Pure Appl. Chem., Vol. 85, No. 1, pp. 159–199, 2013.
http://dx.doi.org/10.1351/PAC-CON-12-06-03

Multiscale modeling of solvation in chemical and biological nanosystems and in nanoporous materials*

Andriy Kovalenko‡

National Institute for Nanotechnology, National Research Council, and Department of Mechanical Engineering, University of Alberta, 11421 Saskatchewan Dr., Edmonton, AB, T6G 2M9, Canada

Abstract: Statistical–mechanical, 3D-RISM-KH molecular theory of solvation (3D reference interaction site model with the Kovalenko–Hirata closure) is promising as an essential part of multiscale methodology for chemical and biomolecular nanosystems in solution. 3D-RISM-KH explains the molecular mechanisms of self-assembly and conformational stability of synthetic organic rosette nanotubes (RNTs), aggregation of prion proteins and β-sheet amyloid oligomers, protein-ligand binding, and function-related solvation properties of complexes as large as the *Gloeobacter violaceus* pentameric ligand-gated ion channel (GLIC) and GroEL/ES chaperone. Molecular mechanics/Poisson–Boltzmann (generalized Born) surface area [MM/PB(GB)SA] post-processing of molecular dynamics (MD) trajectories involving SA empirical nonpolar terms is replaced with MM/3D-RISM-KH statistical–mechanical evaluation of the solvation thermodynamics. 3D-RISM-KH has been coupled with multiple time-step (MTS) MD of the solute biomolecule driven by effective solvation forces, which are obtained analytically by converging the 3D-RISM-KH integral equations at outer time-steps and are calculated in between by using solvation force coordinate extrapolation (SFCE) in the subspace of previous solutions to 3D-RISM-KH. The procedure is stabilized by the optimized isokinetic Nosé–Hoover (OIN) chain thermostatting, which enables gigantic outer time-steps up to picoseconds to accurately calculate equilibrium properties. The multiscale OIN/SFCE/3D-RISM-KH algorithm is implemented in the Amber package and illustrated on a fully flexible model of alanine dipeptide in aqueous solution, exhibiting the computational rate of solvent sampling 20 times faster than standard MD with explicit solvent. Further substantial acceleration can be achieved with 3D-RISM-KH efficiently sampling essential events with rare statistics such as exchange and localization of solvent, ions, and ligands at binding sites and pockets of the biomolecule. 3D-RISM-KH was coupled with ab initio complete active space self-consistent field (CASSCF) and orbital-free embedding (OFE) Kohn–Sham (KS) density functional theory (DFT) quantum chemistry methods in an SCF description of electronic structure, optimized geometry, and chemical reactions in solution. The (OFE)KS-DFT/3D-RISM-KH multiscale method is implemented in the Amsterdam Density Functional (ADF) package and extensively validated against experiment for solvation thermochemistry, photochemistry, conformational equilibria, and activation barriers of various nanosystems in solvents and ionic liquids (ILs). Finally, the replica RISM-KH-VM molecular theory for the solvation structure, thermodynamics, and electrochemistry of electrolyte solutions sorbed in nanoporous materials reveals the molecular mechanisms of sorption and supercapacitance in

Pure Appl. Chem.* **85, 1–305 (2013). A collection of invited papers based on presentations at the 32nd International Conference on Solution Chemistry (ICSC-32), La Grande Motte, France, 28 August–2 September 2011.
‡E-mail: andriy.kovalenko@nrc-cnrc.gc.ca; http://www.inrf.ualberta.ca/research/kovalenko/

nanoporous carbon electrodes, which is drastically different from a planar electrical double layer.

Keywords: chemical physics; computational chemistry; computer-aided molecular design; electrochemistry; electrolyte solution; electronic structure in solution; electrosorption in nanoporous electrodes; molecular recognition; molecular theory of solvation; multiscale modeling; solvation; supercapacitor devices.

INTRODUCTION: A CHALLENGE OF MULTISCALE THEORY, MODELING, AND SIMULATION OF NANOSYSTEMS

Nanoscale properties, phenomena, and processes are profoundly different from the macroscopic laws governing the behavior of continuous media and materials. Functional features of nanostructures all stem from microscopic properties of the atoms and chemical groups they are built with, but manifest on length scale from one to hundreds of nanometers and time scale up to microseconds and more. By changing size, composition, and fabrication protocol, the properties of nanostructures and processes involving them can be tuned up in a wide range. Predictive modeling of nanosystems should operate at length scales from an ångström to hundreds of nanometers and microns and time scales from femtoseconds to milliseconds and seconds (e.g., in the description of various biological cellular systems acting as nanomachines operating in a crowded environment), and yet derive their properties from the chemical functionalities of the constituents. Explicit molecular modeling of such nanosystems involves millions and billions of molecules and is by far not feasible in a "brute force" approach employing just ab initio quantum chemical methods and/or molecular simulations. A proper way thus requires multiscale methods coupling several levels of description, from electronic structure methods for building blocks and classical molecular simulations for critical aggregates in the system, to statistical–mechanical theories for their large assemblies and mean properties in a statistical ensemble over characteristic size and time scales, to eventually come up with macroscopic-scale properties of the nanostructures and related processes showing up in the "real observable world". A true, genuine challenge of multiscale modeling is a theoretical framework that couples 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 an example of such a theoretical coupling between microscopic molecular variables and thermodynamic, macroscopic properties.

STATISTICAL–MECHANICAL, 3D-RISM-KH MOLECULAR THEORY OF SOLVATION

Integral equation theory of liquids [1] is becoming increasingly popular, as it provides a firm platform to handle complex chemical and biomolecular systems in solution. The methodology that has shown substantial success for a number of systems in solution is based on the first-principles foundation of statistical mechanics and Ornstein–Zernike (OZ)-type integral equation theory of molecular liquids [1], also known as reference interaction site model (RISM) molecular theory of solvation [1,2]. As distinct from molecular simulations which explore the phase space by direct sampling, RISM theory operates with spatial distributions rather than trajectories of molecules and is based on analytical summation of the free energy diagrams, which yields the solvation structure and thermodynamics in the statistical–mechanical ensemble. It yields the solvation structure by solving the RISM integral equations for the correlation functions and then the solvation thermodynamics analytically as a single integral in terms of the correlation functions obtained. Its three-dimensional (3D) version, 3D-RISM theory gives the 3D maps of distributions of solvent around a solute macromolecule of arbitrary shape [3–11]. An important component of 3D-RISM theory has been the closure proposed by Kovalenko and Hirata (KH approximation) [8,11]. For simple and complex solvents and solutions of a given composition, including buffers, salts, polymers, ligands, and other cofactors at a finite concentration, the 3D-RISM-KH

molecular theory of solvation properly accounts for chemical functionalities by representing in a single formalism both electrostatic and nonpolar features of solvation, such as hydrogen bonding, structural solvent molecules, salt bridges, solvophobicity, and other electrochemical, associative, and steric effects. For real systems, solving the 3D-RISM-KH integral equations is far less computationally expensive than running molecular simulations which must be long enough to sample all relevant exchange and binding events. This enables handling complex systems and processes occurring on large space and time scales, problematic and frequently not even feasible for molecular simulations. The 3D-RISM-KH theory provides a successful description of both simple and complex associating liquids with various chemical functionalities [12–16] in the whole range of fluid thermodynamic conditions [17,18] and a variety of local environments in different systems, such as carbon nanotubes [14], synthetic organic rosette nanotubular architectures [19–23], and biomolecular systems [24–33]. The latter range from structural water and xenon bound in the pocket of lysozyme protein [24,25], permeation of water and ions through aquaporin channels [25], ligand efflux in the multidrug transporter AcrB [26], aggregation of amyloid (A)β oligomers and fibrils [27–29] and prion proteins [27,29], binding modes for inhibitors of the pathological conversion, and aggregation of prion proteins [29] and thiamine against the extracytoplasmic thiamine binding lipoprotein MG289 [30,31], to biomolecular systems as large as the ligand-gated ion channel (GLIC) in a lipid bilayer [29,30] and the GroEL/ES chaperone complex [32]. The RISM/3D-RISM-KH approach provided an insight into a number of experimentally observed phenomena in soft matter systems, including the structural transitions and related thermodynamic anomalies for the formation of micromicelles and tetrahedral-to-zigzag transformation of the hydrogen bonding network in water–alcohol mixtures in the whole range of concentrations [34,35], microscopic structure of interfaces of nonpolar and polar hydrogen bonding forming molecular liquids [36,37], and microscopic structure of gels formed by oligomeric polyelectrolyte gelators in different solvents [33].

Integral equations for the solvation structure

The solvation structure is represented by the probability density $\rho_\gamma g_\gamma(\mathbf{r})$ of finding interaction site γ of solvent molecules at 3D space position $\mathbf{r}$ around the solute molecule (which can be both a macromolecule and supramolecule), as determined by the average number density ρ_γ in the solution bulk times the normalized density distribution, or 3D distribution function, $g_\gamma(\mathbf{r})$. The values of $g_\gamma(\mathbf{r}) > 1$ or $g_\gamma(\mathbf{r}) < 1$ indicate areas of density enhancement or depletion, respectively, relative to the average density at a distance from the solute in the solution bulk where $g_\gamma \to 1$. The 3D distribution functions of solvent interaction sites around the solute molecule are obtained from the 3D-RISM integral equation [3–11]

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

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, $c_\gamma(\mathbf{r}) \sim -u_\gamma(\mathbf{r})/(k_\mathrm{B}T)$; 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. The KH closure approximation [8,11] accounts in a consistent manner for both electrostatic and nonpolar features (associative and steric effects) of solvation in simple and complex liquids, non-electrolyte and electrolyte solutions, and complex macromolecular and supramolecular solutes in chemical [8–18], synthetic organic [19–23], biomolecular [24–33], and soft matter [33–37] systems. The 3D-KH closure reads

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

$$d_\gamma(\mathbf{r}) = -u_\gamma(\mathbf{r})/k_\mathrm{B}T + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$$

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 $k_\mathrm{B}T$ is the Boltzmann constant times the solution temperature. The 3D-KH closure (eq. 2) couples in a nontrivial way the so-called mean spherical approximation (MSA) and the hypernetted chain (HNC) approximation [1], the former being applied to spatial regions of solvent density enrichment $g_\gamma(\mathbf{r}) > 1$ such as association peaks and critical enhancement long-range tails, and the latter to regions of density depletion $g_\gamma(\mathbf{r}) < 1$, including the repulsive core. The distribution function and its first derivative are continuous at the joint boundary $g_\gamma(\mathbf{r}) = 1$ [or equivalently $d_\gamma(\mathbf{r}) = 1$] by construct. A substantial advantage of the 3D-KH closure approximation is that it properly accounts for the solvation structure of complex solvated systems with significant association effects. For comparison, the 3D-HNC closure strongly overestimates such associative effects and therefore the 3D-RISM-HNC equations diverge in many practical applications for macromolecules with considerable site charges solvated in polar solvents or electrolyte solutions. Other approximations such as, for example, the Percus–Yevick (PY), modified Verlet (VM), Martynov–Sarkisov (MS), and Ballone–Pastore–Galli–Gazzillo (BPGG) closures do not properly account for the electrostatic asymptotics of the interaction potential.

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

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

where the intramolecular correlation function

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

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 intermolecular, radial total correlation function between sites α and γ enumerating all sites on all sorts of molecules in bulk solvent. The site–site total correlation functions $h_{\alpha\gamma}(r)$ to be input in eq. 3 and then (1) are obtained in advance to the 3D-RISM-KH calculation from the dielectrically consistent RISM theory [38] coupled with the KH closure (DRISM-KH approach [11]) which can be applied to the bulk solution of a given composition, including polar solvent, cosolvent, electrolyte, and ligands at a given concentration. The DRISM integral equation reads as

$$\tilde{h}_{\alpha\gamma}(r) = \tilde{\omega}_{\alpha\mu}(r) * c_{\mu\nu}(r) * \tilde{\omega}_{\nu\gamma}(r) + \tilde{\omega}_{\alpha\mu}(r) * c_{\mu\nu}(r) * \rho_\nu \tilde{h}_{\nu\gamma}(r) \tag{5a}$$

where $c_{\alpha\gamma}(r)$ is the site–site direct correlation function of bulk solvent, an asterisk "*" means convolution in the direct space, and summation over repeating site indices is implied. Both the intramolecular correlation functions $\tilde{\omega}_{\alpha\gamma}(r)$ and total correlation functions $\tilde{h}_{\alpha\gamma}(r)$ are renormalized due to a dielectric bridge correction in a particular analytical form [38] that ensures consistency and given value of the dielectric constant obtained for solvent–solvent, solvent–ion, and ion–ion effective interactions in electrolyte solution

$$\tilde{\omega}_{\alpha\gamma}(r) = \omega_{\alpha\gamma}(r) + \rho_\alpha \zeta_{\alpha\gamma}(r) \tag{5b}$$

$$\tilde{h}_{\alpha\gamma}(r) = h_{\alpha\gamma}(r) - \zeta_{\alpha\gamma}(r) \tag{5c}$$

The renormalized dielectric correction enforcing the given phenomenological value of the dielectric constant and the proper orientational behavior and consistency of the dielectric response in

the electrolyte solution are obtained in the analytical form specified in the reciprocal k-space as follows [38]

$$\zeta_{\alpha\gamma}(k) = j_0(kx_\alpha)j_0(ky_\alpha)j_1(kz_\alpha)h_c(k)j_0(kx_\gamma)j_0(ky_\gamma)j_1(kz_\gamma) \tag{6}$$

where $j_0(x)$ and $j_1(x)$ are the zeroth- and first-order spherical Bessel functions, $\mathbf{r}_\alpha = (x_\alpha,y_\alpha,z_\alpha)$ and $\mathbf{r}_\gamma = (x_\gamma,y_\gamma,z_\gamma)$ are the Cartesian coordinates of partial site charges q_α and q_γ at sites α and γ of the same solution species s with respect to its molecular origin, and its dipole moment $\mathbf{d}_s = \Sigma_{\alpha\in s}q_\alpha\mathbf{r}_\alpha$ is oriented along the z-axis, $\mathbf{d}_s = (0,0,d_s)$. Note that the renormalized dielectric correction (6) is nonzero only for polar solvent species of sorbed electrolyte solution which possess a dipole moment and thus are responsible for the dielectric response in the DRISM approach. The envelope function $h_c(k)$ has the value at $k = 0$ determining the dielectric constant of the solution and is assumed in the smooth non-oscillatory form quickly falling off at wavevectors k larger than those corresponding to characteristic size l of liquid molecules [38]

$$h_c(k) = A\exp\left(-l^2k^2/4\right) \tag{7}$$

where the amplitude A is related to the dielectric constant ε of the electrolyte solution

$$A = \frac{1}{\rho_{\text{polar}}}\left(\frac{\varepsilon}{y} - 3\right) \tag{8}$$

The form (6–8) is extended to mixed solvents [39] by using the total number density of solution polar species

$$\rho_{\text{polar}} = \sum_{s\in \text{ polar}} \rho_s \tag{9}$$

and the solution dielectric susceptibility

$$y = \frac{4\pi}{9k_BT}\sum_{s\in \text{ polar}} \rho_s\left(d_s\right)^2 \tag{10}$$

The parameter l specifies the characteristic separation from a liquid molecule below which the dielectric correction (6) is switched off so as not to distort the short-range solvation structure. It can be chosen to be about $l = 1$ Å for water solvent; however, in solvent of larger molecules such as octanol or in the presence of such cosolvent it should be increased to about $l = 10$ Å so as to avoid "ghost" associative peaks appearing in the radial distributions if the dielectric correction (5) interferes with the intramolecular structure of the large solvent species.

The DRISM integral eqs. 5 with the site–site version of the KH closure [11]

$$g_{\alpha\gamma}(r) = \begin{cases} \exp\left(d_{\alpha\gamma}(r)\right) & \text{for} \quad d_{\alpha\gamma}(r) \le 0 \\ 1 + d_{\alpha\gamma}(r) & \text{for} \quad d_{\alpha\gamma}(r) > 0 \end{cases} \tag{11}$$

$$d_{\alpha\gamma}(r) = -u_{\alpha\gamma}(r)/k_BT + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r)$$

keep the same dielectrically consistent asymptotics (6 and 7) as the originally derived DRISM-HNC theory [38] but extend the description to solutions with strong associative species in a wide range of composition and thermodynamic conditions not amenable to HNC due to its overestimation of associative forces and phase transition phenomena.

Analytical expressions for the solvation thermodynamics

As mentioned above, 3D-RISM is an integral equation theory of OZ type based on analytical summation of the free energy diagrams, which yields the 3D solute–solvent site correlation functions and the solvation thermodynamics. The KH as well as HNC closure approximations to the 3D-RISM integral equation have an exact differential of the solvation free energy, and allow one to analytically perform Kirkwood's thermodynamic integration gradually switching on the solute–solvent interaction. The solvation free energy of the solute macro- or supramolecule in multicomponent solvent following from the 3D-RISM-KH integral eqs. 1 and 2 is thus given in the closed analytical form as a single integral of the correlation functions [8,11]

$$\Delta\mu = k_{\mathrm{B}}T\sum_{\gamma}\rho_{\gamma}\int \mathrm{d}\mathbf{r}\left[\frac{1}{2}h_{\gamma}^{2}(\mathbf{r})\Theta(-h_{\gamma}(\mathbf{r})) - c_{\gamma}(\mathbf{r}) - \frac{1}{2}h_{\gamma}(\mathbf{r})c_{\gamma}(\mathbf{r})\right] \tag{12}$$

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

Other thermodynamics quantities can be derived from the solvation free energy (12) by differentiation. In particular, the solvation chemical potential is decomposed into the energetic and entropic contributions at constant volume

$$\Delta\mu = \Delta\varepsilon^{\mathrm{uv}} + \Delta\varepsilon^{\mathrm{vv}} - T\Delta s_{\mathrm{V}} \tag{13}$$

by calculating the solvation entropy at constant volume as

$$\Delta s_{\mathrm{V}} = -\frac{1}{T}\left(\frac{\partial\Delta\mu}{\partial T}\right)_{\mathrm{V}} \tag{14}$$

and the internal energy of the solute ("u")–solvent ("v") interaction as

$$\Delta\varepsilon^{\mathrm{uv}} = k_{\mathrm{B}}T\sum_{\gamma}\rho_{\gamma}\int \mathrm{d}\mathbf{r}\; g_{\gamma}(\mathbf{r})u_{\gamma}(\mathbf{r}) \tag{15}$$

with the remaining term $\Delta\varepsilon^{\mathrm{vv}}$ giving the energy of solvent reorganization around the solute. Further, the partial molar volume of the solute macromolecule is obtained from the Kirkwood–Buff theory [40] extended to the 3D-RISM formalism [25,41,42]

$$V = k_{\mathrm{B}}T\chi_{\mathrm{T}}\left(1 - \sum_{\gamma}\rho_{\gamma}\int \mathrm{d}\mathbf{r}\; c_{\gamma}(\mathbf{r})\right) \tag{16}$$

where χ_{T} is the isothermal compressibility of bulk solvent obtained in terms of the site–site direct correlation functions of bulk solvent as

$$\rho k_{\mathrm{B}}T\chi_{\mathrm{T}} = \left(1 - 4\pi\sum_{\alpha\gamma}\rho_{\alpha}\int_{0}^{\infty} r^{2}\,\mathrm{d}r\; c_{\alpha\gamma}(r)\right)^{-1} \tag{17}$$

where $\rho = \Sigma_s \rho_s$ is in general the total number density of bulk solvent mixture of molecular species s.

Note that the solvation free energy (12) and its derivatives (13–15), as well as the partial molar volume (16) can be decomposed into partial contributions of the interaction sites of the solute macromolecule, providing a basis for spatial decomposition analysis (SDA) of association effects in solution [43,44], as illustrated for supramolecular complexation [43] and proteins [44].

