The multi-scale polarizable pseudo-particle solvent coarse-grained approach: From NaCl salt solutions to polyelectrolyte hydration

Michel Masella, Fabien Léonforté

The multi scale polarizable pseudo particle solvent coarse grained approach : from NaCl salt solutions to polyelectrolyte hydration.

Michel Masella*,† and Fabien Léonforté‡

†*Laboratoire de Biologie Structurale et Radiobiologie, Service de Bioénergétique, Biologie Structurale et Mécanismes, Institut de Biologie et de Technologies de Saclay, CEA Saclay, F-91191 Gif sur Yvette Cedex, France*

‡*L'Oréal Group, Research & Innovation, Aulnay-Sous-Bois,France*

E-mail: michel.masella@cea.fr

Abstract

We discuss key parameters that affect the reliability of hybrid simulations in the aqueous phase based on an efficient multi-scale coarse grained polarizable pseudo-particle approach, denoted as ppp^l, to model the solvent water, whereas solutes are modeled using an *all atom* polarizable force field. Among those parameters the extension of the solvent domain SD at the solute vicinity (domain in which each solvent particle corresponds to a single water molecule) and the magnitude of solute/solvent short range polarization damping effects are shown to be pivotal to model NaCl salty aqueous solutions and the hydration of charged systems like the hydrophobic polyelectrolyte polymer that we recently investigated [J Chem Phys, 114903 (155) 2021]. Strong short range damping is pivotal to simulate aqueous salt NaCl solutions at moderate concentration (up to 1.0M). The SD domain extension (as well as short range damping) has a weak effect on the polymer conformation, however it plays a pivotal role to compute accurate polymer/solvent interaction energies. As the ppp^l approach is up to two order of magnitude computationally more efficient than *all atom* polarizable force field methods, our results show it to be an efficient alternative route to investigate the equilibrium properties of complex charged molecular systems in extended chemical environments.

Keywords Polarization. Coarse grained model. Salt solutions. Polyelectrolyte polymer.

1 Introduction

Coarse Grained, CG, approaches are efficient molecular modeling scheme, which are more and more commonly used to investigate the properties of very large molecular systems that are beyond the computational capacity of standard *all atom* force fields (*i.e.* interaction potentials that explicitly take into account all the atoms of a molecular system). We may cite among others (see the recent reviews Refs.[1–3]) the MARTINI force field to model lipids, proteins, carbohydrates and other biomolecules,[4] the CG approach devoted to study complex aggregates made of huge chitosan chains (at the 1 000 unit size scale) in salty aqueous solutions,[5] as well as the water CG approach mW.[6] CG approaches are of particular interest in the industrial research and development field (from chemistry, oil manufacturers to personal care/cosmetics)[7–12] because of the complexity of most of the solutions that these industries use and the size of those solutions ingredients, like chitosan chains. Most of CG approaches rely on simple pair wise interaction potentials whose parameters are assigned from macroscopic properties (top down schemes, see for instance Chen et al.[13]) and/or from microscopic features of a molecular systems (bottom up schemes, see among others Shell[14]) using different numerical methods, like force matching[15,16] and machine learning[17–20] techniques. Recently sophisticated CG approaches accounting for many-body effects (by approximating the total potential energy of a molecular system as a sum N-body energy components) have emerged.[21,22]

In a series of articles[23–25] we detailed a multi-scale version, denoted as ppp^l, of the original polarizable hybrid CG scheme of Ha-Duong et al.[26] that was built to simulate the hydration of explicit solutes modeled using *all atom* force fields and as dissolved in a fluid of polarizable pseudo particles (denoted as *ppp*'s). Such kind of CG scheme belongs to the wide class of solvent implicit continuum approaches, like the popular Polarizable Continuum Model, PCM.[27] However, as the original *ppp* approach, its multi scale ppp^l version retains the notion of particle as it relies on a hierarchical representation of a solvent surrounding a solute: at short range from that solute the solvent is modeled by polarizable pseudo particles whose size corresponds to that of a single solvent molecule, whereas at longer and longer range from the solute larger and larger local volumes of the solvent are modeled by means of larger and larger ppp^l polarizable particles. Despite its sophistication, the computational complexity of the ppp^l approach scales as $O(N)$ and it allows

one to readily simulate using reduced computational ressources the hydration of complex solutes in extended water environments, as we recently showed for medium-size chitosan chains[28] as well as for a hydrophobic polyelectrolyte polymer dissolved in an aqueous environment comprising the equivalent of 10 M water molecules.[29]

To model water and regardless of their size, the (isotropic) polarizability α_s of the ppp^l particles obeys the Clausius-Mosotti relation[23,24]

$$\alpha_s = \frac{\epsilon_s - 1}{4\pi\rho_s\epsilon_s}. \tag{1}$$

Here ϵ_s and ρ_s are the dielectric constant and the density of liquid water, respectively. The natural choice is to set the particle volume to that of a water molecule. However the above relation allows us to consider a particle size that corresponds to a solvent local volume comprising a cluster of water molecules. We used that feature to propose the multi-scale version ppp^l depicted in Figure 1. α_s do not correspond to a standard microscopic (atomic or molecular) polarizability: it also allows to account for solvent molecular orientational polarization (*i.e.* water orientational perturbation arising from the solute presence).[26] For instance, by setting the particle volume to that of a single water molecule, the above relation yields $\alpha_s = 2.35$ Å^3, a value that is 60% larger than the isotropic polarizability of a water molecule (1.45 Å^3).

An important assumption of the ppp^l approach is the neglect of intra-solvent polarization effects: the magnitude of the particle induced dipole moments $\mathbf{p}_s$ is only a function of the solute electric field $\mathbf{E}_{\text{solute}}$ acting on the particle centers. However that '*local*' approximation may yield solute/solvent over polarization phenomena that can be circumvent by allowing the dipoles of the particles to saturate according to[26]

$$\mathbf{p}_s = \mu_s \mathcal{L}\left(\frac{3\alpha_s \mathbf{E}_{\text{solute}}}{\mu_s}\right)\frac{\mathbf{E}_{\text{solute}}}{|\mathbf{E}_{\text{solute}}|}, \tag{2}$$

here $\mathcal{L}$ is the Langevin function and μ_s is the particle saturation dipole value. The corresponding solvent/solute polarization energy for a system comprising N_s particles is then

$$U_{ps}^{pol} = -\frac{\mu_s^2}{3\alpha_s}\sum_{j=1}^{N_s}\ln\left[\frac{\sinh\left(3\alpha_s\left|\mathbf{E}_{\text{solute}}^{j}\right|/\mu_s\right)}{3\alpha_s\left|\mathbf{E}_{\text{solute}}^{j}\right|\mu_s}\right]. \quad (3)$$

All modern *all atom* polarizable force fields (*i.e.* interaction potentials that model water as an explicit tri-atomic polarizable molecule) account for short range damping of the electric fields.[30–32] As the *ppp* particles are expected to behave as real water molecules at short range from a solute, a reasonable assumption is also to damp the solute electric fields $\mathbf{E}_{\text{solute}}$ acting on the particles at short range. As they both weaken the solute electric field, dipole saturation and electric field short range damping are redundant effects, and it is far from obvious to disentangle them in order to build a transferable and reliable CG approach able to model complex charged microscopic systems.

The aim of the present study is to further discuss the ability of the ppp^l approach to simulate large, complex and highly charged solutes in the aqueous phase, in particular the behavior of the hydrophobic polyelectrolyte polymer that we recently investigated using a ppp^l model.[29] In that study, we showed the polymer (simulated with its counter ions) to rapidly collapse towards a globular form. Even if the goal of that original study was to discuss the strength of dynamic microscopic polarization effects on complex polyelectrolyte polymers, the nature of such a counter intuitive globular form for a heavily charged polymer needs to be further investigated to assess the reliability of the ppp^l approach. However as salts are usual components of complex solutions (as used in the personal care/cosmetic industry, for instance), we also assess in the present study the ability of the ppp^l approach to model NaCl salty aqueous solutions (within the 0.2 - 1.0M salt concentration range) as well as the effects of the presence of an opposite charge ion pair (whose cation/anion absolute charge varies from +1 to +4 e) on water at medium range. To this end we compare ppp^l simulation results to available data from accurate *all atom* polarizable force fields[33,34] and/or to new *all atom* polarizable force field simulations. Even if CG procedures from projection techniques that focus on dynamics properties have been proposed (see for instance Ref.[35] and the references mentioned therein), most of the CG approaches are essentially developed to study the equilibrium properties of large systems (in particular their structure). As our ppp^l approach is based on explicit particles, we also estimate a temporal property (namely, the ion diffusion coefficient) to discuss possible future routes of development for CG approaches.

First we will shortly detail the ppp^l approach and the main features of its latest version. Regarding the results, we focus our discussion on three key parameters that strongly affect the ppp^l approach accuracy, namely the dipole saturation μ_s, the solute/solvent electric field short range damping, and the solvent granularity to compute accurate solute/solvent energies. Below we label as '*ppp*' and 'ppp^l' the CG approaches, simulations and data generated using only small particles (whose volume is that of real water molecule) and set of particles of different size, respectively.

2 Computational details

2.1 Long range interaction truncation scheme

Solute/*ppp* particle electrostatic (polarization) interactions are truncated according to a shell-based cut off scheme: a *ppp* particle undergoes the electric field from all the solute electrostatic charges if the smallest distance among that particle and all the non-hydrogen solute atoms is smaller than a reference cut off distance R_{cut}. For other kinds of pair interactions, truncation is achieved using a standard spherical radius-based cut off scheme. Truncation is performed by scaling a pair potential energy component by means of a function G that smoothly downscales it for distances that span between R_{cut} and $R_{cut} + \delta R$. We systematically set δR to 0.5 Å and the function G is defined from a fifth order B-spline polynôme P_5, see Section S1 of Supplementary Material.

2.2 The *ppp* approach

The electric field generated by an atom j on a *ppp* particle i in Equations (2) and (3) obeys

$$\mathbf{E}^j_{\text{solute}} = \frac{1}{4\pi\epsilon_0 r_{ij}^3}\left[\lambda_{3,i}(q_i(\mathbf{r}_i - \mathbf{r}_j) - \mathbf{p}_i) - 3\lambda_{5,i}\frac{(\mathbf{p}_i \cdot \mathbf{r}_i) \times \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^2})\right], \tag{4}$$

$\mathbf{r}_i$ and $\mathbf{r}_j$ are the vector positions of the atom or particle i and j. q_i and $\mathbf{p}_i$ are the static Coulombic charge and the induced dipole moment of atom i. $\lambda_{3,i}$ and $\lambda_{5,i}$ are two damping functions that alter the magnitude of the electric field damping at short range[36] (up to 5.0 Å) :

$$\lambda_{3,i} = 1 - \exp\left(-a_i \times r_{ij}^3\right) \quad \text{and} \quad \lambda_{5,i} = 1 - (1 + a_i \times r_{ij}^3)\exp\left(-a_i \times r_{ij}^3\right). \tag{5}$$

a_i is a parameter, expressed in Å^{-3}, which depends only on the nature of atom i. The weaker is a_i the stronger are the corresponding damping effects.

On the other hand an atom i undergoes the electric field $\mathbf{E}^i_{\rm ppp}$ generated by only the induced dipoles of the *ppp* particles. We set the saturation parameter μ_s for atom dipole/atom dipole interactions to a large value (12 Debye) to recover the standard dipole/dipole interaction potential.

