Hybrid polarizable simulations of a conventional hydrophobic polyelectrolyte. Toward a theoretical tool for green science innovation

Michel Masella, Alina Crudu, Fabien Léonforté

Hybrid polarizable simulations of a conventional Hydrophobic PolyElectrolyte. Towards a theoretical tool for Green Science innovation.

Michel Masella,*,† Alina Crudu,‡ and Fabien Léonforté*,‡

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

‡*L'Oréal Research & Innovation, France*

E-mail: michel.masella@cea.fr; fabien.leonforte@rd.loreal.com

Abstract

Hybrid modeling approaches based on *all-atom* force fields to handle a solute and coarse grained models to account for the solvent are promising numerical tools that can be used to understand the properties of large and multi components solutions and thus to speed up the development of new industrial products that obey the standard of green and sustainable chemistry. Here we discuss the ability of a full polarizable hybrid approach coupled to a standard Molecular Dynamics scheme to model the behavior in aqueous phase and at infinite dilution conditions of a standard hydrophobic polyelectrolyte polymer whose charge is neutralized by explicit counter ions. Beyond the standard picture of a polyelectrolyte behavior governed by an interplay between opposite intra polyelectrolyte and inter polyelectrolyte/counter ion Coulombic effects, our simulations show the key role played by both intra solute polarization effects and long range solute/solvent electrostatics to stabilize compact globular conformations of that polyelectrolyte. Our full polarizable hybrid modeling approach is thus a new theoretical tool well suited to be used in digital strategies for accelerating innovation for Green Sciences, for instance.

1 Introduction

Developing new products that obey the standards of green and sustainable chemistry with eco-sustainable ingredients is presently a major goal of many industries ranging from chemistry, oil manufacturers, personal care or cosmetics. However, because of the complexity of most of the solutions/emulsions that these industries use, the development of new, safe and biodegradable products is a particularly challenging task. For instance, standard oil-in-water emulsions, for cosmetic purpose for instance, comprise several components like neutral/charged flexible macromolecules, anionic/zwitterionic surfactants as well as salts and alcohols, and lots of efforts focus today in making such formula more and more natural by substituting historical compounds by naturally-derived molecules, *e.g.* polysaccharides[1]. However these kinds of emulsions correspond to complex multi-phasic systems whose properties are highly sensitive to the substitution/addition of new compounds and most of the inherent mechanisms that are involved in systems under such conditions are still poorly understood.

To substitute a component in a solution, a strategy consists in understanding how it interacts with its partners and to select a substitute able to play a similar role. This can be *a priori* achieved from experimental data, in particular microscopic structural ones or from physico-chemically equivalent parameters matching. However investigating experimentally the microscopic structures of complex molecular assemblies in solution is still a challenging task in modern chemistry and materials science[2,3]. Regarding macromolecular polymeric systems for instance, Atomic Force Microscope, AFM, imaging technique is able to provide mesoscopic pictures of them in various environments at best at the 10 nm scale[4–7]. Small Angle Neutron and X-ray Scattering, SANS/SAXS, techniques are also commonly used to investigate polymer-based solutions[8–11] even if the interpretation of their data may be controversial.[12] To tackle out the lack of experimental data, molecular simulations are a

valuable alternative. Since the nineties, computer Monte-Carlo, MC, and Molecular Dynamics, MD, simulations started to play an important role to understand the properties of polymer-based solutions, see Refs[13,14] for instance. These simulations are performed using different theoretical approaches, from pure *all-atom* schemes for which all the system components (solute, solvent and counter ions) are modeled at the same level of accuracy to hybrid approaches for which the solvent can be accounted for implicitly. For instance in 2009 Yethiraj and co-workers reported simulations performed using pure *all-atom* and hybrid approaches of simple Hydrophobic PolyElectrolyte, HPE, systems comprising at most 384 beads[15,16]. Today with the on-going increase of the available computational resources, more realistic *all-atom* simulations are commonly performed to investigate the properties of polyelectrolytes in aqueous phase. Besides simulations based on standard Lennard-Jones, LJ, potentials (see the recent studies Refs.[17–20]), MD simulations at the 100 *ns* scale of large hydrated polyelectrolyte systems (comprising up to 250k atoms) performed using standard pairwise force fields (and lattice-based Ewald summation techniques to account for long range electrostatics) are now regularly reported, see Refs.[21–23] for instance.

Despite the impressive progress of molecular modeling techniques, simulations of complex multi-components solutions/emulsions are still not routinely performed. That arises from the atomic size of realistic emulsions that are at the very least one to two order of magnitude larger than the systems simulated in the above studies. Moreover most of the simulations performed today rely on standard additive force fields that are acknowledged for their inability in accurately modeling charged systems in a polar medium like water, from single ion[24–27] to ionic liquids[28,29] and all hydrated polyelectrolyte systems[30]. That results from the inability of additive force fields to model the inherent non-pairwise character of microscopic polarization, a pivotal force to understand the interaction involving charged species, see Fig. 1.

Here we present a simulation work based on a new polarizable hybrid molecu-

lar simulation approach a priori well suited for accelerating innovation with digital strategies for Green Sciences. That approach relies on modeling interactions among non-solvent atoms using the *all atoms* polarizable force field TCPEp, which is based on a polarizable induced dipole moment approach[31], like the one of Fig. 1. However polarizable force fields are computationally demanding, not only because of the polarization computational cost (in particular as using Drude oscillators[32]) but also because usually such force fields also consider more sophisticated interaction potentials to model electrostatic effects (like the force field AMOEBA that accounts for atomic static dipoles and quadrupoles[33]) or short range electronic density reorganization phenomena, like TCPEp[31].

To perform efficiently MD simulations in aqueous phase by means of polarizable force fields, one of us proposed to handle water using the multi-level Polarizable Pseudo Particle, $\mathbf{PPP}^l$, coarse-grained, CG, approach[34–36]. $\mathbf{PPP}^l$ belongs to the macroscopic Density Functional Theory field. However as it retains the notion of particles, it also allows for an explicit solvent representation that is recommended for a realistic modeling of polyelectrolytes[18]. Moreover as $\mathbf{PPP}^l$ relies on a polarizable induced dipole moment approach, it can be readily coupled to *all-atom* induced dipole moment-based polarizable force fields like AMOEBA and TCPEp[37]. Two important features of the $\mathbf{PPP}^l$ approach are (1) to readily allow simulations in infinite dilution conditions and (2) to permit the modulation of the spatial range of the solvent electrostatic effects on a solute. That second feature allows the modulation the apparent solvent dielectric constant by taking constant the intensity of local interactions between the solvent particles and the solute[36]. As compared to standard *all-atom* simulations based on periodic condition schemes, the $\mathbf{PPP}^l$ approach thus allows one to investigate the properties of a HPE in water-like solutions that are fully disentangled from solute concentration effects, as well as to discuss the role of solvent macroscopic properties (in particular the magnitude of its dielectric constant).

The current work aims at showing the ability of the hybrid $\mathbf{PPP}^l$/TCPEp approach to model the complex components of industrial solutions and thus at paving the way in developing new digital workflows allowing to speed up the development of Natural products based on the knowledge of conventional ones. To this aim, we first focus on investigating the behavior of a generic mid-size HPE copolymer that is commonly used in conventional chemical industry for water/wastewater treatments, personal care as well as for clean and renewable energy production[38–42].

The HPE we consider, that is schematically depicted in Fig. 2, possesses a fraction f_c of charged groups per copolymer unit equals to 0.7, and independent from usual pH conditions. We denote that Polyelectrolyte as **P70** (because of its total charge : $+70\ e$) and we simulate it in two water-like solutions at infinite dilution conditions by means of MD simulations and of our hybrid approach : **P70** and its Cl^- counter ions are modeled by the *all-atom* polarizable TCPEp force field whereas the water-like environments are handled by the $\mathbf{PPP}^l$ CG approach. In particular we simulated the **P70** collapse from a linear conformation as dissolved in pure (*i.e.* salt free) $\mathbf{PPP}^l$ box that comprised about 0.9M solvent particles. The dielectric constant of that environment is one order of magnitude smaller than for liquid water. Then we investigated the effect of the solvent dielectric constant on the stability of the **P70** collapsed structures by simulating them embedded in a eight times more extended $\mathbf{PPP}^l$ medium whose dielectric constant is close to the liquid water one.

Less attention has been paid to investigate the role of polarization on HPEs than on bio-molecular polyelectrolytes like DNA/RNA[43–45] and charged polypeptides[46]. HPE properties in aqueous phase are usually inferred to arise from a interplay of microscopic interactions, namely intra HPE Coulombic repulsion and effective attraction resulting from hydrophobic effects[47,48]. We thus simulated a non-polarizable **P70** in our two $\mathbf{PPP}^l$ media in order to further discuss the role of polarization on hydrated HPEs beyond the standard Coulombic picture[49–53].

As far as we know no experimental data regarding the conformational behav-

ior of **P70**-like HPEs in water are available contrary to sulfonated polystyrenes for instance, see in particular Refs.[11,54]. To assess the reliability of our computational protocol we will also briefly present simulation results regarding hydrated carboxylated polystyrenes whose degree of carboxylation (charge fraction) varies from to 0 to 1 an whose properties may be reasonably expected to be close to those of sulfonated polystyrenes.

2 Theoretical methods

2.1 The all-atom force field

The *all-atom* polarizable force field TCPEp[31] is based on a total potential energy U of a molecular system decomposed into a sum of six contributions

$$U = U^{rep} + U^{qq'} + U^{pol} + U^{lh} + U^{disp} + U^{intra}. \quad (1)$$

The last energy term is the sum of the standard stretching, bending, improper torsional and dihedral torsional energy terms modeling the interactions among covalently bonded atoms. The first two terms correspond to the inter atomic short range repulsion and to Coulombic interactions

$$U^{rep} = \sum_{i=1}^{N} \sum_{j>i}^{N^*} A_{ij} \exp\left(-B_{ij} r_{ij}\right),$$
$$U^{qq'} = \sum_{i=1}^{N} \sum_{j>i}^{N^*} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}.$$

r_{ij} is the distance between atoms i and j, the q_is are the static charges located on atomic centers, and A_{ij} and B_{ij} are adjustable parameters. N is the total number of atoms within the molecular system and the superscript * indicates that the corresponding sum includes only pairs of atoms separated by more than two chemical

bonds. The repulsive term is truncated for distances r_{ij} greater than 5 Å.

The polarization energy term U^{pol} is based on an induced dipole moment approach. Induced dipole moments $\{\mathbf{p}_i\}_{1\leq i\leq N_\mu}$ are introduced, one per polarizable non-hydrogen atom. They obey

$$\mathbf{p}_i = \alpha_i \cdot \left(\mathbf{E}_i^q + \sum_{j=1}^{N_\mu^*} \mathbf{T}_{ij} \cdot \mathbf{p}_j \right). \tag{2}$$

The static charge electric field $\mathbf{E}_i^q$ acting on a polarizable atom i is generated only by the charges of the $N_\mathbf{E}$ surrounding atoms belonging to charged or dipolar chemical moieties. The N_μ and $N_\mathbf{E}$ sets do not necessarily match to each other. Let us consider the molecule $HO-CH_2-CH_2-CH_2-CH_2-NH_3^+$: only its first three atoms HO−C (alcoholic moiety) and its last four ones $C-NH_3^+$ (cationic head) belong to the $N_\mathbf{E}$ set, and only its carbon and oxygen atoms are polarizable centers. α_i is the isotropic polarizability of the polarizable atom i and $\mathbf{T}_{ij}$ is the dipolar interaction tensor. The static electric fields and dipolar tensors include short range Thole-like damping functions[55] that vanish for interatomic distances greater than 5 Å. The set of Eqs. 2 is iteratively solved and the resulting polarization energy term U^{pol} is

$$U^{pol} = \frac{1}{2} \sum_{i=1}^{N_\mu} \frac{\mathbf{p}_i^2}{\alpha_i} - \sum_{i=1}^{N_\mu} \mathbf{p}_i \cdot \mathbf{E}_i^q - \frac{1}{2} \sum_{i=1}^{N_\mu} \sum_{j=1}^{N_\mu^*} \mathbf{p}_i \mathbf{T}_{ij} \mathbf{p}_j. \tag{3}$$

The term U^{disp} models microscopic dispersion effects that are pivotal to describe hydrophobic alkyl chains. It is a sum of basic C_{ij}/r_{ij}^6 terms where the $C_{ij}(<0)$ are ajustable parameters. For our purpose only non-hydrogen atoms are considered as dispersion centers. To model particular electronic density reorganization effects (as those occurring within acrylamide hydrogen bonds), TCPEp also includes a set of short-range cooperative energy terms denoted U^{lh} that are very close to the terms we introduce to model accurately water hydrogen bond networks[56]. For the present study such energy terms are taken into account to model the interactions among

the **P70** acrylamide moieties. However they play a negligible role in our simulations and they will not be further discussed.

2.2 Force field parameters

TCPEp parameters are adjusted to reproduce high-level *ab-initio* quantum data regarding a training set of small molecular systems (like the atomic charges that are assigned to reproduce the dipole moment of neutral molecules and the charges of charged molecules to meet those computed from the quantum Natural Population Analysis scheme[57]). **P70** can be decomposed in a set of chemical groups for which we have already assigned accurate parameters[35,36,58–60], at the exception of a reduced set of parameters regarding the torsional degrees of freedom of the dimethyldiallyl ammonium five membered ring that were specifically assigned for the present study according to the protocol detailed in Ref.[59].

Regarding alkyl groups, their parameters (mainly repulsion and dispersion) were assigned to reproduce the quantum binding energy (-0.5 kcal mol^{-1}) and geometry (a carbon/carbon equilibrium distance of 3.5 Å) of the methane dimer in gas phase as predicted by quantum CCSD(T) computations at the Complete Basis Set limit[61]. Because of the very weak permanent dipole in alkanes, we assume in the present study that only short range repulsion and dispersion govern the interactions involving alkyl groups in absence of external electric fields. We thus systematically neglect Coulombic interactions and static charge electric fields arising from alkyl moieties. However we consider the alkyl carbons as polarizable centers (their isotropic polarizability is set to 2.1 $Å^3$) and they undergo the static charge and/or induced dipole electric fields generated by all the chemical moieties surrounding them.

Lastly, we use no long range truncation scheme as computing Coulombic interactions and static charge electric fields regarding **P70** and its counter ion cloud, whereas we truncate all their induced dipole/induced dipole interactions corresponding to distances larger than 12 Å.

2.3 The multi-level coarse-grained water model PPP^l

The main feature of the water coarse-grained approach **PPP**^{l} [35–37] is to model a single three atomic water molecule or a set of water molecules by a single polarizable pseudo-particule (denoted as PPP) whose polarizability α_s obeys a Clausius-Mosotti relation:

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

here, ϵ_s and ρ_s are the dielectric constant and the particle density of liquid water (a quantity that is proportional to the inverse of the water molecular volume). α_s is thus a linear function of the PPP volume. Another important assumption is the use of the *local* approximation : the intensity of the induced dipole moment $\mathbf{p}_j^s$ corresponding to PPP j is modulated only by the solute electric field $\mathbf{E}^j_{\text{solute}}$ acting on j, *i.e.* the PPPs don't polarize each other. However that can yield to largely overestimate solute/PPP polarization effects up to unphysical values in particular situations. To prevent such artefacts, **PPP**^{l} allows the $\mathbf{p}_j^s$s to saturate according to[34]

$$\mathbf{p}_j^s = \mu_s \mathcal{L}\left(\frac{3\alpha_s \mathbf{E}^j_{\text{solute}}}{\mu_s}\right) \frac{\mathbf{E}^j_{\text{solute}}}{\left|\mathbf{E}^j_{\text{solute}}\right|}, \tag{5}$$

here $\mathcal{L}$ is the Langevin function and μ_s is the PPP saturation dipole value. The corresponding PPP/solute polarization energy is then

$$U_{ps}^{pol} = -\frac{\mu_s^2}{3\alpha_s} \sum_{j=1}^{N_s} \ln\left[\frac{\sinh\left(3\alpha_s \left|\mathbf{E}^j_{\text{solute}}\right| / \mu_s\right)}{3\alpha_s \left|\mathbf{E}^j_{\text{solute}}\right| \mu_s}\right]. \tag{6}$$

Here N_s is the number of PPPs. Note that for both the above equations and weak solute electric fields, the linear regime corresponding to Eqs. 2-3 is recovered.

The polarizability α_s of the PPPs accounts not only for the solvent electronic polarization induced by the solute but also for the solvent orientational polarization[34].

As such α_s does not correspond to a standard atomic or a molecular polarizability (which models only the perturbation of an electronic cloud by an external electric field). Moreover the α_s allows one to consider PPPs of different sizes. As modeling a HPE by means of an *all-atom* force field, the natural choice is to set the size of all the PPPs to that of a single water molecule (that yields a PPP polarizability of 2.35 Å^3). Moreover, within the $\mathbf{PPP}^l$ framework, we assume solute/solvent long range electrostatic contributions arising from neutral solute groups to be negligible. We thus systematically truncate both the induced dipole and static charge components of the corresponding electric fields $\mathbf{E}^j_{\text{solute}}$ for solute atom/PPP distances greater than the cut off distance $R^{\text{pol},1}_{\text{cut}}$.