The potential of mean force (PMF) $W_{\gamma}(\mathbf{r})$ acting on site $\gamma$of solvent species near the solute macro- or supramolecule can be defined in terms of the 3D site distribution function as

$$W_{\gamma}(\mathbf{r}) = -k_{\mathrm{B}}T\ln g_{\gamma}(\mathbf{r}) \tag{18}$$

The form (18) gives a 3D map of the effective potential between each solvent species and the solute macromolecule, and determines the binding strength and most probable locations (binding modes) of structural solvent molecules at the solute macromolecule, averaged over the statistical mechanical ensemble of mutual arrangements and orientations.

Analytical treatment of electrostatics and accelerated numerical solution of the integral equations

To properly treat electrostatic forces in electrolyte solution with polar molecular solvent and ionic species in 3D-RISM/1D-RISM theory, we analytically handle the long-range electrostatic asymptotics of the radial site–site total and direct correlation functions of bulk solvent in the DRISM-KH equations (5–11) and the 3D site total and direct correlation functions in the 3D-RISM-KH equations (1 and 2), as well as in the thermodynamics expressions for the solvation free energy (12) and its derivatives [9–11,15,46]. The analytical forms for the nonperiodic electrostatic asymptotics are separated out in the direct and reciprocal space from all the correlation functions, and the remaining short-range terms are discretized for the DRISM-KH equations on a uniform radial grid with resolution 0.01–0.1 Å and for the 3D-RISM-KH equations on a uniform 3D rectangular grid with resolution 0.1–0.5 Å in a 3D box large enough to ensure decay of the short-range terms at the 3D box boundaries [15,46]. The convolution of the short-range terms in the 3D-RISM integral equation (1) is calculated using 3D fast Fourier transform (3D-FFT) in the 3D box of size including at least 2 to 3 solvation shells around the solute. This cancels out the aliasing effects since the resulting short-range solvation shells of the total correlation functions usually decay at this distance and the remaining electrostatic asymptotics are separated out and handled analytically [46]. Accordingly, the electrostatic asymptotics terms in the thermodynamic integral (4) are handled analytically and reduced to 1D integrals easy to compute [15,46].

The 3D-RISM-KH integral equations (1 and 2) are converged typically to a relative root mean square accuracy of 10^{-4}–10^{-5} and the DRISM-KH equations (5 and 11) to an accuracy of 10^{-8}–10^{-10} by using the modified direct inversion in the iterative subspace (MDIIS) numerical solver which is an iterative procedure achieving accelerated convergence for integral equations of liquid state theory by optimizing each iterative guess in a Krylov subspace of typically last 10–20 successive iterations [9–11,47]. It is closely related to Pulay's DIIS approach for quantum chemistry equations [48] and other similar algorithms like the generalized minimal residual (GMRes) solver [49]. Note that the GMRes method applied on the fine grid discretizing the 3D-RISM integral equations has also been coupled with a Newton–Raphson-type numerical scheme on a coarse grid [50] which can be limited to the solute repulsive core area as well [51]. Direct methods of Newton–Raphson type to solve integral equations provide a quadratic convergence when approaching the solution. However, they require calculating and inverting the Jacobian matrix in one or another way, which is prohibitively wasteful for fine 3D grids and very slow for coarse 3D grids even if performed just once and used for preconditioning [51]. The MDIIS solver combines the simplicity and relatively small memory usage of an iterative approach with the efficiency of a direct method. Compared to damped (Picard) iterations, MDIIS provides substantial acceleration with quasiquadratic convergence practically throughout the entire range of root mean square residual, and is robust and stable. Of particular importance is that MDIIS ensures convergence (provided a solution exists) for complex charged systems with strong associative and steric effects, which is usually not achievable by Picard iterations and constitutes a challenging task in the case of 3D integral equations. Further, a core-shell-asymptotics technique coupling MDIIS for the excluded volume core with iteration of the solvation shells has been developed which provides 6- to 16-fold memory reduction and corresponding CPU load decrease in MDIIS [47]. Although being of benefit for solutes of any size, this memory reduction becomes critical in 3D-RISM calculations for large solvated systems, such as macromolecules in solution with ions, ligands, and other cofactors.

Examples of 3D-RISM-KH calculations of solvation structure and thermodynamics

One of the basic cases of electrolyte solution systems is aqueous solution of sodium chloride. Figures 1–3 show the 3D-RISM results for the solvation structure and PMF of the Na^+ and Cl^- ion pair in aqueous solution at infinite dilution and at a high concentration of 1 mol/l [9,10]. This generic system presents a very illustrative test for a solvation theory to reproduce most of the essential features of a variety of chemical and biomolecular effects in solution, and the 3D-RISM theory succeeds in that.

Figure 1 shows the 3D site distributions of the water oxygen (O) and hydrogen (H) site distributions around the pair of the Na^+ and Cl^- ions in aqueous solution at infinite dilution. Both the O and H distributions form high crowns of the first solvation shell around the ions, with the high O and lower H peaks near the contacts of the crowns corresponding to water molecules bridging the ions. This structure is then followed by the shallow second solvation shells. The corresponding PMF is shown in Fig. 2.

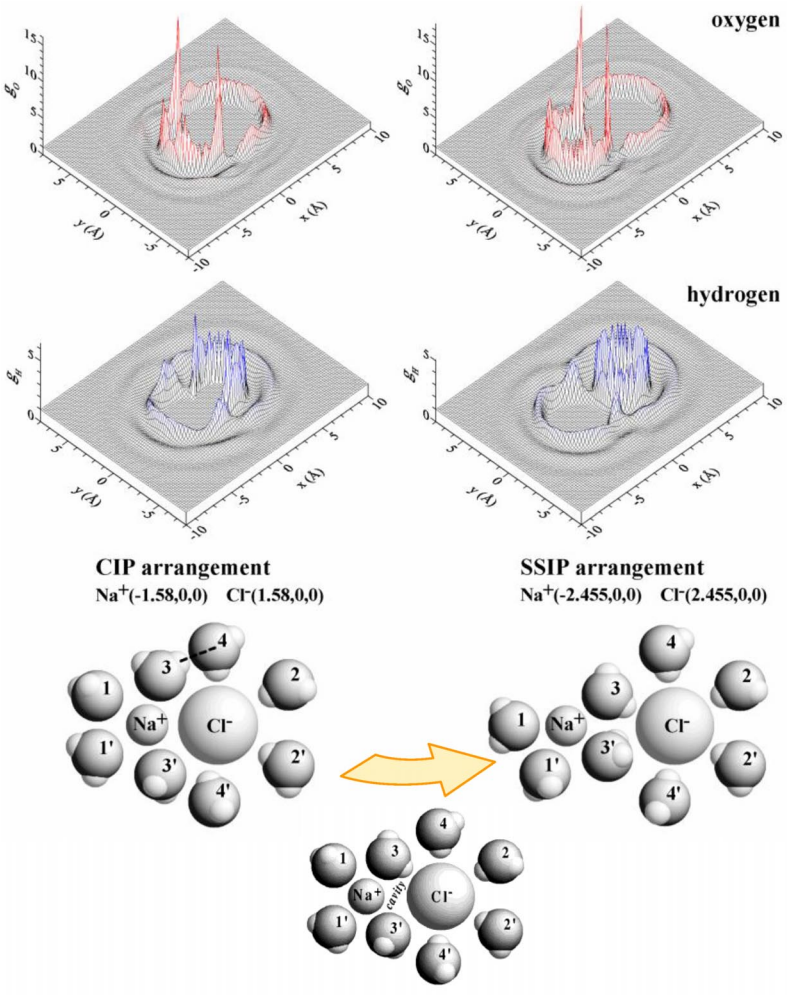

Fig. 1 Solvation structure of the CIP and SSIP of the Na^+ and Cl^- ion pair in aqueous solution at infinite dilution (left and right columns, respectively), obtained from 3D-RISM theory [9,10]. Section of the 3D distribution functions of water oxygen (O) and hydrogen (H) in the plane passing through the ion–ion axis (3D graphs in the upper part). Visualization of the water solvent arrangements around the CIP and SSIP, as well as the ions separated by a gap corresponding to the barrier at their PMF (cartoons in the lower part).

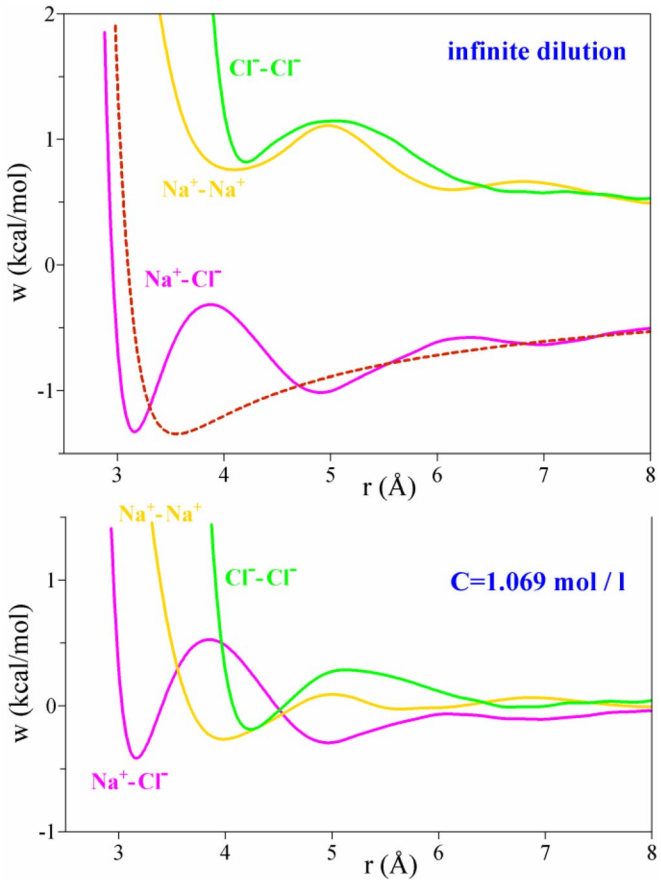

Fig. 2 PMF between Na^+ and Cl^- ions in aqueous solution at infinite dilution (upper part) and concentration 1 mol/l (lower part), obtained from 3D-RISM theory [9,10]. Section of the 3D distribution functions of water oxygen (O) and hydrogen (H) in the plane passing through the ion–ion axis (3D graphs in the upper part). Shown in the upper part for comparison is also the PMF obtained with the primitive model of water solvent as structureless dielectric continuum (dashed line).

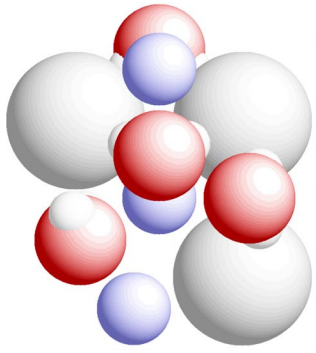

Fig. 3 Cl^- ions (in gray) in a stabilized CIP arrangement bridged by Na^+ ions (in blue) as well as by hydrogen-bonding bridges of water molecules (O in red, H in white) in aqueous electrolyte solution of concentration 1 M. The visualization is made based on the peaks positions on the 3D site distribution functions between the species of this system obtained from 3D-RISM theory [9,10].

The H distribution has two crowns around the Cl^- ion; their separation from the ion shows that the inner H crown gives the water hydrogens in contact with the Cl^- ion while the outer H crown corresponds to the other water hydrogens looking outwards but at the angle determined by the tetrahedral hydrogen-bonding structure of the water. On the other hand, there is a single H crown around the Na^+ ion and the separations of the O and H crowns from Na^+ show that both the water hydrogens are looking outward Na^+, tilted at the same angle. These are typical arrangements of dipole-like oriented water around a cation and water hydrogen-bonded with one hydrogen to an anion. The arrangements are visualized in the cartoons in the lower part of Fig. 1, with the left and right columns corresponding to the 3D water site distributions around the Na^+ and Cl^- contact ion pair (CIP) and solvent-separated ion pair (SSIP), respectively. Water molecules in contact with both the cation and anion form the dipole-like association with the former and the hydrogen bonding with the latter, thus creating a water bridge of strongly associated water molecules located in a ring between the ions. This bridge strongly deepens the solvation contribution to the PMF at the CIP arrangement; interplayed with the ion–ion core repulsion, this results in a significant shift of the first minimum on the PMF to a shorter ion–ion separation, compared to the primitive solvation model just uniformly reducing the Coulomb attraction between the ions by the water dielectric constant $\varepsilon = 80$ (dashed line in Fig. 2, upper part). At the SSIP arrangement, the water bridge strengthens the association between the ions and results in the second minimum on the ion–ion PMF (Fig. 2). The oscillations diminish with distance, and the PMF goes to the limit of the pure dielectrically screened electrostatic potential (dashed line). At an intermediate separation between the ions, a desolvated gap forms because of the steric effect of expulsion of solvent molecules by the repulsive cores of the ions (middle cartoon in Fig. 1). The work against the solvent environment to expel the solvent and create the desolvation gap results in a barrier between the PMF first and second minima corresponding to the CIP and SSIP arrangements (Fig. 2, upper part).

The PMFs in the Na^+–Na^+ and Cl^-–Cl^- pairs of like ions have the same features of the first and second minima and the barrier between them due to the interplay of the associative forces and molecular structure of the ions and solvent molecules. However, at infinite dilution the strength of the solvent bridges is not sufficient to overcome the electrostatic repulsion and to stabilize the like ion pairs, and both the first and second minima are local (Fig. 2, upper part). (Note that this can be very different for large molecular ions with weaker electrostatic attraction at the separation determined by their size.) The picture changes for the electrolyte solution at a high concentration of 1 mol/l; numerous salt bridges form in addition to water hydrogen bridges, and the like ion pairs get stabilized in both CIP and SSIP arrangements (Fig. 2, lower part). For example, the Cl^-–Cl^- ion pair in the aqueous solution at this concentration is bridged by several Na^+ ions and water molecules, forming a cluster depicted in Fig. 3. The structure of such a cluster follows from the analysis of the 3D-RISM results for the 3D distribution functions of solution species and the corresponding coordination numbers of Na^+, Cl^-, and water around each ion pair (Na^+–Cl^-, Cl^-–Cl^-, and Cl^-–Cl^-).

A further important example of crucial importance in chemistry and biomolecular nanosystems is formation of nanostructures in solution driven by hydrophobic attraction. Figure 4 illustrates the RISM theory predictions for the structure of the ambient mixtures of water and *tert*-butyl alcohol (TBA) in the whole range of concentrations [34,35]. TBA is a generic example of primitive surfactant with a hydrophobic head of four carbons and a hydrophilic "tale" represented by the hydroxyl group. The solvation structure of this system successively goes through several stages with TBA concentration in water changing from infinite dilution pure TBA: a separate TBA molecule embedded in a water tetrahedral hydrogen bonding cage at infinite dilution changes to micromicelles of four to six TBA molecules in the head-to-head arrangement incorporated in a water hydrogen-bonding cage at about 4 % TBA molar fraction; then, the tetrahedral hydrogen bonding structure of water gets disrupted at about 40 % TBA molar fraction to be replaced by the zigzag hydrogen-bonding structure of alcohol, with separate water molecules embedded in it at infinite dilution of water in TBA. The RISM theory predicted both the structure and thermodynamics of these mixtures in full agreement with experiment, in partic-

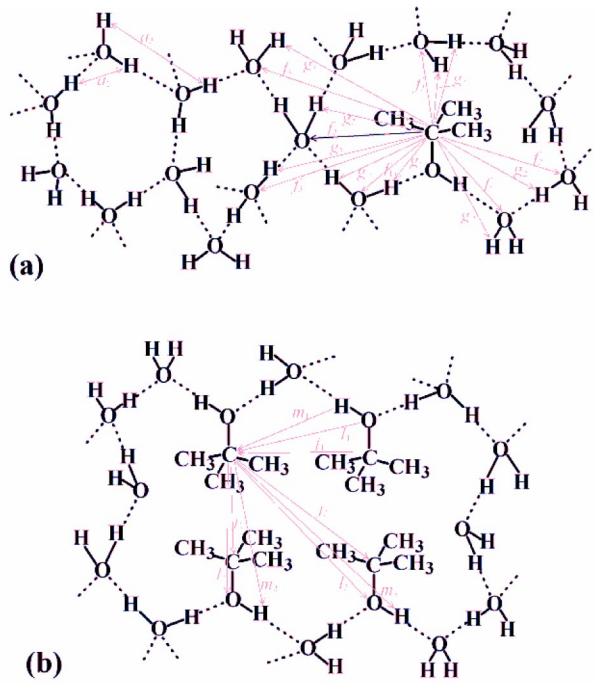

Fig. 4 Structural transitions with concentration in a mixture of water and TBA and predicted by the RISM-KH molecular theory of solvation: (a) TBA molecule at infinite dilution incorporated in a cage of water tetrahedral hydrogen bonding; (b) micromicelle of four TBA molecules in a cage of the water hydrogen bonding.

ular, the concentration and temperature dependence of the compressibility, including the isosbestic point and minimum corresponding to the formation of micromicelles [34,35].

One complex example from supramolecular organic chemistry is synthetic organic supramolecular rosette nanoarchitectures, or rosette nanotubes (RNTs) [19–23]. The 3D-RISM-KH molecular theory of solvation revealed the molecular mechanisms of self-assembly and conformational stability of RNTs forming in different solvents and held through noncovalent forces. The molecular building block, cytosine/\guanine motif decorated with various functional groups, undergoes an hierarchical self-assembly process in solution to form a six-membered supermacrocycle (rosette ring) maintained by 18 hydrogen bonds between its complementary sites, which in turn self-organizes into a linear stack (a nanotube with an open central channel), an aggregate which in general is highly stable and readily withstands boiling and drying (Fig. 5). Any functional group covalently attached to the motif, for example, crown ether, alkyl chains, and lysine tails, could be expressed on the RNT surface, thereby offering a general "built-in" strategy for tailoring the physico-chemical properties of RNTs. The 3D-RISM-KH theory uncovered the pathways of step-by-step self-assembly of RNTs from motifs into rosettes and then nanotubes, showing that the thermodynamically preferred mechanism of RNT growth is attachment of motifs to the nanotube end to form and complete a new rosette ring, rather than growth of a separate ring in solution and then its attachment to the nanotube end [22]. Figure 6 depicts the inner and outer hydration structure of the RNT with crown ether functionalities on its outer surface. The theory exhibited that the RNT channel is covered with a wetting monolayer of structural water molecules strongly bound to its surface and crucially contributing to the RNT stability, holds a chain of loosely bound water molecules at the channel center, and can also hold host molecules of inert gases or drugs inside [22]. This suggests potential use of RNTs as artificial channels for molecular transport in nano-engineered bioorganic systems and as drug delivery vehicles. The 3D-RISM-KH theory predicted that

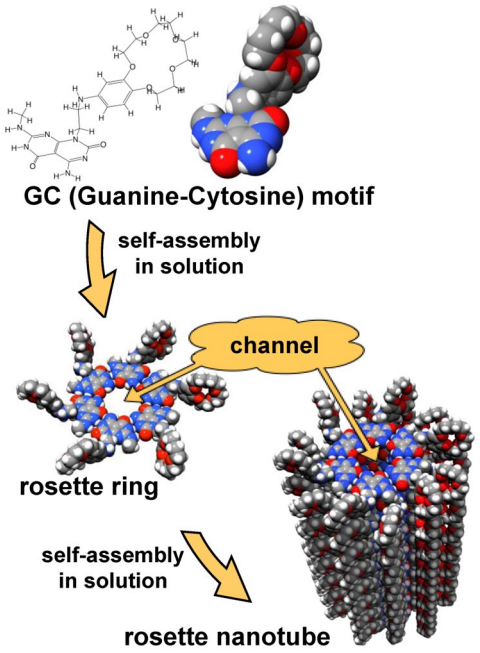

Fig. 5 Formation of supramolecular RNTs [19–23]. RNT functionalized with crown ethers expressed on the outer RNT surface is shown.

a G/\C base bearing two C12 alkyl chains undergoes a solvent-controlled multistep hierarchical self-assembly process into lamellar prolate nanospheroids [21]. The theory explained how the stability of the helical rosette nanotubes (HRNs) self-assembling from the C/\G motif decorated with the lysine tail can be tuned (significantly increased) by adding a covalent linker pairing adjacent G/\C bases in neighboring rings, and how HRNs undergo further hierarchical self-organization into superhelices [19]. The theory predicted the molecular mechanism of solvent-driven supramolecular chirality in HRNs: structural solvent molecules localized in the pockets between the lysine tails on the HRN surface play a role of molecular switches causing the tails to form (i) a right-hand supramolecular helix in water, the conformation with the lowest free energy (thermodynamic formation), and (ii) a left-hand supramolecular helix in methanol, the conformation preferred due to a kinetic barrier for the right-hand one at the beginning of HRN self-assembly in methanol (kinetic formation), but undergoing subsequent conversion to the right-hand one under heating, in full agreement with experiment [20]. The 3D-RISM-KH theory predicted the most stable conformation of the RNTs made of twin G/\C module bearing the lysine side chain, with the "nests" consisting of four side chains used for one-pot nucleation, growth, morphogenesis, and passivation of 1.4 nm Au nanoparticles, one of the possible applications of RNTs [23].

