



HAL
open science

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

Michel Masella, Fabien Léonforté

► **To cite this version:**

Michel Masella, Fabien Léonforté. The multi-scale polarizable pseudo-particle solvent coarse-grained approach: From NaCl salt solutions to polyelectrolyte hydration. *The Journal of Chemical Physics*, 2024, 160 (20), pp.204902. 10.1063/5.0194968 . hal-04590021

HAL Id: hal-04590021

<https://hal.science/hal-04590021v1>

Submitted on 28 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

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.

20 1 Introduction

21 Coarse Grained, CG, approaches are efficient molecular modeling scheme, which are more and more
 22 commonly used to investigate the properties of very large molecular systems that are beyond the
 23 computational capacity of standard *all atom* force fields (*i.e.* interaction potentials that explicitly
 24 take into account all the atoms of a molecular system). We may cite among others (see the re-
 25 cent reviews Refs.¹⁻³) the MARTINI force field to model lipids, proteins, carbohydrates and other
 26 biomolecules,⁴ the CG approach devoted to study complex aggregates made of huge chitosan chains
 27 (at the 1 000 unit size scale) in salty aqueous solutions,⁵ as well as the water CG approach mW.⁶
 28 CG approaches are of particular interest in the industrial research and development field (from
 29 chemistry, oil manufacturers to personal care/cosmetics)⁷⁻¹² because of the complexity of most
 30 of the solutions that these industries use and the size of those solutions ingredients, like chitosan
 31 chains. Most of CG approaches rely on simple pair wise interaction potentials whose parameters
 32 are assigned from macroscopic properties (top down schemes, see for instance Chen et al.¹³) and/or
 33 from microscopic features of a molecular systems (bottom up schemes, see among others Shell¹⁴)
 34 using different numerical methods, like force matching^{15,16} and machine learning¹⁷⁻²⁰ techniques.
 35 Recently sophisticated CG approaches accounting for many-body effects (by approximating the to-
 36 tal potential energy of a molecular system as a sum N -body energy components) have emerged.^{21,22}

37 In a series of articles²³⁻²⁵ we detailed a multi-scale version, denoted as ppp^l , of the original
 38 polarizable hybrid CG scheme of Ha-Duong et al.²⁶ that was built to simulate the hydration of
 39 explicit solutes modeled using *all atom* force fields and as dissolved in a fluid of polarizable pseudo
 40 particles (denoted as ppp 's). Such kind of CG scheme belongs to the wide class of solvent implicit
 41 continuum approaches, like the popular Polarizable Continuum Model, PCM.²⁷ However, as the
 42 original ppp approach, its multi scale ppp^l version retains the notion of particle as it relies on a
 43 hierarchical representation of a solvent surrounding a solute: at short range from that solute the
 44 solvent is modeled by polarizable pseudo particles whose size corresponds to that of a single solvent
 45 molecule, whereas at longer and longer range from the solute larger and larger local volumes of
 46 the solvent are modeled by means of larger and larger ppp^l polarizable particles. Despite its
 47 sophistication, the computational complexity of the ppp^l approach scales as $O(N)$ and it allows

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

48 one to readily simulate using reduced computational resources the hydration of complex solutes
49 in extended water environments, as we recently showed for medium-size chitosan chains²⁸ as well
50 as for a hydrophobic polyelectrolyte polymer dissolved in an aqueous environment comprising the
51 equivalent of 10 M water molecules.²⁹

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

$$\alpha_s = \frac{\epsilon_s - 1}{4\pi\rho_s\epsilon_s}. \quad (1)$$

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

63 An important assumption of the ppp^l approach is the neglect of intra-solvent polarization effects:
64 the magnitude of the particle induced dipole moments \mathbf{p}_s is only a function of the solute electric field
65 $\mathbf{E}_{\text{solute}}$ acting on the particle centers. However that 'local' approximation may yield solute/solvent
66 over polarization phenomena that can be circumvented by allowing the dipoles of the particles to
67 saturate according to²⁶

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

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

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

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

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

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

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

102 2 Computational details

103 2.1 Long range interaction truncation scheme

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

112 2.2 The ppp approach

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

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

114 \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
 115 charge and the induced dipole moment of atom i . $\lambda_{3,i}$ and $\lambda_{5,i}$ are two damping functions that
 116 alter the magnitude of the electric field damping at short range³⁶ (up to 5.0 Å) :

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

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

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

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

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

125 Accounting for the smoothing function G , that relation yields $\delta G_{\text{lr}} = -13.5 \text{ kcal mol}^{-1}$ for a
126 $R_{\text{cut},1}^{\text{pol}} = 12 \text{\AA}$ and $|Q| = 1 e$. For complex non-symmetric solutes, we proposed a multi-scale
127 approach to compensate truncation, see Section 2.3 below.

128 Besides solute/solvent polarization, the ppp model also accounts for solute/solvent non electro-
129 static interactions using a Lennard-Jones-like potential truncated for distances larger than $R_{\text{cut},1}^{\text{pol}}$:

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

130 N_a and N_s are the numbers of atoms and ppp 's, respectively. As in our original study,²⁵ we set
131 (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
132 reproduce (within $0.1 \text{ kcal mol}^{-1}$) the available experimental hydration Gibbs free energy for a set
133 a target entities, like ions Na^+ and Cl^- , as well as the mean solute/water distances in the entity
134 first hydration shell. We don't pay attention here to reproduce ion first shell coordination numbers.

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

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

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

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

141 The parameters of \tilde{U}_{ppp}^{LJ} and U^{dens} are assigned to reproduce (at ambient conditions) the hydra-
142 tion Gibbs free energy of a water molecule (6.3 kcal mol⁻¹), the water density (0.0335 molecules
143 per Å³) and the two regimes of the free energy corresponding to the creation of an empty cavity in
144 neat water according to the Lum-Chandler-Weeks theory of hydrophobicity.³⁷ The latter conditions
145 are met by truncating the terms \tilde{U}_{ppp}^{LJ} and U^{dens} for inter ppp distances larger than $R_{ppp}^{cut} = 7$ Å,
146 whereas a larger cut off distance is needed (at least about 15 Å) when only taking into account
147 \tilde{U}_{ppp}^{LJ} . As the computational time to estimate interactions among ppp 's scales as $(R_{ppp}^{cut})^3$, the use
148 of the term U^{dens} provides an efficient way to compute them.

149 The use of the energy term U^{dens} in conjunction with \tilde{U}^{LJ} yields ppp particles that diffuse fastly
150 : their self diffusion coefficient is about seven times larger (17.2 10⁻⁵ cm² s⁻¹ using the present
151 ppp parameter set) than for real water at ambient conditions (2.6 10⁻⁵ cm² s⁻¹). That allows for a
152 faster sampling of the potential energy surfaces of hydrated systems (at least for the solvent part).