However the above spherical truncation scheme not only yields large underestimation of ion hydration energies but also large artefacts in modeling ion association (see Ref.[36] and references cited therein). To remediate both these drawbacks we proposed the multi-level scheme $\mathbf{PPP}^l$ to model efficiently the long-range electrostatic interactions between a charged solute and farther (and farther) solvent domains[36]. Within a shell lying at the vicinity of the charged solute (that shell extends to no more than $R^{\text{pol},1}_{\text{cut}}$ from any non hydrogen solute atom) that multi-level approach considers PPPs whose size matches that of a water molecule. These first solvent level PPPs ($l = 1$) undergo the static charge electric field generated by all the solute charged groups (and the truncated static and dipole electric field components arising from neutral groups). To model solute/solvent interactions at longer distances than $R^{\text{pol},1}_{\text{cut}}$, larger (and larger) PPPs are introduced according to the scheme shown in Fig. 3. These larger PPPs define higher level ($l > 1$) solvent shells. They undergo the static charge electric field generated by all the solute charged groups if they lie within a shell surrounding the solute and extending from $R^{\text{pol},l-1}_{\text{cut}}$ to $R^{\text{pol},l}_{\text{cut}}$ from it. The solute/solvent polarization term $U^{\text{pol}}_{ps,l>1}$ corresponding to each $l > 1$ PPP level is taken under the linear form

$$U^{\rm pol}_{ps,l>1} = \frac{1}{2}\sum_{i=1}^{N_\mu^{l>1}} \frac{\mathbf{p}_i^2}{\alpha_s^{l>1}} - \sum_{i=1}^{N_\mu^{l>1}} \mathbf{p}_i \cdot \mathbf{E}_i^q. \quad (7)$$

Here, $N_\mu^{l>1}$ are the number of $l > 1$ PPPs, $\alpha_s^{l>1}$ is their polarizability, according to Eq. 4, and $\mathbf{E}_i^q$ is the static charge electric field generated by the solute charged groups on these PPPs.

Besides largely improving the modeling of ion hydration and ion association, we also showed that multi-level approach to allow an efficient modeling of particularly large and extended solvent domains at the μm scale and above[36]. For the present study, we consider two **PPP**^{l} solvent approaches. The first, denoted Short Range PPP, **SRP**, models the solvent by means of only first level PPPs and by setting the shell-based cut off distance $R_{\rm cut}^{\rm pol,1}$ to 12 Å. The second approach, denoted Long Range PPP, **LRP**, considers also a second level of PPPs made of particles that are 8 times larger than the first level ones. Their polarizability is thus $\alpha_{s,2} = 18.8$ Å^3. The second level PPPs undergo the static charge electric field generated the **P70** charged groups if these PPPs are located within a shell encompassing the HPE from $R_{\rm cut}^{\rm pol,1} = 12$ to $R_{\rm cut}^{\rm pol,2} = 143$ Å. In an earlier study[36] we showed the **SRP** approach to reinforce the association of oppositely charged ions and the repulsion of charge-like ions as compared to liquid water (the apparent dielectric constant of a first level PPP solution is one order of magnitude weaker than that of liquid water), whereas the **LRP** approach already provides a description of ion pairing agreeing with that expected in liquid water. The dielectric constant of the **LRP** medium is thus close to the liquid water one.

Within the multi level **PPP**^{l} approach, only electrostatic solute/$l > 1$ PPPs interactions are taken into account. Regarding first level PPPs, their non-electrostatic interactions with solute atoms are modeled by means of a Lennard-Jones-like potential corresponding to a stronger repulsion[37] to which is added a specific short range many-body term to prevent over populated first hydration shell at the vicinity of

charged species[36]. The corresponding force field parameters for first level PPPs are assigned to reproduce both the first hydration shell structure and the hydration Gibbs energy of a training set of small solutes and ions[36]. For the present purpose we used an updated parameters that will be discussed elsewhere.

Interactions among first level PPPs are modeled using a standard additive Lennard-Jones, LJ, energy term and a many-body term that is a function of the solvent local density at the vicinity of a PPP[35]. Both these terms are truncated for inter particle distances greater than $R_{\text{cut}}^{\text{pp}} = 7$ Å and their parameters were assigned to reproduce the liquid water density (0.0331 molecule per Å^3) and the two regimes of the energy corresponding to the creation of an empty cavity in water at ambient conditions according to the Lum-Chandler-Weeks theory of hydrophobicity[62].

Interactions among second order PPPs are modeled using only a LJ term that is truncated for distances greater than 10 Å. The corresponding LJ radius is twice as large as for first order PPPs and the LJ intensity (denoted ϵ_2) is one order of magnitude weaker than for the first level PPPs as in our original study[36]. In that original study we showed second level PPPs to be slightly over concentrated at their lower cut off distance as interacting with a single monovalent ion like Na^+ and Cl^-. Because of the large total charge of the **P70** ($+70\ e$) and because of the overall large number of counter ions Cl^-, our choice regarding the LJ parameter ϵ_2 can lead to drawbacks like over organized **P70**/counter ions/second order PPPs structures. We will discuss that particular issue in the forthcoming Section 4.

Contrary to the recent approach proposed by Chremos and Douglas[17,20,63], the use of only first level PPPs or in conjunction with second level ones allows the modulation of the apparent dielectric constant of the solvent without altering the short range solute/solvent interactions. Moreover our shell-based cutoff scheme also allows the modeling of solvation at infinite dilution conditions. In our $\mathbf{PPP}^l$ approach, a solute is embedded in a cubic box fulfilled with first level PPPs and that system is then embedded in a larger box made of second level PPPs and so on[36]. To maintain

the PPP density within each box, the PPPs are allowed to interact with their own periodic images. As the long range interactions among PPPs are truncated, only the first 26 periodic box images are usually needed for that task. Regarding the solute, to model its solvation at infinite dilution conditions, one just needs to neglect the interactions between the solute and its own periodic images. That choice is made for both the solvent approaches **SRP** and **LRP**. However note that solute atoms can interact with PPP images if they lie close to the box boundaries.

Regarding **SRP**, ignoring the interactions between the solute and its own periodic images corresponds technically to model it as infinitively diluted. However we may also interpret that solute as embedded in a confined water cavity delimited by the shell of first order PPPs. That interpretation is supported by the modeling of the short range solute/PPP interactions using the same potential energy terms in both the **SRP** and **LRP** approaches.

We assume here first level PPPs (those who interact at short range with the solute) to be symmetric. Other choices are possible as considering a triatomic molecule holding a permanent dipole like the isolated water molecule. However we may note that the isotropic polarizability of Equation 4 allows to account for both the solvent orientational and electronic polarization induced by a solute[34]. Moreover modeling the solvent using triatomic molecules will lead the approach to be less computationally efficient, at least by factor 3.

2.4 Molecular Dynamics details

MD simulations are performed at ambient conditions using the code POLARIS(MD)[64]. The Newtonian equations of motion are solved using a multiple-time-steps algorithm with two time steps: 0.25 *fs* for intra-solute stretching, bending and improper torsional energy terms and 2 *fs* for all the other interactions. All the covalent X−H bonds and H-X-H angles are constrained to their force field reference values by means of the iterative RATTLE procedure (the convergence criterion is set to 10^{-5}

Å). The system temperature (and pressure) in NVT and NPT simulations are monitored by means of the Langevin dynamics approach detailed in Ref.[65] and of the Nosé-Hoover barostat[66], respectively. Solute atom and PPP induced dipole moments are iteratively solved until the mean difference in these dipoles between two successive iterations is less than 10^{-6} Debye.

The center of mass of the **P70**/counter ions system is regularly reset to the simulation box center (and the PPP coordinates updated accordingly) along our MD simulations. However to prevent any solute atom to leave that box we also consider the repulsive potential $U_{\rm box}^{rep}$ acting only at the simulation box boundaries in order to restrain the solute atoms within the simulation box core. Let us consider a solute atom whose cartesian coordinates are (x, y, z) and a simulation box whose dimensions are (L_x, L_y, L_z), $U_{\rm box}^{rep}$ is defined as

$$U_{\rm box}^{rep} = \sum_{\xi=x,y,z} k \left[\max\left(2|\xi| - (L_\xi - \delta L), 0\right)\right]^2, \tag{8}$$

here k and δL are two constant parameters that are set for the present study to 5 kcal mol^{-1} Å^{-2} and 3 Å, respectively. $U_{\rm box}^{rep}$ and all its derivatives vanish for solute atoms located within the simulation box and farther than δL from its boundaries. Because of the dimensions of our cubic boxes (at least 90 Å) the effect of that potential is very weak (to not say negligible) on the **P70** and counter ions dynamic.

3 Results and discussion

3.1 The collapse of P70 in the SRP medium

3.1.1 Simulation starting structure

The simulation Linear Starting Structure, LSS, of **P70** is shown in Fig. 2. The **P70** total electrostatic charge is neutralized by adding 70 Cl^- counter ions that are successively set on the nodes of a cubic grid in which the LSS is set (the grid node

dimension is 3 Å). A counter ion is set on a node (1) to minimize the Coulombic interaction energy between it and **P70** and all the already added counter ions, and (2) so that the shortest distance between it and any **P70** non hydrogen atom is larger than 3.5 Å. The resulting **P70**/counter ions structure is then set in a cubic box fulfilled with PPPs that are set on all the nodes of the above grid that are distant by more than 3 Å from any non-hydrogen/counter ion atom. To solvate **P70** in its LSS conformation, a large cubic box whose volume is 301^3 Å^3 and comprising about 0.9M PPPs is needed to ensure that there is at least a distance of 12 Å between any solute atom and the box boundaries, see Fig. 2.

3.1.2 P70 structural evolution

We performed 8 independent MD simulations at the 2 *ns* scale of the **P70**/counter ions system in the NPT ensemble as dissolved in the **SRP** medium. Each simulation corresponds to a different set of starting atom/particle velocities randomly set and corresponding to a temperature of 300 K. Along the first 200 *ps* simulation segments, we constrained **P70** to its LSS using a harmonic potential restraining the position of its nitrogens. That potential was then removed. We plot in Fig. 4 the temporal evolution of the polymer gyration radius R_g and of the ratio R^{inertia} between the largest and smallest **P70** inertia moment values.

As soon as the harmonic potential preventing polymer structural transitions is removed, R^{inertia} evolves to reach a value included within 1 and 2 and R_g converges towards a value of 14 $\pm$ 0.2 Å within at most 1.2 *ns*, regardless of the simulation. **P70** collapses thus rapidly towards a Compact Quasi Spherical, CQS, conformation, see Fig. 2. Note that we performed also these simulations however in the NVT ensemble (see our protocol detailed in Section 2.4) from a starting structure relaxed by performing a NPT MD run along which the linear **P70** structure is harmonically constrained. Along these NVT trajectories, **P70** collapses again towards a CQS conformation but usually at a slower rate (at least not before 2 ns). To our opinion

that difference in the collapsing time scales arises from the thermostat and barostat that we consider. That may also explain the much longer time needed by **P70** to collapse along our **LRP** simulations that are performed in the NVT ensemble (see below).

The **P70** compact conformations at the end of each simulation differ at the atomic level. For instance the Root Mean Square Deviation, RMSD, of the ammonium nitrogens N_a, among the 8 final simulation snapshots amounts to 7 Å. However the structural properties of the ammonium groups within these final conformations are close on average. In Fig. 4 we plot the radial N_a/N_a pair distribution functions g_{NN} as computed along the last 500 *ps* simulation segments as well as the mean numbers $N_{N_a}(r)$ of N_a atoms lying within a sphere of radius r from the Polymer Center of Mass PCOM (*i.e.* the integrals of N_a/PCOM radial distribution functions). These plots show all the N_a conformations to be very close at the end of each independent simulation. Hence **P70** evolves towards different compact conformations whose cationic charge distribution (that governs the electrostatic interactions within the polymer and between the polymer and its chemical environment) are close, suggesting these conformations to be iso-energetic in the **SRP** medium.

Regarding counter ions, we plot in Fig. 4 their mean radial distribution functions g_{ci} from PCOM, computed from the last 500 *ps* simulation segments, as well as their integrals yielding the mean number of counter ions N_{ci} within a sphere of radius r from PCOM. The functions g_{ci} are close : they present two peaks located at 15 and 23 $\pm$ 1 Å from PCOM and encompassing 5 and 68 $\pm$ 1 counter ions, respectively. About 8% of the counter ions are thus absorbed on the **P70** surface whereas 90% of them are located in a more distant shell extending from 4 to 21 Å from the polymer surface. Lastly, at most 2 counter ions are present within the polymer CQS core and only one is released far away from **P70** along a single simulation. Counter ions were distributed all along the **P70** LSS. The counter ions undergo relatively large displacements at the early stages of our simulations before to form a spherical

cloud. In the starting linear structure the counter ions mainly interact only with a few cationic charges of **P70** whereas they interact all with all the **P70** cationic charges located in rather small spherical volume at the end of the simulations.

In Fig. 4 we also plot the radial distribution functions $g_{\mathrm{PPP}}(r)$ between the PPPs and PCOM as well as their integrals $N_{\mathrm{PPP}}(r)$. The functions $g_s(r)$ are converged to their expected bulk value for PPP/PCOM distances $r \leq 30$ Å, *i.e.* at about 16 Å from the **P70** surface. We note also (1) a depletion of the PPP density at the vicinity of the polymer surface, that may be interpreted as resulting from the hydrophobic nature of **P70**, and (2) by a PPP density reinforcement in between the **P70** surface and the external counter ions shell. Lastly a far from negligible number of PPPs (about 270) are trapped within the core of the **P70** CQS conformations : their density within the **P70** spherical volume is 0.024 PPPs per Å^3, a value 40 % weaker than within the solvent bulk.

To assess the stability of a **P70** CQS conformation on longer simulation times, we select a final CQS structure (together with its counter ion cloud) from one the above independent simulations. That structure is embedded in a new cubic box whose volume is 90^3 Å^3 and comprising 23k PPPs. We then performed a 500 *ns* scale NPT MD simulation of that solvated system. Along the new simulation, **P70** jumps from its starting structure towards more and more different ones (the RMSD($\mathrm{N_a}$) value increases up to 5 Å) but the new structures all correspond to CQS conformations, see the superimposition of **P70** structures in Fig. 5. Along that 500 *ns* simulation, the counter ion cloud and solvent structural properties all agree with those discussed above. Hence, in a salt free **SRP** medium and at infinite dilution conditions, the **P70** structure fluctuates between different CQS conformations whose properties are close on time intervals close to the μs scale. However we can not conclude from our simulations on the thermodynamical nature (global minima of the potential energy surface or transient structures[67]) of these CQS conformations.

3.1.3 Energies

To discuss the energetic data, we performed a single 1 *ns* NPT simulation of **P70** constrained in its LSS conformation and embedded together with its counter ion cloud within a 0.9M first order PPP cubic box. The mean values of all the interaction energy components computed from the last 800 *ps* segment of that simulation are taken as the reference energy values at the beginning of the HPE collapse process.

In Fig. 6, we plot the temporal evolutions of the total intra-**P70** and intra-counter ions energies, of the **P70**/PPPs and counter ions/PPPs polarization energies, and of the total **P70**/counter ions interaction energy as soon as the polymer constraints are removed along our independent simulations. The dispersion in these energy values among the simulations is weak (at most 1 %) showing all the final **P70**/counter ions structures to be close in energy, as expected from the similarity of the **P70** charge distribution at the end of all the simulations. Because of the large **P70** electrostatic positive charge (+70 e) and thus from the large negative charge of the counter ion cloud, there is a large increase in both the intra polymer ($\Delta\bar{U}_{\rm p}$) and intra counter ions ($\Delta\bar{U}_{\rm ci}$) energies during the polymer collapse process : $\Delta\bar{U}_{\rm p}$ and $\Delta\bar{U}_{\rm ci}$ amount to about +25 and +15 10^3 kcal mol^{-1}, respectively. However these large energy increases are largely counter balanced by strong stabilizing **P70** and counter ions polarization interactions with the solvent, as well as by strong stabilizing **P70**/counter ions interactions : during the polymer collapse process, the decreases $\Delta\bar{U}_{\rm ps}$, $\Delta\bar{U}_{\rm cs}$ and $\Delta\bar{U}_{\rm pc}$ in the energies corresponding to the latter interactions amount to -5, -16 and -35 10^3 kcal mol^{-1}, respectively.