Atom/*ppp* interactions are truncated for distances larger than $R^{pol}_{\rm cut,1} = 12$ Å. Assuming the *ppp* density to be constant at long range from a non polarizable point charge Q, the long range electrostatic energy not taken into account because of truncation is

$$\delta G_{\rm lr} = -\frac{\alpha_s \rho_s}{4\pi\epsilon_0} \int_{R^{pol}_{\rm cut,1}}^{\infty} \left(\frac{Q}{r^2}\right)^2 4\pi r^2 dr = -\frac{Q^2}{8\pi\epsilon_0 R^{pol}_{\rm cut,1}}. \tag{6}$$

Accounting for the smoothing function G, that relation yields $\delta G_{\rm lr} = -13.5$ kcal mol^{-1} for a $R^{pol}_{\rm cut,1} = 12$ Å and $|Q| = 1$ e. For complex non-symmetric solutes, we proposed a multi-scale approach to compensate truncation, see Section 2.3 below.

Besides solute/solvent polarization, the *ppp* model also accounts for solute/solvent non electrostatic interactions using a Lennard-Jones-like potential truncated for distances larger than $R^{pol}_{\rm cut,1}$:

$$\tilde{U}^{LJ} = \sum_{i=1}^{N_a} \sum_{j=1}^{N_s} \epsilon^*_{i,ppp} \left[\left(\frac{\sigma^*_{i,ppp}}{r_{ij}}\right)^m - \left(\frac{m}{n}\right) \times \left(\frac{\sigma^*_{i,ppp}}{r_{ij}}\right)^n \right] \times G(r_{ij}). \tag{7}$$

N_a and N_s are the numbers of atoms and *ppp*'s, respectively. As in our original study,[25] we set (m,n) to (18,6). For given μ_s and a_i values, the parameters $\sigma_{i,ppp}$ and $\epsilon_{i,ppp}$ are assigned to best reproduce (within 0.1 kcal mol^{-1}) the available experimental hydration Gibbs free energy for a set a target entities, like ions Na^+ and Cl^-, as well as the mean solute/water distances in the entity first hydration shell. We don't pay attention here to reproduce ion first shell coordination numbers.

Interactions among *ppp*'s are modeled using Lennard-Jones-like term $\tilde{U}^{LJ}_{ppp}$ and a many-body term U^{dens} that is a function of the measures n_0 and n_1 of the instantaneous coordination numbers in first and second hydration shells of a *ppp* :

$$U^{dens} = \sum_{j=1}^{N_s} \sum_{p=1,2} \nu_p \times \max\left[(n_p - \bar{n}_p, 0)\right]^2. \tag{8}$$

Here ν_p and $\bar{n}_p$ are adjustable parameters. Our numerical tests show the choice of $\tilde{U}^{LJ}_{ppp}$ to have negligible effects on simulation results. All the data discussed below are based on the above potential $\tilde{U}^{LJ}$ using $m = 12$ and $n = 3$.

The parameters of $\tilde{U}^{LJ}_{ppp}$ and U^{dens} are assigned to reproduce (at ambient conditions) the hydration Gibbs free energy of a water molecule (6.3 kcal mol^{-1}), the water density (0.0335 molecules per Å^3) and the two regimes of the free energy corresponding to the creation of an empty cavity in neat water according to the Lum-Chandler-Weeks theory of hydrophobicity.[37] The latter conditions are met by truncating the terms $\tilde{U}^{LJ}_{ppp}$ and U^{dens} for inter *ppp* distances larger than $R^{cut}_{ppp} = 7$ Å, whereas a larger cut off distance is needed (at least about 15 Å) when only taking into account $\tilde{U}^{LJ}_{ppp}$. As the computational time to estimate interactions among *ppp*'s scales as $\left(R^{cut}_{ppp}\right)^3$, the use of the term U^{dens} provides an efficient way to compute them.

The use of the energy term U^{dens} in conjunction with $\tilde{U}^{LJ}$ yields *ppp* particles that diffuse fastly : their self diffusion coefficient is about seven times larger (17.2 10^{-5} cm^2 s^{-1} using the present *ppp* parameter set) than for real water at ambient conditions (2.6 10^{-5} cm^2 s^{-1}). That allows for a faster sampling of the potential energy surfaces of hydrated systems (at least for the solvent part).

In the present implementation of the *ppp* model, we use cubic periodic conditions to maintain constant the *ppp* density along NPT simulations, on average. Because of the truncation of long range inter *ppp* interactions, we usually need to only account for the 27 first *ppp* periodic images. To simulate a solute at the infinite dilution conditions we do not account for solute interactions with its own images. For a single molecule system (even a large protein or a polymer) the latter approach does not introduce artifacts if the *ppp* box in which that molecule is dissolved is large enough (*i.e.* as all the distances from solute atoms to the box boundaries are larger than $R^{pol}_{\mathrm{cut},1}$). Nevertheless we allow solute atoms to interact with *ppp* periodic images. That can yield artifacts like over accumulation of charge like ions at the *ppp* box boundaries as simulating a salt solution, for instance (that ion configuration maximizes the polarization of *ppp*'s located close to the box boundaries). To prevent such spurious effects, we enforce all solute atoms to be confined in a sub cubic volume of the *ppp* box by means of a harmonic potential detailed in Section S2 of Supplementary Material.

2.3 Long range electrostatics and the ppp^l approach

We showed the effect of truncating the solute electric field for atom/*ppp* distances larger than $R^{pol}_{\text{cut},1}$ to largely overestimate ion pair association in a *ppp* fluid as compared to real water.[25] To remediate that artifact we proposed the 'Russian doll' multi-scale ppp^l approach depicted in Figure 1. We add $l \geq 2$ levels of larger and larger *ppp*-like particles (denoted as ppp^l, ppp^1 particles are the original *ppp*'s). Here we set the ppp^l radius to that of the original *ppp*'s scaled by an integer $n = 2^{l-1}$. A single ppp^l particle thus models a cluster of 8^{l-1} original *ppp*'s.

Particles of a given level l (they are all of the same size) do not interact with the particles of a different level, and they only electrostatically interact with a solute if at least one atom/ppp^l distance r is included between $R^{pol}_{\text{cut},l-1}$ and $R^{pol}_{\text{cut},l}$. The solute/ppp^l electrostatic energy term for $l > 1$ corresponds to the linear version (for weak electric fields) of the polarization term of Equation (3) and dipole/dipole interactions among atoms and ppp^l particles are omitted. We maintain the ppp^l density along a simulation in the corresponding level box using periodic conditions. The numerical precision of a ppp^l scheme is estimated from Equation (6) and the largest $R^{pol}_{\text{cut},l}$ truncation distance.

For a given level $l > 1$, ppp^l's interact with each other according to the $\tilde{U}^{LJ}_{ppp}$ potential detailed above, with specific σ^*_l and ϵ^*_l parameters. For efficiency reason we truncate the ppp^l interactions for inter ppp^l distances larger than $5 \times l$ Å and we assign the parameters σ^*_l to reproduce the water density at ambient conditions for a neat ppp^l fluid (accounting for a ppp^l to correspond to a cluster of 8^{l-1} real water molecules). Setting all the ϵ^*_l's to the value ϵ^*_{ppp} of the original *ppp*'s, the latter condition is met using $\sigma^*_l = \gamma \times 2^{l-1} \times \sigma^*_{ppp}$, where $\gamma = 1.245615$.

The ppp^l masses are all set to 2^3 times the mass of an original *ppp* (*i.e.* the mass of a real water molecule), regardless of l. Such a large mass value slow down the ppp^l diffusion and that allows us to account for solute/ppp^l long range forces as slow fluctuating interactions in the multiple time step algorithm to solve the Newtonian equations of motion detailed below.

For readability purpose, simulations performed using only first level *ppp* particles are denoted below as performed in a neat *ppp* fluid, whereas they are denoted as performed in a ppp^l fluid as using higher level ppp^l particles.

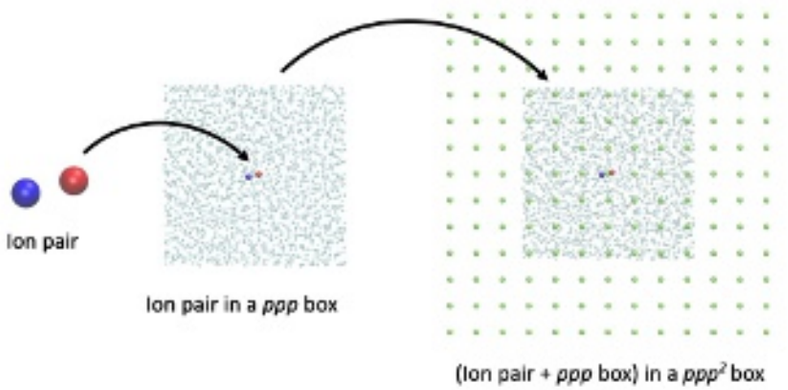

Figure 1: The multi-scale ppp^l approach. Here an ion pair (blue and red spheres) is dissolved in a box of *ppp*'s (each *ppp* models a single water molecule). Then the ion pair and the *ppp* box is set at the center of a new box made of ppp^2 particules, each of them models a local spherical volume of the solvent water. The new ion pair + *ppp*'s + ppp^2's systems may be then set in a larger ppp^3 box, up to reach a desired precision as computing long range solute/solvent electrostatic interactions.

2.4 Intra solute interactions

In the present study we simulate explicit ions $\mathbf{X}^{+Q}$ and $\mathbf{Y}^{-Q}$ as polarizable mono atomic centers. The monovalent ions are Na^+ and Cl^- whose interactions among them and explicit water molecules are modeled by means of the *all atom* polarizable force field detailed in an earlier study.[34] Ions with higher Coulombic charges ($Q \geq 2\ e$) are modeled using the Na^+/Cl^- *all atom* force field parameter set. However we reinforce ion/water short range repulsion to prevent unrealistic too short ion/water distances in MD simulations. Simulations of a hydrophobic polyelectrolyte polymer are performed using the polarizable *all atom* force field detailed in our earlier study.[29]

2.5 MD simulation details

Preliminary (relaxation) MD simulations are performed only in a neat *ppp* fluid and in the NPT ensemble, whereas production MD runs in *ppp* and ppp^l fluids are performed in the NVT ensemble. All simulations based on a *all atom* force field are performed in the NPT enesemble. Temperature and pressure are monitored along NPT runs using the Nosé-Hoover barostat[38] (the barostat coupling constant is set to 2.5 ps), whereas temperature is monitored along NVT runs according to a Langevin thermostat approach.[39] Induced dipole moments are solved iteratively until the mean

difference in their values between two successive iterations is smaller than 10^{-6} Debye and the maximum difference for a single dipole is smaller than 25×10^{-6} Debye. However for comparison purpose with earlier studies,[29,34] the dipole moments for salt solutions and hydrated **HPP** systems are iteratively solved until the mean dipole moment difference between two successive iterations is only smaller than 10^{-6} Debye.

The equations of motion are solved using a Multiple Time Steps, MTS, algorithm devoted to polarizable force field based on induced dipole moments.[40] For NaCl simulations, two time steps are used: 2 and 6 fs for short range and long range electrostatic/polarization forces, respectively. To simulate the hydration of more complex entities (with intra molecular chemical bonds like **HPP**) we used three time steps, namely 0.25 fs for intra molecular stretching and bending forces, and 1.0/5.0 fs (small systems) or 2.0/6.0 fs (large systems) for short and long range electrostatic and dispersion forces. 1-4 dihedral forces are considered as short range forces and the cutoff distance to compute electrostatic/dispersion short range forces is set to 8 Å. As discussed above, the forces corresponding to solute/ppp^l interactions are considered as long range polarization forces.