To conclude this part, the above examples represent generic forces acting in chemical and biomolecular systems in solution. They demonstrate how this theory captures and resolves in 3D detail the molecular origin of the solvation structure shells and the PMFs between the solution species. Note that the molecular picture provided by the 3D-RISM theory is at the level of molecular simulation, and is not achievable otherwise by a continuum solvation semi-empirical treatment; the latter can be parameterized to represent, for example, the effect of hydrogen bonding in pure water but is not transferable to solution with other components like electrolyte, cosolvent, not speaking of such complex cofactors as different ligand compounds.

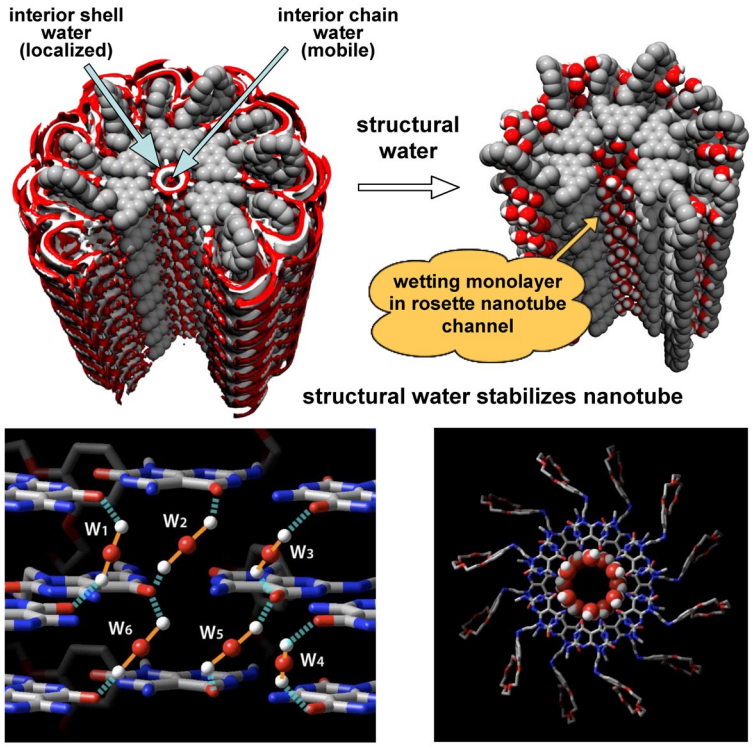

Fig. 6 Hydration structure of the RNT functionalized with crown ethers in aqueous electrolyte solution with Cl^- counterions [22]. 3D distributions of water oxygen (in red) and hydrogen (in white) in the inner and outer hydration shells of the RNT predicted by the 3D-RISM-KH molecular theory of solvation (upper left part). Visualization of the structural molecules in the channel and outer pockets of the RNT (upper right part). Structural water molecules of the interior shell water in the RNT channel attach to the RNT motifs with hydrogen bonds stabilizing the RNT (lower left part) and form a wetting monolayer (lower right part).

Molecular recognition and protein-ligand binding with 3D-RISM-KH

The PMF (eq. 18) defined in terms of the 3D site distribution functions as a 3D map of the effective potential between each solvent species and the solute biomolecule determines the binding strength and most probable locations (binding modes) of structural solvent molecules at the solute macromolecule, averaged over the statistical mechanical ensemble of mutual arrangements and orientations. With the solvent mixture including also ions and other cofactors such as cosolvent and ligands, this constitutes an approach to predict molecular recognition in supramolecules and biomolecules from the first principles of statistical mechanics, with full account of their molecular specificities [25]. It gives a new computational method of mapping ligand molecules on protein surfaces for protein-ligand binding and fragment-based drug design [25,26]. This approach has been realized in the 3D-RISM-based ligand mapping (3D-RISM-LM) algorithm [45] applied to identify the drug efflux pathway in the membrane protein and multidrug transporter AcrB [26], and in the ligand docking (3D-RISM-Dock) algorithm [29,31] implemented in the AutoDock package and validated on the experimental data for the binding modes and binding free energy of the antiprion compound GN8 against mouse PrP protein [29] and those of thiamine against the extracytoplasmic thiamine binding lipoprotein MG289 [44,45]. The 3D-RISM-Dock approach properly accounts for molecular specificity of the ligand and solvent and allows one to study concentration effects on protein-ligand binding in fragment-based rational drug design.

Post-processing of the thermodynamics of MD trajectories with 3D-RISM-KH

The popular MM/PB(GB)SA post-processing of the thermodynamics of MD trajectories uses the Poisson–Boltzmann (PB) or generalized Born (GB) approach combined with the solvent-accessible surface area (SA) empirical term to account for nonpolar interactions in solution, such as hydrophobic hydration and hydrophobic attraction. There is growing evidence that the PB(GB)SA continuous solvation models cannot describe the nonelectrostatic effects accurately [52–54], especially for small proteins. This follows from the fact that for small peptides there is no proportionality between solvent-accessible SA of a protein and the nonpolar part of the solvation free energy [55]. In addition, such models do not account correctly for the dispersion interactions and excluded volume effects [52,53]. This raises questions about the applicability of PB(GB)SA to describe quantitatively the hydrophobic effects in biomolecular systems. An alternative to account accurately for nonpolar effects is to use the following combined approach: molecular dynamics (MD) is used to generate trajectories, followed by molecular mechanics (MM) to calculate the peptides internal energy and conformational entropy, and the 3D-RISM-KH molecular theory of solvation to characterize the solvation free energy. It worth noting that the continuous solvation models such as PB(GB)SA require phenomenological parameters such as ionic radii and surface tension coefficients to be used as input or modeling [54], which makes them less reliable, compared to the molecular theory of solvation.

The MM/3D-RISM-KH post-processing method was applied to study the thermodynamics and volumetrics of the hydration and aggregation of the HET-s prion and amyloid-β fibril [27,29], and the conformational stability and association thermodynamics of small wild-type Aβ(17–42) oligomers with different protonation states of Glu(22) vs. the E22Q (Dutch) mutants [28,29].

MULTISCALE COUPLING OF 3D-RISM-KH MOLECULAR THEORY OF SOLVATION WITH ELECTRONIC STRUCTURE METHODS

The 3D-RISM-KH molecular theory of solvation has been coupled with quantum chemistry methods, including Kohn–Sham density functional theory (KS-DFT) [8,11,13–15] and ab initio complete active space self-consistent field (CASSCF) method [11,12], in a multiscale approach to electronic structure and chemical reactions in solution [11,13–15] and at solid–liquid interfaces [8]. Note that 3D-RISM-KH can be combined in an SCF approach with any multireference electronic structure theory. The closed analytical form for the solvation free energy stemming from the 3D-RISM-KH equations allows one to obtain its analytical gradients with respect to nuclei coordinates, or analytical solvation forces acting on nuclei in solution, which gives access to geometry optimization and evaluation of reaction pathways in solution. The SCF, KS-DFT/3D-RISM-KH multiscale theory, including analytical gradients, for electronic structure in solution is implemented in the Amsterdam density functional (ADF) computational chemistry package [13–15]. Further, the orbital-free embedding (OFE) KS-DFT method with Wesolowski and Warshel's two-density functionals separately representing the electronic structure of the solute and that of environment [55] was coupled with 3D-RISM-KH for an improved description of effect of solution environment on the solute electronic structure properties such as photochemistry in solution; the OFE-DFT/3D-RISM-KH multiscale method is also implemented in the ADF package [15]. The KS-DFT(OFE-DFT)/3D-RISM-KH multiscale theory of electronic structure in solution has been extensively validated against experiment for thermochemistry, conformational equilibria and activation barriers for different solutes in various solvents [13–15], solvated carbon nanotubes [14], chemical reactions [14] and photochemistry [15] in solution, structure, and workfunction of metal–water interface [8], and structure of ionic liquids (ILs) [16].

Self-consistent field coupling of KS-DFT with 3D-RISM-KH

KS-DFT in the presence of solvent

The electronic structure of the solute is calculated from the self-consistent KS-DFT equations modified to include the presence of solvent. The whole system of the solute and solvent has the Helmholtz free energy defined as [8,11]

$$A[n_e(\mathbf{r}),\{\rho_\gamma(\mathbf{r})\}] = E_{solute}[n_e(\mathbf{r})] + \Delta\mu[\{\rho_\gamma(\mathbf{r})\};n_e(\mathbf{r})] \tag{19}$$

where E_{solute} is the internal energy of the solute macromolecule including the standard components of electronic structure theory, $\Delta\mu$ is the solvation free energy determined by the expression (12) which comes from the solute–solvent interaction and solvent reorganization around the solute macromolecule, $n_e(\mathbf{r})$ is the electron density distribution of the solute macromolecule, and $\rho_\gamma(\mathbf{r}) = \rho_\gamma g_\gamma(\mathbf{r})$ is the classical density distributions of solvent interaction sites γ of solvent molecules obtained from the 3D-RISM-KH integral equation. In the KS-DFT of electronic structure, the electronic internal energy of the solute is written in atomic units as

$$E_{solute}\left[n_e(\mathbf{r})\right] = T\left[n_e(\mathbf{r})\right] + \int d\mathbf{r} n_e(\mathbf{r})\upsilon_i(\mathbf{r}) + \int d\mathbf{r}\, d\mathbf{r}' \frac{n_e(\mathbf{r})n_e(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}\left[n_e(\mathbf{r})\right] \tag{20}$$

where $T[n_e(\mathbf{r})]$ is the kinetic energy of a non-interacting electron gas in its ground state with density distribution $n_e(\mathbf{r})$, $\upsilon_i(\mathbf{r})$ is the attractive potential of the nuclei, and $E_{xc}[n_e(\mathbf{r})]$ is the exchange-correlation energy. Minimization of the free energy functional (19) by functional variation with respect to the electron density

$$\frac{\delta A[n_e(\mathbf{r}),\{\rho_\gamma(\mathbf{r})\}]}{\delta n_e(\mathbf{r})} = 0 \tag{21}$$

subject to the normalization condition for N_e valence electrons of the solute

$$\int d\mathbf{r} n_e(\mathbf{r}) = N_e \tag{22}$$

leads to the self-consistent KS-DFT equation modified due to the presence of solvent [8,11,13,14]

$$\left[-\frac{1}{2}\Delta + \upsilon_i(\mathbf{r}) + \upsilon_H(\mathbf{r}) + \upsilon_{xc}(\mathbf{r}) + \upsilon_{solv}(\mathbf{r})\right]\psi_j(\mathbf{r}) = \varepsilon_j\psi_j(\mathbf{r}) \tag{23}$$

where the Hartree potential of the electrostatic interaction with the electron cloud is

$$\upsilon_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n_e(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{24}$$

the electron density is determined by summation over the N_e lowest occupied eigenstates with account of their double occupancy by electrons with opposed spins

$$n_e(\mathbf{r}) = \sum_{j=1}^{N_e} |\psi_j(\mathbf{r})|^2 \tag{25}$$

the exchange-correlation potential is the functional derivative

$$\upsilon_{xc}(\mathbf{r}) = \frac{\delta E_{xc}\left[n_e(\mathbf{r})\right]}{\delta n_e(\mathbf{r})} \tag{26}$$

and the solvent potential acting on the valence electrons of the solute is defined as

$$\upsilon_{\mathrm{solv}}(\mathbf{r}) = \frac{\delta \Delta\mu[n_{\mathrm{e}}(\mathbf{r}), \{\rho_\gamma(\mathbf{r})\}]}{\delta n_{\mathrm{e}}(\mathbf{r})} \tag{27}$$

In the ADF package implementation, calculation of $\upsilon_{\mathrm{H}}(\mathbf{r})$ is simplified by using the fitted density

$$n_{\mathrm{e}}(\mathbf{r}) \approx \sum_{j=1}^{N_{\mathrm{e}}} c_j f_j(\mathbf{r}) \tag{28}$$

expressed in terms of the single-center Slater functions f_j with the coefficients c_j determined by least-square fitting [56,57]. Substituting the expansion (28) in the Hartree form (24) and using the locality properties [58] dramatically reduces the computational load for evaluation of the potentials and matrix elements [13,14].

Finally, the total free energy is calculated as

$$A_{\mathrm{total}} = \sum_{j=1}^{N_{\mathrm{e}}} \varepsilon_j T - \frac{1}{2}\int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \frac{n_{\mathrm{e}}(\mathbf{r}) n_{\mathrm{e}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{\mathrm{xc}}[n_{\mathrm{e}}(\mathbf{r})] - \int \mathrm{d}\mathbf{r} n_{\mathrm{e}}(\mathbf{r}) \upsilon_{\mathrm{xc}}(\mathbf{r}) + \Delta\mu[\{\rho_\gamma(\mathbf{r})\}; n_{\mathrm{e}}(\mathbf{r})] - \int \mathrm{d}\mathbf{r} n_{\mathrm{e}}(\mathbf{r}) \upsilon_{\mathrm{solv}}(\mathbf{r}) \tag{29}$$

Effective potentials coupling the electronic and classical subsystems

With the free energy functional in the form (19) allowing functional differentiation analytically, the effective potentials of the SCF coupling the electronic and classical subsystems are obtained in the closed analytical form.

The classical effective potential of the whole solute acting on solvent site γ

$$u_\gamma(\mathbf{r}) = u_\gamma^{\mathrm{sr}}(\mathbf{r}) + q_\gamma(\phi^{\mathrm{n}}(\mathbf{r}) + \phi^{\mathrm{e}}(\mathbf{r})) \tag{30}$$

is broken up [8,11,13,14] into the short-range interaction term $u_\gamma^{\mathrm{sr}}(\mathbf{r})$ and the electrostatic energy of the solvent site charge q_γ in the electrostatic potential of the solute valence electrons $\phi^{\mathrm{e}}(\mathbf{r})$ and that of nuclei charges Z_i

$$\phi^{\mathrm{n}}(\mathbf{r}) = \sum_i \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} \tag{31}$$

The short-range interaction is given by the sum of the pairwise contributions from the solute nuclei represented by the 12-6 Lennard–Jones potential

$$u_\gamma^{\mathrm{sr}}(\mathbf{r}) = \sum_i 4\varepsilon_{i\gamma}\left[\left(\frac{\sigma_{i\gamma}}{r_{i\gamma}}\right)^{12} - \left(\frac{\sigma_{i\gamma}}{r_{i\gamma}}\right)^{6}\right] \tag{32}$$

where $\sigma_{i\gamma}$ and $\varepsilon_{i\gamma}$ are the Lennard–Jones cross-term diameter and energy parameters obtained from those of solute nucleus i and solvent site γ by using the standard mixing riles $\sigma_{i\gamma} = (\sigma_i + \sigma_\gamma)/2$ and $\varepsilon_{i\gamma} = (\varepsilon_i \varepsilon_\gamma)^{1/2}$, and $r_{i\gamma} = |\mathbf{r}_\gamma - \mathbf{R}_i|$ is the nucleus-site separation. Within the ADF package implementation, the potential of valence electrons acting on a single solvent site $\phi^{\mathrm{e}}(\mathbf{r})$ is calculated from the valence electron density $n_{\mathrm{e}}(\mathbf{r})$ in the density fitting procedure (28) [57].

With the solvation free energy given by the 3D-RISM-KH analytical expression (12), the effective potential of the solvent acting on the solute valence electrons given by the functional derivative (27) is derived as [8,11,13,14]

$$\upsilon_{\mathrm{solv}}(\mathbf{r}) = \frac{\delta\Delta\mu}{\delta n_{\mathrm{e}}(\mathbf{r})} = \sum_{\gamma} \rho_{\gamma} \int \mathrm{d}\mathbf{r}' h_{\gamma}(|\mathbf{r}-\mathbf{r}'|)\upsilon_{\gamma}^{\mathrm{ps}}(\mathbf{r}') \tag{33}$$

where $\upsilon_{\gamma}^{\mathrm{ps}}(\mathbf{r})$ is the contribution of solvent site γ to the pseudopotential of a solvent molecule acting on an external electron, which is in general given by the variational derivative of the classical site potential with respect to the valence electron density [8,11]

$$\upsilon_{\gamma}^{\mathrm{ps}}(\mathbf{r}-\mathbf{r}') = \frac{\delta u_{\gamma}(\mathbf{r})}{\delta n_{\mathrm{e}}(\mathbf{r}')} \tag{34}$$

The potential (33) implies the mean field approximation, which follows essentially from the use of the solvation free energy in the form (12).

Analytical gradients in solution

The analytical first derivative with respect to the coordinates of the solute nuclei $\mathbf{R}_i$ is obtained by differentiation of the whole system free energy (19)

$$\frac{\mathrm{d}A[n_{\mathrm{e}}(\mathbf{r}),\{\rho_{\gamma}(\mathbf{r})\}]}{\mathrm{d}\mathbf{R}_i} = \frac{\mathrm{d}E_{\mathrm{solute}}[n_{\mathrm{e}}(\mathbf{r})]}{\mathrm{d}\mathbf{R}_i} + \frac{\mathrm{d}\Delta\mu[\{\rho_{\gamma}(\mathbf{r})\};n_{\mathrm{e}}(\mathbf{r})]}{\mathrm{d}\mathbf{R}_i} \tag{35}$$

where the former term has the same structure as in the gas-phase case [60] and the latter term is derived from the solvation free energy (12) as follows

$$\frac{\mathrm{d}\Delta\mu}{\mathrm{d}\mathbf{R}_i} = \sum_{\gamma} \rho_{\gamma} \int \mathrm{d}\mathbf{r} g_{\gamma}(\mathbf{r}) \frac{\partial u_{\gamma}^{\mathrm{sr}}(\mathbf{r})}{\partial \mathbf{R}_i} + \int \mathrm{d}\mathbf{r} \upsilon_{\mathrm{solv}}(\mathbf{r}) \frac{\partial n_{\mathrm{e}}(\mathbf{r})}{\partial \mathbf{R}_i} + q_i \left.\frac{\partial \upsilon_{\mathrm{solv}}(\mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{R}_i} \tag{36}$$

Calculation of the expression (36) requires little computational effort. The second term is calculated together with the gradients of the exchange-correlation potential [60], and the rest is calculated in the 3D-RISM procedure. Notice that the first term in the square brackets does not contribute much to the gradients because the large magnitude of the derivatives of the short-range potential $u_{\gamma}^{\mathrm{sr}}(\mathbf{r})$ at the repulsive core edge are suppressed by the distribution function $g_{\gamma}(\mathbf{r})$ exponentially decaying there.