Within our hybrid modeling framework, the collapse of **P70** towards CQS conformations in a **SRP** medium is driven by a competition between (1) strong destabilizing intra-**P70** and intra-counter ion cloud interactions and (2) strong stabilizing **P70**/solvent and **P70**/counter ions ones. By analyzing in details the main components of the interaction energies (namely the repulsion, Coulombic, polarization and dispersion components), interactions among counter ions are largely dominated by

repulsive Coulombic effects. Intra-**P70** interactions are also dominated by repulsive Coulombic effects and by far from negligible attractive polarization interactions that favor **P70** CQS conformations : the variation of the intra-**P70** polarization energy during the collapse process amounts to about 20% of the Coulombic one in absolute values. We may also note that intra-**P70** dispersion and atomic repulsion equally favor/disfavor the CQS conformations, however the corresponding energy variations during the collapse process are one order of magnitude smaller than for polarization. Lastly we identified no more than 7 intra-**P70** hydrogen bonds among the acrylamide groups along the simulations. That corresponds to an interaction energy of about 35 kcal mol^{-1}, which is negligible compared to the main intra-**P70** Coulombic and polarization energy components.

The attractive polarization forces within the **P70** CQS conformations are tied to strong alkyl carbon induced dipole moment values μ_C that can reach values as large as 5 Debye. In Fig. 7, we plot the mean values of μ_C for each carbon computed along all the final 500 *ps* simulation segments as a function of the mean distance of the alkyl carbons from PCOM. In line with the expected intensity of electric fields generated by quasi spherical positive charge distributions, the values μ_C increase as the alkyl carbons are closer to the surface of **P70** in a CQS conformation. Regarding the strong μ_C values, we may note here that strong induced dipole moments (larger than 2 Debye) are also predicted for water lying at the vicinity of the heavy cation Th(IV) or the halide anion F^- in liquid water by both polarizable *all-atom* force field and quantum Car-Parinello simulations[68–70].

Contrary to the present results, a large charge fraction f_c of 0.7, as for **P70**, is experimentally shown to prevent a HPE to collapse towards a compact globular conformation and to favor instead elongated structures from pear-necklace to wormlike ones in different environments like water and water/acetone mixtures[6,9–11]. First a non negligible fraction of the counter ions (about 8 %) are adsorbed on the **P70** surface. That discharging effect yields thus a slightly weaker apparent HPE $\tilde{f}_c$ value

(about 0.65) than f_c. Moreover the intra-molecular stabilizing polarization effects within a **P70** CQS conformation counter balance about 20 % of the destabilizing intra-polymer Coulombic interactions. We may also interpret these polarization effects to be responsible for a further decrease of $\tilde{f}_c$ to about 0.5. Lastly note also (1) the ionic charge to be more diluted in quaternary ammonium groups than in the $-NH_3^+$ and $-SO_3^-$ ionic heads that were used to investigate the structural properties of HPEs as a function of f_c in Refs.[6,9–11]; and (2) the inter ionic head distances between adjacent polymer units in the highly charged poly(vinyl amine) HPEs of Ref.[6] are about twice shorter than in **P70** (see the g_{NN} data in Fig. 4 for instance). That yields a still weaker $\tilde{f}_c$ value and that explains why a HPE with a large f_c value like **P70** can collapse towards a CQS conformation in a **SRP** medium.

3.1.4 Comparison to carboxylated polystyrenes $\mathbf{PSC}^{f_c}$

To further discuss the reliability of our multi level simulation scheme, we investigated the behavior of HPEs corresponding to carboxylated polystyrenes $\mathbf{PSC}^{f_c}$ as dissolved in first order PPP boxes. Here f_c is the degree of carboxylation/charge fraction that we set for the present purpose to 0.00, 0.25, 0.50 and 1.00, respectively. The length of the $\mathbf{PSC}^{f_c}$ is 80 and their negative charge is neutralized by adding Na^+ counter ions. For our purpose we consider all the carboxylates as deprotonated anionic heads that are regularly spaced along the $\mathbf{PSC}^{f_c}$ chains. Our $\mathbf{PSC}^{f_c}$ are similar to the sulfonated polystyrenes experimentally investigated by Boué and co workers[11] even if their length is about one order of magnitude shorter. Note the distance between two adjacent ionic heads in the linear and fully carboxylated $\mathbf{PSC}^1$ is about 2.5 Å whereas it is at least twice as large in any **P70** CQS conformation.

To fully solvate the linear starting structures of the $\mathbf{PSC}^{f_c}$ we need smaller PPP boxes as compared to **P70** : their dimensions are now about 200 Å and they comprise about 260k first order PPPs. For each solvated $\mathbf{PSC}^{f_c}$ we performed a single MD simulation at the 100 ns scale according to our MD protocol detailed

above. As for **P70** all the force field parameters to model intra $\mathbf{PSC}^{f_c}$ and $\mathbf{PSC}^{f_c}$ /counter ion interactions are assigned only from quantum *ab initio* computations and taken from Ref.[31].

The temporal evolutions of the gyration radii R_g and the final snapshots of our $\mathbf{PSC}^{f_c}$ are provided as Supplementary Material. Contrary to **P70** the fully carboxylated PSC^1 does not collapse towards a compact conformation along the 100 ns trajectory: rather it slowly evolves towards a helicoidal structure stabilized by salt bridges between carboxylate heads and Na^+ ions. The three other $\mathbf{PSC}^{f_c}$ collapse towards more and more compact (and less organized) conformations. In particular $\mathbf{PSC}^0$ collapse fastly (within a few ns) towards a stable compact conformation close to the **P70** CQS ones. Moreover the $\mathbf{PSC}^{f_c}$ gyration radii decrease as f_c decreases, from 22 ($f_c = 1$) down to 11 ($f_c = 0$) Å. Even if Boué and co workers[11] experimentally investigated longer sulfonated polystyrene chains at 0.1-0.3 M concentrations, our simulation results regarding our $\mathbf{PSC}^{f_c}$ are in line with their experimental conclusion[11,54] : the $\mathbf{PSC}^{f_c}$ gyration radius R_g depends increasingly on f_c.

In all our **P70** and $\mathbf{PSC}^{f_c}$ simulations clearly show the variability of HPE behaviors in water, behaviors that can not be modeled by considering only the HPE charge fraction f_c. In the particular case of **P70**, we will discuss below another pivotal parameter to understand HPE behaviors, *i.e.* the polarizability of the neutral HPE moieties.

3.2 The collapsed conformations of P70 in the LRP medium

We performed a new set of 8 independent NVT simulations at the 10 *ns* scale of **P70** surrounded by its counter ion cloud in the **LRP** medium. The simulation starting structures correspond the final ones of the **SRP** simulations discussed above : the first level box comprising **P70** in a CQS conformation, the counter ions and the 0.9M first level PPPs is embedded in a second level cubic box whose volume (about 216 nm^3) is 8 times larger than the relaxed volume of the first level PPP box and

that is fulfilled by about 0.9M second level PPPs. The molecular size of that new **LRP** system is thus equivalent to a 8M water molecules system.

3.2.1 Structural properties

In Fig. 8 we plot the temporal evolution of the **P70** gyration radius R_g along the new simulations. These plots show a non-negligible and systematic contraction of R_g along these simulations from 14.1 (**SRP**) to 12.4 (**LRP**) $\pm$ 0.1 Å. In **LRP**, the **P70** CQS conformations are thus more compact than in **SRP**, by about 20 % in volume. However, the most striking difference between **SRP** and **LRP** simulations concerns the counter ion cloud. In Fig. 8 we plot the functions $g_{\rm ci}$ and $N_{\rm ci}(r)$ computed along the last 2 *ns* **LRP** simulation segments. If about 10 % of the counter ions are adsorbed on the polymer surface as in **SRP**, only 24% of them now form a first shell lying at the vicinity of the polymer surface (as in the **SRP** case) and 66% of them form a second shell away from **P70** (it is centered at about 60 Å from PCOM).

Regarding first order PPPs, their structural features at the vicinity of the **P70** surface are close to the **SRP** case, as well as their number trapped within the **P70** CQS conformations (again about 270, see Fig. 4). In all the **P70** CQS conformations are stable in **LRP**. However they are more contracted and their counter ion cloud is strongly altered compared to **SRP**, with a large fraction of the counter ions lying now away from **P70**.

3.2.2 Energies

The energy components $\Delta\bar{U}_{\rm p}$, $\Delta\bar{U}_{\rm ci}$, $\Delta\bar{U}_{\rm pc}$, $\Delta\bar{U}_{\rm ps}$ and $\Delta\bar{U}_{\rm cs}$ computed along the last 2 *ns* segments of the **LRP** simulations are summarized in Tab. 1. As expected the largest differences between **SRP** and **LRP** data arise from interactions involving the counter ions and the PPPs. Because of the release of the counter ions away from **P70**, the magnitude of the **P70**/counter ions interaction energy $\Delta\bar{U}_{\rm pc}$ and of the intra counter ion cloud energy $\Delta\bar{U}_{\rm ci}$ decreases from **SRP** to **LRP** by about 90%

and 80 %, respectively. On the other hand long range solute/solvent interactions strongly reinforce the stability of the polymer CQS conformations : the magnitudes of $\Delta\bar{U}_{\rm ps}$ and $\Delta\bar{U}_{\rm cs}$ are from 3 to 4 times stronger along **LRP** simulations compared to **SRP** ones, and that arises mainly from the second level PPPs.

From Eq. 7, we may estimate the order of magnitude of the hydration long range electrostatic free energy $\Delta G_{\rm lr}$ of a solute of charge Q (modeled as a point charge and whose gyration radius is R_g) dissolved in a homogeneous second level PPPs shell extending from $R_{\rm cut}^{\rm pol,1}+R_g$ to $R_{\rm cut}^{\rm pol,2}$ according to

$$\Delta G_{\rm lr} = -\frac{\alpha_{s,2}}{2}\frac{Q^2}{4\pi\epsilon_0}\int_{R_{\rm cut}^{\rm pol,1}+R_g}^{R_{\rm cut}^{\rm pol,2}}\frac{\rho_s}{R^4}4\pi R^2 dR = -\frac{Q^2}{8\pi\epsilon_0}\left(\frac{1}{R_{\rm cut}^{\rm pol,1}+R_g}-\frac{1}{R_{\rm cut}^{\rm pol,2}}\right). \quad (9)$$

For **P70** in a **SRP** CQS conformation, that yields $\Delta G_{\rm lr} \approx -26\ 10^3$ kcal mol^{-1}, a value in line with the $\Delta\bar{U}_{\rm ps}$ data of Tab. 1. That relation can be readily rewritten to estimate the hydration short range solute/solvent electrostatic free energy $\Delta G_{\rm sr}$ (arising from first level PPPs) as

$$\Delta G_{\rm sr} = -\frac{Q^2}{8\pi\epsilon_0}\left(\frac{1}{R_g}-\frac{1}{R_{\rm cut}^{\rm pol,1}+R_g}\right). \quad (10)$$

These relations suggest that the more compact is a HPE CQS conformation the stronger it is stabilized by both short and long range HPE/solvent electrostatic interactions. For instance the gyration radius R_g smaller by 1.7 Å in **LRP** than in **SRP** is responsible for an overall large decrease of $\Delta G_{\rm sr}$ by about -5 700 kcal mol^{-1} according to the above relation. Within the **PPP**^{l} framework, long range **P70**/solvent electrostatic effects favor the contraction of a CQS conformation and thus the increase of the local charge densities within **P70**, a phenomenon which in turn reinforces both the short and long range solvent/polymer electrostatic interactions. However the contraction of a CQS conformation also yields to strengthen both the intra HPE repulsive Coulombic and attractive polarization effects, by about 25

and 40 %, respectively (see Tab. 1).

We assess the stability of one of the final **LRP** CQS conformations by performing a 500 *ns* simulation of it embedded together with its counter ion cloud in a first level cubic box whose volume is 21.9 nm^3 (and comprising about 80k first level PPPs), and in a 8 times larger second level box (comprising also about 80k second level PPPs). In that particular case $R_{\text{cut}}^{\text{pol},2}$ is 62.5 Å. We observe here also the **P70** structure to jump among different CQS conformations whose structural and energetic properties are close. In particular, their gyration radii are smaller than in **SRP** by about 1.7 Å, on average, see Fig. 5.

3.2.3 Second order PPPs and the distant counter ion shell

The **P70** PCOM/second order PPP radial distribution functions reported in Fig. 8 (d) show over concentration of second order PPPs in between **P70** and the distant counter ion shell, a domain in which the electric field components generated by opposite charged solutes are expected to be strong (whereas they mostly cancel each other for like charged solutes, see Fig. 1 and Fig. 3-B). Hence the organized **P70**/second order PPPs/counter ions shell structures observed at the end of our **LRP** simulations yields to reinforce the polarization interactions among the second order PPPs, **P70** and a set of counter ions. In our original study, we already showed the density of second order PPPs to be slightly increased at the lower boundary of their cut off shell domain[36] as they interact with a single monovalent ion like Cl^- or Na^+. Here the large electrostatic charge of **P70** (and of its counter ions cloud) yields to overestimate that phenomenon. In Section 4 we show how the decrease its intensity by reinforcing the repulsion of second order PPPs at short range. However even in that case the above conclusion regarding the stability and the contraction (and its magnitude) of the **P70** CQS conformations arising from long range solvent effects remains valid.

3.2.4 Comparison to carboxylated polystyrene $\mathbf{PSC}^{f_c}$

We also simulated the carboxylated $\mathbf{PSC}^{0.5}$ and $\mathbf{PSC}^{1}$ in the **LRP** medium. The linear $\mathbf{PSC}^{f_c}$ starting structures (comprising Na^+ counter ions) correspond to those embedded in a 260k first order PPP box as detailed above. These hydrated structures are here embedded in second order PPP boxes comprising also about 260k particles. For each $\mathbf{PSC}^{f_c}$ we performed a single 100 ns MD simulations. Anticipating the conclusions of Section 4, we set the parameter ϵ_2 to 1 kcal mol^{-1}. The temporal evolutions of the gyration radii and the final snapshots at the end of these simulations for both the $\mathbf{PSC}^{f_c}$ are provided as Supplementary Materials.

In **LRP**, the final conformation of $\mathbf{PSC}^{1}$ is more elongated than in **SRP**: its gyration radius R_g is 31 Å in **LRP**, larger by 41 % than in **SRP**. A weaker fraction of counter ions Na^+ interact at short range from carboxylate heads in **LRP** than in **SRP**: as stated in Ref.[54] that yields to strengthen the intra HPE destabilizing forces arising from the Coulombic interactions among ionic heads and that explains the more elongated conformation of $\mathbf{PSC}^{1}$ in **LRP** than in **SRP**. Regarding $\mathbf{PSC}^{0.5}$ its final conformation in **LRP** is also more elongated than in **SRP** ($R_g = 26$ Å) and we note the formation of two "pearls" at both its extremities and whose cores are stabilized by clusters of neutral phenyl groups. Such a conformation is in line with the interpretation of experimental data provided in Ref.[54]. As for $\mathbf{PSC}^{1}$ a weaker fraction of counter ions interact at short range from carboxylate heads in the final conformation of $\mathbf{PSC}^{0.5}$ in **LRP** than in **SRP**. That also favors its more elongated conformation in **LRP**.

Hence our simulations show long range solvent effects to favor different kind of HPE conformations : for **P70** they reinforce the stability of its CQS conformations whereas for $\mathbf{PSC}^{f_c}$ with a high charge fraction ($f_c \geq 0.5$) they favor elongated conformations. Hence our multi level PPP approach does not systematically favor CQS-like conformations for highly charged HPEs.

3.3 Intra-polymer polarization

3.3.1 SRP medium

To further investigate the role of intra molecular polarization on **P70** conformations solvated in a salt free **SRP** medium at infinite dilution conditions, we performed two sets of 8 independent NVT simulations at the 10 *ns* scale of the **P70**/counter ions system embedded in a 0.9 M first order PPPs cubic box, however by setting to zero all the atomic polarizabilities. For these two simulation sets, their starting structures correspond to those obtained at the end

1. of the initial 200 *ps* segments of the polarizable **P70** simulations in **SRP**, segments along which **P70** is harmonically constrained in its LSS conformation;

2. of the polarizable **P70** simulations in **SRP**, *i.e.* **P70** is in a CQS conformation.

These starting structures are relaxed by performing preliminary 200 *ps* NVT simulations along which the **P70** structure is constrained and corresponding to different starting velocity sets. The simulations are then restarted by removing the constraints on **P70**.

In Fig. 9 we plot the temporal evolutions of the **P70** gyration radius R_g and ratio R^{inertia} along both sets of simulations. As shown by these plots, the **P70** structures converge within less than 8 *ns* towards quasi spherical conformations whose R_g value is almost equal, about 21.5 $\pm$ 1.0 Å, regardless of the simulation starting structure. These final conformations, denoted as QS^{nopol}, all present relatively large cavities within them. Compared to the CQS conformations of polarizable **P70** in **SRP**, they are less compact by a factor 3 in volume, and they correspond more to coil/helicoidal conformations than to compact globular ones, see Fig. 2.