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

166 2.3 Long range electrostatics and the ppp^l approach

167 We showed the effect of truncating the solute electric field for atom/ ppp distances larger than $R_{\text{cut},1}^{\text{pol}}$
 168 to largely overestimate ion pair association in a ppp fluid as compared to real water.²⁵ To remediate
 169 that artifact we proposed the 'Russian doll' multi-scale ppp^l approach depicted in Figure 1. We add
 170 $l \geq 2$ levels of larger and larger ppp -like particles (denoted as ppp^l , ppp^1 particles are the original
 171 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
 172 single ppp^l particle thus models a cluster of 8^{l-1} original ppp 's.

173 Particles of a given level l (they are all of the same size) do not interact with the particles
 174 of a different level, and they only electrostatically interact with a solute if at least one atom/ ppp^l
 175 distance r is included between $R_{\text{cut},l-1}^{\text{pol}}$ and $R_{\text{cut},l}^{\text{pol}}$. The solute/ ppp^l electrostatic energy term for $l >$
 176 1 corresponds to the linear version (for weak electric fields) of the polarization term of Equation (3)
 177 and dipole/dipole interactions among atoms and ppp^l particles are omitted. We maintain the ppp^l
 178 density along a simulation in the corresponding level box using periodic conditions. The numerical
 179 precision of a ppp^l scheme is estimated from Equation (6) and the largest $R_{\text{cut},l}^{\text{pol}}$ truncation distance.

180 For a given level $l > 1$, ppp^l 's interact with each other according to the $\tilde{U}_{ppp}^{L,J}$ potential detailed
 181 above, with specific σ_l^* and ϵ_l^* parameters. For efficiency reason we truncate the ppp^l interactions
 182 for inter ppp^l distances larger than $5 \times l \text{ \AA}$ and we assign the parameters σ_l^* to reproduce the water
 183 density at ambient conditions for a neat ppp^l fluid (accounting for a ppp^l to correspond to a cluster
 184 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
 185 condition is met using $\sigma_l^* = \gamma \times 2^{l-1} \times \sigma_{ppp}^*$, where $\gamma = 1.245615$.

186 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
 187 molecule), regardless of l . Such a large mass value slow down the ppp^l diffusion and that allows us
 188 to account for solute/ ppp^l long range forces as slow fluctuating interactions in the multiple time
 189 step algorithm to solve the Newtonian equations of motion detailed below.

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

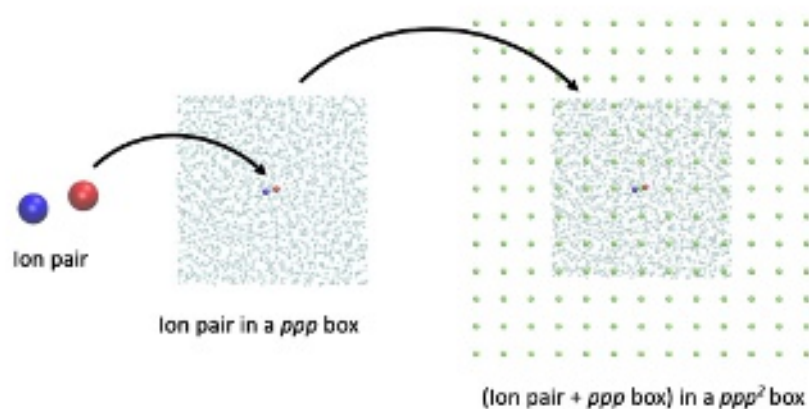


Figure 1: The multi-scale ppp^l approach. Here an ion pair (blue and red spheres) is dissolved in a box of ppp 's (each ppp models a single water molecule). Then the ion pair and the ppp box is set at the center of a new box made of ppp^2 particles, each of them models a local spherical volume of the 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.

193 2.4 Intra solute interactions

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

201 2.5 MD simulation details

202 Preliminary (relaxation) MD simulations are performed only in a neat ppp fluid and in the NPT
 203 ensemble, whereas production MD runs in ppp and ppp^l fluids are performed in the NVT ensemble.
 204 All simulations based on a *all atom* force field are performed in the NPT ensemble. Temperature
 205 and pressure are monitored along NPT runs using the Nosé-Hoover barostat³⁸ (the barostat cou-
 206 pling constant is set to 2.5 ps), whereas temperature is monitored along NVT runs according to
 207 a Langevin thermostat approach.³⁹ Induced dipole moments are solved iteratively until the mean

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

213 The equations of motion are solved using a Multiple Time Steps, MTS, algorithm devoted to
214 polarizable force field based on induced dipole moments.⁴⁰ For NaCl simulations, two time steps are
215 used: 2 and 6 fs for short range and long range electrostatic/polarization forces, respectively. To
216 simulate the hydration of more complex entities (with intra molecular chemical bonds like **HPP**)
217 we used three time steps, namely 0.25 fs for intra molecular stretching and bending forces, and
218 1.0/5.0 fs (small systems) or 2.0/6.0 fs (large systems) for short and long range electrostatic and
219 dispersion forces. 1-4 dihedral forces are considered as short range forces and the cutoff distance
220 to compute electrostatic/dispersion short range forces is set to 8 Å. As discussed above, the forces
221 corresponding to solute/*ppp*^l interactions are considered as long range polarization forces.

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

228 2.6 Hydration free energy and PMF computations

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

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

240 The Potential of Mean Force, PMF, of an ion pair is computed using an umbrella sampling MD
241 protocol. The degree of freedom R restrained along MD simulations, using the harmonic potential
242 $k_c (R - R_c)^2$, is the distance between the ion centers of mass. The target distances R_c usually span
243 from 2 to 17 Å and they are regularly spaced by 0.5 Å. If not otherwise stated, the constant k_c is
244 set to 5 kcal mol⁻¹ Å⁻². The MD duration is set to 10 ns and the distances R are sampled each
245 50 fs once a starting phase of 1 ns is achieved. We post process the set of sampled distances to
246 compute the PMF's according to the Umbrella Integration method.⁴³ The PMF's account for the
247 entropic correction $2RT \ln R$.

248 3 Results

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

253 3.1 *ppp* response to solute medium range electrostatics

254 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*
255 *atom* simulation in bulk water at ambient conditions. The distance between the ions \mathbf{X}^{+Q} and \mathbf{Y}^{-Q}
256 is harmonically restrained to a target distance of 14 Å (the harmonic constant is set to 50 kcal
257 mol⁻¹ Å⁻²). We computed the mean water normalized density $\bar{\rho}_s$ and the mean dipole moment $\bar{\mu}_{\mathbf{X}}$
258 (projected on the ion pair axis and expressed in Debye per water molecule/*ppp*) within a cylindrical
259 volume centered at the ion pair center and whose length and radius are 2 and 0.5 Å, respectively.

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

262 are those of the model corresponding to the damping parameters $a_i = 0.3 \text{ \AA}^{-3}$ detailed below. As
 263 the ion/*ppp* electric field damping vanishes beyond 5.0 \AA , the magnitudes of the quantities $\bar{\rho}_s$ and
 264 $\bar{\mu}_{\mathbf{X}}$ do not depend on the damping parameters a_i 's.