Example of multiscale KS-DFT/3D-RISM-KH calculations

Novel applications of green chemistry involve ILs as tunable solvent environment, including organic synthesis and other applications such as electrodeposition. Properties of molecules solvated in ILs, including ions constituting IL itself are strongly affected by solvent environment. In bulk liquid, the solvent environment affects physicochemical characteristics of IL constituents, including NMR chemical shifts, relaxation times, IR frequencies, as well as such important chemical behavior as acidity of aromatic protons. To give reliable results, ab initio calculations for such systems have to self-consistently account for the change in both electronic and classical solvation structure. This can be achieved in principle by using a Car–Parrinello molecular dynamics (CPMD) method in which MD is driven by the forces obtained from electronic structure DFT for the whole system or at least by using a quantum mechanics/molecular dynamics (QM/MD) approach in which QM treats one selected ion and MD handles the classical solvation structure of IL. However, both the approaches, particularly the former, are extremely computationally demanding, especially to obtain the solvation thermodynamics. A simpler alternative is a QM/MM approach; however, for many systems, MM cannot adequately represent the solvation structure and thermodynamics in the statistical ensemble average. Meanwhile, the 3D-RISM-KH molecular theory of solvation properly resolves a 3D spatial map of solvation structure and reliably describes solvation effects for macromolecules of complex geometry and different chemical specificities, such as hydrogen bonding and/or solvophobic solvation and solvophobic interaction. The SCF KS-DFT/3D-RISM-KH multiscale method provides a first-principle physical view on electronic struc-

ture in solution and thus offers an accurate and computationally efficient procedure to perform ab initio calculations on species solvated in ILs.

Figure 7 depicts the KS-DFT/3D-RISM-KH results for the solvation structure of the methyl-methyl imidazolium ion in bulk liquid of [mmim][Cl] at T = 400 K [16]. The method predicts the IL properties in remarkable agreement with the conclusions drawn from CPMD simulations for the 3D solvation structure and the solvent environment effect of the IL constituents. This includes such effects as the polarization due to electronic effects in the IL environment, the dipole moment enhancement from μ = 2.10 D for the isolated cation to 2.59 D in the bulk IL (vs. μ = 2.10 D to 2.59 D from CPMD), the antiparallel stacking of adjacent cations with the central CR hydrogens of neighboring sites pointing in the opposite directions, the coordination number of 7.4 anions around the cation (vs. 7.5 from CPMD), and the solvation free energy of –22.1 kcal/mol for the $[mmim]^+$ cation and –49.1 kcal/mol for the Cl^- anion in the IL [16].

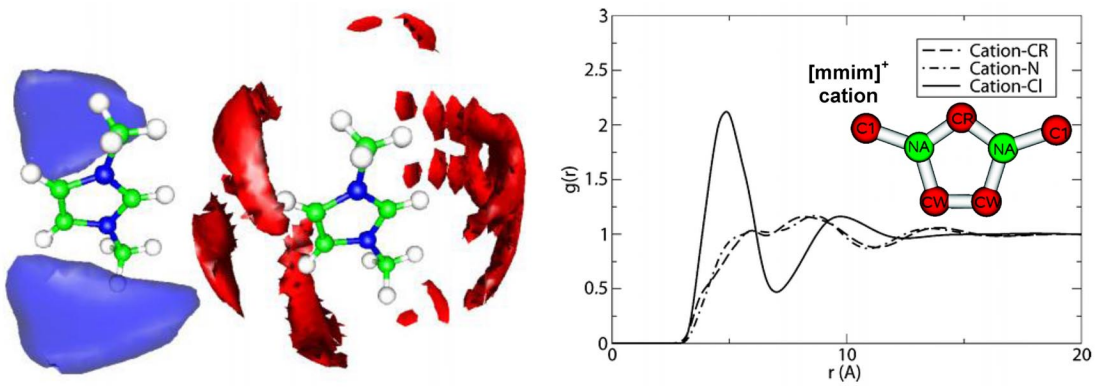

Fig. 7 Solvation structure of the methyl–methyl imidazolium ion in bulk liquid of [mmim][Cl] at T = 400 K, obtained by the SCF KS-DFT/3D-RISM-KH multiscale theory of electronic structure in solution. Left part: Isosurfaces of the nitrogen of $[mmim]^+$ cations at $g_N(r) > 2$ (in blue) and of Cl^- anions at $g_{Cl}(r) > 5$ (in red). Right part: RDFs of Cl^- anions (solid line), cation CR site (dashed line), and cation N site (dash–dotted line) around the center of mass of the $[mmim]^+$ cation in the IL.

Another example of the performance of the KS-DFT/3D-RISM-KH method at the level of CPMD simulation is the self-consistent electronic and classical solvation structure of a (100) copper–water interface [8]. The predicted shift of the Fermi level of the metal due to the presence of water matches the values typically found in experiment. Dense water substantially affects the electron distribution around a water molecule adsorbed at the metal surface and changes the metal–water effective potential. The latter follows the shapes of the metal effective electrostatic potential which for this interface is strongest at the hollow site adsorption positions. Therefore, both CPMD and KS-DFT/3D-RISM-KH predict the 3D water density distributions with the maxima at the hollow and bridge adsorption positions, unlike classical MD predicting the maxima at the on-top site positions. The layering and orientations of water molecules near the surface found with DFT/3D-RISM-KH are in agreement with experiment.

The above examples demonstrate the capability of the KS-DFT/3D-RISM-KH multiscale method to predict electronic structure in solution, including solid–liquid interfaces, essentially at the level of CPMD simulation.

MULTIPLE TIME-STEP MOLECULAR DYNAMICS OF BIOMOLECULES IN THE EFFECTIVE SOLVATION POTENTIAL OBTAINED BY 3D-RISM-KH

The statistically averaged effective forces of solvent acting of each solute atom produced by the 3D-RISM-KH molecular theory of solvation can then be used together with the direct solute–solute interactions to drive MD simulation of the solute macromolecule. Solving the 3D-RISM integral equations is computationally expensive compared to a single MD step, and it naturally leads to resorting to multiple time-step (MTS) MD techniques. One of the first attempts toward the hybrid MD/3D-RISM method [61] when studying acetylacetone in aqueous solution, the consideration was limited to the reference system propagator algorithm (RESPA) in the microcanonical ensemble, making it possible to perform outer time-steps up to 5 fs. The outer time-step in microcanonical simulations of ambient liquids cannot exceed 4–8 fs due to resonance instabilities arising due to an interplay in an MTS procedure between strong intramolecular forces handled with an inner spacing of 0.5–1 fs and weak long-ranged intermolecular interactions. It was proven that in the microcanonical ensemble the maximum allowed length of the outer time-step cannot exceed the theoretical limit of 20 fs [62]. Therefore, such an ensemble is very inefficient for complex systems, including large macromolecules in solution, characterized by time scales spanning up to micro- and milliseconds.

It is often much more convenient for compatibility with experiment to sample configurations from the canonical (constant temperature) ensemble, instead of the microcanonical (constant energy) one. Many thermostats have been devised to handle the canonical distributions within MTS dynamics. In the popular Langevin (LN) approach [63–67], artificial friction and random forces are incorporated to stabilize the solutions. Although these forces satisfy the fluctuation-dissipative theorem, the target temperature is not guaranteed. On the other hand, employing a large viscosity coefficient to ensure the stability can deviate the system from the true canonical state corresponding to real interactions between particles. In addition, the LN dynamics does not possess any conserved quantity, and so it is difficult to examine the quality of the trajectories obtained. The latter drawback is absent in the well-established Nosé–Hoover (NH) chain method [68–71]. Here, the canonical behavior is reproduced by introducing extra phase-space variables associated with a thermostat, while the ergodicity is ensured by its chain counterparts. A famous feature of this method is that the temperature can be controlled without involving random numbers. Moreover, the NH equations of motion can be derived within the Hamiltonian formalism using the extended energy function which is conserved during the time evolution. A great advantage of the canonical description is the possibility to postpone the appearance of unphysical MTS effects to outer time-steps significantly longer than in the case of the microcanonical ensemble. For instance, the maximal applicable time-steps can reach 50 fs [72] within the NH thermostat and can further increase up to 75 fs [72] in the isokinetic ensemble [73]. Here, employing Gauss' principle of least constraint, the total kinetic energy is kept strictly constant. This causes a modification of the real dynamics but holds the correct canonical distributions of position-dependent functions. A better efficiency can be observed by combining the canonical NH chain method [68] with the isokinetic ensemble [73]. This yields the so-called isokinetic NH chain RESPA (INR) approach [74,75] in which the heat baths are coupled individually to each degree of freedom in the system. The INR algorithm was applied to MD simulations of water to certify that the outer time-step size of 100 fs is workable [74].

The coupled MD/3D-RISM-KH approach was extended to the canonical ensemble by employing the LN thermostat [76]. Using the solvent-induced forces calculated by 3D-RISM-KH at outer time-steps and the solvation force coordinate extrapolation (SFCE) at inner time-steps of LN, it was demonstrated on an example of alanine dipeptide in aqueous solution that larger outer time-steps up to 20 fs are possible with the LN/SFCE/3D-RISM-KH approach. Such steps, however, are still smaller than 100 fs ones achievable within MD simulation using the INR algorithm [74].

A more general formulation of the canonical–isokinetic NH chain approach [77] consists of a special splitting of the total kinetic energy into its partial components. Such components are first collected

into groups with a chosen number of either translational degrees of freedom of atoms or translational, orientational, and vibrational degrees of molecules. Then each of these groups is coupled to its own set of chain thermostats characterized with given lengths and relaxation times. This allows us to optimize and further significantly increase in efficiency the integration of motion, resulting in the optimized isokinetic Nosé–Hoover (OIN) chain algorithm [77]. The standard isokinetic [73] or INR [74,75] ensembles appear from the generalized formulation as particular cases. With the stabilizing effect of OIN thermostatting in MTS-MD, gigantic outer time-steps up to picoseconds can be employed to accurately calculate equilibrium and conformational properties, both in pure OIN and multiscale OIN/SCFE/3D-RISM-KH simulations [77]. In the atomic version of OIN for MTS-MD of a biomolecule in a solvent PMF, the solvation forces are obtained analytically by converging the 3D-RISM-KH integral equations once per several OIN outer time-steps, and are calculated in between by using SFCE in the subspace of previous successive solutions to 3D-RISM-KH. While the computational rate of solvent sampling in OIN/SFCE/3D-RISM-KH is already about 20 times faster than standard MD with explicit solvent, further substantial acceleration of sampling stems from making solute evolution steps in a statistically averaged PMF obtained from 3D-RISM-KH. The latter efficiently samples the phase space for essential events with rare statistics such as exchange and localization of solvent and ligand molecules in confined spaces, pockets, and at binding sites of the solute macromolecule, as distinct from MD with explicit solvent, which requires enormous computational time and number of steps in such cases.

Calculation of solvation forces by 3D-RISM-KH coupled with MTS-MD

In the coupled MTS-MD/3D-RISM-KH method implemented in the Amber MD package [76,77], MD is applied to the biomolecule driven by effective solvation forces which are obtained analytically by the 3D-RISM-KH molecular theory of solvation. 3D-RISM-KH yields the solvation structure of the biomolecule, the solvation free energy landscape dependent on its conformation, and the corresponding effective solvation forces acting on its interaction sites. The latter are derived by differentiating the Kirkwood thermodynamic integration formula analytically, also valid for the solvation free energy (12) in the 3D-KH approximation (2)

$$\mathbf{f}_i(\mathbf{R}_i) \equiv -\frac{\mathrm{d}\Delta\mu}{\mathrm{d}\mathbf{R}_i} = \sum_\gamma \rho_\gamma \int \mathrm{d}\mathbf{r}\; g_\gamma(\mathbf{r}) \frac{\partial u_{i\gamma}(\mathbf{r}-\mathbf{R}_i)}{\partial \mathbf{R}_i} \tag{37}$$

where $u_{i\gamma}(\mathbf{r} - \mathbf{R}_i)$ is the pairwise interaction potential between biomolecule site i located at position $\mathbf{R}_i$ and solvent site γ at position $\mathbf{r}$. To obtain meaningful sampling of solute conformations, the computational expense of the 3D-RISM-KH calculations is reduced with several optimization strategies: (i) creating a high-quality initial guess for the direct correlation function $c_\gamma(\mathbf{r})$ from previous successive solutions; (ii) accelerating pre- and post-processing of the solute–solvent potentials, long-range asymptotics, and forces by using a cutoff scheme and a varying solvation box of smallest size containing two to three solvation shells around the current conformation of the protein; and (iii) avoiding direct calculation of solvation forces (37) with solving 3D-RISM-KH at every inner time-step δt but rather doing that at each outer time-step $h >> \delta t$ and extrapolating the solvation forces at inner time-steps in between based on previous successive vectors of solvation forces calculated with 3D-RISM-KH.

The extrapolation of solvation forces can be performed by using the so-called force-coordinate approximation [76,77], or the SFCE. The main idea is that solvation forces acting on each atom of the solute macromolecule do not have strong repulsive cores and vary smoothly in space and time with solute evolution, and therefore can be extrapolated at subsequent inner time-steps $\delta t << h$ until the next 3D-RISM-KH calculation at outer time-step h with sufficient accuracy by using the solvation forces and solute atomic coordinates at previous outer time-steps h. Let $\mathbf{f}^{(k)} = \{\mathbf{f_i}^{(k)}\}$ be a $3M$-dimensional vector of effective solvation forces acting on all solute atoms $i = 1, \ldots, M$ at locations specified by a $3M$-dimensional vector of coordinates $\mathbf{R}^{(k)} = \{\mathbf{R_i}^{(k)}\}$ for each of $k = 1, \ldots, K$ previous successive outer

time-steps of length h at which the solvation forces were calculated by using 3D-RISM-KH. Then the solvation forces $\mathbf{f}(t)$ at current time t within the next outer time-step interval can be extrapolated in the subspace of the previous $k = 1, \ldots, K$ vectors of forces by the linear combination

$$\mathbf{f}(t) = \sum_{k=1}^{K} a_k(t)\mathbf{f}^{(k)} \tag{38}$$

where the weight coefficients $a_k(t)$ are obtained by best representing the vector of solute atomic coordinates $\mathbf{R}(t)$ at time t in terms of its projections onto the "basis" of the previous ones $\mathbf{R}^{(k)}$, $k = 1, \ldots, K$ to minimize the norm of the residual

$$\min \left| \mathbf{R}(t) - \sum_{k=1}^{K} a_k(t)\mathbf{R}^{(k)} \right| \tag{39}$$

A further improvement of SFCE [77] over its genuine version [76] is achieved with imposing an additional condition of normalization on the weight coefficients of the expansion (38)

$$\sum_{k=1}^{K} a_k = 1 \tag{40}$$

The solute–solvent forces (36) can be expanded in the power series of a deviation of the current coordinate vector $\mathbf{R}$ from the closest basis point $\mathbf{R}^{(*)} \in \{\mathbf{R}^{(k)}\}$

$$\mathbf{f}(\mathbf{R}) = \mathbf{f}(\mathbf{R}^{(*)}) - \left.\frac{\partial \mathbf{f}}{\partial \mathbf{R}}\right|_{\mathbf{R}^{(*)}} \left(\mathbf{R} - \mathbf{R}^{(*)}\right) + O\left(\mathbf{R} - \mathbf{R}^{(*)}\right)^2 \tag{41}$$

where $\partial\mathbf{f}/\partial\mathbf{R}$ denotes the Hessian matrix, and the second and higher-order spatial inhomogeneities are neglected. Imposing the condition (40) automatically reproduces the zero-order term of the expansion, while the first-order one linear in coordinate is extrapolated by minimizing the coordinate norm (39). Note that putting $K = 1$ reduces to the simplest case of constant-force extrapolation. The minimization procedure (39) with the normalization condition (39) leads to the set of $K + 1$ linear equations for the weight coefficients

$$\begin{pmatrix} S_{11} & \cdots & S_{1K} & -1 \\ \vdots & & \vdots & \vdots \\ S_{K1} & \cdots & S_{KK} & -1 \\ -1 & \cdots & -1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_K \\ \lambda \end{pmatrix} = \begin{pmatrix} S_{1t} \\ \vdots \\ S_{Kt} \\ -1 \end{pmatrix} \tag{42}$$

where S_{kl} is the inner scalar product of the $3M$-dimensional vectors of atomic coordinates for solute conformations $\mathbf{R}^{(k)}$ and $\mathbf{R}^{(l)}$ at outer time-steps k and l, as well as S_{kt} is that for $\mathbf{R}^{(k)}$ and running conformation $\mathbf{R}(t)$

$$\begin{aligned} S_{kl} &= \left\langle \mathbf{R}^{(k)} \middle| \mathbf{R}^{(l)} \right\rangle \equiv \sum_{i=1}^{M} \left(\mathbf{R}_i^{(k)} \cdot \mathbf{R}_i^{(l)}\right) \\ S_{kt} &= \left\langle \mathbf{R}^{(k)} \middle| \mathbf{R}(t) \right\rangle \equiv \sum_{i=1}^{M} \left(\mathbf{R}_i^{(k)} \cdot \mathbf{R}_i(t)\right) \end{aligned} \tag{43}$$

and λ is a Lagrangian multiplier yielding the squared norm of the minimized approximation residual (38). At each inner step δt within the outer time-step interval h, eqs. 42 are being solved by using standard numerical methods of solving a set of linear equations and the extrapolation (38) is being computed. After the next 3D-RISM-KH calculation, the basis set is updated with the new vector of forces

and coordinates and the oldest one is discarded. In practical calculations, we found optimal to perform extrapolation with about $K = 20$ coordinate-force pairs. Further increase of the basis size gives no significant improvement in the extrapolation accuracy since the current conformation and forces get too different from the earliest ones in a larger basis set, but eventually lead to accumulation of numerical errors when eqs. 42 become ill-conditioned for larger K.

Illustrations for the MTS-MD/SFCE/3D-RISM-KH method on biomolecular solvation

A simple system that nevertheless has some essential biomolecular properties and thus allows one to test and validate simulation and modeling methods is alanine dipeptide in aqueous solution. This simple molecule has often been used as a model in theoretical studies of backbone conformational equilibria in proteins [83]. For hydrated alanine dipeptide, there are several transitions between different conformational states characterized by the mean life time of order varying from 30 ps, through 250 ps, and up to 10 ns [84]. Thus, even for such a relatively simple organic molecule as alanine dipeptide, a long observation time of at least $t \approx 100$ ns is necessary to gain proper statistic and average out the statistical noise in explicit solvent MD simulation. The accuracy of the MTS-MD/SFCE/3D-RISM-KH algorithms was estimated by measuring the dipole moment distribution function $f(p)$ of the hydrated alanine dipeptide [77]. Such a function is very sensitive to the conformational sampling and the effect of solvent, and comparing $f(p)$ with its "exact" counterpart is a reliable way of testing an MD integrator. To estimate the accuracy of the MTS thermostats and SFCE approach, the "expected" values of $f(p)$ were obtained by using the LN/3D-RISM-KH run with no solvent force extrapolation and a small friction coefficient $\gamma = 5$ fs with small inner time-step $\delta t = 1$ fs for MD and outer time-step $\Delta t = 4$ fs for solvation forces from 3D-RISM-KH in order to minimize all possible uncertainties.