In Fig. 9 we also plot the **P70** $g_{\text{NN}}(r)$ and $N_{\text{Na}}(r)$, the counter ion $g_{\text{ci}}(r)$ and $N_{\text{ci}}(r)$, and the solvent $g_{\text{PPP}}(r)$ and $N_{\text{PPP}}(r)$ functions computed along the last 2 *ns* segments of the non-polarizable simulations. Regardless of the simulation starting

structure, the main features of the $g_{\mathrm{NN}}(r)$ and $N_{\mathrm{N_a}}(r)$ functions are close. The cationic charge distributions within all the $\mathrm{QS^{nopol}}$ conformations are thus close, suggesting their main energetic properties to be also close (see below). We may also note that, in line with the larger gyration radius of the $\mathrm{QS^{nopol}}$ conformations, the present functions $g_{\mathrm{NN}}(r)$ start to be non zero for inter nitrogen $\mathrm{N_a}$ distances r larger by about 1.5 Å compared to their polarizable CQS counter parts, and they vanish for distances r larger than 45 Å, *i.e.* 10 Å farther than for the polarizable CQSs.

The counter ion $g_{\mathrm{ci}}(r)$ and $N_{\mathrm{ci}}(r)$ functions are also close, regardless of the starting simulation structure. Contrary to the **SRP** simulations of polarizable **P70**, a large fraction of the counter ions (about 20 %) is now adsorbed within the polymer core and the remaining ones form a spherical shell centered at about 28 Å from PCOM and extending up to about 18 Å from the **P70** 'surface' (*i.e.* the surface of a sphere whose radius is R_g). Lastly, the solvent functions $g_{\mathrm{PPP}}(r)$ show the first level PPP density to be almost constant and to already reach 90% of its bulk value within the **P70** core domain. That yields the hydration index (defined as the mean number of first order PPPs lying within the **P70** core) of the $\mathrm{QS^{nopol}}$ conformations to be 4 times larger than for CQS conformations (about 1200 and 270, respectively). Hence the large cavities within the $\mathrm{QS^{nopol}}$ conformations are full filled by counter ions and first level PPPs.

As expected from the close charge distributions within the $\mathrm{QS^{nopol}}$ conformations and from the close counter ion and PPP structural properties, there are weak dispersions among the values of the mean energy components $\Delta\bar{U}_{\mathrm{p}}$, $\Delta\bar{U}_{\mathrm{ci}}$, $\Delta\bar{U}_{\mathrm{pc}}$, $\Delta\bar{U}_{\mathrm{ps}}$ and $\Delta\bar{U}_{\mathrm{cs}}$ computed along all the final 2 *ns* non-polarizable simulation segments. Compared to polarizable data, the less compact $\mathrm{QS^{nopol}}$ conformations are tied to twice weaker intra-**P70** Coulombic and **P70**/PPPs interaction energies, see Tab. 1. According to Eqs. (9-10), the latter results may be interpreted as arising from the more diluted cationic charge within the less compact $\mathrm{QS^{nopol}}$ conformations. Interestingly the $\mathrm{QS^{nopol}}$ magnitudes of the mean intra counter ion cloud $\Delta\bar{U}_{\mathrm{ci}}$, **P70**/counter ions

$\Delta\bar{U}_{\rm pc}$ and counter ions/solvent energies $\Delta\bar{U}_{\rm cs}$ agree with the CQS one.

3.3.2 LRP medium

We performed a new set of 8 independent NVT simulations in **LRP** whose starting **P70**/counter ions structures correspond to the final ones of the non-polarizable simulations in **SRP**. However these structures are embedded in two 0.9M first and second level PPP cubic boxes and they are simulated at the 20 *ns* scale. As above all the **P70** atomic polarizabilities are zeroed. The main **P70**, counter ions and solvent structural and energetic data, averaged over the last 2 *ns* simulation segments, are reported in Fig. 10 and in Tab. 1.

Accounting for second level PPPs, and thus for long range solvent electrostatic effects, yields again (1) a large fraction of the counter ions to be largely repelled away from **P70** (they form again a distant shell lying at about 50-60 Å from PCOM), and (2) a strong contraction of the starting $\rm QS^{nopol}$ conformations : their mean gyration radius R_g decrease from 21.5 ± 1 Å in **SRP** towards 14.2 ± 0.2 Å in **LRP**. The structural details of these new **LRP** compact and quasi spherical conformations, denoted as $\rm CQS^{nopol}$, are overall close to the CQS ones corresponding to polarizable **P70** in **SRP**, see Fig. 10(b). However, regarding the counter ions, the noticeable fraction of them (about 20%) that were trapped within the starting $\rm QS^{nopol}$ conformations are still trapped within the final $\rm CQS^{nopol}$ ones. All these results may be interpreted as above for the polarizable **P70** embedded in the **LRP** medium. However the large fraction of the counter ions trapped within the $\rm CQS^{nopol}$ final conformations are responsible for a strong HPE discharging effect that stabilizes these conformations as intra-**P70** polarization attractive forces favor CQS conformations.

As for polarizable **P70**, the **P70**/counter ion radial distribution functions reported in Fig. 10 show here also the existence of a distant counter ion shell lying at 60 Å from PCOM. As discussed for polarizable **P70**, the origin of that shell arises here also from over concentration of second order PPPs in between that shell and

P70. As it will be discussed in Section 4, reinforcing the short repulsion among second order PPPs will yield that distant shell to vanish. However that will not lead to invalidate the present conclusions about the contraction of the non polarizable **P70** conformations arising from long range solvent effects.

At the exception of polarization and counter ion energy data, the energetic data reported in Tab. 1 for non-polarizable **P70** in **LRP** are in line with their polarizable **P70** counter parts. In particular we note again the strong stabilization of the CQS^{nopol} conformations arising from long range solvent electrostatic effects. Regarding counter ions, long range solvent electrostatic effects favors their structure as for polarizable **P70**, but more strongly as suggested by the magnitude of the energy component $\Delta\bar{U}_{cs}$. As that component only measures the strength of the counter ion/solvent interaction, its strong intensity here may result from the large fraction of counter ions trapped within the overall small volume corresponding to CQS^{nopol} conformations, as suggested by Eqs. 9 and 10.

Besides confirming the role of electrostatic long range solvent effects, our non-polarizable **P70** simulations clearly show the strong effect of intra-molecular polarization on the conformation of a HPE in liquid water-like **SRP** and **LRP** media. According to our modeling framework, these intra-molecular polarization effects are not tied to the strongest interaction energies within our HPE/counter ions/solvent systems, but they are a key factor favoring HPE compact globular conformations.

4 HPE collapse and the solvent coarse grained approach

Coarse-graining the solvent yields approximations in modeling solute/solvent interactions. As even conclusions from all atom simulations regarding the association of standard ion pairs in aqueous phase vary noticeably when considering different sets of all atom force field parameters[71–73], one may wonder about the reliability of

simulations based on a solvent coarse grained approach to investigate the behavior of a heavily charged HPE like **P70**.

In an earlier study[37], we showed the $\mathbf{PPP}^l$ approach to meet the predictions of the lattice numerical method implemented in the APBS code[74] to solve the Poisson Boltzman equations as computing the electrostatic component of the hydration free energy of proteins, a particular kind of polyelectrolytes. Here we further investigate the $\mathbf{PPP}^l$ accuracy by performing an additional set of eight simulations of **P70** embedded in 1M first level PPP box. These simulations correspond to polarizability values α_s of the PPPs that range from 1.2 up 2.7 Å^3 (the α_s value computed from Equation (4) is 2.35 Å^3). Each simulation was performed as detailed above up to 10 ns from the **P70** linear starting structure. In Fig. 11 we plot the mean **P70** gyration radius R_g and the mean **P70**/solvent interaction energy $U_{\rm ps}$ as a function of α_s (these mean quantities are computed over the final simulation segment on which the **P70** structure is stable). From these plots it appears that **P70** collapses within 10 ns towards a CQS conformation as soon as $\alpha_s = 1.8$ Å^3, a value 30% smaller than the one we used in the above simulations. Regarding the $U_{\rm ps}$ values, their magnitude for the **P70** CQS conformation is a decreasing linear function of α_s : it ranges from -20 ($\alpha_s = 1.8$ Å^3) down to -39 ($\alpha_s = 2.7$ Å^3) 10^3 kcal mol^{-1}.

We simulated **P70** in its CQS conformation as embedded in a spherical cavity whose radius is 43 Å and comprising about 7 500 explicit and polarizable water molecules. **P70** and water were simulated using an all atom polarizable force field whose accuracy is discussed in Refs.[56,75] (see also Supplementary Material for details). That all atom simulation yields a $U_{\rm ps}$ value that ranges from -26 to -28 10^3 kcal mol^{-1} as accounting for water molecules lying within a shell extending up to 12 Å from the **P70** boundary (as defined in Section 2.3). Such a $U_{\rm ps}$ value is fully in line with our $\mathbf{PPP}^l$ estimate for $\alpha_s = 2.35$ Å^3.

Defining HPE structures by analogy with proteins, the secondary structure of **P70** for low values α_s (up to 1.8 Å^3) corresponds to long helices where the ammo-

nium groups point towards the solvent and they interact at short range with the counter ions. These helices can adopt neck-lace, wormlike or more compact folded conformations (see Supplementary Material) showing the wide range of potential **P70** tertiary structures in low to moderately polarizable solvents. For instance the **P70** tertiary structure along our simulations is a neck lace for $\alpha_s = 1.2$ Å^3 and wormlike for $\alpha_s = 1.4$ Å^3 (that explains their respective gyration radius values in Fig. 11). However both kind of structures are very close in energy, regardless of the latter α_s values.

For large α_s values, the helix secondary structure vanishes and **P70** collapses towards more disordered, compact and quasi spherical conformations as discussed in the above section. Around these CQS conformations we note the counter ions to form a more and more diffuse cloud that extends farther and farther from the **P70** boundary as α_s increases. Whereas salts were considered as fully dissociated in water at low to moderate concentrations, large salt/water aggregates are now inferred to be abundant even in dilute solutions[76–79]. The stability of such edifices may be explained from similar arguments as those schematized in Fig. 1(b) : the stability of polarizable water molecules intercalated between two species of opposite charges is reinforced by the resulting non zero electric field generated by the latter charges on them. These effects are responsible for the stability of the well known Solvent Separated Ion Pair, SSIP, conformations that are pivotal in the ion pair dissociation process in aqueous phase (see Ref.[80] and the references mentioned therein). Our hybrid simulations show SSIP-like **P70**/counter ions to be more and more favored as α_s increases.

Lastly we have to take in mind that our simulations are performed in infinite dilution conditions. The HPE solutions that are experimentally investigated comprise an overall large number of HPE chains, each of them generating local electric fields on the solvent. Even if these fields undergo large fluctuations, they will on average less polarize the solvent located in between two HPE chains than the solvent lying

at the vicinity of a single and isolated HPE chain as in the present study. That arises form the electric field cancellation effect detailed in Fig. 1(b) (that effect is thus opposite to the above one favoring SSIP-like conformations). Hence within our $\mathbf{PPP}^l$ frame work, increasing the concentration of HPE is expected to weaken HPE/solvent interactions and thus to destabilize HPE CQS conformations and to favor more elongated HPE ones.

Regarding second order PPPs and the magnitude of the ϵ_2 parameter modulating the intensity of the LJ interactions among them, we performed two new series of eight MD simulations by considering **P70** as a polarizable and a non polarizable entity. Instead of using a parameter ϵ_2 one order of magnitude smaller than the one used for first level PPPs (*i.e.* $\epsilon_2 = 0.05$ kcal mol^{-1}), we set ϵ_2 to a value ranging from 1 to 8 kcal mol^{-1} to perform these new simulations. Besides reinforcing the interactions among second order PPPs, these larger ϵ_2 values also noticeably reinforce the strength of the short range repulsion among second order PPPs (see Supporting Material). That will thus prevent (at least in part) over concentration of second order PPPs at the vicinity of charged solutes. The **P70** starting structure (and its counter ion cloud) correspond to one of the final CQS conformations of a polarizable simulation in **SRP**. That structure is embedded in a first order PPP box whose dimension is 200 Å. That yields a system comprising about 260k first order PPPs that is embedded in a second order PPP box comprising also about 260k particles. The simulations were performed as detailed above up to 30 ns. The ion/**P70** PCOM radial distribution functions $g_{\mathrm{ci}}(r)$ and their integrals $N_{\mathrm{ci}}(r)$ are provided as Supplementary Material. These functions show a large fraction of the counter ions to not be any more organized as a distant shell lying at about 60 Å from the **P70** PCOM, but rather to belong to a diffuse counter ion cloud extending within the entire simulation box. That behavior is observed for any value of $\epsilon_2 \geq 1$ kcal mol^{-1}. Regarding polarizable **P70**, its CQS conformation is stable and still more contracted as compared to simulations in **SRP** (the **P70** gyration radii at

the end of these new simulations are weaker by 0.6 to 1.5 Å than in **SRP**). For non polarizable **P70**, its structure is also more contracted than in **SRP** (the **P70** gyration radii range from 15 to 18 Å). For both polarizable and non polarizable **P70**, we do not observe a dependence of their gyration radius on the magnitude of $\epsilon_2 \geq 1$ kcal mol^{-1}. Hence, even if there is room to improve the modeling of inter particle interactions for high order PPPs, all our results show solvent long range effects (as modeled by our second order PPPs) to reinforce the stability of the **P70** CQS conformations and to favor more contracted CQS conformations than do only short range solvent effects.

In all the above discussions suggest our **PPP**^{l} approach to capture the main effects regarding HPE hydration. In particular our hybrid approach does not overestimate the solute/solvent interactions that stabilize CQS-like HPE conformations in the present study. It is thus a priori able to provide reliable quantitative conclusions regarding the behavior of HPEs in aqueous phase. However and contrary to our original study[36] we recommend the use of large LJ ϵ_k parameters (*i.e.* $\epsilon_k \geq 1$ kcal mol^{-1}) to model the interactions among PPPs whose order k is higher than one.

5 Conclusion

We presented MD simulation results about the HPE **P70** solvated in salt free liquid water-like **SRP** and **LRP** media, in presence of counter ions and at infinite dilution conditions. The **SRP** and **LRP** media are modeled by means of polarizable particles PPPs according to the multi-level CG **PPP**^{l} approach. In the **SRP** medium, only PPPs whose size matches that of a single water molecule are taken into account and they interact with the solutes if they lie within a shell extending to no more than 12 Å from any solute non hydrogen atom. For the **LRP** medium, we take into account these first level PPPs as well as a new set of larger PPPs (whose volume

is 8 times larger than that of a single water molecule) that interact only with the HPE charged groups and the counter ions if they lie within 12 to 143 Å from the solutes. The two regimes behavior of the free energy corresponding to the creation of an empty cavity within liquid water (as predicted by to the Lum-Chandler-Weeks theory of hydrophobicity[62]) is reproduced by these two media. The properties of the **LRP** medium are close to the liquid water one[36] whereas the **SRP** medium may be interpreted as the solvent water confined in a cavity. The HPE and its counter ion cloud are taken into account explicitly and modeled according to the *all atom* polarizable force field TCPEp approach[31].

From a **P70** linear starting structure surrounded by a thin layer of counter ions, we performed multiple simulations of that system embedded in the **SRP** medium modeled as a cubic box comprising about 0.9 M first level PPPs (and whose volume is 27 nm^3). Our simulations show **P70** to collapse in that medium towards compact and quasi-spherical conformations surrounded by a counter ion shell lying a short range (about 8-10 Å) from the **P70** surface. Even if the atomic details of these conformations differ among the simulations, the spatial distributions of their cationic groups are close suggesting these conformations to be close in energy in **SRP** as supported by the analysis of different energy components performed along the simulations, like intra-**P70** and **P70**/counter ions interaction energies. The **P70** compact and quasi spherical conformations relaxed in the **SRP** medium were then simulated in the **LRP** medium by embedding the final structures of the **P70**/counter ions/0.9 M first level PPPs systems in a new cubic box comprising 0.9 M second level PPPs and whose volume is 216 nm^3. That new system accounts thus for an equivalent of 8M water molecules. The **SRP** compact and quasi spherical conformations are structurally stable in **LRP**, even if we note a noticeable contraction of them (by about 20 % in volume) as well as a large fraction of the counter ions to be repelled at farther distances from **P70** than in **SRP**. Both the latter results arise from long range solvent electrostatic effects that weakens the association of oppositely charged

systems and favors the contraction of heavy charged flexible solutes.

For comparison purposes we also simulated at the 100 ns scale four carboxylated polystyrenes $\mathbf{PSC}^{f_c}$ (whose fraction of charge varies from 0 to 1) in both the **SRP** and the **LRP** medium by means of our MD protocol. These HPEs are shown by our simulations to behave as the sulfonated polystyrenes experimentally investigated by Boué and co workers[11,54] : their gyration radius increases as their charge fraction increases. In particular and contrary to our **P70**, the fully carboxylated $\mathbf{PSC}^1$ evolves along a 100 ns MD trajectory towards a helicoidal structure in **SRP** (structure that is stabilized by local carboxylate/counter ion salt bridges) and towards a more elongated structure in **LRP**.