The starting simulation structure of a solute dissolved in a *ppp* box is built by setting the solute center of mass to the box center. Then the *ppp*'s are set on a cubic grid (the grid node dimension is 2.8 Å) if their distance to any solute non-hydrogen atom is < 3.5 Å. For ppp^l simulations, the starting coordinates of solute atoms and *ppp*'s correspond to the final point of a preliminary 5 ns NPT simulation in a neat *ppp* fluid. All molecular modeling computations and simulations were performed with our own code POLARIS(MD).[41]

2.6 Hydration free energy and PMF computations

Ion Gibbs hydration free energies ΔG_{hyd} are computed in two steps using a 32 windows Thermodynamical Integration, TI, scheme.[42] The first step consists in linearly downscaling to zero the ion charge and polarizability, and during the second step, the uncharged and non polarizable ion is then linearly transformed into a ghost entity fully decoupled from the solvent. To prevent numerical instabilities during the second step, we add the quantity $1 - \lambda$ to all ion/*ppp* distances as computing solute/solvent interactions (λ is the scaling parameter monitoring the progressive ion/*ppp* decoupling). Each TI MD simulation is performed at the 2.5 ns scale and the ion/*ppp* potential

energy derivatives are computed each 250 fs once a starting relaxation phase of 0.5 ns is achieved. The two free energy components computed from the two main steps (whose sum yields $\Delta G_{\rm hyd}$) are denoted $\Delta G_{\rm pol}$ and $\Delta G_{\rm np}$, respectively. The simulated systems are single ions dissolved in a cubic box comprising 1 000 *ppp*'s.

The Potential of Mean Force, PMF, of an ion pair is computed using an umbrella sampling MD protocol. The degree of freedom R restrained along MD simulations, using the harmonic potential $k_c\left(R-R_c\right)^2$, is the distance between the ion centers of mass. The target distances R_c usually span from 2 to 17 Å and they are regularly spaced by 0.5 Å. If not otherwise stated, the constant k_c is set to 5 kcal mol^{-1} Å^{-2}. The MD duration is set to 10 ns and the distances R are sampled each 50 fs once a starting phase of 1 ns is achieved. We post process the set of sampled distances to compute the PMF's according to the Umbrella Integration method.[43] The PMF's account for the entropic correction $2\mathrm{RT}\ln R$.

3 Results

In the following discussions, we denote as '*all atom*', '*ppp*' and '*ppp*^{l}' data computed along *all atom* simulations and simulations performed in a *ppp* or *ppp*^{l} fluids, respectively. We denote as Coulombic potential the classical effective potential $qq'/4\pi\epsilon r$ of two charges q and q' lying at a distance r from each other and dissolved in a fluid whose dielectric constant is ϵ.

3.1 *ppp* response to solute medium range electrostatics

For each of the four ion pairs [$\mathbf{X}^{+Q}$,$\mathbf{Y}^{-Q}$] defined in Section 2.4, we performed a 200 ns scale *all atom* simulation in bulk water at ambient conditions. The distance between the ions $\mathbf{X}^{+Q}$ and $\mathbf{Y}^{-Q}$ is harmonically restrained to a target distance of 14 Å (the harmonic constant is set to 50 kcal mol^{-1} Å^{-2}). We computed the mean water normalized density $\bar{\rho}_s$ and the mean dipole moment $\bar{\mu}_{\mathbf{X}}$ (projected on the ion pair axis and expressed in Debye per water molecule/*ppp*) within a cylindrical volume centered at the ion pair center and whose length and radius are 2 and 0.5 Å, respectively.

We computed accordingly the quantities $\bar{\rho}_s$ and $\bar{\mu}_{\mathbf{X}}$ along 200 ns *ppp* simulations performed using three dipole saturation μ_s values: 1.2, 2.0 and 12.0 Debye. The remaining ion/*ppp* parameters

are those of the model corresponding to the damping parameters $a_i = 0.3$ Å^{-3} detailed below. As the ion/*ppp* electric field damping vanishes beyond 5.0 Å, the magnitudes of the quantities $\bar{\rho}_s$ and $\bar{\mu}_{\mathbf{X}}$ do not depend on the damping parameters a_i's.

The $\bar{\rho}_s$ and $\bar{\mu}_{\mathbf{X}}$ values for *all atom* and *ppp* simulations are reported in Table S1 of Supplementary Material. The magnitude of μ_s has a weak effect on the densities $\bar{\rho}_s$ along *ppp* simulations. They depend mostly on the ion charge Q: they increase from 1.03 ($Q = 1$) to 1.20 ($Q = 4$) $\pm$ 0.01, regardless of μ_s. Along *all atom* simulations, the ion charge has a weaker but opposite effect on $\bar{\rho}_s$: it decreases from 0.97 to 0.90 $\pm$ 0.01 as Q increases.

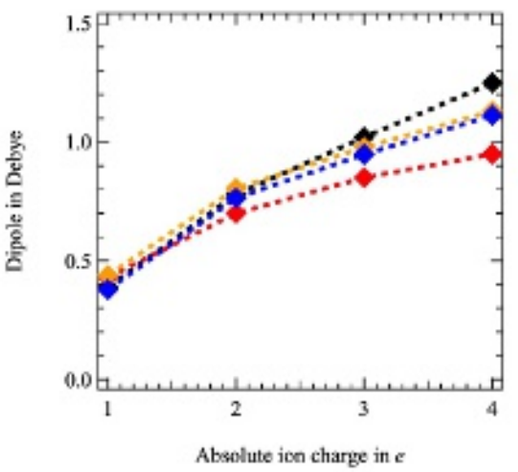

Figure 2: Water and *ppp* mean dipole projections $\bar{\mu}_{\mathbf{X}}$ at the center of the ion pair [$\mathbf{X}^{+Q}$,$\mathbf{Y}^{-Q}$] axis as a function of Q (*ppp* data correspond to $\mu_s = 1.2$ Debye). Black and red data : $\bar{\mu}_{\mathbf{X}}$ values from 200 ns scale simulations in an *all atom* water environment and in a *ppp* fluid, respectively. Blue and orange data : mean dipole projections $\bar{\mu}_{\mathbf{X}}$ scaled by the solvent mean density $\bar{\rho}_s$ at the ion pair center from *all atom* and *ppp* simulations, respectively.

μ_s has a much more significant effect on the *ppp* dipole projections $\bar{\mu}_{\mathbf{X}}$: they increase by a factor ranging from 2 ($\mu_s = 1.2$ Debye) up to 4 ($\mu_s = 12.0$ Debye) as Q increases from 1 to 4. We also note a good agreement between the *ppp* $\bar{\mu}_{\mathbf{X}}$ data corresponding to $\mu_s = 2.0$ Debye and *all atom* values. However as the *ppp* polarizability depends explicitly on the solvent density, we need to compare the $\bar{\mu}_{\mathbf{X}}$ values scaled by the local solvent densities to draw sound conclusions. In that case, the *ppp* values $\bar{\mu}_{\mathbf{X}} \times \bar{\rho}_s$ for $\mu_s = 1.2$ Debye nicely match the corresponding *all atom* data.

The parameter μ_s was originally introduced to prevent potential over polarization effects at

short range from a solute. In our *ppp* approach, that parameter allows us to readily compensate *ppp* over concentration effects at medium range from heavily charged solutes. Note that our estimate of the isothermal compressibility of a neat *ppp* fluid, about 16 ±1 10^{-6} atm^{-1} for pressures ranging from 1 to 1000 atm and T = 300 K, is three times smaller than for neat water.

3.2 *ppp*^{l} domain boundary effects on long range electrostatics

In the multi-scale *ppp*^{l} approach, solute/solvent long range electrostatics is taken into account by a sum of solvent domain effects. Each domain corresponds to set of particles of increasing size, particles that can over accumulate at the domain boundaries to strengthen solute/solvent interactions.[25] To investigate such potential drawbacks, we performed two series of simulations of the cations Na^+ and $\mathbf{X}^{2+}$, as well as of a single [Na^+,Cl^-] ion pair as dissolved alone in a two level *ppp*^{l} fluid. Each *ppp*^{l} box comprises about 2 000 particles and the largest cutoff distance $R^{pol}_{\mathrm{cut},2}$ is set to 31 Å. The two simulation series correspond to $R^{pol}_{\mathrm{cut},1}$ set to 12 and 15 Å, respectively. Regarding the ion pair and for each $R^{pol}_{\mathrm{cut},1}$ value, we performed three simulations along which we harmonically restrain the [Na^+,Cl^-] distance to 3, 5 and 8 Å, respectively.

From the particle normalized radial distribution functions $g^l_{ppp}(r)$ computed from the cations along the latter simulations, we compute the functions $\Delta\rho^l(r) = g^l_{ppp}(r) - 1$ (expressed in %, see their plots in Figure S1 of Supplementary Material). Below we denote as boundary domain **BD** the solvent shell domain corresponding to $R^{pol}_{\mathrm{cut},1} \pm 0.5$ Å, domain in which the ion electric field is progressively zeroed (original *ppp*'s) or increased from zero to its unaltered value (*ppp*^{2}'s).

For single cations, the magnitude of *ppp* and *ppp*^{2} particle over accumulation within **BD** increases as the cation charge increases and/or as the $R^{pol}_{\mathrm{cut},1}$ distance decreases. The *ppp* over accumulation is overall weak ($\Delta\rho^1(r)$ never exceeds 5 %), whereas the *ppp*^{2} over accumulation is much more accented, as $\Delta\rho^2(r)$ can be as large as +90% ($\mathbf{X}^{2+}$) and +15% (Na^+) using $R^{pol}_{\mathrm{cut},1}$ = 12 Å (and as large as 20% ($\mathbf{X}^{2+}$) and +6% (Na^+) for $R^{pol}_{\mathrm{cut},1}$ = 15 Å). Moreover the $\Delta\rho^2(r)$'s for single cations oscillate for distances above the boundary domain, showing the cation effects on the *ppp*^{l} fluid structure to extend at long range from them. For [Na^+,Cl^-] pairs and regardless of the cation/anion target distance, the magnitude of the differences $\Delta\rho^2(r)$ is much weaker compared to single cations. It amounts at most to +3 % within **BD** and then it converges rapidly to zero.

ppp over accumulation is taken into account as assigning solute/solvent parameters (see below). Its effect on hydration energies is thus removed. Regarding ppp^2's we estimate the energetic error arising from their over accumulation at **BD** according to :

$$\Delta U^2_{ps} = -\frac{\alpha_{s,2}\rho_{s,2}}{\epsilon_0}\int_{R^{pol}_{\mathrm{cut},1}}^{R^{pol}_{\mathrm{cut},2}} \Delta\rho^2(r)\frac{f(r)^2}{r^2}dr. \tag{9}$$

Here, $\alpha_{s,2}$ and $\bar{\rho}_{s,2}$ are the ppp^2 polarizability and mean density, respectively. f is the truncation function G at the ppp/ppp^2 boundary domain. The ppp^2 over accumulation at **BD** yields $\Delta U^2_{\mathrm{pol}}$ to amount from 4% (Na^+) up to 14% ($\mathbf{X}^{2+}$) of the expected ideal values computed from Equation 6 for $R^{pol}_{\mathrm{cut},1}$ = 12 Å, and by about 1.5 and 4% for the latter cations for $R^{pol}_{\mathrm{cut},1}$ = 15 Å. From the Na^+/ppp^2 radial distribution functions corresponding to our [Na^+,Cl^-] pair simulations, ΔU^2_{ps} amounts at most to only 0.05% of the expected U^2_{ps} value. Domain boundary effects thus yield strong drawbacks only for highly charged solutes as modeled at infinite dilution conditions.