Figure 8 presents the dipole moment distribution $f(p)$ of the alanine dipeptide molecule in cSPC/E water [76] obtained with the MTS-MD/SPCE/3D-RISM-KH integrators vs. the reference run with LN/3D-RISM-KH [77]. The "expected" dipole moment distribution function $f(p)$ has two clear peaks on the left at $p \approx 3$ D and on the right at $p \approx 7.5$ D, as well as an enhancement in the intermediate region at $p \approx 5.5$ D; a similar pattern is observed experimentally in the real system of hydrated alanine dipeptide, which stems from the coexistence of different conformational states [84]. The SFCE performed within the OIN and INR ensembles for up to $\kappa = 20$ time-steps provides an accuracy high enough at $h \leq 200$ fs for OIN and up to 1 ps for INR. The deviations become visible for OIN at $h = 400$ fs (solid cyan curve) and achieve a value of about 10 %, which can still be acceptable. When OIN/3D-RISM-KH with constant force extrapolation (referred to as CIN) is employed, i.e., $\kappa = 1$ and the solvation forces stay the same during the full outer time interval h instead of the SFCE with $\kappa = 20$, a much larger uncertainty of 30 % is exhibited in CIN/3D-RISM-KH already at $h = 200$ fs (dashed magenta curve). Such uncertainties exceed even those in the INR/SFCE algorithm at $h = 2000$ fs (dashed black curve). Approximately the same level of accuracy of 30 % is inherent in the LN/SFCE/3D-RISM-KH integration at $h = 96$ fs, despite the use of the extrapolation with the same $\kappa = 20$. This indicates an evident advantage of the generalized canonical–isokinetic OIN/INR approach over the LN scheme. Note that the maximal acceptable outer time-step reported earlier [76] was $h = 20$ fs, as achieved in the LN/SFCE/3D-RISM-KH scheme with the extrapolation. This time-step size is more than one order of magnitude smaller than steps of $h = 200$–400 fs feasible with the OIN/SFCE/3D-RISM-KH algorithm. It is worth recalling that the outer time-step in the explicit microcanonical MTS-MD/3D-RISM simulations cannot exceed 5 fs [50,51]. Thus, the use of the generalized canonical–isokinetic OIN ensemble coupled with the 3D-RISM-KH molecular theory of solvation supplemented with the SFCE in the multiscale OIN/SFCE/3D-RISM-KH integrator allowed this value to be exceeded by at least 40–80 times.

Ion channels belong to an important class of biomolecular systems where solvent (ions, protons, water, and ligands) performs the biological functions in confinement. Ion translocation through the channel is a rare event on the time scale of bulk solvent relaxation, and explicit solvent simulations

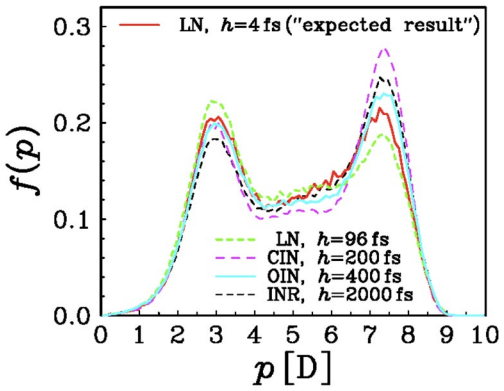

Fig. 8 Distribution of conformations of hydrated alanine dipeptide over its dipole moment. Results of the multiple time scale integrators: LN/SFCE/3D-RISM-KH at outer time-step h = 96 fs (solid green); CIN/3D-RISM-KH at h = 200 fs (dashed magenta); OIN/SFCE/3D-RISM-KH at h = 400 fs (solid cyan), and INR/SFCE/3D-RISM-KH at h = 2000 fs (dashed black), vs. the "expected result" by LN/3D-RISM-KH (solid red curve).

require extensive computations to achieve statistically reliable results. Furthermore, study of selectivity mechanisms and permeation properties of channels under physiological conditions may require modeling of complex solvent conditions, including low ionic concentration environments, which is currently beyond the scope of explicit solvent MD simulations. To alleviate these difficulties, coarse-grained models of a channel were proposed and used to study the physical properties and functioning of ionic channels [80–82], such as the importance of protein polarization and side chains mobility on the selectivity of sodium channels [80]. However, they neither account for the atomic structure of the real ion channels nor allow one to include in a transferable manner more complex solvent compositions with different cofactors, ligands, and other small molecules. This can be done with the 3D molecular theory of solvation which provides a natural bridge between the coarse-grained description and all-atom explicit solvent MD simulations.

The structure of the bacterial *G. violaceus* pentameric ligand-gated ion channel (GLIC) homologue in an open conformation was resolved in the X-ray experiments [78]. It was also demonstrated that this channel is sensitive to general anesthetics [79]. Figures 9 and 10 present the results for the solvation structure of the GLIC channel inserted in the 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer obtained by the LN/SFCE/3D-RISM-KH modeling [29,30]. The 3D density distributions obtained from 3D-RISM-KH for water oxygen (O) and sodium (Na^+) and chloride (Cl^-) ions of solvent around the channel protein embedded in the cellular membrane are shown in Fig. 9; and those of Na^+ and Cl^- are detailed in Fig. 10 for the top, bottom, and side views of the channel with isosurfaces of $g(\mathbf{r}) \geq 5$. It is seen that the Cl^- anions are mostly expelled from the pore of this cationic channel, whereas there are nonzero densities of Na^+ ions throughout the channel pore, in agreement with experimental data. The solvent distributions in the proximity of the protein are characterized by some features. For example, the preferable solvation sites for Cl^- ions are located along the external ring of the channel funnel, facing the extracellular space above the membrane. Inside the funnel, right before the entrance to the channel pore, the distributions of water and Na^+ ions are complementary. There is a pocket of Cl^- ions which separates the upper and lower parts of the funnel, the area above the level of the lipid bilayer. The lower part being adjacent to the channel pore is characterized by the preferable solvation of Na^+ ions much as the rest of the channel pore. The exit from the channel pore from the intracellular side below the membrane is surrounded by the solvation sites of Cl^- ions. In general, both the extra- and intracellular surfaces of the lipid bilayer are characterized by the enhanced densities of the anions. The sodium distribution inside the channel pore correlates with the channel radius and the pattern in the distributions of the hydrophobic and hydrophilic residues in the channel pore. The above

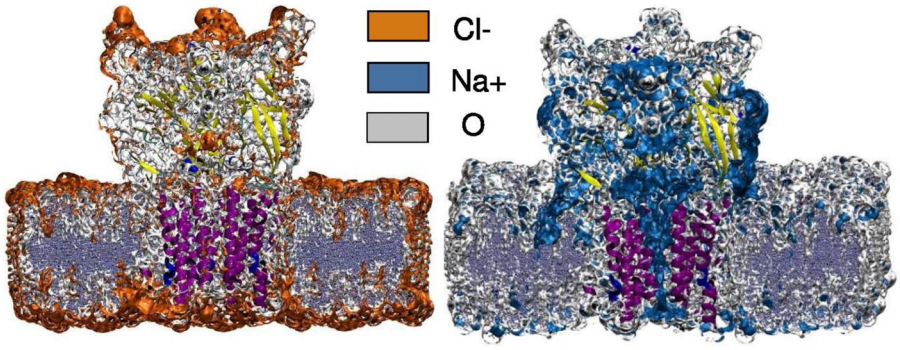

Fig. 9 3D distributions of water oxygen (in gray, both parts), Cl^- (in orange, left part) and Na^+ (in blue, right part) ions around the GLIC channel embedded in the POPC lipid bilayer membrane. The channel protein is shown as β-sheet (in yellow) and α-helices (in magenta), and the lipids are shown in the ball-and-stick representation.

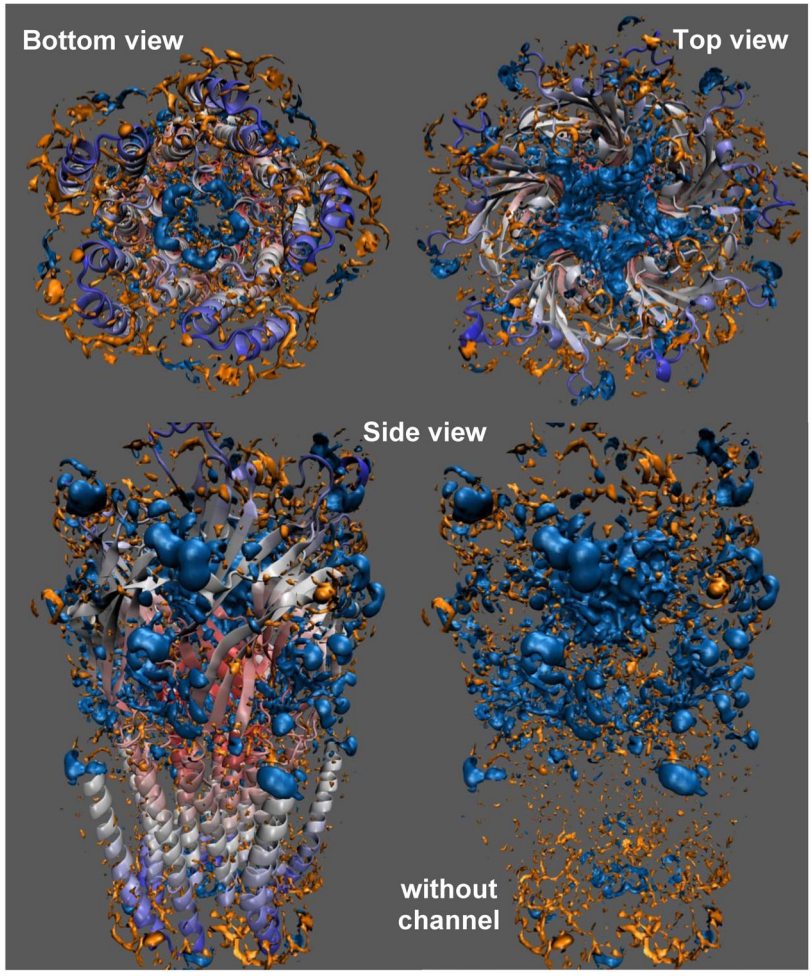

Fig. 10 3D distribution functions of Na^+ (in blue) and Cl^- (in orange) ions inside the GLIC ion channel, obtained from 3D-RISM-KH. Isosurfaces of $g(\mathbf{r}) \geq 5$. Bottom and top view of the channel (upper left and right parts), and side view with and without the channel protein (lower left and right parts, respectively).

illustrates how the ion charge distribution over the pore space contributes to the selectivity and permeation properties of the ion channel.

MOLECULAR THEORY OF SOLVATION AND ELECTRICAL DOUBLE LAYER IN NANOPOROUS MATERIALS

The properties of an electric double layer (EDL) formed in inner spaces of nanoporous electrodes are very different from a conventional, planar electrochemical capacitor of equal surface area due to an overlap of adjacent EDLs. The EDL at the surface of carbon nanopores gets substantially distorted, compared to that at a planar electrode in contact with electrolyte solution, resulting in the specific capacitance per surface area by up to 1–2 orders of magnitude less than for the planar electrode. At present, molecular simulation description of these effects is virtually unfeasible due to the interplay of long-range electrostatic and short-range steric interactions on large-space and long time scales, and the necessity to satisfy the conditions of chemical and mechanical balance between the species in the bulk electrolyte solution and those sorbed in the nanoporous electrodes.

A generalization of RISM molecular theory of solvation to solutions sorbed in disordered nanoporous materials using the replica method for statistical mechanics of quenched–annealed systems, the so-called replica RISM-KH-VM theory, provides full microscopic details of the solvation structure and thermodynamics averaged over the thermal motion of sorbed solution and over the quenched distribution of host nanoporous material (morphology of nanopores) [18,85,86]. This theory enables predictive molecular modeling of thermochemistry and electrochemistry of electrolyte solutions sorbed in functionalized nanoporous materials. In particular, it reveals the mechanisms of sorption and supercapacitance in nanoporous carbon electrodes [18,85–89]. The replica RISM-KH-VM theory predicts, from the first principles of statistical mechanics, such effects as solvent-specific wetting [18] and water depletion in hydrophobic carbon nanopores [87–89], asymmetry in solvation and adsorption of cations and anions [86–89], desalination of simple ions in hydrophobic nanopores and its reversal with external voltage applied [86–89], efficient removal of ionic impurities from an aqueous waste stream by a nanoporous carbon electrosorption cell [88], and specific adsorption at chemical functionalities on the surface of carbon nanopores [85].

Replica formalism of statistical mechanics for fluid sorbed in a disordered matrix

The replica formalism of statistical–mechanical, integral equation theory of quenched–annealed systems treats "annealed" fluid with equilibrium temperature T_1 (species 1) sorbed in a porous matrix of "quenched" particles with a spatial distribution corresponding to an equilibrium ensemble with temperature T_0 (species 0). Note that in general $T_0 \neq T_1$ because the matrix and liquid subsystems are not in equilibrium and do not exchange with energy. This has implications on the thermodynamic expressions in the replica RISM theory involving density and temperature derivatives of matrix material [85,86]; however, for the purpose of typical systems and calculations with rigid unchangeable matrix, T_0 affects only its structure obtained beforehand and can be reduced to $T_0 = T_1 = T$ at the beginning by scaling the interaction potential used for the matrix model (if this is not a purely hard core model). The mean free energy of the sorbed annealed fluid is obtained as a statistical average of the free energy with the canonical partition function $Z_1(\mathbf{q}_0)$ of the fluid sorbed in the matrix with a particular spatial configuration of quenched particles $\mathbf{q}_0$ over the ensemble of all realizations of matrix configurations $\mathbf{q}_0$

$$\overline{A}_1 = A_1^{\mathrm{id}} - k_{\mathrm{B}}T_1 \left\langle \ln Z_1\left(\mathbf{q}_0\right)\right\rangle_{\mathbf{q}_0} \tag{44}$$

where A_1^{id} is the ideal gas free energy. The statistical average of a logarithm is not amenable to standard evaluation and is obtained by using the so-called replica identity similar to theory of spin glasses that

relates the logarithm to the analytic continuation of moments Z^s as follows: $\ln Z_1 = \lim_{s\to 0} \mathrm{d}Z^s/\mathrm{d}s$. For integer values of s, the statistical average of the moments takes the form of the equilibrium canonical partition function of a fully annealed $(s + 1)$-component liquid mixture consisting of matrix species 0 and s equivalent replicas of fluid species 1, with no interaction between the fluid replicas. The average free energy of the annealed fluid is then obtained in the assumption of no replica symmetry breaking in the analytic continuation of the free energy of the annealed replicated system $A_{\mathrm{rep}}(s)$ as

$$\overline{A}_1 = \lim_{s\to 0} \frac{\mathrm{d}A_{\mathrm{rep}}(s)}{\mathrm{d}s} \tag{45}$$

Using the above formalism to evaluate in terms of Meyer diagrams the free energy and statistical sum of "mobile" atomic fluid sorbed in a "frozen" matrix of spherical obstacles generates the so-called replica OZ integral equations for a quenched-annealed atomic system [90–93]. This formalism was generalized to replica RISM for simple and complex associating molecular systems [18,85,86].

Replica DRISM-KH-VM molecular theory for electrolyte solution sorbed in nanoporous material

Applying this replica formalism to the dielectrically consistent reference interaction site model (DRISM) integral equations of Perkyns and Pettitt [38], we get after some algebra the replica DRISM integral equations for annealed molecular liquid (or equally for liquid mixture) sorbed in a quenched matrix

$$h^{00}_{\alpha\gamma}(r) = \omega^{00}_{\alpha\mu}(r) * c^{00}_{\mu\nu}(r) * \omega^{00}_{\nu\gamma}(r) + \omega^{00}_{\alpha\mu}(r) * c^{00}_{\mu\nu}(r) * \rho^0_\nu h^{00}_{\nu\gamma}(r) \tag{46a}$$

$$h^{10}_{\alpha\gamma}(r) = \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{10}_{\mu\nu}(r) * \omega^{00}_{\nu\gamma}(r) + \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{10}_{\mu\nu}(r) * \rho^0_\nu h^{00}_{\nu\gamma}(r) + \tilde{\omega}^{11}_{\alpha\mu} r) * c^{(\mathrm{c})}_{\mu\nu}(r) * \rho^1_\nu h^{10}_{\nu\gamma}(r) \tag{46b}$$

$$\begin{aligned} h^{01}_{\alpha\gamma}(r) &= \omega^{00}_{\alpha\mu}(r) * c^{01}_{\mu\nu}(r) * \tilde{\omega}^{11}_{\nu\gamma}(r) + \omega^{00}_{\alpha\mu}(r) * c^{00}_{\mu\nu}(r) * \rho^0_\nu h^{01}_{\nu\gamma}(r) + \\ &\omega^{00}_{\alpha\mu}(r) * c^{01}_{\mu\nu}(r) * \rho^1_\nu \tilde{h}^{(\mathrm{c})}_{\nu\gamma}(r) \end{aligned} \tag{46c}$$

$$\begin{aligned} \tilde{h}^{11}_{\alpha\gamma}(r) &= \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{11}_{\mu\nu}(r) * \tilde{\omega}^{11}_{\nu\gamma}(r) + \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{10}_{\mu\nu}(r) * \rho^0_\nu h^{01}_{\nu\gamma}(r) + \\ &\tilde{\omega}^{11}_{\alpha\mu}(r) * c^{(\mathrm{c})}_{\mu\nu}(r) * \rho^1_\nu h^{10}_{\nu\gamma}(r) + \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{(\mathrm{b})}_{\mu\nu}(r) * \rho^1_\nu \tilde{h}^{(\mathrm{c})}_{\nu\gamma}(r) \end{aligned} \tag{46d}$$

$$\tilde{h}^{(\mathrm{c})}_{\alpha\gamma}(r) = \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{(\mathrm{c})}_{\mu\nu}(r) * \tilde{\omega}^{11}_{\nu\gamma}(r) + \tilde{\omega}^{11}_{\alpha\mu}(r) * c^{(\mathrm{c})}_{\mu\nu}(r) * \rho^1_\nu \tilde{h}^{(\mathrm{c})}_{\nu\gamma}(r) \tag{46e}$$

where ρ^1_γ is the number density of interaction site γ of liquid species and ρ^0_γ is that of matrix nanoparticles, $h^{ij}_{\alpha\gamma}(r)$ and $c^{ij}_{\alpha\gamma}(r)$ are, respectively, the total and direct correlation functions between site α of species i and site γ of species j ($i,j = 0$ for matrix nanoparticles and 1 for liquid molecules). An asterisk "*" means convolution in the direct space, and summation over repeating site indices is implied. In the replica formalism, the liquid–liquid total and direct correlation functions are subdivided into the so-called connected and blocking (or disconnected) parts

$$\tilde{h}^{11}_{\alpha\gamma}(r) = \tilde{h}^{(\mathrm{c})}_{\alpha\gamma}(r) + h^{(\mathrm{b})}_{\alpha\gamma}(r) \tag{46f}$$

$$c^{11}_{\alpha\gamma}(r) = c^{(\mathrm{c})}_{\alpha\gamma}(r) + c^{(\mathrm{b})}_{\alpha\gamma}(r) \tag{46g}$$

with the connected correlations denoted by superscript (c) stemming following from the corresponding correlations between the particles of the same replica of the liquid and the blocking ones denoted by (b) stemming from those between different replicas of the liquid in the analytical continuation limit of $s \to \infty$. In terms of Mayer diagrams, the blocking correlations are identified as a subset of diagrams in which all paths between the two root vortices pass through at least one field vortex of matrix, that is,

such liquid–liquid diagrams that are completely blocked/disconnected by matrix vortices. In other words, this is the indirect, matrix-mediated part of the liquid–liquid correlations. The remaining portion constitutes the connected part of the liquid–liquid correlations. Further, within the DRISM approach both the intramolecular and total liquid–liquid correlation functions are renormalized due to a dielectric bridge correction introduced to ensure the consistency and given value of the dielectric constant obtained for solvent–solvent, solvent–ion, and ion–ion effective interactions in electrolyte solution

$$\tilde{\omega}^{11}_{\alpha\gamma}(r) = \omega^{11}_{\alpha\gamma}(r) + \rho^1_\alpha \zeta^{11}_{\alpha\gamma}(r) \tag{47a}$$

$$\tilde{h}^{11}_{\alpha\gamma}(r) = h^{11}_{\alpha\gamma}(r) - \zeta^{11}_{\alpha\gamma}(r) \tag{47b}$$

$$\tilde{h}^{(c)}_{\alpha\gamma}(r) = h^{(c)}_{\alpha\gamma}(r) - \zeta^{11}_{\alpha\gamma}(r) \tag{47c}$$

The renormalized dielectric correction enforcing the given phenomenological value of the dielectric constant and the proper orientational behavior and consistency of the dielectric response in the electrolyte solution is obtained in the same analytical form (6) as for bulk electrolyte solution but now for annealed liquid species

$$\zeta^{11}_{\alpha\gamma}(k) = j_0(kx^1_\alpha) j_0(ky^1_\alpha) j_1(kz^1_\alpha) h_c(k) j_0(kx^1_\gamma) j_0(ky^1_\gamma) j_1(kz^1_\gamma), \quad \alpha,\gamma \in s \tag{48}$$

where $\mathbf{r}^1_\alpha = (x^1_\alpha, y^1_\alpha, z^1_\alpha)$ and $\mathbf{r}^1_\gamma = (x^1_\gamma, y^1_\gamma, z^1_\gamma)$ are the Cartesian coordinates of partial site charges q^1_α and q^1_γ at sites α and γ of the same solution species s with respect to its molecular origin, the dipole moment $\mathbf{d}^1_s = \Sigma_{\alpha\in s}\, q^1_\alpha \mathbf{r}^1_\alpha$ is oriented along the z-axis, $\mathbf{d}^1_s = (0,0,d^1_s)$, the envelope function $h_c(k)$ is assumed in the same form

$$h_c(k) = A \exp\left(-l^2 k^2/4\right)$$

the amplitude A is related to the dielectric constant ε of the electrolyte solution

$$A = \frac{1}{\rho^1_{\text{polar}}}\left(\frac{\varepsilon}{y} - 3\right)$$

the total number density of solution polar species is

$$\rho^1_{\text{polar}} = \sum_{s\in \text{ polar}} \rho^1_s$$

and the solution dielectric susceptibility is

$$y^1 = \frac{4\pi}{9k_B T} \sum_{s\in \text{ polar}} \rho^1_s \left(d^1_s\right)^2$$

The parameter l specifying the characteristic separation from which the dielectric correction is switched on is chosen according to the maximal repulsive core size of liquid species (e.g., l = 1 Å for water solvent, and l = 10 Å for octanol).