Even if the behavior of **P70** in our water-like **SRP** and **LRP** media does not prejudge of the behavior of any kind of HPEs in such environments (in particular their collapse towards compact conformations as discussed for carboxylated polystyrenes $\mathbf{PSC}^{f_c}$ with a high charge fraction) we may draw general conclusions from our simulations. The behavior of a HPE in any solvent is governed by a complex interplay among microscopic interactions occurring within it and between the HPE and its environment (counter ions and solvent). In particular, in salt free solutions and at infinite dilution conditions, our simulations show that

1. in line with the standard picture of the effects governing the behavior of a HPE in solution (see Ref.[81] for instance), HPE compact conformations are disfavored by strong Coulombic intra HPE effects and favored by strong Coulombic HPE/counter ions ones, especially in a **SRP** medium in which the association of oppositely charged ions is more favored than in **LRP**;

2. solute/solvent electrostatic interactions (as described within the $\mathbf{PPP}^l$ framework) and particularly long range ones, favor HPE compact conformations : that yields the reinforcement of the HPE local density of charge which in turn strengthens the electrostatic response of a polarizable solvent like water to the

HPE presence;

3. the existence of strong intra HPE attractive polarization effects (arising from the polarization of alkyl carbons by the HPE ionic charges, for instance) also favors HPE compact conformations.

We also simulated a non-polarizable **P70** in both the **SRP** and **LRP** media. These new simulations support the key role played by intra HPE polarization effects to understand the stability of HPE compact globular conformations as in the particular case of **P70**. Not accounting for them in our simulations yields **P70** to adopt less compact conformations : their gyration radii increase by a factor of about 50% (**SRP**) and 20% (**LRP**) compared to the compact structures predicted by accounting for these polarization effects. In the present case, intra HPE polarization effects correspond to attractive forces. They may thus be modeled using a basic additive dispersion-like energy term. However polarization is a non-linear phenomenon that can be attractive, repulsive and even vanish for particular molecular geometries, and might be extremely relevant for studying Natural ingredients like polysaccharides as already shown for polypeptides. Hence microscopic polarization has to be explicitly taken into account not only to model HPE systems at infinite dilution conditions (as in the present study) but also in more complex and concentrated chemical environments for which it is not obvious to infer the role of polarization (like to favor/disfavor short range interactions among HPEs for instance). Our works emphasizes the importance of considering polarizability for the future of digitally-enabled formulation for Green Sciences. Lastly our hybrid modeling approach is well suited to be used in conjunction with efficient algorithms to simulate explicitly very large solutes (or mixtures of solutes), like a Fast Multipole Method devoted to induced dipole-based polarizable force fields.[82]

Acknowledgments

This work was granted access to the TGCC HPC resources under the allocation 2019-2020 [x2016081859] and the Grand Challenge allocation [GC0429] made by GENCI.

Data Availability

Data openly available in a public repository that does not issue DOIs. A full binary version of our simulation code POLARIS(MD) (version 1.2.0.1) and the input files that were used for the present study to perform simulations on a multi core processor are freely available for download[64].

Supplementary Material

See supplementary material for the details regarding (1) the all atom simulation of **P70** embedded in an aqueous cavity and the its final conformations as a function of the solvent polarizability; (2) the simulation data of the $\mathbf{PSC}^{f_c}$ in the **SRP** and **LRP** media; (3) the simulation data of **P70** in **LRP** using a LJ parameter ϵ_2 ranging from 1 to 8 kcal mol^{-1}. A short video is also provided that shows the collapse of the polyelectrolyte along a simulation in the 1M first level pseudo particle box.

Conflicts of interest

A. C. and F. L. are full employees of L'Oréal involved in research activities.

Table 1: Decomposition in Coulombic $\Delta\bar{U}^{qq}$, polarization $\Delta\bar{U}^{pol}$ and non electrostatic $\Delta\bar{U}^{ne}$ energy components of the intra-**P70** and intra-counter ion cloud mean energies $\Delta\bar{U}_{\rm p}$ and $\Delta\bar{U}_{\rm ci}$ and of the **P70**/counter ions, **P70**/PPPs and counter ions/PPPs interaction energies $\Delta\bar{U}_{\rm pc}$, $\Delta\bar{U}_{\rm ps}$ and $\Delta\bar{U}_{\rm cs}$. Data in italic correspond to the root mean square deviations of these values among the independent simulations. For **LRP**, the $\Delta\bar{U}_{\rm ps}$ and $\Delta\bar{U}_{\rm cs}$ data in brackets corresponds to the first level PPP contributions. The non electrostatic component $\Delta\bar{U}^{ne}$ corresponds to the energy terms U^{disp}/U^{coop} or to the PPP/solute Lennard-Jones-like term (see Section 2). All values are expressed in 10^3 kcal mol^{-1}. Values smaller than 100 kcal mol^{-1} are not reported.

	$\Delta\bar{U}^{qq}$	$\Delta\bar{U}^{pol}$	$\Delta\bar{U}^{ne}$
SRP	medium and polarizable **P70**		
$\Delta\bar{U}_{\rm p}$	23.9 *0.4*	-2.6 *0.1*	-0.3 *0.02*
$\Delta\bar{U}_{\rm ci}$	9.8 *0.4*	-0.1 *0.02*	
$\Delta\bar{U}_{\rm pc}$	-24.1 *0.6*		0.1 *0.05*
$\Delta\bar{U}_{\rm ps}$		-11.6 *0.3*	0.1 *0.05*
$\Delta\bar{U}_{\rm cs}$		-3.2 *0.1*	
LRP	medium and polarizable **P70**		
$\Delta\bar{U}_{\rm p}$	30.2 *0.6*	-4.6 *0.3*	-0.8 *0.1*
$\Delta\bar{U}_{\rm ci}$	2.4 *1.3*		0.1 *0.05*
$\Delta\bar{U}_{\rm pc}$	-2.8 *1.6*	0.3 *0.1*	
$\Delta\bar{U}_{\rm ps}$		-52.8 *2* [-15 *0.3*]	0.6 *11*
$\Delta\bar{U}_{\rm cs}$		-13.8 *14* [-9.3 *1.8*]	
SRP	medium and non polarizable **P70**		
$\Delta\bar{U}_{\rm p}$	11.9 *0.3*		
$\Delta\bar{U}_{\rm ci}$	10.4 *0.3*		
$\Delta\bar{U}_{\rm pc}$	-23.8 *0.7*		
$\Delta\bar{U}_{\rm ps}$		-5.1 *0.2*	-0.1
$\Delta\bar{U}_{\rm cs}$		-3.9 *0.1*	
LRP	medium and non polarizable **P70**		
$\Delta\bar{U}_{\rm p}$	30.3 *0.5*		-0.5 *0.1*
$\Delta\bar{U}_{\rm ci}$	14.5 *4.8*		
$\Delta\bar{U}_{\rm pc}$	-5.8 *3.2*	0.1 *0.02*	-0.1 *0.02*
$\Delta\bar{U}_{\rm ps}$		-52.5 *1.5* [-10.3 *0.3*]	0.8 *0.1*
$\Delta\bar{U}_{\rm cs}$		-31.8 *7.7* [-7.7*1.9*]	0.2 *0.01*

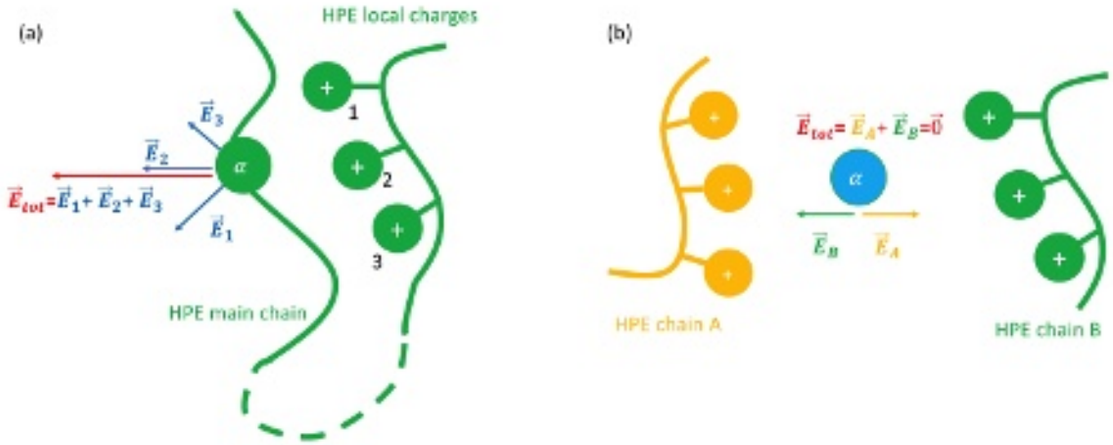


Figure 1: Microscopic polarization effects. (a) An atom, whose isotropic polarizability is α, belonging to a HPE (in green) is located at the vicinity of three positively charged groups. All the other atoms are non-polarizable centers. That atom undergoes the total electric field $\mathbf{E}_{tot} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3$ arising from the three positive charges. From an induced dipole moment approach, the dipole moment generated on that atom obeys $\alpha\mathbf{E}_{tot}$ and it is at the origin of non additive polarization energy component $-\frac{1}{2}\alpha\mathbf{E}_{tot}^2$ that can not be modeled using standard additive potentials. (b) A solvent molecule whose polarizability is α is located in between two HPE chains A and B. The electric fields arising from each chain cancel out. If the charges of chain A are opposite to those of chain B, then the resulting electric field on the solvent is $\mathbf{E}_{tot} = \mathbf{E}_A + \mathbf{E}_B = 2\mathbf{E}_A = 2\mathbf{E}_B \neq \mathbf{0}$, and the corresponding polarization energy on the solvent is four times greater than as arising from each single chain taken alone.

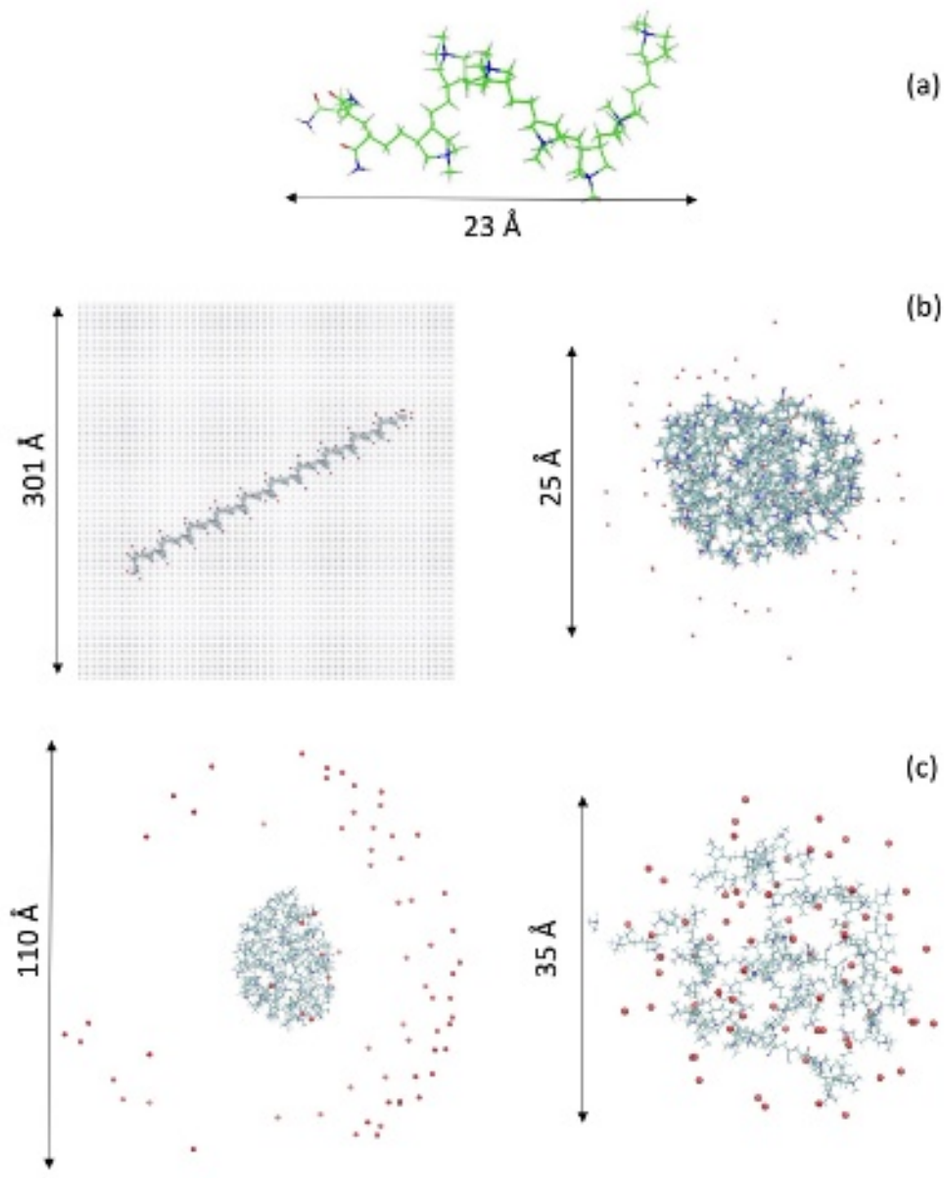


Figure 2: (a) A three dimensional representation of the main **P70** unit of the modeled, generic, HPE copolymer. The HPE is made of 10 units each comprising 7 adjacent cyclo butyl ammonium cationic heads interconnect by ethyl groups, that are connected to 3 more adjacent neutral groups equivalent to a triade of asparagine proteinic residues[38–42]. The HPE total charge is thus +70 e. Here the carbon, hydrogen, nitrogen and oxygen atoms are shown in light blue, grey, dark blue and red. (b) Left : **P70** in its LSS conformation embedded in a 0.9M PPPs cubic box and whose electrostatic charge is neutralized by 70 Cl^- counter ions (in red, PPPs are shown in blue). Right : a CQS conformation of **P70** in the **SRP** medium. (c) Left : a CQS conformation of **P70** in the **LRP** medium. Right : an inflated conformation of the non polarizable **P70** in the **SRP** medium.

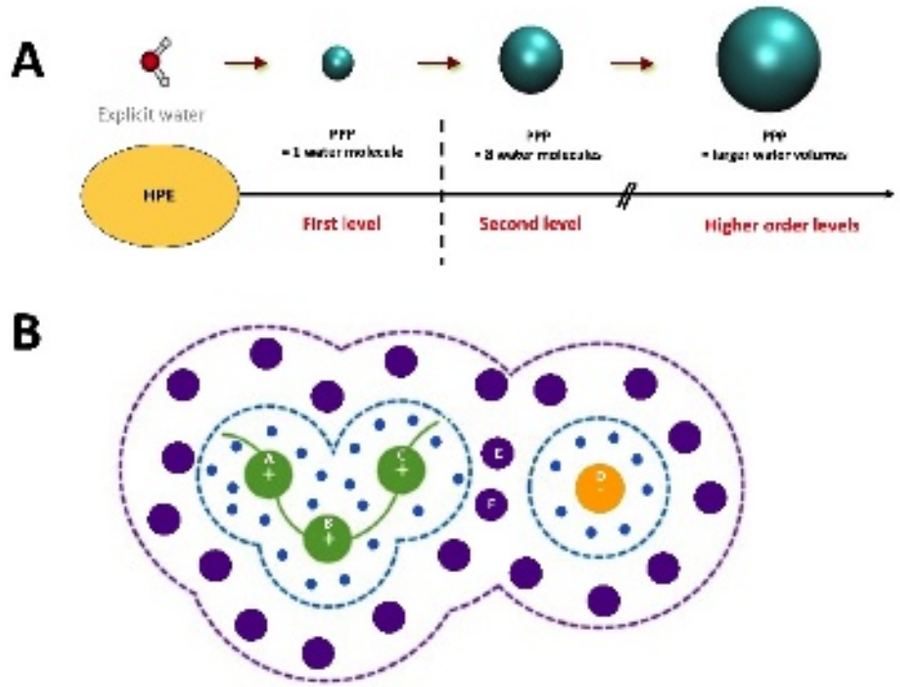


Figure 3: A The multi-level solvent CG approach $\mathbf{PPP}^l$. The blue spheres are PPP particles that differ by their size and thus by their polarizability. The vertical line correspond to the cutoff distance $R_{\text{cut}}^{\text{pol}}$ mentioned in the text. Smooth transitions exist between the different PPP level shells, *i.e.* the PPP/HPE interactions are scaled by B-spline functions that smoothly vary between 1 and 0 (lower PPP level) and between 0 and 1 (higher PPP level) at solvent shell boundaries. B Example of a HPE (in green) interacting with a counter ion D (in orange) solvated by first (in blue) and second order (in violet) PPPs. The dashed lines defined the first and second order shell-based cut off domains. Any first level PPP located within the domain delimited by the blue dashed lines undergoes the total electric field generated by all the HPE and counter ion charges (denoted here as A, B, C and D). Any second level PPPs located within the domain delimited by the blue and the violet dashed lines also undergoes the latter electric field.

