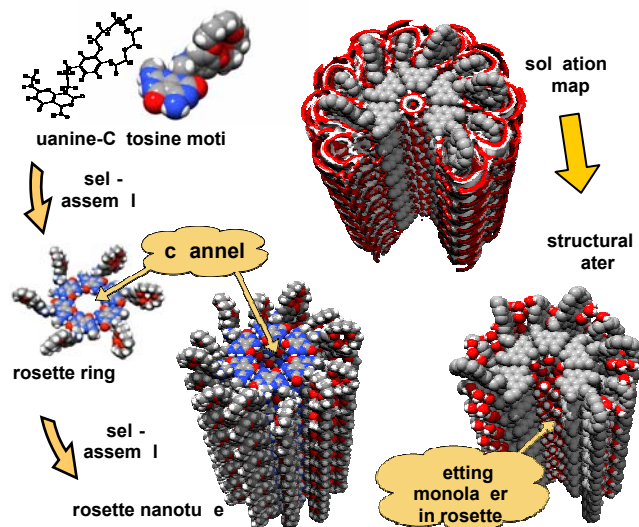


Figure 3: Self-assembly pathway (left part) and 3D maps of the solvation shell (right part) of the organic rosette nanotube (RNT), predicted by the 3D-RISM-KH molecular theory of solvation; water O (red) and H (white) [9].

functions [9]. The theory reveals that structural water molecules penetrate the pockets on the RNT outer surface, form a wetting monolayer in the RNT channel and bridge RNT rosettes. Both inner and outer structural water is hydrogen-bonded to the rosettes and crucially contributes to the stability of the RNT supramolecular architecture. The immobile wetting monolayer, or “outer hydration shell” of the channel, is different from that in a carbon nanotube in that respect that the latter hydrophobically hydrates the CNT channel. The chain of mobile water molecules in the center of the RNT channel is very similar to that in the CNT channel, and shows a possibility of using RNTs as a synthetic water channel in nanodevices.

Acknowledgments. This work was supported by the Prion Core Project of the Alberta Prion Research Institute (APRI) and by the National Research Council (NRC) of Canada.

REFERENCES

- [1] A. Kovalenko, Three-dimensional RISM theory for molecular liquids and solid-liquid interfaces, in: *Molecular Theory of Solvation*, F. Hirata (ed.). Series: *Understanding Chemical Reactivity*, vol.24, (Kluwer, Dordrecht, 2003) pp.169-275, (and references therein).
- [2] S. Gusarov, T. Ziegler, A. Kovalenko, *J. Phys. Chem. A*, 110, 6083, 2006.
- [3] D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler, *J. Chem. Theory Comput.*, 3, 458, 2007.
- [4] J. W. Kaminski, S. Gusarov, T. A. Wesolowski, A. Kovalenko, *J. Phys. Chem. B*, 2010, (in press).
- [5] M. Malvaldi, S. Bruzzone, C. Chiappe, S. Gusarov, A. Kovalenko, *J. Phys. Chem. B*, 113, 3536, 2009.
- [6] J. G. Moralez, J. Ruez, T. Yamazaki, R. K. Motkuri, A. Kovalenko, H. Fenniri, *J. Am. Chem. Soc.*, 127, 8307, 2005.
- [7] R. S. Johnson, T. Yamazaki, A. Kovalenko, H. Fenniri, *J. Am. Chem. Soc.* 129, 5735, 2007; Editors' choice of Science, 316, 663, 2007.
- [8] G. Tikhomirov, T. Yamazaki, A. Kovalenko, H. Fenniri, *Langmuir*, 24, 4447, 2007.
- [9] T. Yamazaki, H. Fenniri, A. Kovalenko, *Chem Phys Chem.*, 11, 361, 2010.
- [10] T. Yamazaki, N. Blinov, D. Wishart, A. Kovalenko, *Biophys. J.*, 95, 4540, 2008.
- [11] N. Blinov, L. Dorosh, D. Wishart, A. Kovalenko, *Biophys. J.* 98, 282, 2010.
- [12] N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, F. Hirata, *J. Phys. Chem. B Feature Article*, 113, 873, 2009.
- [13] T. Yamazaki, A. Kovalenko, *J. Chem. Theory Comput.*, 5, 1723, 2009.
- [14] A. Tanimura, A. Kovalenko, F. Hirata, *Chem. Phys. Lett.*, 378, 638, 2003.
- [15] A. Kovalenko, *J. Comput. Theor. Nanosci.*, 1, 398, 2004.
- [16] A. Tanimura, A. Kovalenko, F. Hirata, *Langmuir*, 23, 1507, 2007.