The average number densities of sorbed electrolyte solution species ρ^1_s substantially differ from those in bulk solution ρ^{bulk}_s and so does the dielectric constant of sorbed solution ε compared to that of bulk solution $\varepsilon^{\text{bulk}}$. This is taken into account by smoothly interpolating in density between the gas form $\varepsilon = 1 + 3y$ and the bulk solution value $\varepsilon^{\text{bulk}}$ as follows

$$\varepsilon = 1 + \left(3 + \rho^1_{\text{polar}} A^{\text{bulk}}\right) y \tag{49}$$

where the correction amplitude for the sorbed solution with densities ρ_s^1 keeps the same value as for the bulk solution with ρ_s^{bulk}

$$A = A^{\text{bulk}} = \frac{1}{\rho_{\text{polar}}^{\text{bulk}}}\left(\frac{\varepsilon^{\text{bulk}}}{y^{\text{bulk}}} - 3\right) \tag{50}$$

The replica DRISM integral equations (46) have to be complemented with appropriate closure relations that properly account for both electrostatic and non-electrostatic features of molecular specificities of the species of both the matrix nanoparticles and the sorbed liquid. The closure proposed by Kovalenko and Hirata (KH approximation) [8,11] accounts for both electrostatic and nonpolar features of solvation, such as hydrogen bonding, solvophobicity, structural solvent molecules, salt bridges, and other steric, associative, and electrochemical effects in simple liquids and complex solutions of a given composition, including buffers, salts, polymers, ligands, and other cofactors at a finite concentration. Further, the KH closure describes associating molecular fluids and solutions in the whole density range from gas to liquid [17,18], and is thus suitable for such a system with strong associative effects as electrolyte solution sorbed in nanoporous materials [18,85,86], in particular, in a charged nanoporous electrode [87–89]. It is used for the matrix–matrix, liquid–matrix, and liquid–liquid correlation functions

$$g_{\alpha\gamma}^{ij}(r) = \begin{cases} \exp\left(d_{\alpha\gamma}^{ij}(r)\right) & \text{for} \quad d_{\alpha\gamma}^{ij}(r) \le 0 \\ 1 + d_{\alpha\gamma}^{ij}(r) & \text{for} \quad d_{\alpha\gamma}^{ij}(r) > 0 \end{cases} \tag{51}$$

$$d_{\alpha\gamma}^{ij}(r) = -u_{\alpha\gamma}^{ij}(r)/k_{\text{B}}T + h_{\alpha\gamma}^{ij}(r) - c_{\alpha\gamma}^{ij}(r)$$

where $g_{\alpha\gamma}^{ij}(r) = h_{\alpha\gamma}^{ij}(r) + 1$ is the radial distribution function (RDF) between interaction sites α and γ of species $i,j = 0,1$ (matrix, liquid), and $u_{\alpha\gamma}^{ij}(r)$ is the site–site pairwise interaction potential, scaled by the Boltzmann constant k_{B} times solution or matrix temperature T. Note that in general $T = T_0$ for $i,j = 0$ and $T = T_0$ otherwise; however, one can put $T_0 = T_1 \equiv T$ simply by scaling the matrix model interaction potential $u_{\alpha\gamma}^{00}(r)$, as explained at the beginning of the previous section.

It was shown that application of linearized closures of MSA type to blocking correlations leads to trivial solutions with $c^{(b)}(r) = 0$ [90–93]. On the other hand, the HNC approximation strongly overestimates the blocking correlations in the presence of charged species and leads to divergence of the replica RISM equations for electrolyte solution sorbed in nanoporous material [18,85–89]. To work at the advanced level with non-trivial blocking correlations, we complement eqs. 46e–g for the blocking correlations with the VM closure

$$g_{\alpha\gamma}^{(\text{b})}(r) = h_{\alpha\gamma}^{(\text{b})}(r) + 1 = \exp\left(h_{\alpha\gamma}^{(\text{b})}(r) - c_{\alpha\gamma}^{(\text{b})}(r) + b_{\alpha\gamma}^{(\text{b})}(r)\right) \tag{52a}$$

$$b_{\alpha\gamma}^{(\text{b})}(r) = -\frac{1}{2}\,\frac{\left(t_{\alpha\gamma}^{(\text{b})}(r)\right)^2}{1 + a\max\left(t_{\alpha\gamma}^{(\text{b})}(r), 0\right)} \tag{52b}$$

where the VM bridge correction (52b) is expressed in terms of the nodal correlation function $t_{\alpha\gamma}^{(\text{b})}(r) = h_{\alpha\gamma}^{(\text{b})}(r) - c_{\alpha\gamma}^{(\text{b})}(r)$, and the parameter value $a = 0.8$ is the same as in the original Verlet correction. (Note that it can be adjusted with liquid density for self-consistency, for example, to satisfy the zero-separation theorem [94].) Note that there is no interaction potential in the closure (52) for the blocking correlations as they stem in the limit $s \to \infty$ from the correlations between different replicas of liquid which do not interact with each other. The VM approximation (52) accounts reasonably well for the nonlinearity of blocking correlations in such strongly associating systems as polar solvents and electrolyte solutions sorbed in nanoporous matrices, both neutral and with external electric charge [18,85–90].

Solving the replica DRISM-KH-VM equations (46–52) at particular force field and thermodynamic parameters requires analytical treatment of all the electrostatic asymptotics and is done similarly to the bulk DRISM-KH equations (5–11). The equations are converged by using by using the MDIIS accelerated numerical solver [9–11,47].

Thermodynamics of sorbed solution

For the replica DRISM-KH-VM integral equations (46), the excess chemical potential $\Delta\mu_s^1$ of sorption of liquid species s in the frozen matrix of nanoparticles is readily decomposed into the contributions from the host matrix due to the liquid-matrix (ij = 10) correlations and from the sorbed liquid due to the liquid–liquid (ij = 11) correlations less an additional term due to the blocking (b) correlations [85,86]

$$\Delta\mu_s^1 = \Delta\mu_s^{10} + \Delta\mu_s^{11} - \Delta\mu_s^{(\mathrm{b})} \tag{53}$$

The KH closure (51) to the replica DRISM integral equations (46) leads to the corresponding components of the excess chemical potential expressed in a closed analytical form in terms of the liquid–matrix and liquid–liquid correlation functions

$$\Delta\mu_s^{1j} = k_\mathrm{B}T \sum_{\alpha\in s}\sum_{\gamma\in j} \rho_\gamma^j 4\pi \int_0^\infty r^2\,\mathrm{d}r \left[\frac{1}{2}\Theta\left(-h_{\alpha\gamma}^{1j}(r)\right)\left(h_{\alpha\gamma}^{1j}(r)\right)^2 - \frac{1}{2}h_{\alpha\gamma}^{1j}(r)c_{\alpha\gamma}^{1j}(r) - c_{\alpha\gamma}^{1j}(r)\right] \tag{54}$$

where the summation is over interaction sites α of liquid species s, and over all sites γ of matrix for j = 0 and liquid for j = 1. The chemical potential term due to the blocking correlations can be obtained from the VM closure (52) to the replica DRISM integral equations (46e–g) by performing thermodynamic integration in an approximate analytical form

$$\Delta\mu_s^{(\mathrm{b})} = k_\mathrm{B}T \sum_{\alpha\gamma\in 1} \rho_\alpha^1 4\pi \int r^2\,\mathrm{d}r \left[\frac{1}{2}\left(h_{\alpha\gamma}^{(\mathrm{b})}(r)\right)^2 - \frac{1}{2}h_{\alpha\gamma}^{(\mathrm{b})}(r)c_{\alpha\gamma}^{(\mathrm{b})}(r) - c_{\alpha\gamma}^{(\mathrm{b})}(r) + \left(h_{\alpha\gamma}^{(\mathrm{b})}(r)+1\right)b_{\alpha\gamma}^{(\mathrm{b})}(r) - s_{\alpha\gamma}^{(\mathrm{b})}(r)\right] \tag{55}$$

where the so-called star function obtained in the assumption of so-called unique functionality of the correlations [94] is written as

$$s_{\alpha\gamma}^{(\mathrm{b})}(r) = \begin{cases} -\dfrac{1}{2a^3}\left[\ln\left(1+at_{\alpha\gamma}^{(\mathrm{b})}(r)\right) - at_{\alpha\gamma}^{(\mathrm{b})}(r) + \dfrac{1}{2}\left(at_{\alpha\gamma}^{(\mathrm{b})}(r)\right)^2\right]\dfrac{t_{\alpha\gamma}^{(\mathrm{b})}(r)}{h_{\alpha\gamma}^{(\mathrm{b})}(r)} & \text{for} \quad t_{\alpha\gamma}^{(\mathrm{b})}(r) > 0 \\ -\dfrac{1}{6}t_{\alpha\gamma}^{(\mathrm{b})}(r)h_{\alpha\gamma}^{(\mathrm{b})}(r) & \text{for} \quad t_{\alpha\gamma}^{(\mathrm{b})}(r) \le 0 \end{cases} \tag{56}$$

Also, the compressibility of the sorbed solution is exactly expressed in terms of the connected part of the correlations [85,86]. Other thermodynamic derivatives of the free energy of the sorbed solution are available as well [85,86].

Electric double layer in the nanopores of the host matrix

The sorbed electrolyte solution forms an EDL at the surface of nanopores of the host material even in the absence of external specific electric charge on the nanoporous electrode due to the asymmetry between the density distributions of cations and anions at the surface, as well as the preferential orientation of polar solvent molecules at the surface. The statistically averaged electrostatic potential of the EDL is obtained as follows [87–89]. The statistical distribution of charge density around a labeled

matrix nanoparticle of sort $c \in 0$ due to interaction site charges q_γ^1 on electrolyte solution species of sort $\gamma \in 1$ as well as q_f^0 on chemical functional groups of sort $f \in 0$ grafted to matrix nanoparticles is obtained as

$$\tau_c^0(r) = \sum_{f\in 0} q_f^0 \rho_f^0 g_{cf}^{00}(r) + \sum_{\gamma\in 1} q_\gamma^1 \rho_\gamma^1 g_{c\gamma}^{01}(r) \tag{57}$$

(Note that matrix nanoparticles and chemical functional groups grafted to them, both belonging to matrix species $j = 0$, are now denoted separately with roman subscript indices of sort c and f, respectively.) The corresponding statistically averages electrostatic potential $\psi_c^0(r)$ of sorbed solution species and nanoporous material surrounding matrix nanoparticle c is determined from the Poisson equation

$$\nabla^2 \psi_c^0(r) = -4\pi\tau_c^0(r) \tag{58}$$

The full local electrostatic potential $\phi_c^0(r)$ around a conducting matrix nanoparticle of radius R_c^0 includes also the contribution from an externally induced charge q_c^0 on the nanoparticle

$$\phi_c^0(r) = \psi_c^0(r) + \frac{q_c^0}{\max\left(r, R_c^0\right)} \tag{59}$$

The external charge of density

$$q_{\text{ex}} = \sum_{c\in 0} q_c^0 \rho_c^0 \tag{60}$$

induced on the electrode and the corresponding opposite external charge $-q_{\text{ex}}$ on the other electrode of the supercapacitor cause separation of cations and anions of the electrolyte diffusing across the separator between the electrodes to the EDLs in their pores, until the bias of the ionic concentrations in each electrode satisfies the condition of electroneutrality for the whole system

$$q_{\text{ex}} + \sum_f q_f^0 \rho_f^0 + \sum_\gamma q_\gamma^1 \rho_\gamma^1 = 0 \tag{61}$$

Note that for symmetric electrodes with charges functional groups of atomic charges q_f^0 and specific densities ρ_f^0 there is a symmetric concentration of counterions (or a bias between the concentrations of cations and anions) sorbed in each electrode initially, at zero voltage of the device. For asymmetric densities of charged functional groups in the two electrodes, the initial concentrations of counterions or the bias between the concentrations of cations and anions are change due to diffusion exchange between the electrodes and the bulk solution bath until the electroneutrality is satisfied in each electrode. For the nanoporous carbon material formed by connected carbon spheres of several sorts α with in general different sizes R_c^0, the external charge q_{ex} is distributed among the charges q_c^0 on each sphere sort subject to the electrostatic potential inside the sphere being equal for all carbon spheres

$$\phi_{\text{c}}(q_{\text{ex}}) \equiv \phi_{\text{c}}^0(r < R_c^0; q_{\text{ex}}) \quad \text{for all} \ \ c \in 0 \tag{62}$$

whereas the electrostatic potential (59) at a distance from each sphere of sort c by definition gives the average potential level in the electrode including both carbon spheres and sorbed solution

$$\phi_{\text{av}}(q_{\text{ex}}) \equiv \phi_{\text{c}}^0(r = \infty; q_{\text{ex}}) \quad \text{for all} \ \ c \in 0 \tag{63}$$

The value of the electrostatic potential on each carbon sphere $\phi_{\text{c}}(q_{\text{ex}})$ gives the potential of the electrode carbon conducting framework. The potential step from the carbon nanoparticle level $\phi_{\text{c}}(q_{\text{ex}})$ to the electrode bulk level $\phi_{\text{av}}(q_{\text{ex}})$ comprises (i) the voltage across the intrinsic EDL in contact with the nanoparticle (surface of the nanopore), and (ii) the electric field of all other surrounding carbon nanoparticles

and their EDLs distributed in the nanoporous material with its spatial correlations. The value of the electrostatic potential statistically averaged over the bulk of the nanoporous electrode, $\phi_{av}(q_{ex})$, is a constant of integration of the Poisson equation (58), or the "ground level" of the whole nanoporous electrode (including both the electrode carbon framework and the sorbed electrolyte solution). The electrode "ground level" is positioned with respect to the vacuum level of the electrostatic potential by an external electric field in the solution outside the electrode, and the solution to the Poisson equation yields the potential of the carbon conducting frame with respect to the electrode "ground level" $\phi_c(q_{ex}) - \phi_{av}(q_{ex})$, caused by all electric charges inside the electrode, that is, the external specific charge on the electrode and EDL charges of ions and polar solvent molecules sorbed inside the nanopores.

Molecular mechanism of electrosorption and capacitance in nanoporous electrodes

The molecular mechanisms determining high specific capacitance of the supercapacitor device and purification efficiency of the electrosorption cell with nanoporous electrodes are much more complex than in a planar EDL of equivalent area. As explained below, they include the intrinsic EDL at the surface of nanopores, the Gouy–Chapman layer statistically averaged over the volume of the nanoporous material, the osmotic term arising due to the difference between the ionic concentrations in the two nanoporous electrodes and in the bulk electrolyte solution outside, and the solvation chemical potentials of sorbed solvated ions statistically averaged over the nanoporous material [87–89].

The chemical potential μ_s of sorbed solution species s of density ρ_s consists of the ideal gas contribution $\mu_s^{id} = k_B T \ln(\rho_s \Lambda_s)$ with the de Broglie thermal wave length $\Lambda_s = \sqrt{2\pi\hbar^2/(m_s k_B T)}$ of ideal monatomic particles with molecular weight m_s, the excess term $\Delta\mu_s$ due to liquid–liquid and matrix–liquid intermolecular interactions inside the nanoporous material, and the electrostatic energy of species s with charge q_s in the electrostatic field between the two electrodes

$$\mu_s = \mu_s^{id} + \Delta\mu_s + q_s\phi_{av} \tag{64}$$

The first term is the osmotic contribution, the second one arises from the interactions inside the nanoporous material, and the third one gives the "ground level" of the electrostatic potential averaged over the whole nanoporous electrode with respect to the gas phase outside the system. The presence of the third term in the chemical potential (64) directly follows from the Nernst equation.

The excess chemical potentials $\Delta\mu_s$ of ionic species s inside a charged electrode are strongly different from those in the bulk solution. For the electrode in contact with the bulk solution, this causes diffusion of ions across the separator until the electric field of ionic dipoles forming at the boundaries of the electrodes counterbalances the difference between the "interior" chemical potential terms $\mu_s^{id} + \Delta\mu_s$ of electrode I, as well as those of electrode II, and thus equalizes the chemical potential in the two electrodes, $\mu_s^I = \mu_s^{II}$. Using the decomposition (65) of the chemical potential inside the electrode yields the bias between the statistically averaged "ground levels" $\phi_{av}^{II}(q_{ex})$ and $\phi_{av}^{I}(-q_{ex})$ of the two electrodes with opposite external charges $\pm q_{ex}$

$$q_s\left(\phi_{av}^{II}(q_{ex}) - \phi_{av}^{I}(-q_{ex})\right) = k_B T \ln(\rho_s^I/\rho_s^{II}) + \Delta\mu_s^I(-q_{ex}) - \Delta\mu_s^{II}(q_{ex}) \tag{65}$$

The same bias of the electrostatic potential levels $\phi_{av}^{II} - \phi_{av}^{I}$ must satisfy the relation (65) for each solution species s, including ions and neutral solvent molecules. A discrepancy between the bias values required to counterbalance the "interior" part of the chemical potential $k_B T \ln(\rho_s \Lambda_s) + \Delta\mu_s$ for cations as well as that for anions causes diffusion of each of the ionic species to the corresponding electrode with the lower chemical potential. This changes the osmotic and excess terms in the chemical potential (64) for cations and anions in the opposite directions until the relation (65) is satisfied for both cations and anions. Similarly, diffusion of solvent molecules between the two electrodes occurs until the osmotic term as well as the excess chemical potentials for the changed densities of solvent in the elec-

trodes, dependent on the densities of both solvent and ions, satisfy the chemical equilibrium condition (65), which for the neutral solvent molecules with $q_s = 0$ reduces to the equality of the "interior" terms $k_B T \ln(\rho_s \Lambda_s) + \Delta\mu_s$.