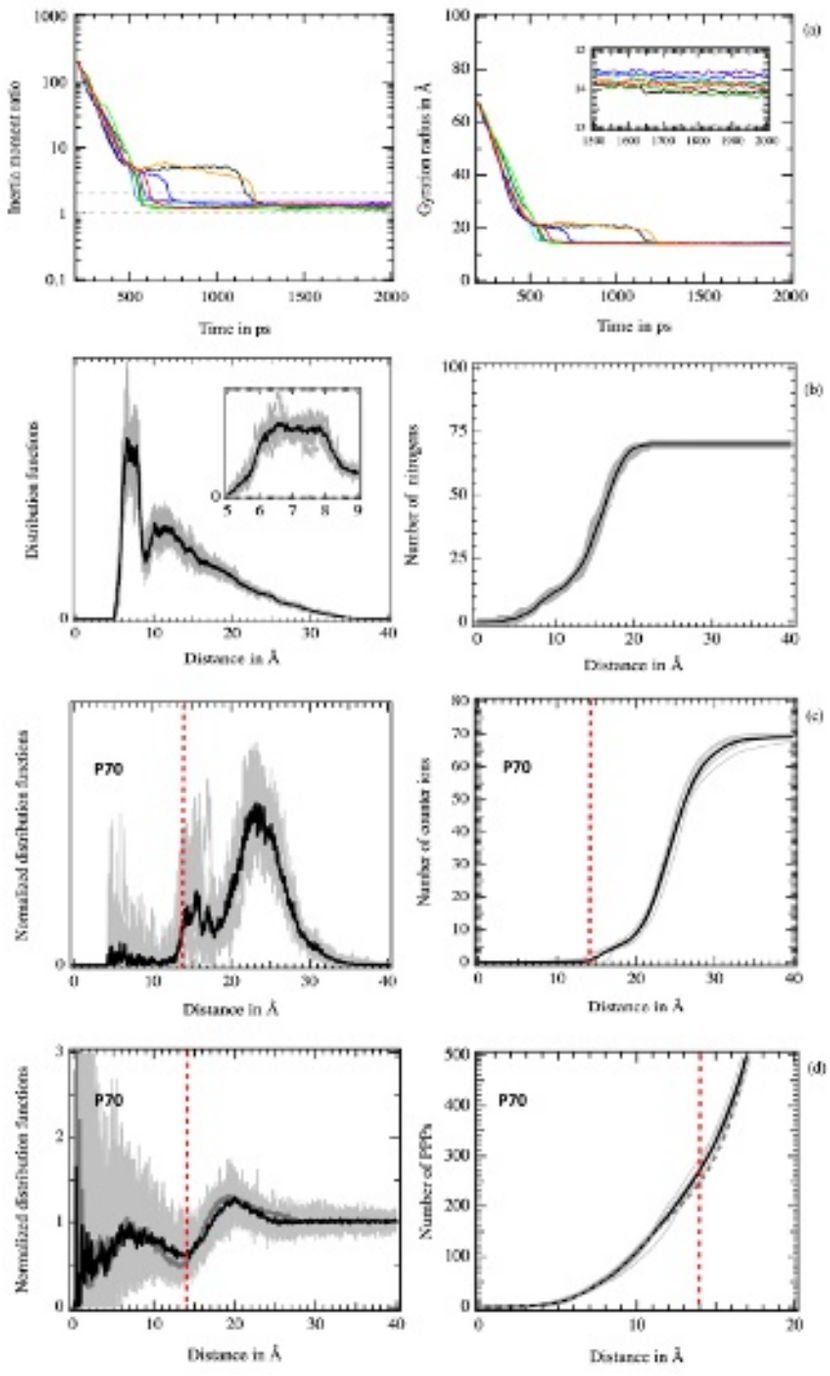


Figure 4: (a) Right : temporal evolution of the ratio between the smallest and largest inertia moment values (left) and of the gyration radius (right) of **P70** along the independent simulations in the **SRP** medium. (b) Mean radial pair distribution function $\bar{g}_{\mathrm{NN}}$ (left, the inset shows the details of its first peak) and mean number $\bar{N}(\mathrm{N_a})$ (right). (c) Mean radial distribution function $\bar{g}_{\mathrm{ci}}(r)$ (left) and its corresponding integral $N_{\mathrm{ci}}(r)$ (right). (d) Mean radial PPP/PCOM distribution function (left) and its corresponding integral (right). From (b) and (d) : the data from each independent simulation are shown by grey lines. The vertical dashed red line is located at the converged **P70** mean gyration radius.

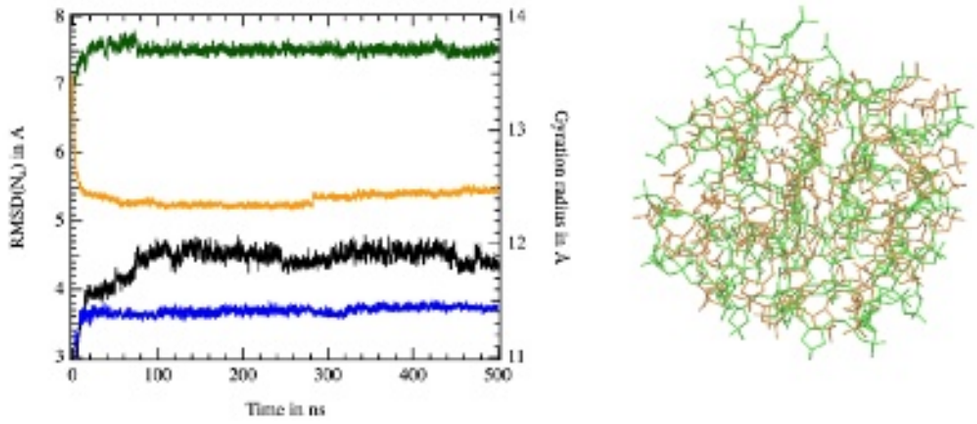


Figure 5: Left : temporal evolution of the **P70** RMSD(N_a) (left axis) and gyration radius (right axis) along the 500 *ns* simulations. Black (**SRP**) and blue (**LRP**) lines : RMSD(N_a) data. Green (**SRP**) and orange (**LRP**) lines : gyration radius data. Right : superposition of the starting (green) and final (orange) **P70** CQS structures from the 500 *ns* **SRP** simulation.

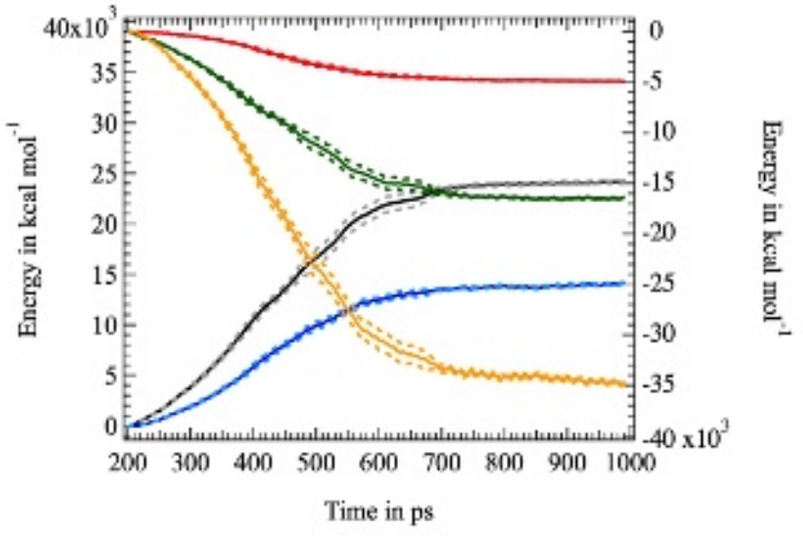


Figure 6: Temporal evolution of the energy component values computed by averaging data from all the independent simulations in the **SRP** medium, in bold lines. In dashed lines, the maximum deviations from the mean data. Black and blue : intra-**P70** and intra-counter ion cloud interaction energies (left axis) ; red, green and yellow : **P70**/PPPs and counter ions/PPPs polarization energies, and **P70**/counter ions interaction energy, respectively (right axis). All the energy components are shifted to be zero at simulation time $t = 200$ ps at which the harmonic constraints preventing **P70** structural transitions are removed.

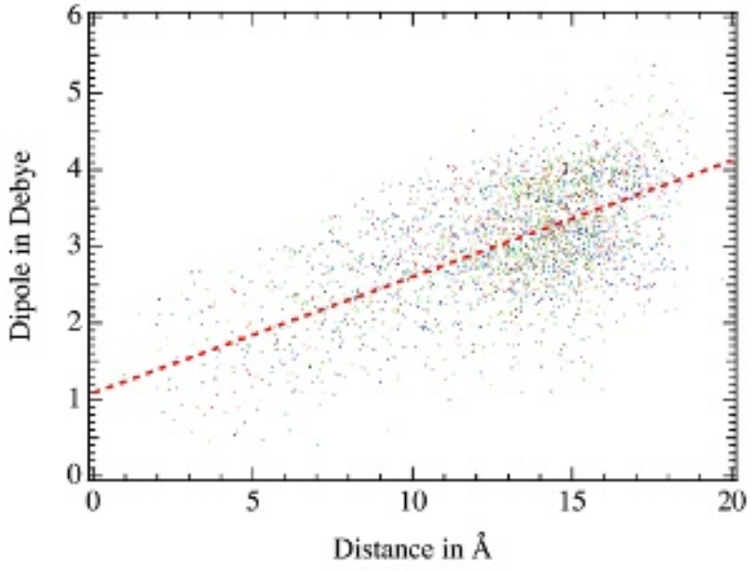


Figure 7: Mean induced dipole moment values of the **P70** alkyl carbons computed along all the independent simulations as a function of the alkyl carbon distance to PCOM. The dashed line is the result of the linear regression fit (the regression coefficient is here of 0.6).

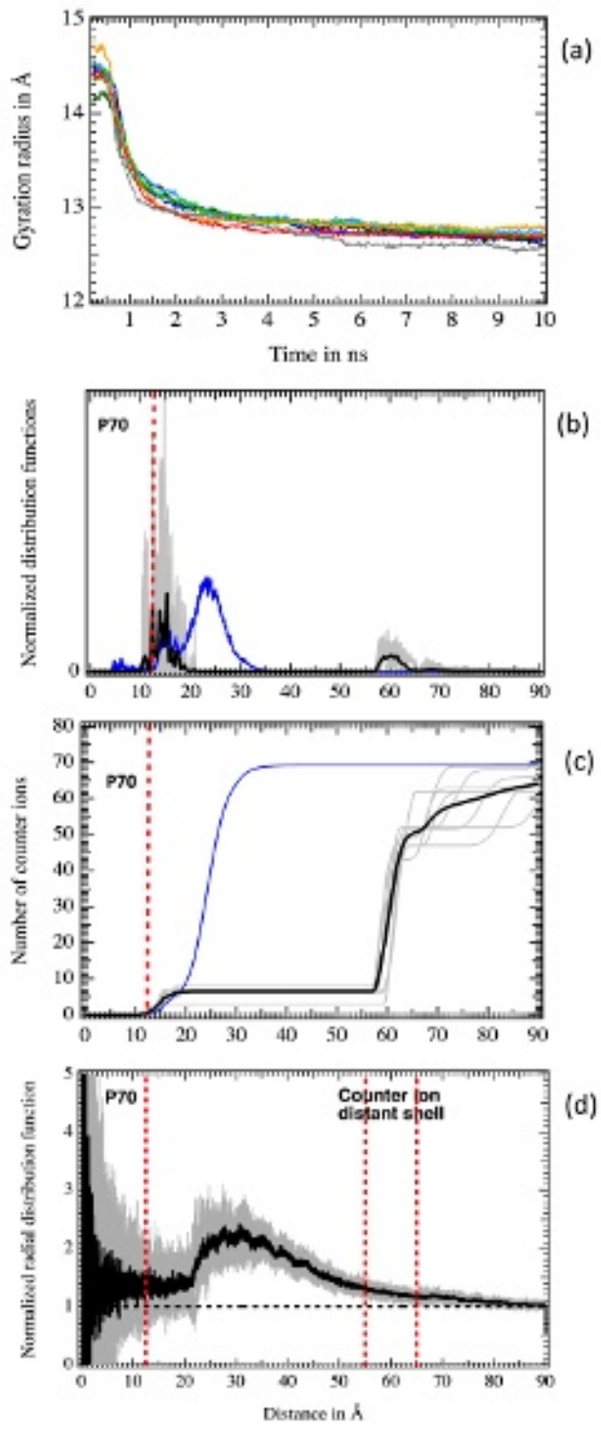


Figure 8: CQS **P70** and counter ion structural data in a **LRP** medium. (a) Temporal evolution of the **P70** gyration radius R_g along the independent simulations. (b) and (c) : counter ion functions $g_{\mathrm{ci}}(r)$ and $N_{\mathrm{ci}}(r)$, respectively (black line : mean functions averaged over all the simulations; grey lines : data from each simulation ; blue thin lines : mean functions from **P70 SRP** simulations; the vertical red line is located at the mean **P70** gyration radius R_g). (d) : **P70** PCOM/second order PPPs radial distribution functions from the independent simulations (in light grey lines) and the corresponding mean function (bold black line).

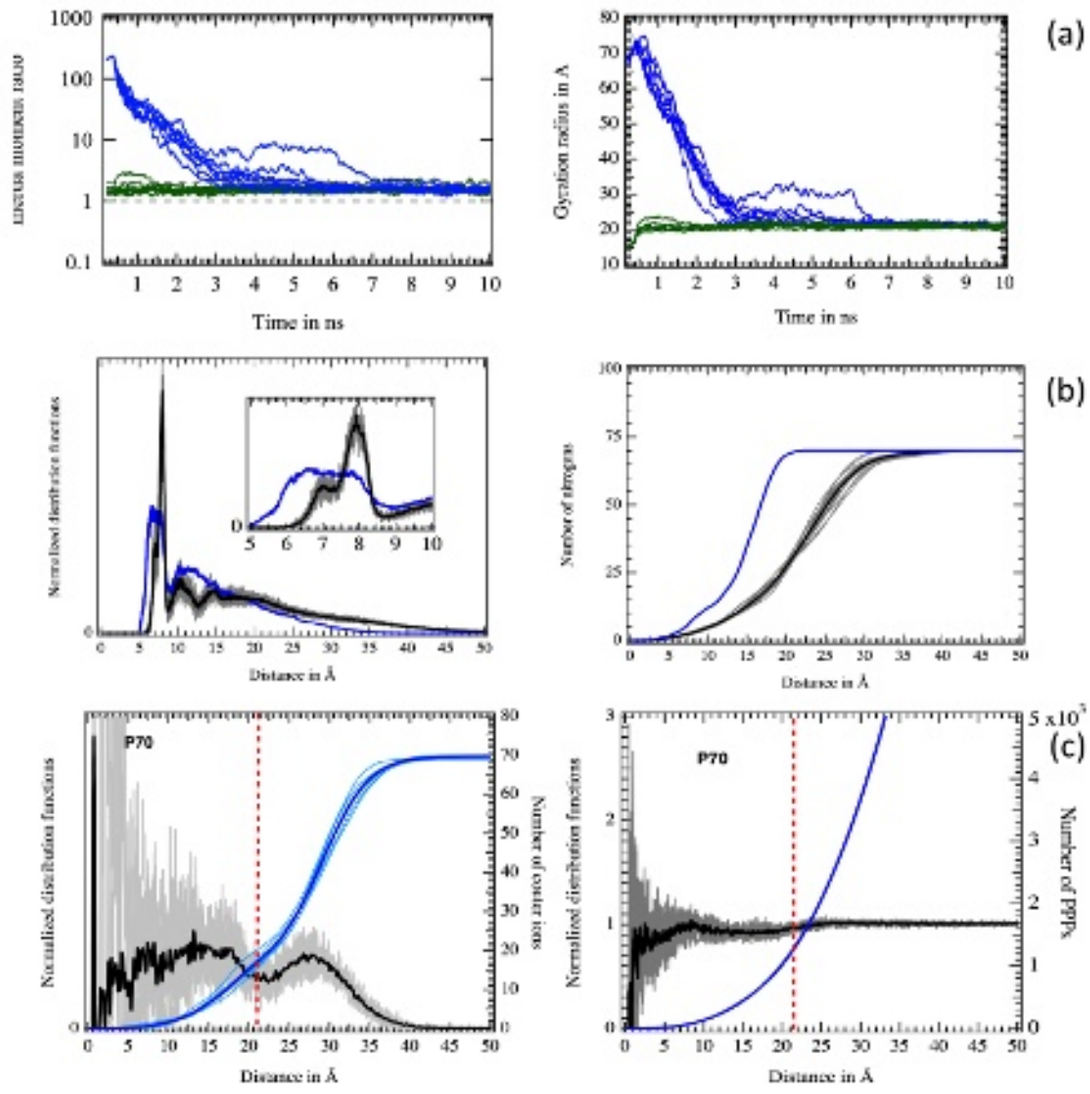


Figure 9: **P70**, counter ion and PPP structural data from non-polarizable **P70** simulations in the **SRP** medium. (a) Temporal evolution of the **P70** ratio R_{inertia} (left) and gyration radius R_g (right). Blue data : from the starting LSS structure; green data : from the starting CQS one. (b) Functions $g_{\text{NN}}(r)$ (left) and $N_{\text{N}_\text{a}}(r)$ (right). Black line : data averaged over all the simulations; grey lines : data from each single simulation ; blue line : mean functions from polarizable **P70** simulations in **SRP**. (c) Left : Counter ion functions $g_{\text{ci}}(r)$ (black) and $N_{\text{ci}}(r)$ (blue), and right : PPP functions $g_{\text{PPP}}(r)$ (black) and $N_{\text{PPP}}(r)$ (blue). Bold lines : data averaged over all the simulations; thin lines : data from each simulation. The vertical red line is located at the mean **P70** gyration radius R_g.