Besides minimizing boundary domain artifacts on ion/ppp^2 electrostatic energies using large enough $R^{pol}_{\mathrm{cut},1}$ values, a computationally more efficient alternative route is to alter the solute/solvent electrostatic energy $u^{l>1}_{ps}$ for each ppp^l particle according to :

$$u^{l>1}_{ps} = -\alpha_{s,l>1}E_s^2 \times (1-\eta E_s^2), \tag{10}$$

here E_s is the electric field magnitude generated by a charged solute on a ppp^l particle. By setting the parameter η to 670.0 $Å^4\ e^{-2}$, the error $\Delta U^2_{\mathrm{pol}}$ for the single cations Na^+ and $\mathbf{X}^{2+}$ is then about 2% and 1% of the U^2_{pol} values as using $R^{pol}_{\mathrm{cut},1}$ = 12 and 15 Å, respectively. We systematically consider that scheme and the above value of the parameter η in all the simulations discussed below.

We discuss more in details the role of accounting for more and more ppp^l levels on ion pair PMF's in the following section. Anticipating our findings, the expected Coulombic potential for an ion pair [Na^+,Cl^-] dissolved in neat water for intermediate 8 to 16 Å ion separation distances is well reproduced by taking in account at least three ppp^l levels. Assuming that result, we investigated possible artifacts arising from the presence of domain boundaries on the long range tail of the PMF's of the ion pairs [Na^+,Cl^-] and [$\mathbf{X}^{2+}$,$\mathbf{Y}^{2-}$], in particular at ion separation distances close to twice the value of $R^{pol}_{\mathrm{cut},1}$. For such ion separation distances, the ppp^2's located in between the ions

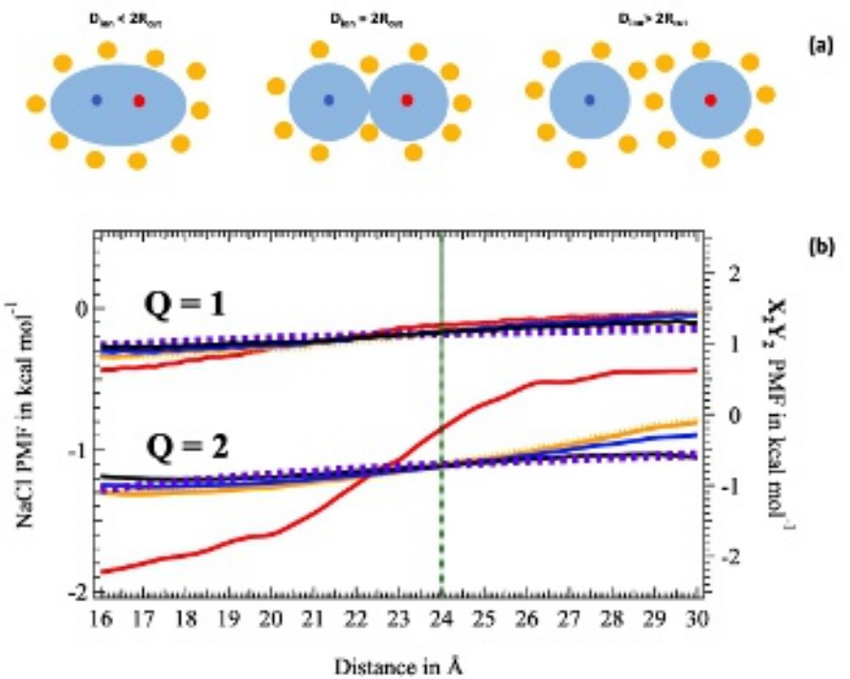

Figure 3: Ion pair PMF's in ppp^l fluids as a function of the truncation distance $R^{pol}_{cut,1}$. (a) schematic representation of the solvent elements explicitly taken into account *wrt* $R^{pol}_{cut,1}$. Blue: *ppp* particles represented as a continuum medium. Yellow spheres: ppp^2 particles. D_{ion} is the ion separation distance and $R_{cut} = R^{pol}_{cut,1}$. (b) PMF's of the [$Na^+$,$Cl^-$] (left axis) and [$\mathbf{X}^{2+}$,$\mathbf{Y}^{2-}$] (right axis) ion pairs as dissolved in ppp^l fluids. Violet dashed lines: expected Coulombic potentials for an opposite ion pair (the absolute charge of both ions is Q). Red, yellow (or yellow cross symbols), blue and black lines: PMF's computed by setting $R^{pol}_{cut,1}$ to 12, 15, 18 and 21 Å, respectively. All PMF's are computed in a four level ppp^l fluid, at the exception of those shown by small cross symbols, which correspond to a six level ppp^l fluid. All PMF's are switched to minimize their numerical difference with the Coulombic potential for ion separation distances spanning from 16 to 30 Å.

start to explicitly interact with them, see Figure 3(a). We performed a new series of 10 ns scale simulations of the ion pairs as dissolved in 7 000 ppp^l's boxes and by setting $R^{pol}_{\mathrm{cut},1}$ to 12, 15, 18 and 21 Å, respectively. Long range solvent effects are accounted for by means of 4 and 6 levels ppp^l schemes. These new PMF's are compared to the expected Coulombic potentials in Figure 3(b).

All the [Na^+,Cl^-] PMF's match the expected Coulombic potential on average, within less than 0.1 kcal mol^{-1} (regardless of $R^{pol}_{\mathrm{cut},1}$) and even 0.03 kcal mol^{-1} (for $R^{pol}_{\mathrm{cut},1} \geq 15$ Å) on the 16-30 Å distance range. However $R^{pol}_{\mathrm{cut},1}$ has a stronger effect on the [$\mathbf{X}^{2+}$,$\mathbf{Y}^{2-}$] PMF's. For instance the PMF computed for $R^{pol}_{\mathrm{cut},1} = 12$ Å (and by taking into account 4 ppp^l levels) obeys two regimes in the distance domain 16-30 Å, with a step transition between them at about 24 Å, *i.e.* at twice the $R^{pol}_{\mathrm{cut},1}$ value, a distance at which higher level ppp^2's located in between both ions start to interact with them. Increasing the cutoff distance $R^{pol}_{\mathrm{cut},1}$ to values > 12 Å yield a better agreement with the expected Coulombic potential (within less 0.05 kcal mol^{-1} for $R^{pol}_{\mathrm{cut},1} = 21$ Å, for instance).

Domain boundary effects on complex ionic solutions are expected to play a marginal role because of our shell truncation scheme. The presence of a 'real' molecular/ionic entities at the vicinity of an ion pair center would usually prevent to account for interactions between the ion pair and the high level ppp^l particles located in between those ions. As modeling the hydration of an assembly made of two independent molecular systems (and whose largest absolute total charges are larger than 1 e), we just need to set $R^{pol}_{\mathrm{cut},1}$ to a value larger than half the smallest inter atomic distance between the two systems to minimize the weight of the potential domain boundary artifacts.

3.3 Short range electric field damping and *ppp* models

To model the hydration of ions Na^+ and Cl^-, we built a series of *ppp* models for which we systematically set μ_s to 1.2 Debye. Those models differ by the value of the ion/*ppp* damping parameter a discussed in Section 2.2, which is taken identical for both ions. Here we consider the set of a values 0.03, 0.05, 0.08, 0.10, 0.15, 0.20 and 0.30 Å^{-3}. For each model, we assign the parameters of the ion/*ppp* energy term $\tilde{U}^{LJ}$ to reproduce the ion/water oxygen distances in ion first hydration shells from our earlier *all atom* simulations[34] (*i.e.* 2.4 and 3.4 Å for Na^+ and Cl^-, respectively) and to reproduce the ion experimental hydration Gibbs free energies $\Delta G^0_{\mathrm{hyd}}$ of Kelly et al.[44] within 0.1 kcal mol^{-1}. Those target energies are -103.8 (Na^+) and -74.5 (Cl^-) kcal mol^{-1}. The *ppp* energies

$\Delta G_{\rm hyd}$ are computed according to the TI protocol detailed in Section 2.6. The hydration energies account for the long range polarization correction of Equation (6) (-13.5 kcal mol^{-1}).

The components $\Delta G_{\rm pol}$ and $\Delta G_{\rm np}$ of the hydration energies $\Delta G_{\rm hyd}$ (see Section 2.6) for all *ppp* models are negative. Their ratios are plotted as a function of the damping parameter a in Figure S2 of Supplementary Material. Because of the shortest ion/*ppp* distances in Na^+ first hydration shell, the ratio for Na^+ varies on a wider range of values (from 1.5 to 16.5) than for Cl^- (from 1.5 to 7.5).

Ion first coordination shells comprise about 5-6 (Na^+) and 15-16 (Cl^-) *ppp*'s, regardless of a. Those coordination numbers n_c are in line with (Na^+) or twice as large (Cl^-) compared to experimental and *all atom* available data.[45–47] The entropic contribution to the hydration free energies $\Delta G_{\rm hyd}$ amounts to 45 cal mol^{-1} K^{-1}, regardless of the ion, a value that is from twice to four times larger compared to experimental-based data.[48,49] The inability of CG approaches to predict well balanced enthalpic and entropic contributions to the hydration free energies have already been discussed, in particular by Noid[3].

We also computed the diffusion coefficient D_i for single Na^+ or Cl^- ions as dissolved in 1 000 *ppp* box from 30 ns scale simulations and the ion velocity auto correlation functions (see Section S6 of Supplementary Material). As the *ppp* self diffusion coefficient (see Section 2.2), the *ppp* D_i estimates, 5.9 (Na^+) and 5.5 (Cl^-) 10^{-5} cm^2 s^{-1}, are largely overestimated compared to experimental data (1.33 and 2.09 10^{-5} cm^2 s^{-1} for both ions, respectively). D_i values from *all atom* simulations are commonly scaled to better match experimental data (see Ref.[50] for instance). However as ion coordination numbers are inferred to be key factors affecting ion diffusion,[51] a first step before building a correction scheme to compute reliable enough D_i values for ions from *ppp* simulations is to further improve the structural modeling of ion first hydration shell in *ppp* fluids (we reported an attempt in Ref.[25]).

We computed the PMF's of an ion pair [Na^+,Cl^-] in a four level *ppp*^{l} fluid from the 7 above models and 10 ns scale simulations. The simulated systems correspond to a single ion pair dissolved in a set of four 2 000 *ppp*^{l}'s boxes. That yields ion/solvent interactions to be accounted for up to 150 Å from the ions. The resulting PMF's are compared to earlier *all atom* data[34] in Figure 4(a). Whereas we are able to well reproduce the *all atom* PMF when using large damping parameter

values ($a \geq 0.2$ Å^{-3}), the use of weaker and weaker damping parameter values yields to more and more over estimate the depth of the first PMF minimum corresponding to an associated ion pair.