The voltage of the supercapacitor device given by the electrostatic potential difference between the conducting carbon matrix frameworks of electrodes I and II is thus obtained by summing up the statistically averaged electrostatic potential steps across the whole device. (Note that all losses due to equivalent serial resistance (ESR), including electric conductivity of the carbon nanoparticles and their contacts as well as electrolyte diffusion resistance are not considered here and can be incorporated separately.) The supercapacitor voltage comprises the potential step $\phi^{\mathrm{I}}_{\mathrm{av}}(-q_{\mathrm{ex}}) - \phi^{\mathrm{I}}_{\mathrm{c}}(-q_{\mathrm{ex}})$ across the intrinsic EDL in nanopores of electrode I from its conducting carbon framework to the electrode bulk, next the potential step $\phi^{\mathrm{II}}_{\mathrm{av}}(q_{\mathrm{ex}}) - \phi^{\mathrm{I}}_{\mathrm{av}}(-q_{\mathrm{ex}})$ from the "ground level" of electrode I across the EDL at its outer edge to the solution bulk and then similarly across the EDL at the outer edge of electrode II to its "ground level", and then $\phi^{\mathrm{II}}_{\mathrm{c}}(q_{\mathrm{ex}}) - \phi^{\mathrm{II}}_{\mathrm{av}}(q_{\mathrm{ex}})$ from the electrode bulk across the intrinsic EDL in nanopores of electrode II to its conducting carbon framework. Using the relation (65) to express $\phi^{\mathrm{I}}_{\mathrm{av}} - \phi^{\mathrm{II}}_{\mathrm{av}}$ in terms of the number densities ρ_s and excess chemical potentials $\Delta\mu_s$ of sorbed ions s, the supercapacitor voltage is written as

$$U(q_{\mathrm{ex}}) = \left(\phi^{\mathrm{I}}_{\mathrm{av}} - \phi^{\mathrm{I}}_{\mathrm{c}}\right) - \left(\phi^{\mathrm{II}}_{\mathrm{av}} - \phi^{\mathrm{II}}_{\mathrm{c}}\right) + \frac{1}{q_s}\left(k_B T \ln(\rho^{\mathrm{I}}_s / \rho^{\mathrm{II}}_s) + \Delta\mu^{\mathrm{I}}_s - \Delta\mu^{\mathrm{II}}_s\right) \tag{66}$$

for any ionic species s (with $q_s \neq 0$). The number densities ρ_s and excess chemical potentials $\Delta\mu_s$ of sorbed species s are obtained by converging the chemical equilibrium conditions (65) for ρ_s by using iterations or any other accelerated solver, in turn with solving at each step the replica DRISM-KH-VM integral equations (46–52) at current ρ_s and calculating the excess chemical potentials from the expressions (53–56). The potential steps $\phi_{\mathrm{av}} - \phi_{\mathrm{c}}$ in electrodes I and II are then calculated at the densities ρ_s for the converged equations (65) as the difference of the boundary values (63) and (62) from solving the Poisson equation (58).

Finally, applying the chemical equilibrium condition between the electrodes and the bulk solution yields the relation for the purification efficiency of the electrosorption cell at voltage $U(q_{\mathrm{ex}})$ which holds sorbed electrolyte at high concentrations ρ^{I}_s and ρ^{II}_s of cations and anions inside the corresponding nanoporous electrodes I and II with the opposite external charge, against the osmotic forces for a significantly lower concentration ρ^{blk}_s of electrolyte in the bulk solution efflux [88]

$$\frac{\rho^{\mathrm{blk}}_s}{\rho^{\mathrm{I}}_s} = \exp\left(-\frac{1}{k_B T}\left(\Delta\mu^{\mathrm{blk}}_s - \Delta\mu^{\mathrm{I}}_s(-q_{\mathrm{ex}}) - q_s \phi^{\mathrm{I}}_{\mathrm{av}}(-q_{\mathrm{ex}})\right)\right) \quad \text{for } q_s > 0 \text{ (cations)} \tag{67a}$$

$$\frac{\rho^{\mathrm{blk}}_s}{\rho^{\mathrm{II}}_s} = \exp\left(-\frac{1}{k_B T}\left(\Delta\mu^{\mathrm{blk}}_s - \Delta\mu^{\mathrm{II}}_s(q_{\mathrm{ex}}) - q_s \phi^{\mathrm{II}}_{\mathrm{av}}(q_{\mathrm{ex}})\right)\right) \quad \text{for } q_s < 0 \text{ (anions)} \tag{67b}$$

where $\Delta\mu^{\mathrm{blk}}_s$ is the excess chemical potential of species s at concentration ρ^{blk}_s in bulk solution efflux.

Illustration for a supercapacitor

Carbonized polyvinylidene chloride (PVDC) material attracted both scientific and industrial attention [95] because of its uniform nanoporous texture with controlled pore size, and the possibility of a cheaper production cycle for applications in high-energy storage and electrochemical separation devices. A model consisting of a statistical–mechanical mix of hard cores and "cavity" spheres at high packing fraction has been developed [87] to represent essential properties of carbonized PVDC material and is so parameterized as to fit the pore size distribution peaks, porosity, pores SA, and the physical density to the experimental data for carbonized PVDC material [95].

The replica DRISM-KH-MV integral equations for the correlation functions of NaCl aqueous electrolyte solution (46–52) with the electrostatic asymptotics separated out and treated analytically were discretized on a uniform radial grid of length 1000 Å. The domain of this significant size is necessary to properly represent the interplay of the slowly decaying oscillations in the correlations of the matrix material and the EDL diffuse layers around matrix nanoparticles, resulting in the solvation structure oscillations with at the size of sorbed solution species which last on many sizes of matrix nanoparticles [87–89]. The replica DRISM-KH-MV integral equations were converged by using the MDIIS accelerated numerical solver [9–11,47]. The force field parameters and densities for the ambient electrolyte solution and the matrix nanoparticles of the present system are summarized in refs. [87–89]. For a given specific charge q_{ex} on the carbonized PVDC material, the number densities of species of the sorbed solutions were obtained by calculating the excess chemical potentials (53–56), obtaining the electrostatic potential from the Poisson equation (58), and solving for the chemical balance equations (65). The supercapacitor voltage for each value of q_{ex} was finally obtained from the expression (66).

Figure 11 exhibits the RDFs of the solvation structure of electrolyte solution species around a carbon nanoparticle of the nanoporous matrix of carbonized PVDC material, which is in equilibrium with the bulk NaCl aqueous electrolyte solution at concentration 1 M at ambient conditions. Shown for comparison is the solvation structure of a single carbon sphere (SCS) of the same size immersed in the bulk electrolyte solution. In both cases, the positions and width of the first peaks of water oxygen and hydrogen sites correspond to the typical solvation structure of a hydrophobic nanosphere in aqueous electrolyte solution, with water hydrogens oriented preferentially in parallel to the surface. The nanoporous PVDC confinement enhances the first solvation peaks but depletes the amplitude of the oscillations in the second solvation peak. The positions of the RDF peaks indicate that Cl^- ions are localized in contact with the surface hydrophobic of carbon nanoparticles, while Na^+ ions located at the surface (the highest peak of the carbon-Na^+ RDF) are surrounded by hydration shells.

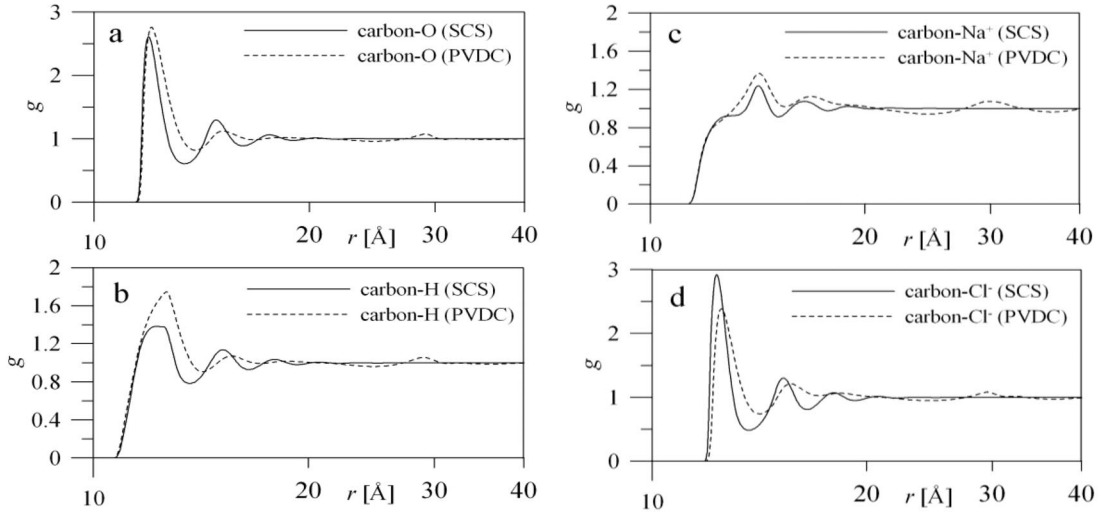

Fig. 11 Solvation structure of the intrinsic EDL formed by NaCl aqueous solution in nanoporous carbonized PVDC material. RDFs of water O and H sites, and of Na^+ and Cl^- ions (parts a–d) at the surface of a carbon nanosphere in the nanoporous matrix framework (solid lines) vs. the solvation structure of an SCS immersed in the bulk aqueous electrolyte solution (dashed lines). The nanoporous material is in contact with the bulk ambient aqueous solution of NaCl at concentration 1M.

Figure 12 shows the change of the carbon-solution RDFs with charging the nanoporous electrode. The carbon-ion RDFs strongly change due to attraction of counterions and repulsion of co-ions according to the external charge of the electrode. The carbon–water oxygen and hydrogen distributions remain almost unchanged with electrode charge. Figure 13 depicts the run of the electrostatic potential with distance from a carbon nanosphere in the nanoporous matrix to the bulk of the nanoporous carbon electrode, which is obtained from the Poisson equation (58) with the electric charge density (57) following from the carbon-solution distributions shown in Figure 12. (Each curve in Fig. 13 is plotted with respect to the average electrostatic potential level ϕ_C inside the nanoporous electrode.) The electrostatic potential is constant for $r < R_c^0$ inside the conducting carbon nanosphere. The carbon nanosphere bears the average external charge which strongly affects the electrostatic potential run next to the nanosphere surface and in the first and second solvation shells. The Stern layer contains no solution charges due to the steric constraints, and the slope of the curves for $r > R_c^0$ right at the nanosphere surface is caused barely by its Coulomb potential. Then follow the potential drop due to the surface dipole with water hydrogens closer to the surface than oxygens and the potential rise due to the surface dipole with Cl^- ions located closer than Na^+. These potential peaks in the first and second solvation layers are identified as the outer Helmholtz layer. The nanosphere electric charge introduces oscillations with a period of 12 Å close to the size of nanospheres in carbonized PVDC, slowly decaying with distance. The whole range of oscillations includes the diffuse layer around the nanoparticle, as well as the statistical average of the EDLs around other nanoparticles which are closely packed and correlated in the nanoporous carbon matrix. The potential drop of the Stern layer is almost completely cancelled out by the electric field of the outer Helmholtz layer and further oscillations which stem mainly from the ionic cloud of the first and second solvation shells screening the external charge of the carbon nanosphere, with Cl^- ions prevailing for positive and Na^+ for negative external charge.

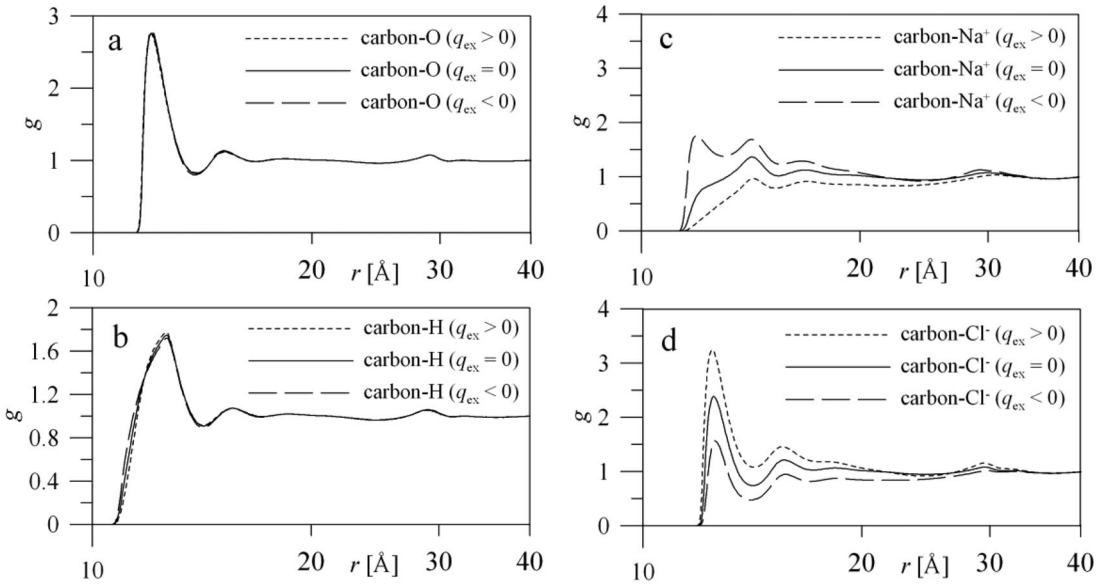

Fig. 12 Solvation structure of the NaCl aqueous electrolyte solution sorbed in the nanoporous carbonized PVDC electrode at zero charge (solid lines), and at positive and negative specific charge $q_{ex} = \pm 12$ [C/cm^3] (dotted and dashed lines, respectively). RDFs of water O and H sites, and Na^+ and Cl^- ions around carbon matrix nanoparticles (parts a–d).

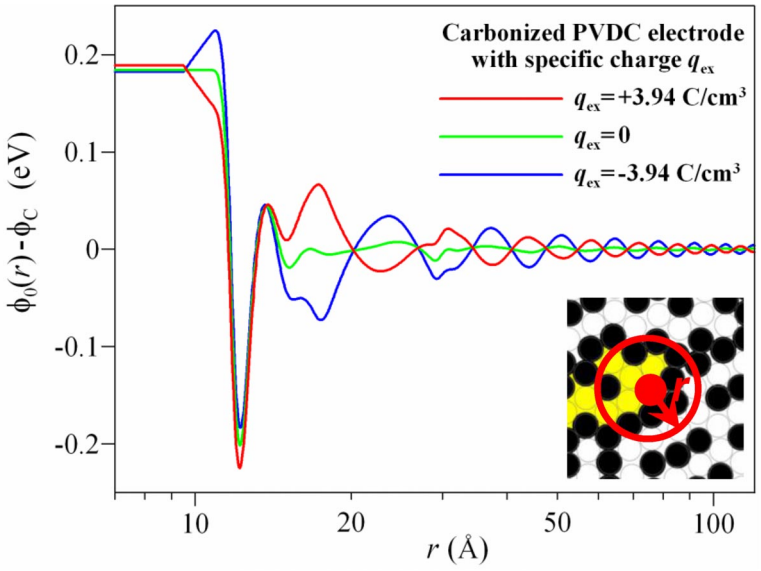

Fig. 13 Statistically averaged electrostatic potential $\phi_0(r)$ around a carbon nanoparticle of the carbonized PVDC electrode with respect to the "ground level" ϕ_C averaged over the whole nanoporous electrode. The electrolyte solution sorbed in the nanoporous electrode is in equilibrium with the bulk ambient aqueous solution of NaCl at concentration 120 ppm. The electrode is at zero charge (green line), and at positive and negative specific charge q_{ex} = ±3.94 [C/cm^3] (red and blue lines, respectively). The inset schematically illustrates statistical–mechanical averaging (red circle and distance vector) around a labeled carbon nanoparticle (red ball) over the whole nanoporous material, including carbon nanoparticles (black balls) and nanopores (white voids).

Finally, Fig. 14 presents a diagram of the overall run of the electrostatic potential across the supercapacitor device. The average potential energy level $q_{ex}\phi_C$ of the nanoporous electrode is shifted due to the electric field of the dipole which is formed by solution species outside the electrode at its macroscopic boundary to satisfy the chemical equilibrium conditions (65). An additional EDL emerges at the macroscopic boundary of each electrode to counterbalance the difference between the "interior" chemical potential part $k_BT \ln(\rho_s\Lambda_s) + \Delta\mu_s$ of ions inside the two nanoporous electrodes. The potential drop across these EDLs outside the electrodes constitutes a major part of the supercapacitor voltage $U(q_{ex})$ and is determined by the conditions of chemical equilibrium between the solution sorbed in each of the electrodes and the bulk solution outside the electrodes. Calculations performed for the carbonized PVDC nanoporous material in contact with the ambient aqueous electrolyte solution of 1 M NaCl reveal that the electrochemical mechanism of the EDL supercapacitor is determined largely by the chemical balance for sorbed ions in the Nernst–Planck equation, rather than just by the EDL potential drop at the surface of a nanopore as in the case of a planar electrode [87–89]. The same molecular forces determine for the specific sorption capacity and purification efficiency (67) of a nanoporous carbon electrosorption cell [88]. As demonstrated above, the specific capacitance and sorption capacity of nanoporous carbon electrodes are determined by the interplay of the EDL potential drop across the Stern layer at the surface of nanopores and the Gouy–Chapman layer statistically averaged over the volume of the nanoporous material, the osmotic term arising due to the difference between the ionic concentrations in the two nanoporous electrodes and in the bulk electrolyte solution outside, and the solvation chemical potentials of sorbed solvated ions statistically averaged over the nanoporous material [87–89]. The latter term is strongly affected by chemical specificity of ions, solvent, surface functional groups, and steric effects for solvated ions confined in nanopores. Note that solvation shells of ions enlarge their effective size and can strongly affect the specific capacitance, which has strong implications for real life supercapac-

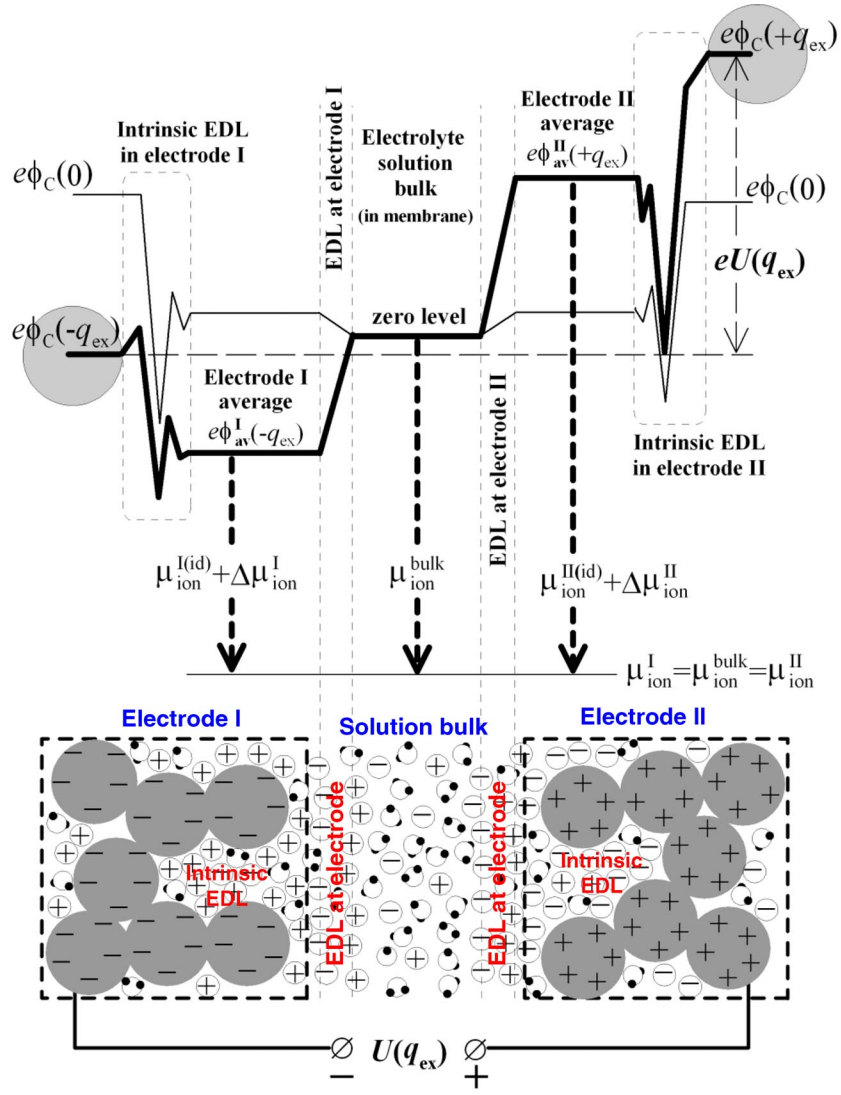

Fig. 14 Diagram of the electrostatic potential across the supercapacitor device: from conducting carbon nanospheres in the nanoporous carbon matrix, across the internal EDL at the surface of nanopores, across the external EDL at the edge of electrode I, to the electrolyte solution bulk, and then across the external EDL and intrinsic EDL to the nanoporous carbon matrix of electrode II.

itor devices. For example, a power circuit supercapacitor should function in a wide range of temperatures down to –40 °C, which stiffens the solvation shells of ions and hinders them from entering small pores, thus decreasing the capacitance. Thus, the chemical potentials of solvated ions in nanopores constitute a major factor driving the specific capacitance, as they bring about two extra EDLs at the outer boundaries of the nanoporous electrodes to offset the chemical potential difference between the electrodes and solution bulk, and therefore substantially contribute to the supercapacitor voltage.