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

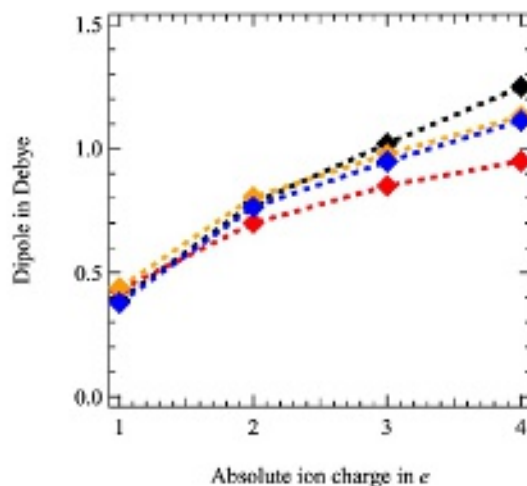


Figure 2: Water and *ppp* mean dipole projections $\bar{\mu}_{\mathbf{X}}$ at the center of the ion pair $[\mathbf{X}^{+Q}, \mathbf{Y}^{-Q}]$ axis as a function of Q (*ppp* data correspond to $\mu_s = 1.2$ Debye). Black and red data : $\bar{\mu}_{\mathbf{X}}$ values from 200 ns scale simulations in an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

270 μ_s has a much more significant effect on the *ppp* dipole projections $\bar{\mu}_{\mathbf{X}}$: they increase by a
 271 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
 272 also note a good agreement between the *ppp* $\bar{\mu}_{\mathbf{X}}$ data corresponding to $\mu_s = 2.0$ Debye and *all*
 273 *atom* values. However as the *ppp* polarizability depends explicitly on the solvent density, we need
 274 to compare the $\bar{\mu}_{\mathbf{X}}$ values scaled by the local solvent densities to draw sound conclusions. In that
 275 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.

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

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

281 3.2 ppp^l domain boundary effects on long range electrostatics

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

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

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

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

$$\Delta U_{ps}^2 = -\frac{\alpha_{s,2}\bar{\rho}_{s,2}}{\epsilon_0} \int_{R_{cut,1}^{pol}}^{R_{cut,2}^{pol}} \Delta\rho^2(r) \frac{f(r)^2}{r^2} dr. \quad (9)$$

308 Here, $\alpha_{s,2}$ and $\bar{\rho}_{s,2}$ are the ppp^2 polarizability and mean density, respectively. f is the truncation
 309 function G at the ppp/ppp^2 boundary domain. The ppp^2 over accumulation at **BD** yields ΔU_{pol}^2
 310 to amount from 4% (Na^+) up to 14% (X^{2+}) of the expected ideal values computed from Equation
 311 6 for $R_{cut,1}^{pol} = 12 \text{ \AA}$, and by about 1.5 and 4% for the latter cations for $R_{cut,1}^{pol} = 15 \text{ \AA}$. From
 312 the Na^+/ppp^2 radial distribution functions corresponding to our $[\text{Na}^+, \text{Cl}^-]$ pair simulations, ΔU_{ps}^2
 313 amounts at most to only 0.05% of the expected U_{ps}^2 value. Domain boundary effects thus yield
 314 strong drawbacks only for highly charged solutes as modeled at infinite dilution conditions.

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

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

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

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

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

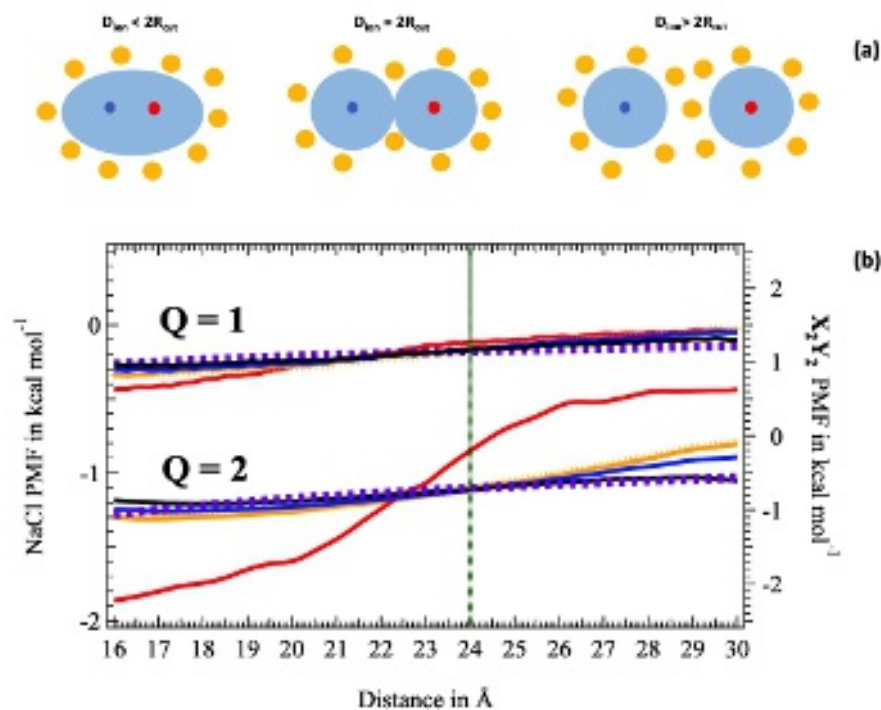


Figure 3: Ion pair PMF's in ppp^l fluids as a function of the truncation distance $R_{cut,1}^{pol}$. (a) schematic representation of the solvent elements explicitly taken into account *wrt* $R_{cut,1}^{pol}$. Blue: ppp particles represented as a continuum medium. Yellow spheres: ppp^2 particles. D_{ion} is the ion separation distance and $R_{cut} = R_{cut,1}^{pol}$. (b) PMF's of the $[Na^+, Cl^-]$ (left axis) and $[X^{2+}, Y^{2-}]$ (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 $R_{cut,1}^{pol}$ 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.

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

333 All the $[\text{Na}^+, \text{Cl}^-]$ PMF's match the expected Coulombic potential on average, within less than
334 $0.1 \text{ kcal mol}^{-1}$ (regardless of $R_{\text{cut},1}^{\text{pol}}$) and even $0.03 \text{ kcal mol}^{-1}$ (for $R_{\text{cut},1}^{\text{pol}} \geq 15 \text{ Å}$) on the 16-30 Å
335 distance range. However $R_{\text{cut},1}^{\text{pol}}$ has a stronger effect on the $[\text{X}^{2+}, \text{Y}^{2-}]$ PMF's. For instance the
336 PMF computed for $R_{\text{cut},1}^{\text{pol}} = 12 \text{ Å}$ (and by taking into account 4 ppp^l levels) obeys two regimes in
337 the distance domain 16-30 Å, with a step transition between them at about 24 Å, *i.e.* at twice the
338 $R_{\text{cut},1}^{\text{pol}}$ value, a distance at which higher level ppp^2 's located in between both ions start to interact
339 with them. Increasing the cutoff distance $R_{\text{cut},1}^{\text{pol}}$ to values $> 12 \text{ Å}$ yield a better agreement with
340 the expected Coulombic potential (within less $0.05 \text{ kcal mol}^{-1}$ for $R_{\text{cut},1}^{\text{pol}} = 21 \text{ Å}$, for instance).