As shown by *all atom* simulations,[52] the magnitude of the ion association process in the aqueous phase is tied to the depth of that first PMF minimum. In particular the deeper is that first minimum the stronger is the expected percentage of ion associated in bulk water. To best reproduce the [Na^+,Cl^-] *all atom* PMF in water as using a *ppp* model corresponding to a weak damping parameter a, we introduce a many-body correction potential δU^a whose parameters are adjusted to best reproduce the *all atom* [Na^+,Cl^-] PMF in bulk water for each *ppp* model (see Section S5 of Supplementary Material). The new *ppp* PMF's based on δU^a are shown in Figure 4(b). They all overall match the *all atom* one, even if the second minimum of the new *ppp* PMF's is shifted to larger ion/ion distances (up to 6 Å), and the height of the energy barrier interconnecting the PMF first and second minimum is increased (up to reach 2.8 kcal mol^{-1}). Note the position of the second minimum in our *all atom* PMF[34] is 5.5 Å and the energy barrier height is 2.4 kcal mol^{-1}.

We compare in Figure 4(c) the PMF's of a [Na^+,Cl^-] pair dissolved in ppp^l fluids to its expected Coulombic potential for ion separation distances that span up to 16 Å and as using the *ppp* model corresponding to the damping parameter $a = 0.15$ Å^{-3}. Those PMF's are switched to minimize their numerical difference with the Coulombic potential for ion separation distances that span from 12 to 16 Å. Note the PMF's computed from all the above *ppp* models to be indistinguishable within the latter ion separation distance range. Moreover, within that distance range, the ions only interact with a single shell of *ppp*'s and with higher level ppp^l's that all lie outside that *ppp* first shell. The new PMF's converge towards the Coulombic potential as soon as a 3 levels ppp^l approach. By best fitting the new PMF's to a Coulombic-like potential $-1/(4\pi\tilde{\epsilon}_0^l r_{\text{ion}})$ (r_{ion} is the ion separation distance), we estimated the apparent dielectric constant $\tilde{\epsilon}_0^l$ of the ppp^l fluids at intermediate ion separation distances. $\tilde{\epsilon}_0^l$ is converged to the expected neat water value (78.35) as soon as 3 levels, within 1.5 % (that corresponds to uncertainty of our fitted values for $l \geq 3$).

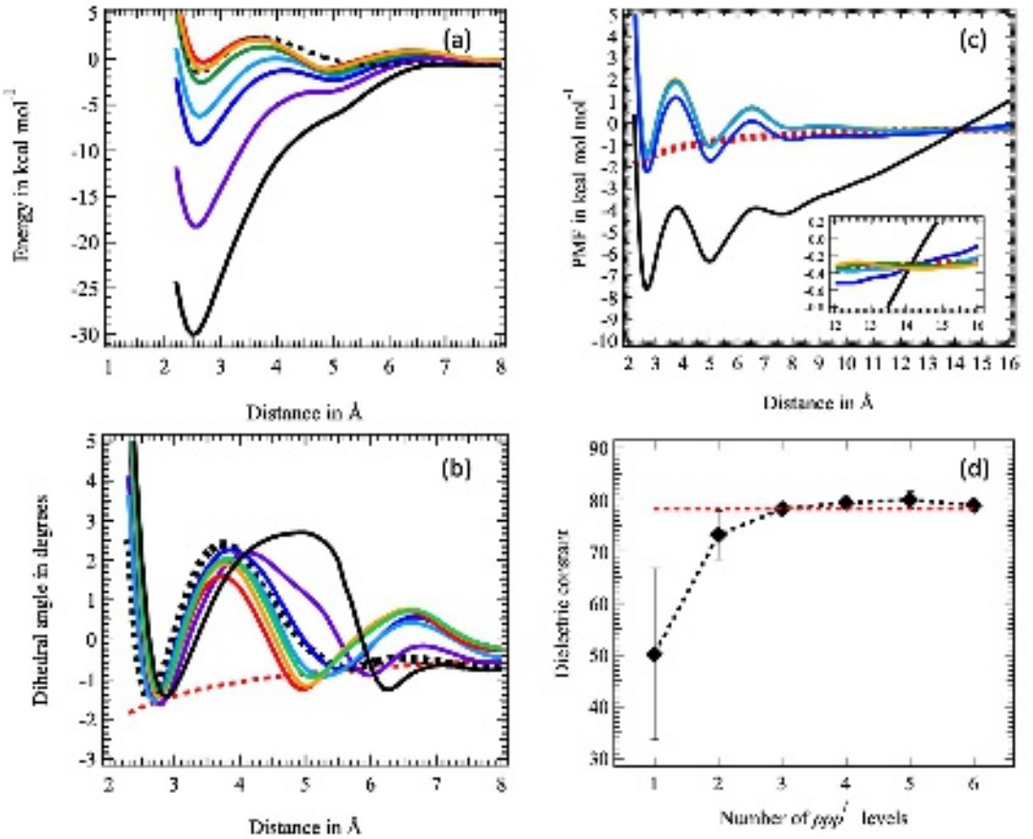

Figure 4: PMF's of the [Na^+,Cl^-] ion pair in ppp^l fluids. All the PMF's are shifted to best reproduce the expected Coulombic potential of the ion pair in neat water for inter ionic distances that span between 12 to 16 Å. From black to red bold lines: PMF's of *ppp* models corresponding to a damping parameter a that increases from 0.03 to 0.3 Å^{-3} (see text). (a) and (b): original ppp^l PMF's and new PMF's computed using the correction energy term δU^a. Black dashed line: [Na^+,Cl^-] PMF in neat water from *all atom* simulations.[34] Red dashed line: Coulombic potential for an opposite ion pair in neat water (the water dielectric constant is here 78.35). (c): PMF's for a damping parameter $a = 0.15$ Å^{-3} *wrt* the number of ppp^l levels taken into account (black to orange: from one to four ppp^l levels, red dashed line: Coulombic potential). The inset shows the ppp^l PMF's for inter ionic distances that span between 12 and 16 Å. (d): convergence of the ppp^l apparent dielectric constant $\tilde{\epsilon}_0^l$ as a function of the number of ppp^l levels taken into account (red horizontal dashed line: the dielectric constant of neat water at ambient conditions, *i.e.* 78.35).

3.4 NaCl association in a ppp^l fluid

We investigate 0.2, 0.6 and 1.0M salty NaCl aqueous solutions using the different *ppp* models detailed above. The simulated systems correspond to 200, 600 and 1 000 [Na^+,Cl^-] pairs dissolved in 100k *ppp* cubic boxes, respectively. For each salt concentration, we performed a 200 ns scale NVT

simulation per *ppp* model. These trajectories are sampled each 12 ps once a starting relaxation phase of 5 ns is achieved. The salt concentration is estimated from the cubic sub volumes in which the ions are confined , see Section 2.2. To account for long range electrostatics, we consider a four level ppp^l scheme. For readability purpose we discuss below energetic results using the unit Mcal, which corresponds to 10^3 kcal mol^{-1}.

We investigate ion clustering as in our study dealing about NaCl salty aqueous droplets.[34] At a time t_0 along a simulation, an ion cluster is a set of Na^+ and Cl^- ions that are all located at distance shorter than $d_{\rm ref}$ from at least one another cluster ion. The cluster survives until one of its ions leaves it or until a new ion is added to it. We set $d_{\rm ref}$ to 4.5 Å, see below. For each cluster size k we estimate the mean cluster survival time t_k^* from the correlation functions of the cluster survival probability at a time t from t_0.[34] ppp^l and *all atom* clustering data are summarized in Table 1 for 0.6 M solutions (and in Table S2 of Supplementary Materialfor 0.2 and 1.0M solutions).

Regardless of the salt concentration, the ion/*ppp* polarization damping parameter a has a strong effect on the Na^+/Cl^- association in a ppp^l fluid. Large values of a (≥ 0.20 Å^{-3}) yield the formation of large clusters comprising most of the ions, whereas weak value of a (≤ 0.08 Å^{-3}) yield the ions to be mostly free in solutions (see also Figure 5). In the latter case from 80% (0.6M and 1.0M salt solutions) and 95% (0.2M solutions) of the ions are dissociated in a ppp^l fluid. That is fully in line with earlier *all atom* simulations of salty aqueous bulk solutions and aqueous droplets performed using polarizable force fields.[33,34]

The order of magnitude of the cluster mean survival times t_k^* for cluster size ranging from 2 to 4 (at the 0.1 ns scale, see Table 1) are in line with available *all atom* estimates, at the remarkable exception of 0.2M and 0.6M salt solutions simulated using a low value of a (≤ 0.05 Å^{-3}). In that case, t_2^* may be as large as 1.3-1.8 ns, a value that is one order of magnitude larger than its *all atom* counterpart. We also note the near total absence of clusters whose size is larger than 2. Regarding the ion/ion radial distribution functions and the salt concentration computed using a damping parameter a set to 0.05 Å^{-3}, the [Na^+,Cl^-] ones are in line with *all atom* simulations (positions and heights of its main maxima and minima), whereas we note the shortest [Na^+,Na^+] and [Cl^-,Cl^-] distances to be shifted to larger distances by about 2 Å as compared to our own *all atom* data[34] (see Figure S4 of Supplementary Material). From visual inspection of the MD

trajectories, the associated [Na^+,Cl^-] pairs in a ppp^l fluid appear to be more uniformly distributed in the solvent than along *all atom* simulations (along such simulations, the associated pairs seem to form locally large and weakly bonded super clusters). The largest mean distance among [Na^+,Cl^-] pairs in a ppp^l fluid explains the largest ion pair mean survival time in a ppp^l fluid as using a low damping parameter a value as compared to *all atom* simulations.

Table 1: Ion association in 0.6 M NaCl ppp^l solutions as a function of the damping parameter a. Free ions: percentage of Na^+ and Cl^- ions not involved in a cluster. Clusters: occurrence of Na_xCl_y clusters (whose size $x+y=$ 2, 3 and 4) identified along the simulations and scaled by the number of sampled simulation snapshots (in parentheses, the cluster mean survival time, in ps). Largest cluster: size and occurrence of the largest cluster identified along the simulations. *na*: the statistical data set is too small to compute meaningful averages. (a) and (b): *all atom* data from our earlier study[34] dealing about of a 0.6M NaCl aqueous droplet that comprise about 100 k water molecules and extrapolated to the bulk limit, respectively.

a (in $Å^{-3}$)	Free ion in %		Clusters			Largest cluster
	Na^+	Cl^-	2	3	4	
0.30	0.56 ± 6.49	22.66 ± 6.93	0.3 (7)	0.1 (125)	0.2 (68)	554 (1)
0.20	0.70 ± 6.77	28.2 ± 5.56	0.3 (8)	0.1 (124)	0.2 (125)	547 (1)
0.15	4.45 ± 7.21	33.04 ± 6.26	0.2 (69)	0.3 (383)	0.2 (101)	526 (1)
0.10	41.52 ± 3.42	62.92 ± 3.42	3.6 (113)	2.5 (295)	0.3 (29)	9 (2)
0.08	66.68 ± 3.59	71.81 ± 3.07	3.7 (221)	2.7 (300)	0.2 (28)	9 (2)
0.05	80.64 ± 2.50	80.61 ± 2.50	1.2 (644)	0.1 (12)	0.0 (*na*)	4 (23)
0.03	80.01 ± 2.82	80.01 ± 2.82	0.5 (1814)	0.0 (*na*)	0 (*na*)	3 (78)
all atom (a)	81.8 ± 2.0	77.5 ± 2.0	6.2 (130)	0.8 (187)	0.2 (326)	10 (1)
all atom (b)	85.0 ± 2.0	82.0 ± 2.0				