CONCLUSIONS

Nanoscale properties and processes are profoundly different from the macroscopic laws in continuous media and materials. Functional features of nanostructures stem from microscopic properties of the atoms and chemical groups they are built of, but manifest on length scale from one to hundreds of nanometers and time scale up to microseconds and more. The properties of nanostructures and processes involving them can be tuned in a wide range by changing size and composition. Predictive modeling of nanosystems should operate at length scales from an ångström to hundreds of nanometers and microns and time scales to milliseconds and seconds, and yet derive their properties from the chemical functionalities of the constituents. Explicit molecular modeling of such nanosystems involves millions and billions of molecules and is by far not feasible in a "brute force" approach employing just ab initio quantum chemical methods and/or molecular simulations. A proper way thus requires multiscale methods coupling electronic structure methods for building blocks, classical molecular simulations for critical aggregates in the system, statistical–mechanical theories for their large assemblies and mean properties over characteristic size and time scales, and macroscopic scale properties.

Integral equation theory of liquids, which is based on the first principles of statistical mechanics, is becoming increasingly popular, as it provides a firm platform to handle the solvation structure and thermodynamics of complex chemical and biomolecular systems in solution. In particular, the statistical–mechanical 3D-RISM-KH molecular theory of solvation is promising as an essential part of the multiscale methodology for chemical and biological nanosystems in solution. As distinct from molecular simulations which explore the phase space by direct sampling of a limited subsystem of molecules on space and time intervals substantially restricted by computational feasibility, 3D-RISM-KH operates with spatial distributions rather than trajectories of molecules and is based on analytical summation of the free energy diagrams which yields the solvation structure and thermodynamics in the statistical–mechanical ensemble. It gives the solvation structure in terms of the 3D maps of distributions of solvent sites around a solute macromolecule of arbitrary shape and then the solvation thermodynamics analytically as a single integral in terms of the correlation functions obtained. The 3D-RISM-KH theory was employed to explain the molecular mechanisms of self-assembly, conformational stability of synthetic organic RNTs, aggregation of prion proteins and β-sheet Amyloid oligomers, protein-ligand binding, and function-related solvation properties of biomolecular complexes as large as the GLIC ion channel in a lipid bilayer and the GroEL/ES chaperone complex.

The 3D-RISM-KH molecular theory of solvation was coupled with ab initio CASSCF, KS and OFE DFT quantum chemistry methods in an SCF description of electronic structure, optimized geometry, and chemical reactions in solution. The (OFE)KS-DFT/3D-RISM-KH multiscale method is implemented in the ADF computational chemistry package and extensively validated against experiment for solvation thermochemistry, photochemistry, conformational equilibria, and activation barriers of various nanosystems in different solvents and ILs.

In biomolecular calculations, MM/3D-RISM-KH statistical–mechanical evaluation of the solvation thermodynamics of MD trajectories replaces their conventional MM/PB(GB)SA post-processing involving empirical nonpolar terms. Recently, the 3D-RISM-KH molecular theory of solvation has been coupled with MTS-MD simulation for a solute biomolecule driven by the effective solvent PMF. The MTS-MD procedure is stabilized using the new algorithm of the OIN chain thermostat. The solvation forces are obtained analytically by converging the 3D-RISM-KH integral equations once per several OIN outer time-steps, and are calculated in between by using SFCE in the subspace of previous successive solutions to 3D-RISM-KH. With the stabilizing effect of OIN thermostatting, gigantic outer time-steps up to picoseconds can be employed to accurately calculate equilibrium and conformational properties. The multiscale OIN/SFCE/3D-RISM-KH integrator algorithm has been implemented in the Amber MD package, and validated and benchmarked on a fully flexible model of alanine dipeptide in aqueous solution. While the computational rate of solvent sampling in OIN/SFCE/3D-RISM-KH is already 20 times faster than standard MD with explicit solvent, further substantial acceleration of sam-

pling stems from the 3D-RISM-KH molecular theory of solvation efficiently sampling the phase space for essential events with rare statistics such as exchange and localization of solvent and ligand molecules in pockets and at binding sites of the biomolecule.

The molecular mechanisms determining high specific capacitance of the EDL supercapacitor with nanoporous electrodes, as well as purification efficiency of the nanoporous electrosorption cell, are much more complex than the naïve picture of just a very large specific surface area of pores densely "folded" or "packed" in the volume of nanoporous material. These mechanisms are very different from a planar EDL of equivalent surface area. This constitutes the reason why the specific capacitance based on such an equivalent planar EDL capacitor with the "insulator" thickness given by the ionic radii turns out to be by an order of magnitude higher that the typical values in real devices. To amend this empirical model, the effective thickness of the EDL insulator is typically assumed to be up to 5 nm instead, resulting in the empirical value of area capacitance of 15–20 $\mu F/cm^3$ for the model of a planar EDL of equivalent area. Meanwhile, the replica RISM-KH-VM molecular theory for the solvation structure, thermodynamics, and electrochemistry of electrolyte solutions sorbed in nanoporous materials reveals that the driving forces of sorption and supercapacitance in nanoporous carbon electrodes are very distinct from a planar EDL capacitor. They are determined by the interplay of the EDL potential drop across the Stern layer at the surface of nanopores and the Gouy–Chapman layer statistically averaged over the volume of the nanoporous material, the osmotic term arising due to the difference between the ionic concentrations in the two nanoporous electrodes and in the bulk electrolyte solution outside, and the solvation chemical potentials of sorbed solvated ions statistically averaged over the nanoporous material. The latter factor is strongly affected by chemical specificity of ions, solvent, surface functional groups, and steric effects for solvated ions confined in nanopores. Note that solvation shells of ions can enlarge their effective size due to stiffening at low operational temperature and hinder ions from entering small pores. The chemical potentials of solvated ions in nanoporous confinement thus constitute a major factor driving the specific capacitance, as they bring about two extra EDLs at the outer boundaries of the nanoporous electrodes to offset the chemical potential difference between the electrodes and solution bulk, which substantially contribute to the supercapacitor voltage.

ACKNOWLEDGMENTS

This work was supported by the National Research Council (NRC) of Canada and by ArboraNano—the Canadian Forest NanoProducts Network.

REFERENCES

1. J.-P. Hansen, I. McDonald. *Theory of Simple Liquids*, 3rd ed., Elsevier, Amsterdam (2006).
2. F. Hirata (Ed.). *Molecular Theory of Solvation*, Series: *Understanding Chemical Reactivity*, P. G. Mezey (Ed.), Vol. 24, p. 360, Kluwer Academic, Dordrecht (2003).
3. D. Chandler, J. McCoy, S. Singer. *J. Chem. Phys.* **85**, 5971 (1986).
4. D. Chandler, J. McCoy, S. Singer. *J. Chem. Phys.* **85**, 5977 (1986).
5. D. Beglov, B. Roux. *J. Chem. Phys.* **103**, 360 (1995).
6. D. Beglov, B. Roux. *J. Phys. Chem. B* **101**, 7821 (1997).
7. A. Kovalenko, F. Hirata. *Chem. Phys. Lett.* **290**, 237 (1998).
8. A. Kovalenko, F. Hirata. *J. Chem. Phys.* **110**, 10095 (1999).
9. A. Kovalenko, F. Hirata. *J. Chem. Phys.* **112**, 10391 (2000).
10. A. Kovalenko, F. Hirata. *J. Chem. Phys.* **112**, 10403 (2000).
11. 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, pp. 169–275, Kluwer, Dordrecht (2003).
12. H. Sato, A. Kovalenko, F. Hirata. *J. Chem. Phys.* **112**, 9463 (2000).

13. S. Gusarov, T. Ziegler, A. Kovalenko. *J. Phys. Chem. A* **110**, 6083 (2006).
14. D. Casanova, S. Gusarov, A. Kovalenko, T. Ziegler. *J. Chem. Theory Comput.* **3**, 458 (2007).
15. J. W. Kaminski, S. Gusarov, T. A. Wesolowski, A. Kovalenko. *J. Phys. Chem. A* **114**, 6082 (2010).
16. M. Malvaldi, S. Bruzzone, C. Chiappe, S. Gusarov, A. Kovalenko. *J. Phys. Chem. B* **113**, 3536 (2009).
17. A. Kovalenko, F. Hirata. *Chem. Phys. Lett.* **349**, 496 (2001).
18. A. Kovalenko, F. Hirata. *J. Theor. Comput. Chem.* **1**, 381 (2002).
19. J. G. Moralez, J. Raez, T. Yamazaki, R. K. Motkuri, A. Kovalenko, H. Fenniri. *J. Am. Chem. Soc.* **127**, 8307 (2005).
20. R. S. Johnson, T. Yamazaki, A. Kovalenko, H. Fenniri. *J. Am. Chem. Soc.* **129**, 5735 (2007).
21. G. Tikhomirov, T. Yamazaki, A. Kovalenko, H. Fenniri. *Langmuir* **24**, 4447 (2007).
22. T. Yamazaki, H. Fenniri, A. Kovalenko. *ChemPhysChem* **11**, 361 (2010).
23. R. Chhabra, J. Moralez, J. Raez, T. Yamazaki, J.-Y. Cho, A. Myles, A. Kovalenko, H. Fenniri. *J. Am. Chem. Soc. Commun.* **132**, 32 (2010).
24. T. Imai, R. Hiraoka, A. Kovalenko, F. Hirata. *J. Am. Chem. Soc. Commun.* **127**, 15334 (2005).
25. N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, F. Hirata. *J. Phys. Chem. B* (Feature Article) **113**, 873 (2009).
26. T. Imai, N. Miyashita, Y. Sugita, A. Kovalenko, F. Hirata, A. Kidera. *J. Phys. Chem. B* **115**, 8288 (2011).
27. T. Yamazaki, N. Blinov, D. Wishart, A. Kovalenko. *Biophys. J*. **95**, 4540 (2008).
28. N. Blinov, L. Dorosh, D. Wishart, A. Kovalenko. *Biophys. J*. **98**, 282 (2010).
29. N. Blinov, L. Dorosh, D. Wishart, A. Kovalenko. *Mol. Simul.* **37**, 718 (2011).
30. A. Kovalenko, N. Blinov. *J. Mol. Liq.* **164**, 101 (2011).
31. D. Nikolic, N. Blinov, D. Wishart, A. Kovalenko. *J. Chem. Theory Comput.* **8**, 3356 (2012).
32. M. C. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, V. S. Pande. *J. Phys. Chem. B* **115**, 205 (2011).
33. A. Kovalenko, A. E. Kobryn, S. Gusarov, O. Lyubimova, X. Liu, N. Blinov, M. Yoshida. *Soft Matter* **8**, 1508 (2012).
34. K. Yoshida, T. Yamaguchi, A. Kovalenko, F. Hirata. *J. Phys. Chem. B* **106**, 5042 (2002).
35. I. Omelyan, A. Kovalenko, F. Hirata. *J. Theor. Comput. Chem.* **2**, 193 (2003).
36. A. Kovalenko, F. Hirata. *Phys. Chem. Chem. Phys*. **7**, 1785 (2005).
37. A. Kovalenko, F. Hirata. "A molecular theory of solutions at liquid interfaces", in *Interfacial Nanochemistry: Molecular Science and Engineering at Liquid-Liquid Interfaces*, H. Watarai (Ed.), Series: *Nanostructure Science and Technology*, D. J. Lockwood (Ed.), pp. 97–125, Springer (2005).
38. J. S. Perkyns, B. M. Pettitt. *J. Chem. Phys.* **97**, 7656 (1992).
39. B. Kvamme. *Int. J. Thermophys.* **16**, 743 (1995).
40. J. G. Kirkwood, F. P. Buff. *J. Chem. Phys.* **19**, 774 (1951).
41. Y. Harano, T. Imai, A. Kovalenko, M. Kinoshita, F. Hirata. *J. Chem. Phys.* **114**, 9506 (2001).
42. T. Imai, Y. Harano, A. Kovalenko, F. Hirata. *Biopolymers* **59**, 512 (2001).
43. T. Yamazaki, A. Kovalenko. *J. Chem. Theory Comput.* **5**, 1723 (2009).
44. T. Yamazaki, A. Kovalenko. *J. Phys. Chem. B* **115**, 310 (2011).
45. T. Imai, K. Oda, A. Kovalenko, F. Hirata, A. Kidera. *J. Am. Chem. Soc.* **131**, 12430 (2009).
46. S. Gusarov, B. S. Pujari, A. Kovalenko. *J. Comput. Chem.* **33**, 1478 (2012).
47. A. Kovalenko, S. Ten-no, F. Hirata. *J. Comput. Chem.* **20**, 928 (1999).
48. P. Pulay. *Chem. Phys. Lett.* **73**, 393 (1980).
49. Y. Saad, M. H. Schultz. *J. Sci. Stat. Comput.* **7**, 856 (1986).
50. J. J. Howard, J. S. Perkyns, N. Choudhury, B. M. Pettitt. *J. Chem. Theory Comput.* **4**, 1928 (2008).
51. N. Minezawa, S. Kato. *J. Chem. Phys.* **126**, 054511 (2007).
52. J. A. Wagoner, N. A. Baker. *Proc. Natl. Acad. Sci. USA* **103**, 8331 (2006).

53. R. M. Levy, L. Y. Zhang, A. K. Felts. *J. Am. Chem. Soc.* **125**, 9523 (2003).
54. H. Gohlke, D. A. Case. *J. Comput. Chem.* **25**, 238 (2004).
55. K. Lum, D. Chandler, J. Weeks. *J. Phys. Chem. B* **103**, 4570 (1999).
56. E. J. Baerends, P. Ros, D. E. Ellis. *Chem. Phys.* **2**, 41 (1973).
57. G. te Velde, F. Bickelhaupt, S. van Gisbergen, C. Guerra, E. Baerends, J. Snijders, T. Ziegler. *J. Comput. Chem.* **22**, 931 (2001).
58. C. F. Guerra, J. Snijders, G. te Velde, E. Baerends. *Theor. Chem. Acc.* **99**, 391 (1998).
59. T. A. Wesolowski, A. Warshel. *J. Phys. Chem.* **97**, 8050 (1993).
60. L. Versluis, T. Ziegler. *J. Chem. Phys.* **88**, 322 (1988).
61. T. Miyata, F. Hirata. *J. Comput. Chem.* **29**, 871 (2008).
62. I. P. Omelyan, A. Kovalenko. *J. Chem. Phys.* **135**, 114110 (2011).
63. E. Barth, T. Schlick. *J. Chem. Phys.* **109**, 1617 (1998).
64. J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, R. D. Skeel. *J. Chem. Phys.* **114**, 2090 (2001).
65. R. D. Skeel, J. A. Izaguirre. *Mol. Phys.* **100**, 3885 (2002).
66. Q. Ma, J. A. Izaguirre. *Multiscale Model. Simul.* **2**, 1 (2003).
67. S. Melchionna. *J. Chem. Phys.* **127**, 044108 (2007).
68. G. J. Martyna, M. E. Tuckerman, D. J. Tobias, M. L. Klein. *Mol. Phys.* **87**, 1117 (1996).
69. A. Cheng, K. M. Merz Jr. *J. Phys. Chem. B* **103**, 5396 (1999).
70. J. Komeiji. *Mol. Struct.: THEOCHEM* **530**, 237 (2000).
71. W. Shinoda, M. Mikami. *J. Comput. Chem.* **24**, 920 (2003).
72. I. P. Omelyan, A. Kovalenko. *J. Chem. Phys.* **135**, 234107 (2011).
73. P. Minary, G. J. Martyna, M. E. Tuckerman. *J. Chem. Phys.* **118**, 2510 (2003).
74. P. Minary, M. E. Tuckerman, G. J. Martyna. *Phys. Rev. Lett.* **93**, 150201 (2004).
75. J. B. Abrams, M. E. Tuckerman, G. J. Martyna. *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*, Vol. 1, Springer, Berlin (2006) [*Lecture Notes in Physics* **703**, 139 (2006)].
76. T. Luchko, S. Gusarov, D. R. Roe, C. Simmerling, D. A. Case, J. Tuszynski, A. Kovalenko. *J. Chem. Theory Comput.* **6**, 607 (2010).
77. I. P. Omelyan, A. Kovalenko. *Mol. Simul.* (2012). http://dx.doi.org/10.1080/08927022.2012.700486
78. N. Bocquet, H. Nury, M. Baaden, C. Le Poupon, J.-P. Changeux, M. Delarue, P.-J. Corringer. *Nature* **457**, 111 (2009).
79. Y. Weng, L. Yang, P.-J. Corringer, J. M. Sonner. *Anesthesia Analgesia* **110**, 59 (2010).
80. D. Boda, W. Nonner, M. Valiskó, D. Henderson, B. Eisenberg, D. Gillespie. *Biophys. J.* **93**, 1960 (2007).
81. D. Boda, M. Valiskó, B. Eisenberg,W. Nonner, D. Henderson, D. Gillespie. *J. Chem. Phys.* **125**, 034901 (2006).
82. D. Boda, M. Valiskó, B. Eisenberg, W. Nonner, D. Henderson, D. Gillespie. *Phys. Rev. Lett.* **98**, 168102 (2007).
83. D. J. Tobias, C. L. Brooks III. *J. Phys. Chem.* **96**, 3864 (1992).
84. D. S. Chekmarev, T. Ishida, R. M. Levy. *J. Phys. Chem. B* **108**, 19487 (2004).
85. A. Kovalenko, F. Hirata. *J. Chem. Phys.* **115**, 8620 (2001).
86. A. Kovalenko, F. Hirata. *Condensed Matter Phys*. **4**, 643 (2001).
87. A. Tanimura, A. Kovalenko, F. Hirata. *Chem. Phys. Lett.* **378**, 638 (2003).
88. A. Kovalenko. *J. Comput. Theor. Nanosci.* **1**, 398 (2004).
89. A. Tanimura, A. Kovalenko, F. Hirata. *Langmuir* **23**, 1507 (2007).
90. J. Given. *Phys. Rev. A* **45**, 816 (1992).
91. J. Given, G. Stell. *J. Chem. Phys.* **97**, 4573 (1992).
92. J. Given, G. Stell. *Physica A* **209**, 495 (1994).
93. J. Given, G. Stell. In *Condensed Matter Theories*, Vol. 8, L. Blum, F. B. Malik (Eds.), pp. 395–410, Plenum, New York (1993).

94. L. L. Lee. *J. Chem. Phys.* **97**, 8606 (1992).
95. M. Endo, T. Takeda, Y. J. Kim, K. Koshiba, K. Ishii. *Carbon Sci.* **1**, 117 (2001).