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

348 3.3 Short range electric field damping and ppp models

349 To model the hydration of ions Na^+ and Cl^- , we built a series of ppp models for which we system-
350 atically set μ_s to 1.2 Debye. Those models differ by the value of the ion/ ppp damping parameter a
351 discussed in Section 2.2, which is taken identical for both ions. Here we consider the set of a values
352 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
353 ion/ ppp energy term \tilde{U}^{LJ} to reproduce the ion/water oxygen distances in ion first hydration shells
354 from our earlier *all atom* simulations³⁴ (*i.e.* 2.4 and 3.4 Å for Na^+ and Cl^- , respectively) and to
355 reproduce the ion experimental hydration Gibbs free energies ΔG_{hyd}^0 of Kelly et al.⁴⁴ within 0.1
356 kcal mol^{-1} . Those target energies are -103.8 (Na^+) and -74.5 (Cl^-) kcal mol^{-1} . The ppp energies

357 ΔG_{hyd} are computed according to the TI protocol detailed in Section 2.6. The hydration energies
358 account for the long range polarization correction of Equation (6) (-13.5 kcal mol⁻¹).

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

364 Ion first coordination shells comprise about 5-6 (Na⁺) and 15-16 (Cl⁻) *ppp*'s, regardless of
365 a . Those coordination numbers n_c are in line with (Na⁺) or twice as large (Cl⁻) compared to
366 experimental and *all atom* available data.⁴⁵⁻⁴⁷ The entropic contribution to the hydration free
367 energies ΔG_{hyd} amounts to 45 cal mol⁻¹ K⁻¹, regardless of the ion, a value that is from twice
368 to four times larger compared to experimental-based data.^{48,49} The inability of CG approaches
369 to predict well balanced enthalpic and entropic contributions to the hydration free energies have
370 already been discussed, in particular by Noid³.

371 We also computed the diffusion coefficient D_i for single Na⁺ or Cl⁻ ions as dissolved in 1
372 000 *ppp* box from 30 ns scale simulations and the ion velocity auto correlation functions (see
373 Section S6 of Supplementary Material). As the *ppp* self diffusion coefficient (see Section 2.2), the
374 *ppp* D_i estimates, 5.9 (Na⁺) and 5.5 (Cl⁻) 10⁻⁵ cm² s⁻¹, are largely overestimated compared to
375 experimental data (1.33 and 2.09 10⁻⁵ cm² s⁻¹ for both ions, respectively). D_i values from *all*
376 *atom* simulations are commonly scaled to better match experimental data (see Ref.⁵⁰ for instance).
377 However as ion coordination numbers are inferred to be key factors affecting ion diffusion,⁵¹ a first
378 step before building a correction scheme to compute reliable enough D_i values for ions from *ppp*
379 simulations is to further improve the structural modeling of ion first hydration shell in *ppp* fluids
380 (we reported an attempt in Ref.²⁵).

381 We computed the PMF's of an ion pair [Na⁺,Cl⁻] in a four level *ppp*^{*l*} fluid from the 7 above
382 models and 10 ns scale simulations. The simulated systems correspond to a single ion pair dissolved
383 in a set of four 2 000 *ppp*^{*l*}'s boxes. That yields ion/solvent interactions to be accounted for up to
384 150 Å from the ions. The resulting PMF's are compared to earlier *all atom* data³⁴ in Figure 4(a).
385 Whereas we are able to well reproduce the *all atom* PMF when using large damping parameter

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

388 As shown by *all atom* simulations,⁵² the magnitude of the ion association process in the aqueous
389 phase is tied to the depth of that first PMF minimum. In particular the deeper is that first
390 minimum the stronger is the expected percentage of ion associated in bulk water. To best reproduce
391 the $[\text{Na}^+, \text{Cl}^-]$ *all atom* PMF in water as using a *ppp* model corresponding to a weak damping
392 parameter a , we introduce a many-body correction potential δU^a whose parameters are adjusted
393 to best reproduce the *all atom* $[\text{Na}^+, \text{Cl}^-]$ PMF in bulk water for each *ppp* model (see Section S5
394 of Supplementary Material). The new *ppp* PMF's based on δU^a are shown in Figure 4(b). They
395 all overall match the *all atom* one, even if the second minimum of the new *ppp* PMF's is shifted
396 to larger ion/ion distances (up to 6 \AA), and the height of the energy barrier interconnecting the
397 PMF first and second minimum is increased (up to reach 2.8 kcal mol⁻¹). Note the position of the
398 second minimum in our *all atom* PMF³⁴ is 5.5 \AA and the energy barrier height is 2.4 kcal mol⁻¹.

399 We compare in Figure 4(c) the PMF's of a $[\text{Na}^+, \text{Cl}^-]$ pair dissolved in *ppp*^{*l*} fluids to its expected
400 Coulombic potential for ion separation distances that span up to 16 \AA and as using the *ppp* model
401 corresponding to the damping parameter $a = 0.15 \text{ \AA}^{-3}$. Those PMF's are switched to minimize
402 their numerical difference with the Coulombic potential for ion separation distances that span from
403 12 to 16 \AA . Note the PMF's computed from all the above *ppp* models to be indistinguishable
404 within the latter ion separation distance range. Moreover, within that distance range, the ions
405 only interact with a single shell of *ppp*'s and with higher level *ppp*^{*l*}'s that all lie outside that *ppp*
406 first shell. The new PMF's converge towards the Coulombic potential as soon as a 3 levels *ppp*^{*l*}
407 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
408 ion separation distance), we estimated the apparent dielectric constant $\tilde{\epsilon}_0^l$ of the *ppp*^{*l*} fluids at
409 intermediate ion separation distances. $\tilde{\epsilon}_0^l$ is converged to the expected neat water value (78.35) as
410 soon as 3 levels, within 1.5 % (that corresponds to uncertainty of our fitted values for $l \geq 3$).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