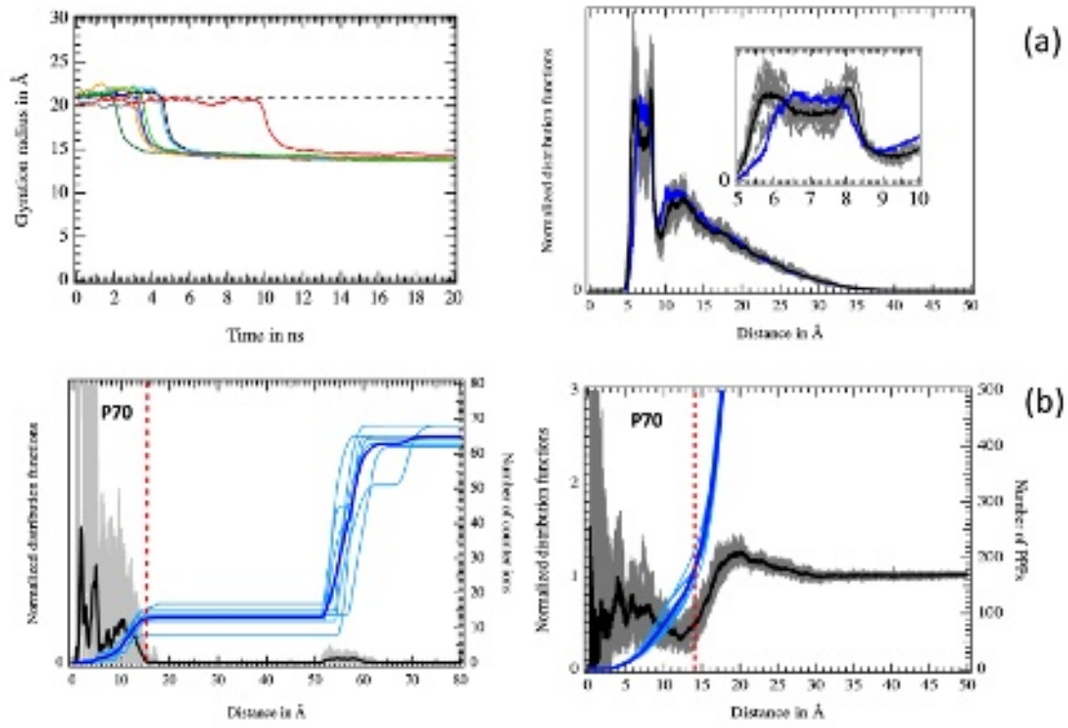


Figure 10: **P70**, counter ion and PPP structural data from non-polarizable **P70** simulations in the **LRP** medium. (a) Left : temporal evolution of the **P70** gyration radius R_g (the horizontal dashed line is located at the non-polarizable **P70** gyration radius converged value in **SRP**). Right : functions $g_{\mathrm{NN}}(r)$ (black line : data averaged over all the **LRP** simulations; grey lines : data from each single **LRP** simulation; blue line : data averaged over the polarizable **P70** simulations in **SRP**). (b) Left : counter ion functions $g_{\mathrm{ci}}(r)$ (black) and $N_{\mathrm{ci}}(r)$ (blue). Right : PPP functions $g_{\mathrm{PPP}}(r)$ (black) and $N_{\mathrm{PPP}}(r)$ (blue). Bold lines : data averaged over all the simulations; thin lines : data from each simulation. The vertical red line is located at the mean **P70** gyration radius R_g.

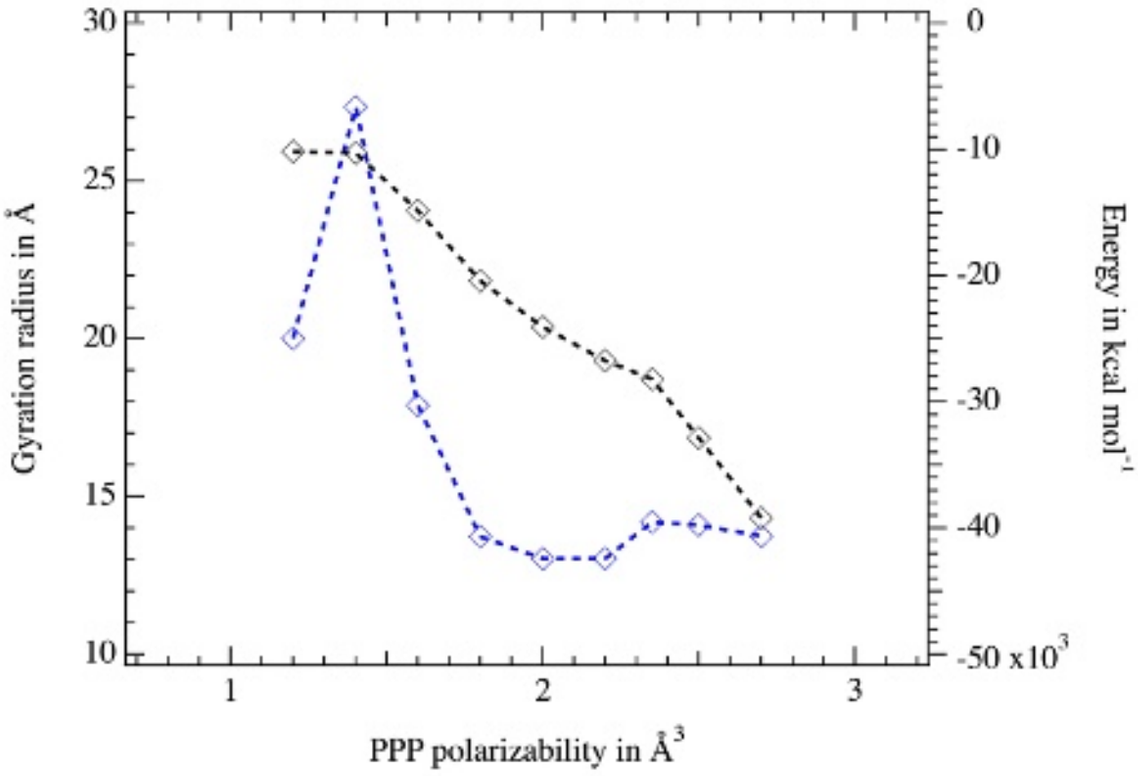


Figure 11: Mean **P70** gyration radius R_g (in blue and left axis) and mean **P70**/solvent interaction energy U_{ps} (in black and right axis) as a function of the PPP polarizability α_s. That polarizability amounts to 2.35 Å^3 according to Equation (4).

References

(1) Fernandez-Pena, L.; Guzman, E.; Léonforte, F.; Serrano-Pueyo, A.; Regulski, K.; Tournier-Couturier, L.; Ortega, F.; R.G., R.; Luengo, G. Effect of molecular structure of eco-friendly glycolipid biosurfactants on the adsorption of hair-care conditioning polymers. *Colloids and Surfaces B: Biointerfaces* **2020**, *185*, 110578.

(2) Manoj Lalwani, S.; Eneh, C. I.; Lutkenhaus, J. L. Emerging Trends in the Dynamics of Polyelectrolyte Complexes. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24157–24177.

(3) Gu, Y.; Zhao, J.; Johnson, J. A. Polymer Networks: From Plastics and Gels to Porous Frameworks. *Angewandte Chemie International Edition* **2020**, *59*, 5022–5049.

(4) Minko, S.; Kiriy, A.; Gorodyska, G.; Stamm, M. Single Flexible Hydrophobic Polyelectrolyte Molecules Adsorbed on Solid Substrate: Transition between a Stretched Chain, Necklace-like Conformation and a Globule. *Journal of the American Chemical Society* **2002**, *124*, 3218–3219.

(5) Kiriy, A.; Gorodyska, G.; Minko, S.; Jaeger, W.; Štěpánek, P.; Stamm, M. Cascade of Coil-Globule Conformational Transitions of Single Flexible Polyelectrolyte Molecules in Poor Solvent. *Journal of the American Chemical Society* **2002**, *124*, 13454–13462.

(6) Kirwan, L. J.; Papastavrou, G.; Borkovec, M.; Behrens, S. H. Imaging the Coil-to-Globule Conformational Transition of a Weak Polyelectrolyte by Tuning the Polyelectrolyte Charge Density. *Nano Letters* **2004**, *4*, 149–152.

(7) Xi, B.; Ran, S.-Y. Formation of DNA Pearl-Necklace Structures on Mica Sur-

face Governed by Kinetics and Thermodynamics. *Journal of Polymer Science Part B: Polymer Physics* **2017**, *55*, 971–979.

(8) Aseyev, V. O.; Klenin, S. I.; Tenhu, H.; Grillo, I.; Geissler, E. Neutron Scattering Studies of the Structure of a Polyelectrolyte Globule in a Water-Acetone Mixture. *Macromolecules* **2001**, *34*, 3706–3709.

(9) Essafi, W.; Spiteri, M.-N.; Williams, C.; Boue, F. Hydrophobic Polyelectrolytes in Better Polar Solvent. Structure and Chain Conformation As Seen by SAXS and SANS. *Macromolecules* **2009**, *42*, 9568–9580.

(10) Essafi, W.; Abdelli, A.; Bouajila, G.; Boué, F. Behavior of Hydrophobic Polyelectrolyte Solution in Mixed Aqueous/Organic Solvents Revealed by Neutron Scattering and Viscosimetry. *The Journal of Physical Chemistry B* **2012**, *116*, 13525–13537.

(11) Ben Mahmoud, S.; Essafi, W.; Brûlet, A.; Boué, F. How Necklace Pearls Evolve in Hydrophobic Polyelectrolyte Chains under Good Solvent Addition: A SANS Study of the Conformation. *Macromolecules* **2018**, *51*, 9259–9275.

(12) Duan, C.; Li, W.; Wang, R. Conformation of a single polyelectrolyte in poor solvents. *The Journal of Chemical Physics* **2020**, *153*, 064901.

(13) Panagiotopoulos, A. Molecular Simulation of Phase Equilibria: Simple, Ionic and Polymeric Fluids. *Fluid Phase Equilibria* **1992**, *76*, 97 – 112.

(14) Escobedo, F. A.; de Pablo, J. J. Molecular Simulation of Polymeric Networks and Gels: Phase Behavior and Swelling. *Physics Reports* **1999**, *318*, 85 – 112.

(15) Reddy, G.; Yethiraj, A. Implicit and Explicit Solvent Models for the Simulation of Dilute Polymer Solutions. *Macromolecules* **2006**, *39*, 8536–8542.

(16) Chang, R.; Yethiraj, A. Dilute Solutions of Strongly Charged Flexible Polyelectrolytes in Poor Solvents: Molecular Dynamics Simulations with Explicit Solvent. *Macromolecules* **2006**, *39*, 821–828.

(17) Chremos, A.; Douglas, J. F. Polyelectrolyte Association and Solvation. *The Journal of Chemical Physics* **2018**, *149*, 163305.

(18) Chremos, A.; Horkay, F. Disappearance of the polyelectrolyte peak in salt-free solutions. *Phys. Rev. E* **2020**, *102*, 012611.

(19) Horkay, F.; Chremos, A.; Douglas, J. F.; L. Jones, R.; Lou, J.; Xia, Y. Systematic investigation of synthetic polyelectrolyte bottlebrush solutions by neutron and dynamic light scattering, osmometry, and molecular dynamics simulation. *The Journal of Chemical Physics* **2020**, *152*, 194904.

(20) Chremos, A.; Douglas, J. F. Polyelectrolyte Association and Solvation. *The Journal of Chemical Physics* **2018**, *149*, 163305.

(21) Mansel, B. W.; Irani, A. H.; Ryan, T. M.; McGillivray, D. J.; Chen, H.-L.; Williams, M. A. K. Resolving Solution Conformations of the Model Semi-Flexible Polyelectrolyte Homogalacturonan Using Molecular Dynamics Simulations and Small-Angle X-ray scattering. *The European Physical Journal E* **2019**, *42*, 19.

(22) Bacle, P.; Jardat, M.; Marry, V.; Mériguet, G.; Batôt, G.; Dahirel, V. Coarse-Grained Models of Aqueous Solutions of Polyelectrolytes: Significance of Explicit Charges. *The Journal of Physical Chemistry B* **2020**, *124*, 288–301.

(23) Mintis, D. G.; Alexiou, T. S.; Mavrantzas, V. G. Effect of pH and Molecular Length on the Structure and Dynamics of Linear and Short-Chain Branched Poly(ethylene imine) in Dilute Solution: Scaling Laws from Detailed Molecular Dynamics Simulations. *The Journal of Physical Chemistry B* **2020**, *124*, 6154–6169.

(24) Jungwirth, P.; Tobias, D. J. Specific Ion Effects at the Air/Water Interface. *Chemical Reviews* **2006**, *106*, 1259–1281.

(25) Tobias, D. J.; Stern, A. C.; Baer, M. D.; Levin, Y.; Mundy, C. J. Simulation and Theory of Ions at Atmospherically Relevant Aqueous Liquid-Air Interfaces. *Annual Review of Physical Chemistry* **2013**, *64*, 339–359.

(26) Houriez, C.; Meot-Ner (Mautner), M.; Masella, M. Solvation of the Guanidinium Ion in Pure Aqueous Environments: A Theoretical Study from an Ab Initio-Based Polarizable Force Field. *The Journal of Physical Chemistry B* **2017**, *121*, 11219–11228.

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

(28) McDaniel, J. G.; Yethiraj, A. Influence of Electronic Polarization on the Structure of Ionic Liquids. *The Journal of Physical Chemistry Letters* **2018**, *9*, 4765–4770.

(29) Bedrov, D.; Piquemal, J.-P.; Borodin, O.; MacKerell, A. D.; Roux, B.; Schräder, C. Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields. *Chemical Reviews* **2019**, *119*, 7940–7995.

(30) Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *The Journal of Physical Chemistry Letters* **2014**, *5*, 3144–3150.

(31) Masella, M.; Cuniasse, P. A Many-Body Model to Study Proteins. I. Applications to MLnm+ Complexes, Mm+=Li+, Na+, K+, Mg2+, Ca2+, and Zn2+, L=H2O, CH3OH, HCONH2, n=1-6, and to Small Hydrogen Bonded Systems. *The Journal of Chemical Physics* **2003**, *119*, 1866–1873.

(32) Rupakheti, C.; Lamoureux, G.; MacKerell, A. D.; Roux, B. Statistical Mechanics of Polarizable Force Fields Based on Classical Drude Oscillators with Dynamical Propagation by the Dual-Thermostat Extended Lagrangian. *The Journal of Chemical Physics* **2020**, *153*, 114108.

(33) Vázquez-Montelongo, E. A.; Vázquez-Cervantes, J. e. E.; Cisneros, G. A. Current Status of AMOEBA-IL: A Multipolar/Polarizable Force Field for Ionic Liquids. *International Journal of Molecular Sciences* **2020**, *21*.

(34) Ha-Duong, T.; Phan, S.; Marchi, M.; Borgis, D. Electrostatic on Particles : Phenomenological and Orientational Density Functional Theory Approach. *The Journal of Chemical Physics* **2002**, *117*, 541–556.

(35) 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.

(36) 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.

(37) 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.

(38) Somasundaran, P.; Chakraborty, S.; Qiang, Q.; Deo, P.; Wang, J.; Zhang, R. Surfactants, Polymers and their Nanoparticles for Personal Care Applications. *J Cosmet Sci.* **2004**, *55*, S1–17.

(39) Cumming, J.; Hawker, D.; Chapman, H.; Nugent, K. *Water, Air and Soil Pollution* **2011**, *216*, 441–450.

(40) Pahal, S.; Gakhar, R.; Raichur, A. M.; Varma, M. M. Polyelectrolyte multilayers for bio-applications: recent advancements. *IET Nanobiotechnology* **2017**, *11*, 903–908.

(41) Besha, A. T.; Tsehaye, M. T.; Aili, D.; Zhang, W.; Tufa, R. A. Design of Monovalent Ion Selective Membranes for Reducing the Impacts of Multivalent Ions in Reverse Electrodialysis. *Membranes* **2019**, *10*, 7.

(42) Beaugeard, V.; Muller, J.; Graillot, A.; Ding, X.; Robin, J.-J.; Monge, S. Acidic Polymeric Sorbents for the Removal of Metallic Pollution in Water: A Review. *Reactive and Functional Polymers* **2020**, *152*, 104599.

(43) Lemkul, J. A.; MacKerell, A. D. Polarizable Force Field for DNA Based on the Classical Drude Oscillator: I. Refinement Using Quantum Mechanical Base Stacking and Conformational Energetics. *Journal of Chemical Theory and Computation* **2017**, *13*, 2053–2071.