As the ions are confined in a cubic sub volume of the simulation box, they do not interact with their periodic images. We may thus wonder about comparing our ppp^l data to earlier *all atom* ones in bulk water. First our earlier results in aqueous droplets show the main features of ion clustering to be almost fully converged to their bulk limit in 100 k water droplets.[34] Moreover ion clustering properties at the air/droplet interface are close to those at the droplet core. The presence of an interface has a priori a weak effect on NaCl association in aqueous finite environments. Moreover the solute/solvent polarization interaction energies corresponding to high level ppp^l domains are weak and they rapidly decrease until to be fully negligible for $l = 4$. Regardless of the damping magnitude and the salt concentration, the ppp^l/salt polarization energies are all larger than -10

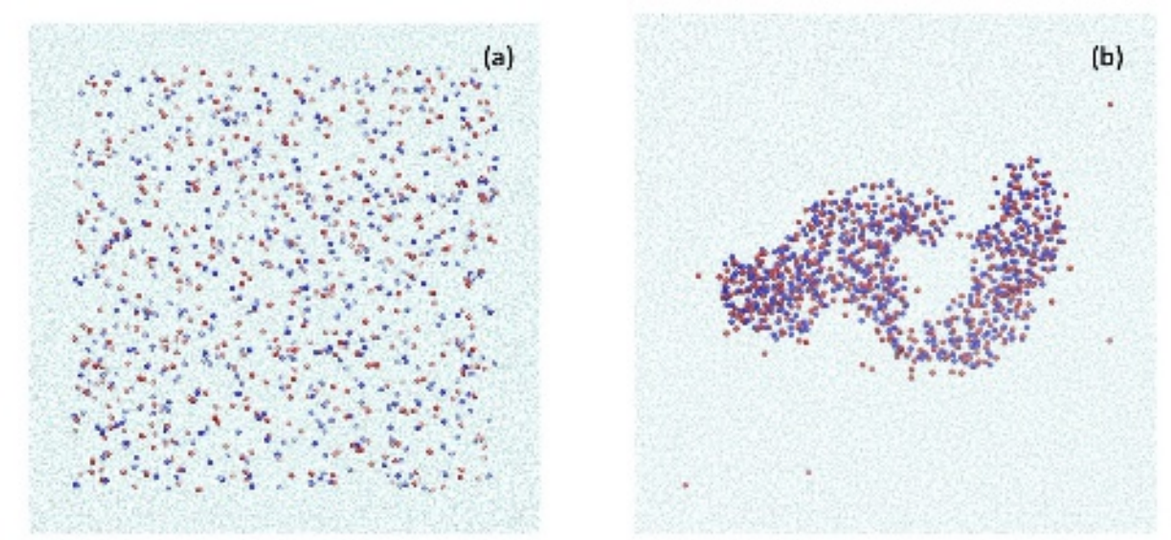

Figure 5: Two snapshots after 200 ns of simulations in a ppp^l fluid of a 1.0M NaCl salt solution (see text for details). (a) and (b): snapshots from a damping parameter a set to 0.05 (left) and 0.30 (right) Å^{-3}. *ppp* particles and ions Na^+ and Cl^- are shown by light blue, dark blue and red spheres, respectively.

Mcal for $l = 1$ (original *ppp*'s), whereas they are from three to five orders of magnitude weaker for $l = 2$ and 3, respectively. We may also note all the ion clustering properties for a 1.0M salt solution computed along a simulation in a neat *ppp* fluid fully match those corresponding to a four level ppp^l fluid. For instance the percentage of free ions along *ppp* simulations ($a = 0.05$ Å^{-3}) agrees with ppp^l data within less than 1 %.

3.5 A hydrophobic polyelectrolyte polymer and at infinite dilution conditions

In an recent study,[29] we discussed simulations dealing about the structural properties of a 10 units hydrophobic polyelectrolyte polymer (denoted as **HPP**) in neat *ppp* and ppp^2 fluids at infinite dilution conditions and in presence of counter ions Cl^-. Each **HPP** unit comprises seven di-allyl di-methyl ammonium cations and three acrylamide groups, see Figure 6(a). The polymer total charge is +70 e. In that former study, the damping parameters a for all the ionic entities was set to the large value 0.3 Å^{-3}. Starting from a quasi-linear conformation, **HPP** rapidly collapses towards a globular form surrounded by a spherical counter ion cloud, a structure that is stabilized by intra

solute polarization forces in both ppp^l fluids.

To discuss the reliability of that result, we built a new *ppp* model as detailed above for NaCl in which we systematically set the damping parameter a for all ionic species to 0.05 $Å^{-3}$, whereas we consider our earlier *ppp*/solute parameter set for the other kinds of polymer moieties. To assess the new *ppp* model, we first computed the PMF of the tetra methyl ammonium/chloride anion [$(CH_3)_4N^+$, Cl^-] ion pair dissolved in a four level ppp^l fluid from 10 ns scale simulations. We don't make use here of any correction potential δU^a. As shown by the plots of 6(c), the new *ppp* model is able to overall reproduce the main features of the *all atom* PMF as computed in neat water from our own polarizable *all atom* approach.[53]

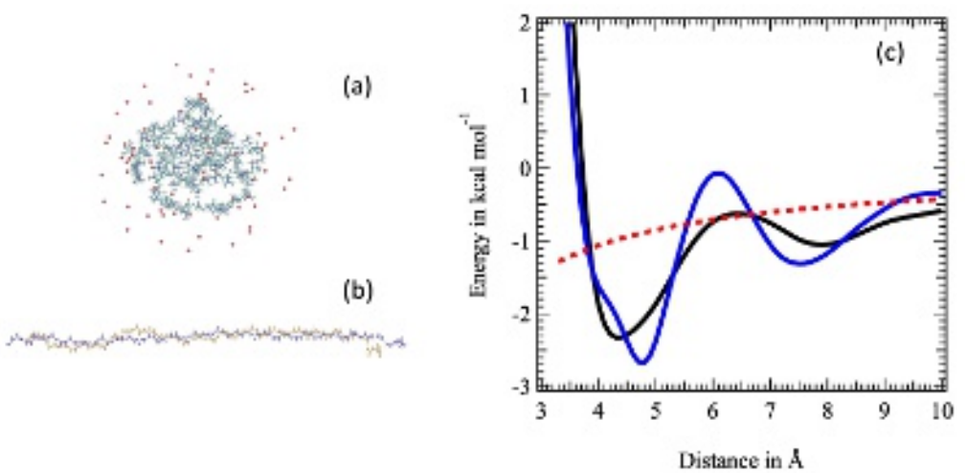

Figure 6: The polymer **HPP** and the new *ppp* model. (a) : example of a simulation final quasi spherical structure of the 10 unit **HPP** polymer with its counter ions Cl^- (in red) embedded in a 0.9M *ppp* cubic box. (b) : final quasi linear structure from simulation without center ions. (c) : tetra methyl ammonium/Cl^- PMF's in a four level ppp^l fluid (in blue) and in neat water using a polarizable *all atom* approach[53] (in black). Both the latter PMF's are shifted to best reproduce the expected Coulombic potential (dashed red line) within the ion separation distance range 12-16 Å. Regarding the *all atom* PMF, the tetra methyl ammonium/Cl^- force field parameters are set to reproduce high level quantum *ab initio* data regarding that ion pair in gas phase, as in our recent study dealing about NaCl salt solutions.[34]

We consider that new *ppp* model to perform two new sets of 10 simulations to investigate the **HPP** hydration, using the same computational protocol as in our original study.[29] We thus performed series of simulations of **HPP** dissolved in a 0.9M *ppp* cubic box using our original simulation starting structure in which the counter ions are located at short range from the **HPP**

cationic groups. For each simulation the atomic/particle starting velocities are randomly set and **HPP** is restrained to its starting linear geometry until an initial relaxation phase of 0.5 ns is achieved. **HPP** is dissolved in a *ppp* and in a ppp^2 fluid in the first and second simulation series. We set $R^{pol}_{\text{cut},1}$ and $R^{pol}_{\text{cut},2}$ to 12 and 150 Å, respectively. ppp^2 particles undergo the electric field arising only from the static charges of ionic moieties (**HPP** cationic groups and counter ions Cl^-). Because of the fast structural evolution of **HPP** as interacting with its counter ion cloud along all the new simulations, we stopped them once the **HPP** structural transition is achieved, usually within a few ns of simulation.

We also accordingly investigate the behavior of **HPP** in absence of counter ions and as dissolved in neat *ppp*, ppp^2 and ppp^3 fluids from simulations at the 100, 30 and 10 ns scale, respectively. The cut off distances $R^{pol}_{\text{cut},1-3}$ are set to 12, 150 and 440 Å. As dissolved in a ppp^3 fluid, we simulate **HPP** as embedded in an aqueous volume comprising the equivalent of 66 M explicit water molecules.

In presence of counter ions and as reported in our earlier study,[29] **HPP** again rapidly evolves towards a globular form (whose radius is about 15 Å) surrounded by a spherical counter ion cloud along all simulations, see Figure 6(a) (and Figure S5 of Supplementary Material). That conformational transition is again driven by polarization forces: intra solute Coulombic interactions disfavor the **HPP** globular form and its spherical counter ion cloud as compared to the starting linear **HPP** structure by about +6.5 Mcal, whereas the system total polarization energy stabilizes the globular form (including Cl^-) by about -10.0 Mcal. Contrary to NaCl solutions, ion/solvent short range electrostatic damping does not thus appear to play a pivotal role as modeling the hydration of **HPP** in presence of counter ions. That may arise from the overall large size of tetra methyl ammonium groups. From a simulation of a tetra methyl ammonium cation dissolved alone in a *ppp* box, we estimate the mean average distance between the *ppp*'s and nitrogen in the cation first hydration shell to be about 4.25 Å, a distance close to the damping maximum range, 5.5 Å. For Na^+ and Cl^-, their first hydration shell radii in a neat *ppp* fluid are about 2.3 and 3.3 Å, respectively.

In absence of counter ions, the **HPP** conformation is stable along the full trajectories in *ppp*, ppp^2 and ppp^3 fluids. It corresponds to an elongated and quasi-linear conformation regardless of the solvent extension, see Figure 6(b). The mean **HPP** end to end distance d_{ee} (measured from the first to the last polymer nitrogen) along the simulations is about 360 ± 10 Å. We may reasonably

assume that magnitude of d_{ee} to be a reliable estimate of the **HPP** persistence length.

To discuss the effects arising from the use of a short cut off distance $R^{pol}_{\mathrm{cut},1}$ and from taking into account only electrostatic interactions between **HPP** charged groups and high level $ppp^{n\geq 2}$ particles, we computed the mean electrostatic energy components $\bar{U}^{ppp^{1,2}}_{\mathrm{pol}}$ corresponding to the interactions of the full solute (**HPP** + counter ions if they are accounted for) with each kind of ppp and ppp^2 particles in simulation segments along which the final polymer conformation is stable (regarding **HPP** in presence of counter ions, the energies $\bar{U}^{ppp^{1,2}}_{\mathrm{pol}}$ are averaged over the final segments of the 10 simulation set). Those $\bar{U}^{ppp^{1,2}}_{\mathrm{pol}}$ energies are computed for $R^{pol}_{\mathrm{cut},1}$ distances that span from 12 to 36 Å (whereas $R^{pol}_{\mathrm{cut},2}$ is systematically set to 150 Å). The sums $\bar{U}^{ppp^{1+2}}_{\mathrm{pol}}$ of those energies are shown in Figure 7.