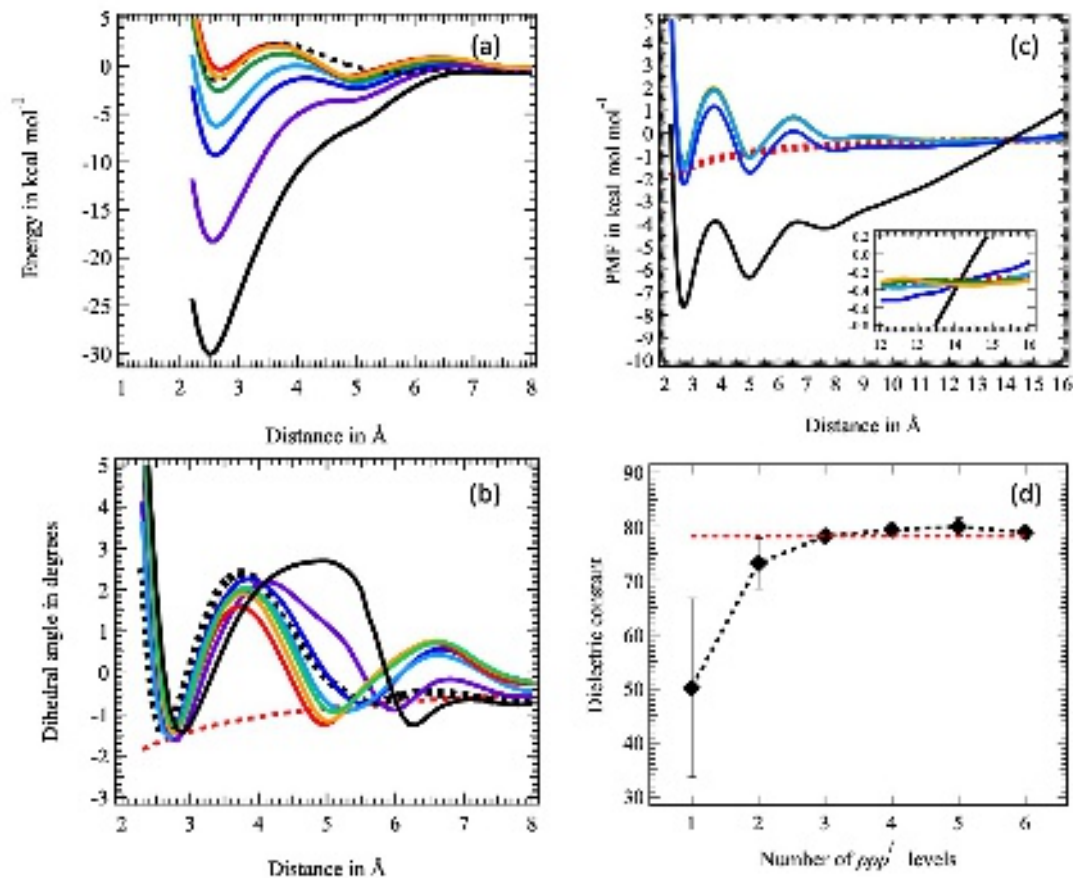


Figure 4: PMF's of the $[\text{Na}^+, \text{Cl}^-]$ ion pair in ppp^l fluids. All the PMF's are shifted to best reproduce the expected Coulombic potential of the ion pair in neat water for inter ionic distances that span between 12 to 16 Å. From black to red bold lines: PMF's of ppp models corresponding 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: $[\text{Na}^+, \text{Cl}^-]$ 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 ϵ_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).

411 3.4 NaCl association in a ppp^l fluid

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

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

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

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

434 The order of magnitude of the cluster mean survival times t_k^* for cluster size ranging from 2 to
435 4 (at the 0.1 ns scale, see Table 1) are in line with available *all atom* estimates, at the remarkable
436 exception of 0.2M and 0.6M salt solutions simulated using a low value of a (≤ 0.05 Å⁻³). In that
437 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*
438 *atom* counterpart. We also note the near total absence of clusters whose size is larger than 2.
439 Regarding the ion/ion radial distribution functions and the salt concentration computed using a
440 damping parameter a set to 0.05 Å⁻³, the [Na⁺,Cl⁻] ones are in line with *all atom* simulations
441 (positions and heights of its main maxima and minima), whereas we note the shortest [Na⁺,Na⁺]
442 and [Cl⁻,Cl⁻] distances to be shifted to larger distances by about 2 Å as compared to our own
443 *all atom* data³⁴ (see Figure S4 of Supplementary Material). From visual inspection of the MD

444 trajectories, the associated $[\text{Na}^+, \text{Cl}^-]$ pairs in a ppp^l fluid appear to be more uniformly distributed
 445 in the solvent than along *all atom* simulations (along such simulations, the associated pairs seem to
 446 form locally large and weakly bonded super clusters). The largest mean distance among $[\text{Na}^+, \text{Cl}^-]$
 447 pairs in a ppp^l fluid explains the largest ion pair mean survival time in a ppp^l fluid as using a low
 448 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 study³⁴ 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 \AA^{-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 (<i>na</i>)	4 (23)
0.03	80.01 ± 2.82	80.01 ± 2.82	0.5 (1814)	0.0 (<i>na</i>)	0 (<i>na</i>)	3 (78)
<i>all atom</i> ^(a)	81.8 ± 2.0	77.5 ± 2.0	6.2 (130)	0.8 (187)	0.2 (326)	10 (1)
<i>all atom</i> ^(b)	85.0 ± 2.0	82.0 ± 2.0				

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

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

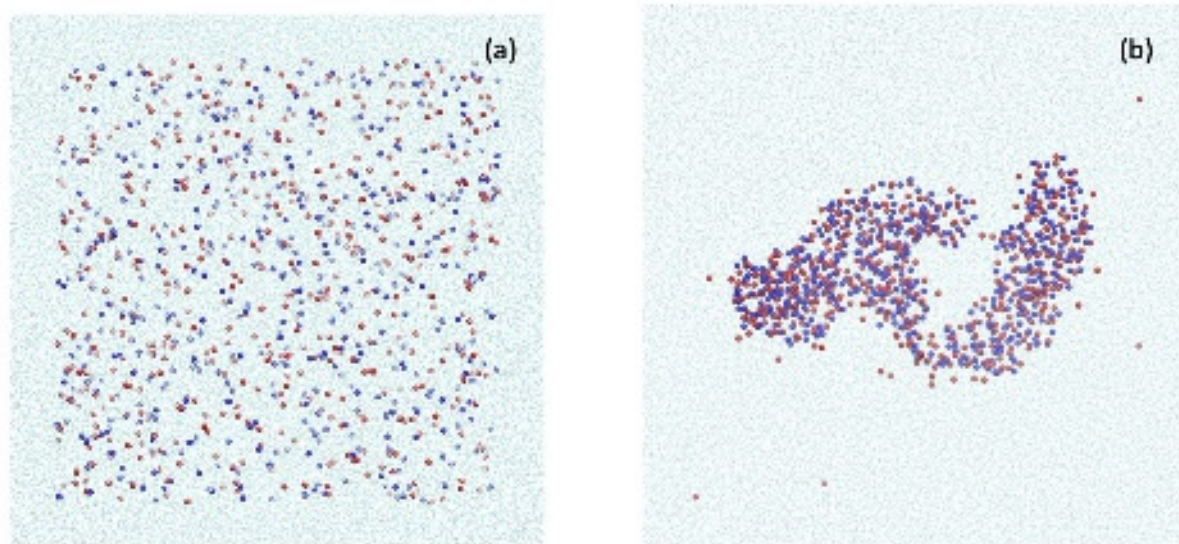


Figure 5: Two snapshots after 200 ns of simulations in a ppp^l fluid of a 1.0M NaCl salt solution (see text for details). (a) and (b): snapshots from a damping parameter a set to 0.05 (left) and 0.30 (right) \AA^{-3} . ppp particles and ions Na^+ and Cl^- are shown by light blue, dark blue and red spheres,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

458 Mcal for $l = 1$ (original ppp 's), whereas they are from three to five orders of magnitude weaker for
 459 $l = 2$ and 3, respectively. We may also note all the ion clustering properties for a 1.0M salt solution
 460 computed along a simulation in a neat ppp fluid fully match those corresponding to a four level
 461 ppp^l fluid. For instance the percentage of free ions along ppp simulations ($a = 0.05 \text{\AA}^{-3}$) agrees
 462 with ppp^l data within less than 1 %.

463 3.5 A hydrophobic polyelectrolyte polymer and at infinite dilution 464 conditions

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

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

472 solute polarization forces in both ppp^l fluids.

473 To discuss the reliability of that result, we built a new ppp model as detailed above for NaCl
474 in which we systematically set the damping parameter a for all ionic species to 0.05 \AA^{-3} , whereas
475 we consider our earlier ppp /solute parameter set for the other kinds of polymer moieties. To assess
476 the new ppp model, we first computed the PMF of the tetra methyl ammonium/chloride anion
477 $[(\text{CH}_3)_4\text{N}^+, \text{Cl}^-]$ ion pair dissolved in a four level ppp^l fluid from 10 ns scale simulations. We don't
478 make use here of any correction potential δU^a . As shown by the plots of 6(c), the new ppp model
479 is able to overall reproduce the main features of the *all atom* PMF as computed in neat water from
480 our own polarizable *all atom* approach.⁵³

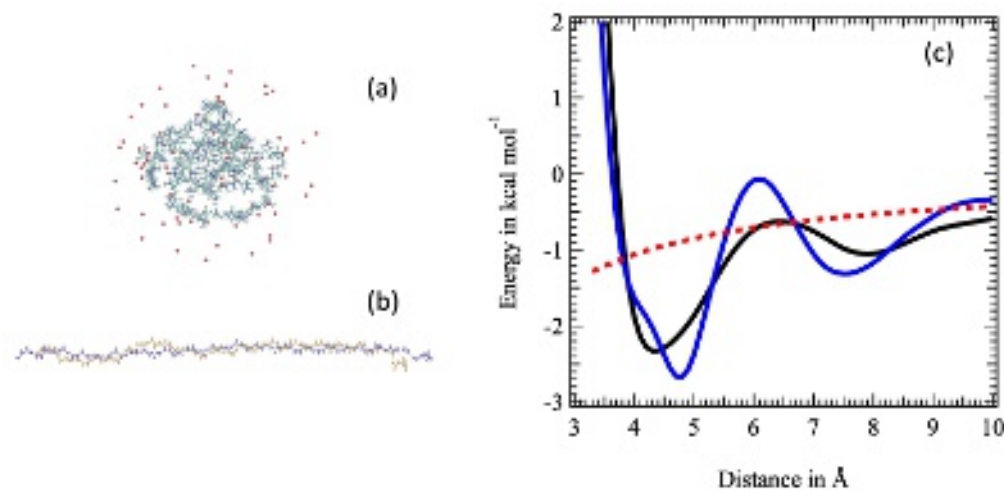


Figure 6: The polymer **HPP** and the new ppp model. (a) : example of a simulation final quasi spherical structure of the 10 unit **HPP** polymer with its counter ions Cl^- (in red) embedded in a 0.9M ppp cubic box. (b) : final quasi linear structure from simulation without center ions. (c) : tetra methyl ammonium/ Cl^- PMF's in a four level ppp^l fluid (in blue) and in neat water using a polarizable *all atom* approach⁵³ (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

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

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

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

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

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

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

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

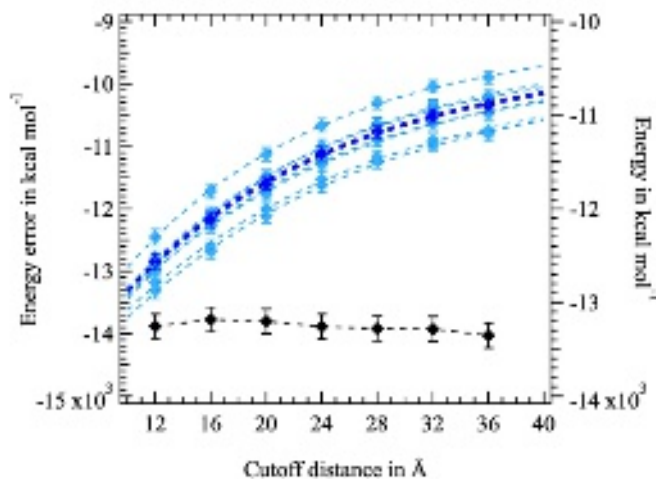


Figure 7: (a) : interaction energies $\bar{U}_{\text{pol}}^{\text{ppp}^{1+2}}$ as a function of the cut off distance $R_{\text{cut},1}^{\text{pol}}$. Black symbols (left axis), light and dark blue symbols (right axis) : energy data for the **HPP** globular and quasi linear conformations, respectively.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

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

527 account the counter ions) noticeably increases as $R_{\text{cut},1}^{\text{pol}}$ increases, from -12.5 ± 0.3 Mcal ($R_{\text{cut},1}^{\text{pol}} =$
528 12 \AA) to -10.7 ± 0.3 Mcal ($R_{\text{cut},1}^{\text{pol}} = 36 \text{ \AA}$). For globular forms, the $\bar{U}_{\text{pol}}^{\text{ppp}^{1+2}}$ data can be accurately
529 reproduced using an exponential function, which yields an extrapolated value for that energy of
530 -9.60 ± 0.2 Mcal at $R_{\text{cut},1}^{\text{pol}} = 150 \text{ \AA}$.

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

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

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

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

560 Simulations in a water droplet using a polarizable *all atom* force field²⁹ showed **HPP** globular
 561 forms surrounded by a spherical counter ion cloud to be stable. Such compact geometries thus
 562 correspond at least to meta stable states in aqueous environments, regardless of the description of
 563 the solute/solvent interactions. Hence stable polyelectrolyte polymer globular forms surrounded
 564 by spherical counter ion clouds as predicted by our multi scale simulation approach for **HPP** do
 565 not a priori arise from artifacts tied to our solvent approach. Note also that globular forms for hy-
 566 drophobic polyelectrolyte polymers in presence of counter ions are not systematically predicted by
 567 our multi scale *ppp*^l approach. In our earlier study²⁹ we also reported simulations of carboxylated
 568 polystyrenes with a high fraction charge and in presence of Na⁺ counter ions (that are also located
 569 close to the polymer ionic heads in the simulation starting structures). Those kinds of polyelec-
 570 trolyte polymer do not evolve towards a globular form as simulated in a *ppp*^l fluid in presence of
 571 Na⁺ counter ions : their linear starting structure is stable along 100 ns scale trajectories.