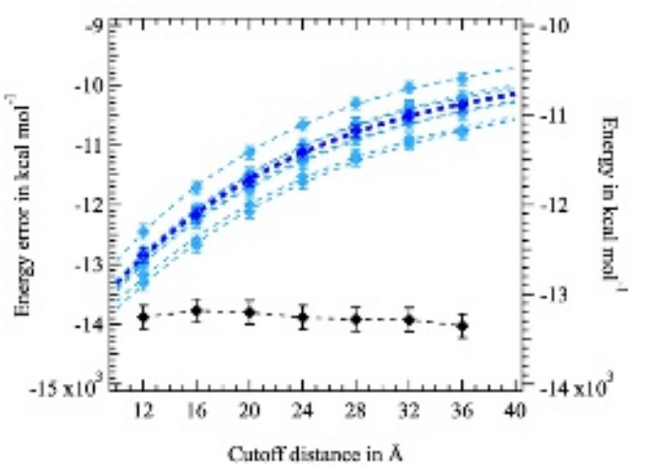

Figure 7: (a) : interaction energies $\bar{U}^{ppp^{1+2}}_{\mathrm{pol}}$ as a function of the cut off distance $R^{pol}_{\mathrm{cut},1}$. Black symbols (left axis), light and dark blue symbols (right axis) : energy data for the **HPP** globular and quasi linear conformations, respectively. The error bars are the root mean square deviations of the averaged data. Dashed blue lines : best exponential functions $U_\infty + b_0 \exp\left(-c_0 \times R^{pol}_{\mathrm{cut},1}\right)$ whose parameters U_∞, b_0 and c_0 are adjusted to reproduce the raw energy data. Regarding the energies $\bar{U}^{ppp^{1+2}}_{\mathrm{pol}}$, their $\bar{U}^{ppp^{1}}_{\mathrm{pol}}$ and $\bar{U}^{ppp^{1}}_{\mathrm{pol}}$ components are in a 3:2 ratio for the linear conformation, whereas that ratio varies from 10:1 up to 100:1 for globular geometries.

Interestingly, the sum $\bar{U}^{ppp^{1+2}}_{\mathrm{pol}}$ for **HPP** in a linear conformation (and simulated in absence of explicit counter ions) is constant within 0.4 %, regardless of the $R^{pol}_{\mathrm{cut},1}$ value. In that case $\bar{U}^{ppp^{1+2}}_{\mathrm{pol}}$ amounts to about -13.25 ± 0.05 Mcal. However that sum for the **HPP** globular forms (taking into

account the counter ions) noticeably increases as $R_{\mathrm{cut},1}^{pol}$ increases, from -12.5 ± 0.3 Mcal ($R_{\mathrm{cut},1}^{pol}$ = 12 Å) to -10.7 ± 0.3 Mcal ($R_{\mathrm{cut},1}^{pol}$ = 36 Å). For globular forms, the $\bar{U}_{\mathrm{pol}}^{ppp^{1+2}}$ data can be accurately reproduced using an exponential function, which yields an extrapolated value for that energy of -9.60 ± 0.2 Mcal at $R_{\mathrm{cut},1}^{pol}$ = 150 Å.

The stronger effect of $R_{\mathrm{cut},1}^{pol}$ on the solute/*ppp* electrostatic interactions for **HPP** in a globular form as compared to a quasi linear one is tied to the magnitude of the **HPP** atomic induced dipole moments. In globular conformations, and as already shown,[29] the **HPP** carbon dipole moment increases as the carbon distance to the polymer center of mass increases up to reach values larger than 4 Debye at the **HPP** globular surface. In a quasi elongated conformation (with no counter ion), the carbon induced dipoles are weaker and they never exceed 1.4 Debye (see Figure S6 of Supplementary Material). The large induced dipole moment values for carbon atoms in globular conformations are in line with the strength of the dipole moments of water molecules interacting at short range from highly charged cations (3 Debye and above), as reported from polarizable *all atom* and quantum Car-Parinello simulations.[54,55] Note that for efficiency reasons the ppp^2 particles do not undergo the electric field arising from the solute induced dipoles in the present study.

Taking into account all the energy contributions, for instance the intra solute mean potential energy $\bar{U}^{\mathbf{HPP}}$ (which strongly favors **HPP** globular forms compared to the linear one, see the data reported in Section S8 of Supplementary Material), the **HPP** linear conformation is more stable (enthalpically) at infinite dilution conditions than the globular forms (interacting with counter ions at short range) by about 0.6 Mcal based on $\bar{U}_{\mathrm{pol}}^{ppp^{1+2}}$ data extrapolated at $R_{\mathrm{cut},1}^{pol}$ = 150 Å. Note also that the difference in intra solvent energies between linear and globular geometries is weak, but still in favor of the linear form by about 0.30 ± 0.05 Mcal. Lastly, as the dissociation of a molecular assembly dissolved in a neat water is usually favored by entropic effects, we may thus reasonably conclude the **HPP** elongated and quasi-linear form to be more stable than globular ones at infinite dilution conditions as taking into account an infinitely large ppp^l environment.

Besides the overall reduced simulation times that can be readily investigated nowadays using standard MD techniques (which prevent an exhaustive exploration of molecular system potential energy surfaces), our new results thus show that simulating a polyelectrolyte polymer with its counter ions to investigate the polymer behavior at infinite dilution conditions is questionable,

especially in the case of polyelectrolyte polymers comprising large polarizable and hydrophobic sub units and whose ionic groups are at a long enough distance from each other, like **HPP**. That arises in particular from the need of explicit accounting for infinitely large solvent domains, which can not be readily simulated even on modern computing systems.

Simulations in a water droplet using a polarizable *all atom* force field[29] showed **HPP** globular forms surrounded by a spherical counter ion cloud to be stable. Such compact geometries thus correspond at least to meta stable states in aqueous environments, regardless of the description of the solute/solvent interactions. Hence stable polyelectrolyte polymer globular forms surrounded by spherical counter ion clouds as predicted by our multi scale simulation approach for **HPP** do not a priori arise from artifacts tied to our solvent approach. Note also that globular forms for hydrophobic polyelectrolyte polymers in presence of counter ions are not systematically predicted by our multi scale ppp^l approach. In our earlier study[29] we also reported simulations of carboxylated polystyrenes with a high fraction charge and in presence of Na^+ counter ions (that are also located close to the polymer ionic heads in the simulation starting structures). Those kinds of polyelectrolyte polymer do not evolve towards a globular form as simulated in a ppp^l fluid in presence of Na^+ counter ions : their linear starting structure is stable along 100 ns scale trajectories.

4 Conclusion

In the present study we discussed potential issues that can alter the reliability of simulations performed using the multi-scale polarizable pseudo-particle ppp^l approach to model the hydration of large and complex polyelectrolyte polymers in salty aqueous solutions. Among the ppp^l parameters that handle the solute/solvent interactions, the magnitudes of two of them, namely the extension of the solvent sub domain SD at the close vicinity of the solute (domain in which each *ppp* particle corresponds to a single water molecule), and the intensity of *ppp*/solute atom short range electrostatic damping, are shown to be pivotal. The extension of sub domain SD is a key parameter to achieve a high accuracy as modeling the hydration of highly charged solutes like the hydrophobic polyelectrolyte polymer **HPP** or the association of highly charged ion pairs. Interestingly that issue is tied to the modeling of medium range solute/solvent interactions (*i.e.* interactions between

ppp's and solute atoms that lie between 10 to 20 Å from each other). However that issue appears to have a weak effect on the structural behavior of a complex polyelectrolyte polymer like **HPP** dissolved in a ppp^l fluid. For efficiency reason we thus recommend to simulate the hydration of complex solutes using a moderately extended *ppp* sub domain SD (*i.e.* a *ppp* shell extending up to 12 Å from any solute atom) and to post process the simulations using a larger extension for SD to better assess the enthalpic stability of the solute conformations observed along a simulation.

Solvent/solute short range electrostatic damping appears to be pivotal as modeling small ions (like Na^+ and Cl^-) and thus common salt solutions. In order to reproduce data from polarizable *all atom* simulations regarding NaCl aqueous solutions at molar/sub molar scale concentrations (data that agree with experiment), we show that solute/solvent short range electrostatic damping effects have to be strong enough to yield a close weight for the electrostatic and non electrostatic components of the hydration Gibbs free energies of small ions. However for large ions like tetra methyl ammonium, damping effects appear to have a marginal role in modeling the hydration of a complex solute like **HPP**. In all we recommend to systematically consider *ppp*/solute strong short range damping to model ion hydration in ppp^l fluids.

Another issue that can affect the conclusions drawn from the ppp^l approach, issue that is not specific to it as it can also impact the conclusions drawn from standard *all atom* simulations, is the explicit accounting of counter ions to model the hydration of a solute at infinite dilution conditions. For instance simulations of **HPP** in presence of counter ions (that are at contact of the **HPP** cationic groups in the starting simulation structure) can lead to compact/globular polymer conformations interacting with an external spherical counter ion cloud, assembly that is stabilized by polarization forces. Such a globular conformation, that is also a priori stable as simulated in a water droplet using a polarizable *all atom* force field, is shown however to be less enthalpically stable than a **HPP** quasi linear conformation fully dissociated from its counter ions by enhancing the precision of the multi scale ppp^l approach (that is achieved by enlarging the upper bound of the sub domain SD away from the solute) and by accounting for the long range solute/solvent interactions not accounted for during the simulations and arising from an infinitely large water environment. We thus recommend to simulate polyelectrolyte polymers, like **HPP**, dissolved alone in water environments in absence of explicit counter ions to investigate their behavior at infinite

dilution conditions. However counter intuitive meta stable conformations, like **HPP** globular forms, which may be observed as simulating a solute in presence of counter ions, are also pivotal as building coarse grained approaches from bottom up schemes in order to account for the solute flexibility.

As most of the CG methods, our ppp^l approach is well suited to investigate the equilibrium properties of large systems in extended aqueous environments. However we briefly discussed in the present study the Na^+ and Cl^- self diffusion coefficients D_i (at ambient and at infinite dilution conditions) in a *ppp* fluid. As the self diffusion coefficient of *ppp* particles, the order of magnitude of ion D_i's are largely overestimated in a *ppp* fluid as compared to experiment. The *ppp* approach (as all CG models) rely on elimination of the fast varying degrees of freedom of the original microscopic (*all atom*) systems (like the degrees of freedom tied to water hydrogens). That is expected to prevent computing accurate temporal properties from *ppp* simulations.[35] As usually proposed as using *all atom* methods,[50] a scaling scheme to post process diffusion coefficients from *ppp* data may be a priori considered. However and to our opinion, the *ppp* approach should be first improved in modeling ion first hydration shells before to propose such a scaling scheme.

In all our results show hybrid simulations based on polarizable pseudo-particles to model extended chemical environments (in particular water) are an efficient alternative route to standard *all atom* approaches to investigate equilibrium properties of large and complex microscopic systems. For instance, whereas our former *all atom* simulations[34] of NaCl aqueous droplets comprising 100k explicit water molecules (using a polarizable force field and a scalable $O(N)$ Fast Multiple Method to compute electrostatic and polarization energies[56]) were run at a rate of about 1 ns per day using 8 AMD 64 cores CPU, we perform the corresponding ppp^l simulations at a rate that spans from 5 to 10 ns using only a single AMD 64 core processor and a basic $O(N^2)$ approach to compute ion/ion interactions. Polarizable pseudo-particle approaches are thus a promising tool to ease the use of sophisticated force fields as modeling large, complex and even heavily charged molecular systems.