572 4 Conclusion

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

583 *ppp*'s and solute atoms that lie between 10 to 20 Å from each other). However that issue appears
584 to have a weak effect on the structural behavior of a complex polyelectrolyte polymer like **HPP**
585 dissolved in a *ppp*^l fluid. For efficiency reason we thus recommend to simulate the hydration of
586 complex solutes using a moderately extended *ppp* sub domain SD (*i.e.* a *ppp* shell extending up to
587 12 Å from any solute atom) and to post process the simulations using a larger extension for SD to
588 better assess the enthalpic stability of the solute conformations observed along a simulation.

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

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

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

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

626 In all our results show hybrid simulations based on polarizable pseudo-particles to model ex-
627 tended chemical environments (in particular water) are an efficient alternative route to standard *all*
628 *atom* approaches to investigate equilibrium properties of large and complex microscopic systems.
629 For instance, whereas our former *all atom* simulations³⁴ of NaCl aqueous droplets comprising 100k
630 explicit water molecules (using a polarizable force field and a scalable $O(N)$ Fast Multiple Method
631 to compute electrostatic and polarization energies⁵⁶) were run at a rate of about 1 ns per day using
632 8 AMD 64 cores CPU, we perform the corresponding ppp^l simulations at a rate that spans from 5
633 to 10 ns using only a single AMD 64 core processor and a basic $O(N^2)$ approach to compute ion/ion
634 interactions. Polarizable pseudo-particle approaches are thus a promising tool to ease the use of
635 sophisticated force fields as modeling large, complex and even heavily charged molecular systems.

636 Supplementary Material

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

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

641 Acknowledgments

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

644 Conflicts of interest

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

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

650 References

- 651 (1) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein
652 Models and Their Applications. *Chemical Reviews* **2016**, *116*, 7898–7936.
- 653 (2) Joshi, S. Y.; Deshmukh, S. A. A review of advancements in coarse-grained molecular dynamics simu-
654 lations. *Molecular Simulation* **2021**, *47*, 786–803.
- 655 (3) Noid, W. G. Perspective: Advances, Challenges, and Insight for Predictive Coarse-Grained Models.
656 *The Journal of Physical Chemistry B* **2023**, *127*, 4174–4207.
- 657 (4) Marrink, S. J.; Monticelli, L.; Melo, M. N.; Alessandri, R.; Tieleman, D. P.; Souza, P. C. T. Two
658 decades of Martini: Better beads, broader scope. *WIREs Computational Molecular Science* **2023**, *13*,
659 e1620.
- 660 (5) Tsereteli, L.; Grafmüller, A. An Accurate Coarse-Grained Model for Chitosan Polysaccharides in Aque-
661 ous Solution. *PLOS ONE* **2017**, *12*, 1–31.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

- 662 (6) Molinero, V.; Moore, E. B. Water Modeled As an Intermediate Element between Carbon and Silicon.
663 *The Journal of Physical Chemistry B* **2009**, *113*, 4008–4016.
- 664 (7) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J.
665 The power of Coarse Graining in Biomolecular Simulations. *WIREs Computational Molecular Science*
666 **2014**, *4*, 225–248.
- 667 (8) Hospital, A.; Goñi, J.; Orozco, M.; Gelpi, J. Molecular Dynamics Simulations: Advances and Applica-
668 tions. *Adv Appl Bioinform Chem. B* **2015**, *8*, 37–47.
- 669 (9) Lu, L.; Yoo, K.; Benyahia, S. Coarse-Grained-Particle Method for Simulation of Liquid-Solids Reacting
670 Flows. *Industrial & Engineering Chemistry Research* **2016**, *55*, 10477–10491.
- 671 (10) Illa-Tuset, S.; Malaspina, D. C.; Faraudo, J. Coarse-Grained Molecular Dynamics Simulation of the
672 Interface Behaviour and Self-Assembly of CTAB Cationic Surfactants. *Phys. Chem. Chem. Phys.* **2018**,
673 *20*, 26422–26430.
- 674 (11) Mehandzhiyski, A. Y.; Rolland, N.; Garg, M.; Wohlert, J.; Linares, M.; Zozoulenko, I. A novel supra
675 coarse-grained model for cellulose. *Cellulose* **2020**, *27*, 4221–4234.
- 676 (12) Nakamura, H.; Takimoto, H.; Kishida, N.; Ohsaki, S.; Watano, S. Coarse-Grained Discrete Element
677 Method for Granular Shear Flow. *Chemical Engineering Journal Advances* **2020**, *4*, 100050.
- 678 (13) Chen, J.; Chen, J.; Pinamonti, G.; Clementi, C. Learning Effective Molecular Models from Experimental
679 Observables. *Journal of Chemical Theory and Computation* **2018**, *14*, 3849–3858.
- 680 (14) Shell, M. S. The relative entropy is fundamental to multiscale and inverse thermodynamic problems.
681 *The Journal of Chemical Physics* **2008**, *129*, 144108.
- 682 (15) Wang, Y.; Noid, W. G.; Liu, P.; Voth, G. A. Effective force coarse-graining. *Phys. Chem. Chem. Phys.*
683 **2009**, *11*, 2002–2015.
- 684 (16) Bacle, P.; Jardat, M.; Marry, V.; Mériguet, G.; Batôt, G.; Dahirel, V. Coarse-Grained Models of
685 Aqueous Solutions of Polyelectrolytes: Significance of Explicit Charges. *J. Phys. Chem. B* **2020**, *124*,
686 288–301.
- 687 (17) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical*
688 *Physics* **2016**, *145*, 170901.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

- 689 (18) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. DeePCG: Constructing Coarse-Grained Models via Deep
690 Neural Networks. *The Journal of Chemical Physics* **2018**, *149*, 034101.
- 691 (19) Wang, J.; Olsson, S.; Wehmeyer, C.; Pérez, A.; Charron, N. E.; de Fabritiis, G.; Noé, F.; Clementi, C.
692 Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. *ACS Central Science* **2019**, *5*,
693 755–767.
- 694 (20) Li, W.; Burkhardt, C.; Polińska, P.; Harmandaris, V.; Doxastakis, M. Backmapping Coarse-Grained
695 Macromolecules: An Efficient and Versatile Machine Learning Approach. *The Journal of Chemical*
696 *Physics* **2020**, *153*, 041101.
- 697 (21) John, S. T.; Csányi, G. Many-Body Coarse-Grained Interactions Using Gaussian Approximation Po-
698 tentials. *The Journal of Physical Chemistry B* **2017**, *121*, 10934–10949.
- 699 (22) Wang, J.; Charron, N.; Husic, B.; Olsson, S.; Noé, F.; Clementi, C. Multi-body effects in a coarse-
700 grained protein force field. *The Journal of Chemical Physics* **2021**, *154*, 164113.
- 701 (23) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Po-
702 larizable Solvent Model: Application to Long Dynamics Simulations of Bovine Pancreatic Trypsin
703 Inhibitor. *Journal of Computational Chemistry* **2008**, *29*, 1707–1724.
- 704 (24) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Po-
705 larizable Solvent Model. II. Accounting for Hydrophobic Effects. *Journal of Computational Chemistry*
706 **2011**, *32*, 2664–2678.
- 707 (25) Masella, M.; Borgis, D.; Cuniasse, P. A multiscale coarse-grained polarizable solvent model for handling
708 long tail bulk electrostatics. *Journal of Computational Chemistry* **2013**, *34*, 1112–1124.
- 709 (26) Ha-Duong, T.; Phan, S.; Marchi, M.; Borgis, D. Electrostatic on Particles : Phenomenological and
710 Orientational Density Functional Theory Approach. *J. Chem. Phys.* **2002**, *117*, 541–556.
- 711 (27) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chemical*
712 *Reviews* **2005**, *105*, 2999–3094.
- 713 (28) Masella, M.; Léonforté, F. Chitosan Polysaccharides from a Polarizable Multiscale Approach. *ACS*
714 *Omega* **0**, *0*, null.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