Supplementary Material

Details about the long range truncation scheme, the way a solute is constrained within a cubic sub volume of the main simulation box, the correction potential used to alter the NaCl PMF in a

ppp fluid and energy components regarding the hydration of the **HPP** polymer may be found in Supplementary Material, as well as the figures and tables mentioned in the main manuscript.

Acknowledgments

This work was granted access to the TGCC HPC resources under the Grand Challenge allocation [GC0429] made by GENCI.

Conflicts of interest

F. L. is a full employee of L'Oréal involved in research activities.

S1 Long range truncation S6 S2 Constraining a solute in cubic sub volume S6 S3 Ion/solvent medium range interactions S7 S4 Domain boundary effects S8 S5 Correction potential to alter the NaCl PMF in water S9 S6 Diffusion coefficients S10 S7 NaCl association in a pppl fluid and radial distribution functions

References

(1) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chemical Reviews* **2016**, *116*, 7898–7936.

(2) Joshi, S. Y.; Deshmukh, S. A. A review of advancements in coarse-grained molecular dynamics simulations. *Molecular Simulation* **2021**, *47*, 786–803.

(3) Noid, W. G. Perspective: Advances, Challenges, and Insight for Predictive Coarse-Grained Models. *The Journal of Physical Chemistry B* **2023**, *127*, 4174–4207.

(4) Marrink, S. J.; Monticelli, L.; Melo, M. N.; Alessandri, R.; Tieleman, D. P.; Souza, P. C. T. Two decades of Martini: Better beads, broader scope. *WIREs Computational Molecular Science* **2023**, *13*, e1620.

(5) Tsereteli, L.; Grafmüller, A. An Accurate Coarse-Grained Model for Chitosan Polysaccharides in Aqueous Solution. *PLOS ONE* **2017**, *12*, 1–31.

(6) Molinero, V.; Moore, E. B. Water Modeled As an Intermediate Element between Carbon and Silicon. *The Journal of Physical Chemistry B* **2009**, *113*, 4008–4016.

(7) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of Coarse Graining in Biomolecular Simulations. *WIREs Computational Molecular Science* **2014**, *4*, 225–248.

(8) Hospital, A.; Goñi, J.; Orozco, M.; Gelpi, J. Molecular Dynamics Simulations: Advances and Applications. *Adv Appl Bioinform Chem. B* **2015**, *8*, 37–47.

(9) Lu, L.; Yoo, K.; Benyahia, S. Coarse-Grained-Particle Method for Simulation of Liquid-Solids Reacting Flows. *Industrial & Engineering Chemistry Research* **2016**, *55*, 10477–10491.

(10) Illa-Tuset, S.; Malaspina, D. C.; Faraudo, J. Coarse-Grained Molecular Dynamics Simulation of the Interface Behaviour and Self-Assembly of CTAB Cationic Surfactants. *Phys. Chem. Chem. Phys.* **2018**, *20*, 26422–26430.

(11) Mehandzhiyski, A. Y.; Rolland, N.; Garg, M.; Wohlert, J.; Linares, M.; Zozoulenko, I. A novel supra coarse-grained model for cellulose. *Cellulose* **2020**, *27*, 4221–4234.

(12) Nakamura, H.; Takimoto, H.; Kishida, N.; Ohsaki, S.; Watano, S. Coarse-Grained Discrete Element Method for Granular Shear Flow. *Chemical Engineering Journal Advances* **2020**, *4*, 100050.

(13) Chen, J.; Chen, J.; Pinamonti, G.; Clementi, C. Learning Effective Molecular Models from Experimental Observables. *Journal of Chemical Theory and Computation* **2018**, *14*, 3849–3858.

(14) Shell, M. S. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *The Journal of Chemical Physics* **2008**, *129*, 144108.

(15) Wang, Y.; Noid, W. G.; Liu, P.; Voth, G. A. Effective force coarse-graining. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2002–2015.

(16) Bacle, P.; Jardat, M.; Marry, V.; Mériguet, G.; Batôt, G.; Dahirel, V. Coarse-Grained Models of Aqueous Solutions of Polyelectrolytes: Significance of Explicit Charges. *J. Phys. Chem. B* **2020**, *124*, 288–301.

(17) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics* **2016**, *145*, 170901.

(18) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. DeePCG: Constructing Coarse-Grained Models via Deep Neural Networks. *The Journal of Chemical Physics* **2018**, *149*, 034101.

(19) Wang, J.; Olsson, S.; Wehmeyer, C.; Pérez, A.; Charron, N. E.; de Fabritiis, G.; Noé, F.; Clementi, C. Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. *ACS Central Science* **2019**, *5*, 755–767.

(20) Li, W.; Burkhart, C.; Polińska, P.; Harmandaris, V.; Doxastakis, M. Backmapping Coarse-Grained Macromolecules: An Efficient and Versatile Machine Learning Approach. *The Journal of Chemical Physics* **2020**, *153*, 041101.

(21) John, S. T.; Csányi, G. Many-Body Coarse-Grained Interactions Using Gaussian Approximation Potentials. *The Journal of Physical Chemistry B* **2017**, *121*, 10934–10949.

(22) Wang, J.; Charron, N.; Husic, B.; Olsson, S.; Noé, F.; Clementi, C. Multi-body effects in a coarse-grained protein force field. *The Journal of Chemical Physics* **2021**, *154*, 164113.

(23) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model: Application to Long Dynamics Simulations of Bovine Pancreatic Trypsin Inhibitor. *Journal of Computational Chemistry* **2008**, *29*, 1707–1724.

(24) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model. II. Accounting for Hydrophobic Effects. *Journal of Computational Chemistry* **2011**, *32*, 2664–2678.

(25) Masella, M.; Borgis, D.; Cuniasse, P. A multiscale coarse-grained polarizable solvent model for handling long tail bulk electrostatics. *Journal of Computational Chemistry* **2013**, *34*, 1112–1124.

(26) Ha-Duong, T.; Phan, S.; Marchi, M.; Borgis, D. Electrostic on Particles : Phenomenological and Orientational Density Functional Theory Approach. *J. Chem. Phys.* **2002**, *117*, 541–556.

(27) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chemical Reviews* **2005**, *105*, 2999–3094.

(28) Masella, M.; Léonforté, F. Chitosan Polysaccharides from a Polarizable Multiscale Approach. *ACS Omega* **0**, *0*, null.

(29) Masella, M.; Crudu, A.; Léonforté, F. Hybrid Polarizable Simulations of a Conventional Hydrophobic Polyelectrolyte. Toward a Theoretical Tool for Green Science Innovation. *The Journal of Chemical Physics* **2021**, *155*, 114903.

(30) Réal, F.; Vallet, V.; Flament, J.-P.; Masella, M. Revisiting a Many-Body Model for Water Based on a Single Polarizable Site. From Gas Phase Clusters to Liquid and Air/Liquid Water Systems. *J. Chem. Phys.* **2013**, *139*, 114502.

(31) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. J. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chemical Reviews* **2016**, *116*, 4983–5013.

(32) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. **2019**, *48*, 371–394.

(33) Soniat, M.; Pool, G.; Franklin, L.; Rick, S. W. Ion Association in Aqueous Solution. *Fluid Phase Equil.* **2016**, *407*, 31–38.

(34) Vallet, V.; Coles, J.; Réal, F.; Houriez, C.; Masella, M. NaCl Salts in Finite Aqueous Environments at the Fine Particle Marine Aerosol Scale. *ACS Earth and Space Chemistry* **2022**, *6*, 1612–1626.

(35) Uneyama, T. Application of Projection Operator Method to Coarse-Grained Dynamics with Transient Potential. *Phys. Rev. E* **2022**, *105*, 044117.

(36) Thole, B. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341–350.

(37) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437*, 640–647.

(38) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit Reversible Integrators for Extended Systems Dynamics. *Molecular Physics* **1996**, *87*, 1117–1157.

(39) Cancès, Eric,; Legoll, Frédéric,; Stoltz, Gabriel, Theoretical and Numerical Comparison of Some Sampling Methods for Molecular Dynamics. *ESAIM: M2AN* **2007**, *41*, 351–389.

(40) Masella, M. The Multiple Time Step r-RESPA Procedure and Polarizable Potentials Based on Induced Dipole Moments. *Mol. Phys.* **2006**, *104*, 415–428.

(41) http://biodev.cea.fr/polaris/.

(42) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press, 2002.

(43) Kästner, J.; Thiel, W. Bridging the Gap between Thermodynamic Integration and Umbrella Sampling Provides a Novel Analysis Method: Umbrella Integration. *J. Chem. Phys.* **2005**, *123*, 144104.

(44) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion: Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006**, *110*, 16066–16081.

(45) Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated ions. *Chemical Reviews* **1993**, *93*, 1157–1204.

(46) Ohtaki, H. Ionic Solvation in Aqueous and Nonaqueous Solutions. *Monatshefte für Chemie / Chemical Monthly* **2001**, *132*, 1237–1268.

(47) Gomez, D. T.; Pratt, L. R.; Asthagiri, D. N.; Rempe, S. B. Hydrated Anions: From Clusters to Bulk Solution with Quasi-Chemical Theory. *Accounts of Chemical Research* **2022**, *55*, 2201–2212.

(48) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; ; Coe, J. V.; Tuttle, T. R. T. J. The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998**, *102*, 7787–7794.

(49) Schmid, R.; Miah, A. M.; Sapunov, V. N. A new table of the thermodynamic quantities of ionic hydration: values and some applications (enthalpy–entropy compensation and Born radii). *Phys. Chem. Chem. Phys.* **2000**, *2*, 97–102.

(50) Zeebe, R. E. On the Molecular Diffusion Coefficients of Dissolved CO2,HCO3-, and CO32- and their Dependence on Isotopic Mass. *Geochimica et Cosmochimica Acta* **2011**, *75*, 2483–2498.

(51) Li, P.; Merz, K. M. J. Metal Ion Modeling Using Classical Mechanics. *Chemical Reviews* **2017**, *117*, 1564–1686.

(52) Debiec, K. T.; Gronenborn, A. M.; Chong, L. T. Evaluating the Strength of Salt Bridges: A Comparison of Current Biomolecular Force Fields. *The Journal of Physical Chemistry B* **2014**, *118*, 6561–6569.

(53) Houriez, C.; Réal, F.; Vallet, V.; Mautner, M.; Masella, M. Ion Hydration Free Energies and Water Surface Potential in Water nano Drops: The Cluster Pair Approximation and the Proton Hydration Gibbs Free Energy in Solution. *J. Chem. Phys.* **2019**, *151*, 174504.

(54) Guárdia, E.; Skarmoutsos, I.; Masia, M. On Ion and Molecular Polarization of Halides in Water. *Journal of Chemical Theory and Computation* **2009**, *5*, 1449–1453.

(55) Réal, F.; Trumm, M.; Schimmelpfennig, B.; Masella, M.; Vallet, V. Further Insights in the Ability of Classical Nonadditive Potentials to Model Actinide Ion-Water Interactions. *Journal of Computational Chemistry 34*, 707–719.

(56) Coles, J. P.; Masella, M. The Fast Multipole Method and Point Dipole Moment Polarizable Force Fields. *J. Chem. Phys.* **2015**, *142*, 024109.