- 715 (29) Masella, M.; Crudu, A.; Léonforté, F. Hybrid Polarizable Simulations of a Conventional Hydrophobic
716 Polyelectrolyte. Toward a Theoretical Tool for Green Science Innovation. *The Journal of Chemical*
717 *Physics* **2021**, *155*, 114903.
- 718 (30) Réal, F.; Vallet, V.; Flament, J.-P.; Masella, M. Revisiting a Many-Body Model for Water Based on a
719 Single Polarizable Site. From Gas Phase Clusters to Liquid and Air/Liquid Water Systems. *J. Chem.*
720 *Phys.* **2013**, *139*, 114502.
- 721 (31) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. J. An Empirical Polarizable Force Field Based
722 on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chemical*
723 *Reviews* **2016**, *116*, 4983–5013.
- 724 (32) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields
725 for Biomolecular Simulations: Recent Advances and Applications. **2019**, *48*, 371–394.
- 726 (33) Soniat, M.; Pool, G.; Franklin, L.; Rick, S. W. Ion Association in Aqueous Solution. *Fluid Phase Equil.*
727 **2016**, *407*, 31–38.
- 728 (34) Vallet, V.; Coles, J.; Réal, F.; Houriez, C.; Masella, M. NaCl Salts in Finite Aqueous Environments at
729 the Fine Particle Marine Aerosol Scale. *ACS Earth and Space Chemistry* **2022**, *6*, 1612–1626.
- 730 (35) Uneyama, T. Application of Projection Operator Method to Coarse-Grained Dynamics with Transient
731 Potential. *Phys. Rev. E* **2022**, *105*, 044117.
- 732 (36) Thole, B. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**,
733 *59*, 341–350.
- 734 (37) Chandler, D. Interfaces and the Driving Force of Hydrophobic Assembly. *Nature* **2005**, *437*, 640–647.
- 735 (38) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit Reversible Integrators for
736 Extended Systems Dynamics. *Molecular Physics* **1996**, *87*, 1117–1157.
- 737 (39) Cancès, Eric.; Legoll, Frédéric.; Stoltz, Gabriel, Theoretical and Numerical Comparison of Some Sam-
738 pling Methods for Molecular Dynamics. *ESAIM: M2AN* **2007**, *41*, 351–389.
- 739 (40) Masella, M. The Multiple Time Step r-RESPA Procedure and Polarizable Potentials Based on Induced
740 Dipole Moments. *Mol. Phys.* **2006**, *104*, 415–428.
- 741 (41) <http://biodev.cea.fr/polaris/>.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

- 742 (42) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic
743 Press, 2002.
- 744 (43) Kästner, J.; Thiel, W. Bridging the Gap between Thermodynamic Integration and Umbrella Sampling
745 Provides a Novel Analysis Method: Umbrella Integration. *J. Chem. Phys.* **2005**, *123*, 144104.
- 746 (44) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous Solvation Free Energies of Ions and Ion: Water
747 Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton.
748 *J. Phys. Chem. B* **2006**, *110*, 16066–16081.
- 749 (45) Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated ions. *Chemical Reviews* **1993**, *93*, 1157–
750 1204.
- 751 (46) Ohtaki, H. Ionic Solvation in Aqueous and Nonaqueous Solutions. *Monatshefte für Chemie / Chemical*
752 *Monthly* **2001**, *132*, 1237–1268.
- 753 (47) Gomez, D. T.; Pratt, L. R.; Asthagiri, D. N.; Rempe, S. B. Hydrated Anions: From Clusters to Bulk
754 Solution with Quasi-Chemical Theory. *Accounts of Chemical Research* **2022**, *55*, 2201–2212.
- 755 (48) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; ; Coe, J. V.;
756 Tuttle, T. R. T. J. The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from
757 Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998**, *102*, 7787–7794.
- 758 (49) Schmid, R.; Miah, A. M.; Sapunov, V. N. A new table of the thermodynamic quantities of ionic
759 hydration: values and some applications (enthalpy–entropy compensation and Born radii). *Phys. Chem.*
760 *Chem. Phys.* **2000**, *2*, 97–102.
- 761 (50) Zeebe, R. E. On the Molecular Diffusion Coefficients of Dissolved CO₂, HCO₃⁻, and CO₃²⁻ and their
762 Dependence on Isotopic Mass. *Geochimica et Cosmochimica Acta* **2011**, *75*, 2483–2498.
- 763 (51) Li, P.; Merz, K. M. J. Metal Ion Modeling Using Classical Mechanics. *Chemical Reviews* **2017**, *117*,
764 1564–1686.
- 765 (52) Debiec, K. T.; Gronenborn, A. M.; Chong, L. T. Evaluating the Strength of Salt Bridges: A Comparison
766 of Current Biomolecular Force Fields. *The Journal of Physical Chemistry B* **2014**, *118*, 6561–6569.
- 767 (53) Houriez, C.; Réal, F.; Vallet, V.; Mautner, M.; Masella, M. Ion Hydration Free Energies and Water
768 Surface Potential in Water nano Drops: The Cluster Pair Approximation and the Proton Hydration
769 Gibbs Free Energy in Solution. *J. Chem. Phys.* **2019**, *151*, 174504.

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0194968

- 770 (54) Guárdia, E.; Skarmoutsos, I.; Masia, M. On Ion and Molecular Polarization of Halides in Water. *Journal*
771 *of Chemical Theory and Computation* **2009**, *5*, 1449–1453.
- 772 (55) Réal, F.; Trumm, M.; Schimmelpfennig, B.; Masella, M.; Vallet, V. Further Insights in the Ability of
773 Classical Nonadditive Potentials to Model Actinide Ion-Water Interactions. *Journal of Computational*
774 *Chemistry* *34*, 707–719.
- 775 (56) Coles, J. P.; Masella, M. The Fast Multipole Method and Point Dipole Moment Polarizable Force
776 Fields. *J. Chem. Phys.* **2015**, *142*, 024109.