(44) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *Journal of Chemical Theory and Computation* **2018**, *14*, 2084–2108.

(45) Inakollu, V. S.; Geerke, D. P.; Rowley, C. N.; Yu, H. Polarisable Force Fields: What do They Add in Biomolecular Simulations? *Current Opinion in Structural Biology* **2020**, *61*, 182 – 190, Theory and Simulation Macromolecular Assemblies.

(46) Singh, A. N.; Yethiraj, A. Driving Force for the Complexation of Charged Polypeptides. *The Journal of Physical Chemistry B* **2020**, *124*, 1285–1292.

(47) Sadman, K.; Wang, Q.; Chen, Y.; Keshavarz, B.; Jiang, Z.; Shull, K. R. Influence of Hydrophobicity on Polyelectrolyte Complexation. *Macromolecules* **2017**, *50*, 9417–9426.

(48) Lopez, C. G.; Colby, R. H.; Cabral, J. T. Electrostatic and Hydrophobic Interactions in NaCMC Aqueous Solutions: Effect of Degree of Substitution. *Macromolecules* **2018**, *51*, 3165–3175.

(49) De Gennes, P.G.,; Pincus, P.,; Velasco, R.M.,; Brochard, F., Remarks on polyelectrolyte conformation. *J. Phys. France* **1976**, *37*, 1461–1473.

(50) Khokhlov, A. R. On the collapse of weakly charged polyelectrolytes. *Journal of Physics A: Mathematical and General* **1980**, *13*, 979–987.

(51) Dobrynin, A. V.; Rubinstein, M.; Obukhov, S. P. Cascade of Transitions of Polyelectrolytes in Poor Solvents. *Macromolecules* **1996**, *29*, 2974–2979.

(52) Muthukumar, M. Double Screening in Polyelectrolyte Solutions: Limiting Laws and Crossover Formulas. *The Journal of Chemical Physics* **1996**, *105*, 5183–5199.

(53) Muthukumar, M. 50th Anniversary Perspective: A Perspective on Polyelectrolyte Solutions. *Macromolecules* **2017**, *50*, 9528–9560.

(54) Spiteri, M. N.; Williams, C. E.; Boué, F. Pearl-Necklace-Like Chain Conformation of Hydrophobic Polyelectrolyte: a SANS Study of Partially Sulfonated Polystyrene in Water. *Macromolecules* **2007**, *40*, 6679–6691.

(55) Thole, B. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chemical Physics* **1981**, *59*, 341–350.

(56) 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. *The Journal of Chemical Physics* **2013**, *139*, 114502.

(57) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *The Journal of Chemical Physics* **1985**, *83*, 735–746.

(58) Houriez, C.; Meot-Ner (Mautner), M.; Masella, M. Simulated Solvation of Organic Ions: Protonated Methylamines in Water Nanodroplets. Convergence toward Bulk Properties and the Absolute Proton Solvation Enthalpy. *The Journal of Physical Chemistry B* **2014**, *118*, 6222–6233.

(59) Houriez, C.; Meot-Ner (Mautner), M.; Masella, M. Simulated Solvation of Organic Ions II: Study of Linear Alkylated Carboxylate Ions in Water Nanodrops and in Liquid Water. Propensity for Air/Water Interface and Convergence to Bulk Solvation Properties. *The Journal of Physical Chemistry B* **2015**, *119*, 12094–12107.

(60) Houriez, C.; Vallet, V.; Réal, F.; Meot-Ner (Mautner), M.; Masella, M. Organic Ion Association in Aqueous Phase and Ab Initio-Based Force Fields: The Case of Carboxylate/Ammonium Salts. *The Journal of Chemical Physics* **2017**, *147*, 161720.

(61) Li, A. H.-T.; Chao, S. D. Interaction Energies of Dispersion-Bound Methane Dimer from Coupled Cluster Method at Complete Basis Set Limit. *Journal of Molecular Structure: THEOCHEM* **2009**, *897*, 90 – 94.

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

(63) Chremos, A.; Douglas, J. F. The Influence of Polymer and Ion Solvation on the Conformational Properties of Flexible Polyelectrolytes. *Gels* **2018**, *4*.

(64) http://biodev.cea.fr/polaris/download.html/.

(65) Cancès, E.; Legoll, F.; Stoltz, G. Theoretical and Numerical Comparison of Some Sampling Methods for Molecular Dynamics. *ESAIM: M2AN* **2007**, *41*, 351–389.

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

(67) Wu, H.; Ting, J. M.; Werba, O.; Meng, S.; Tirrell, M. V. Non-equilibrium phenomena and kinetic pathways in self-assembled polyelectrolyte complexes. *The Journal of Chemical Physics* **2018**, *149*, 163330.

(68) 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.

(69) 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* **2012**, *34*, 707–719.

(70) Trumm, M.; Guerrero Martinez, Y. O.; Réal, F.; Schimmelpfennig, B.; Masella, M.; Vallet, V. Modeling the Hydration of Mono-Atomic Anions From the Gas Phase to the Bulk Phase: The Case of the Halide Ions F^-, Cl^-, and Br^-. *The Journal of Chemical Physics* **2012**, *136*, 044509.

(71) Timko, J.; Bucher, D.; Kuyucak, S. Dissociation of NaCl in Water from Ab Initio Molecular Dynamics Simulations. *The Journal of Chemical Physics* **2010**, *132*, 114510.

(72) Luo, Y.; Jiang, W.; Yu, H.; MacKerell, A. D.; Roux, B. Simulation Study of Ion Pairing in Concentrated Aqueous Salt Solutions with a Polarizable Force Field. *Faraday Discuss.* **2013**, *160*, 135–149.

(73) 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.

(74) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proceedings of the National Academy of Sciences* **2001**, *98*, 10037–10041.

(75) 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. *The Journal of Chemical Physics* **2019**, *151*, 174504.

(76) Georgalis, Y.; Kierzek, A. M.; Saenger, W. Cluster Formation in Aqueous Electrolyte Solutions Observed by Dynamic Light Scattering. *The Journal of Physical Chemistry B* **2000**, *104*, 3405–3406.

(77) Samal, S.; Geckeler, K. E. Unexpected Solute Aggregation in Water on Dilution. *Chem. Commun.* **2001**, 2224–2225.

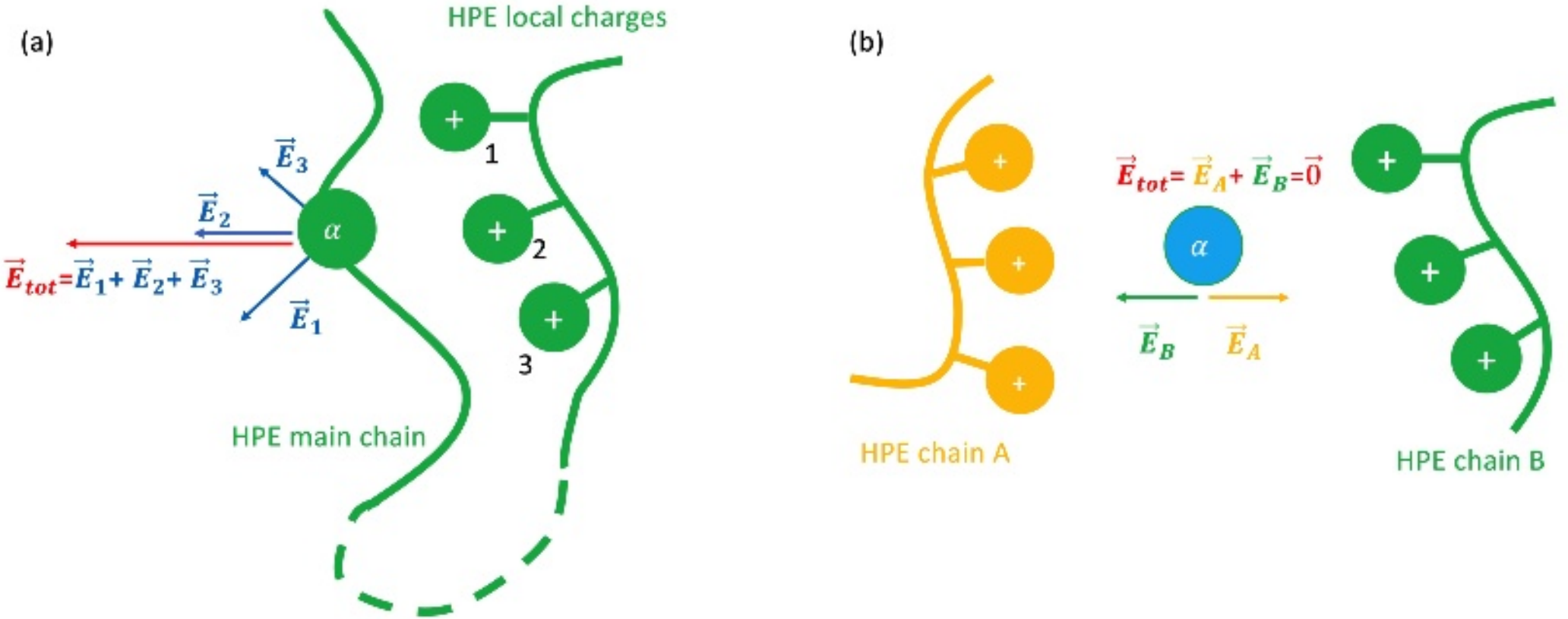
(78) Bharmoria, P.; Gupta, H.; Mohandas, V. P.; Ghosh, P. K.; Kumar, A. Temperature Invariance of NaCl Solubility in Water: Inferences from Salt-Water Cluster Behavior of NaCl, KCl, and NH4Cl. *The Journal of Physical Chemistry B* **2012**, *116*, 11712–11719.

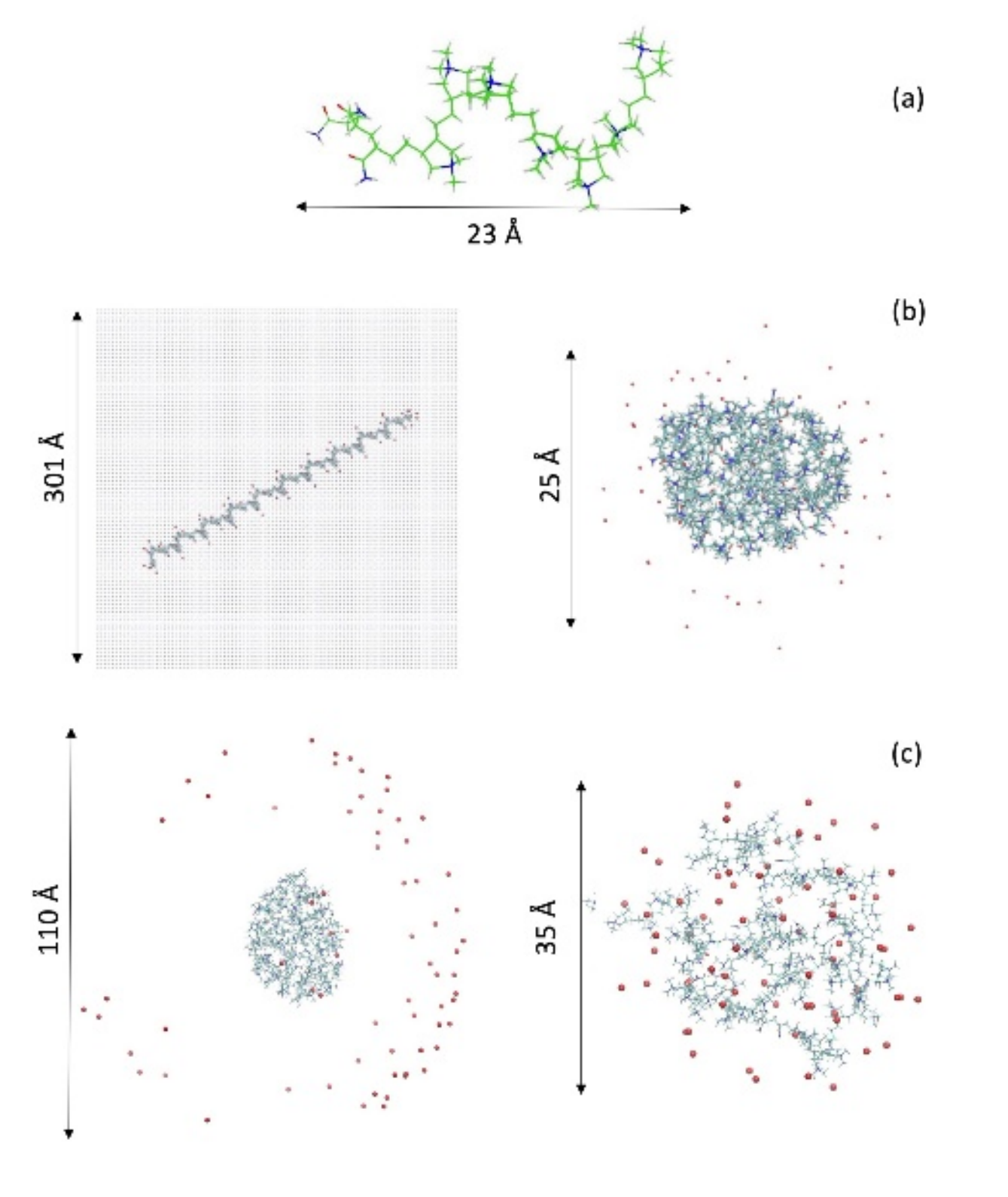
(79) Konovalov, A. I.; Ryzhkina, I. S. Highly Diluted Aqueous Solutions: Formation of Nano-Sized Molecular Assemblies (Nanoassociates). *Geochemistry International* **2014**, *52*, 1207–1226.

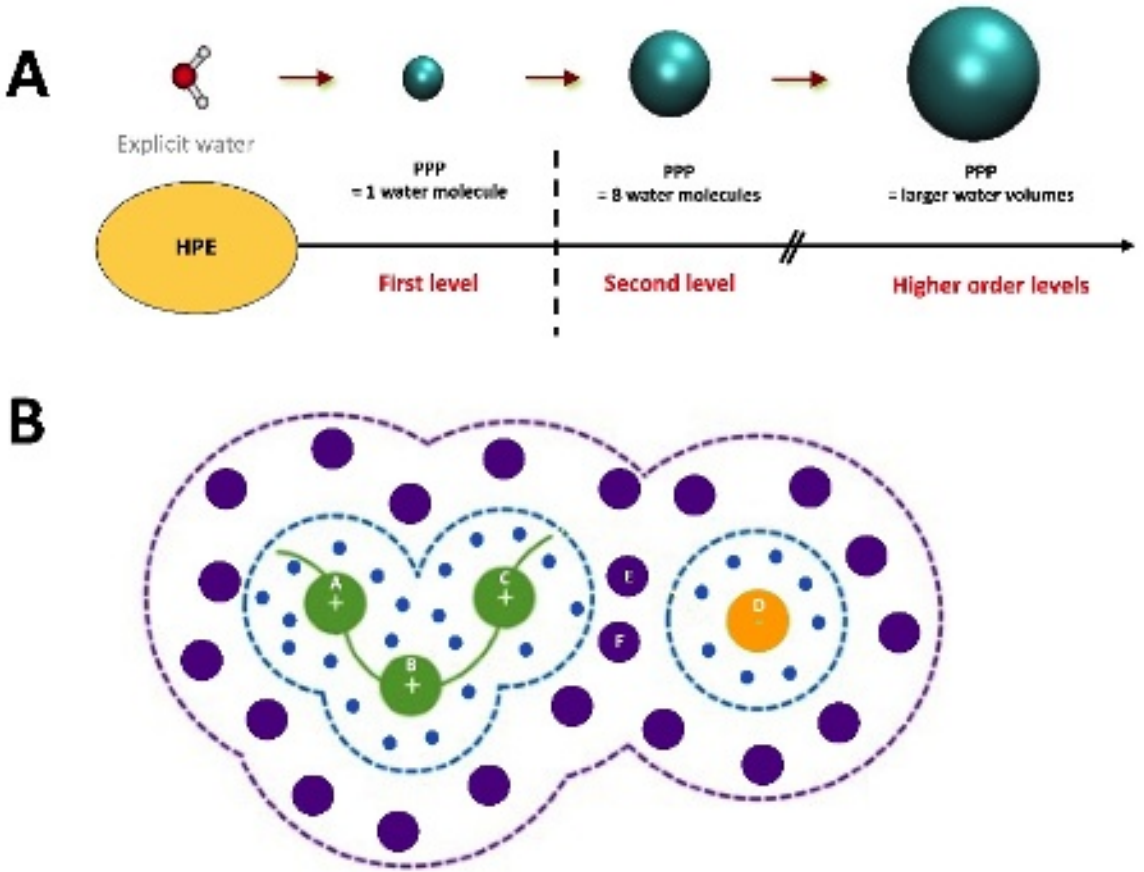
(80) Pliego, J. R. The Role of Intermolecular Forces in Ionic Reactions: the Solvent Effect, Ion-Pairing, Aggregates and Structured Environment. *Org. Biomol. Chem.* **2021**, *19*, 1900–1914.

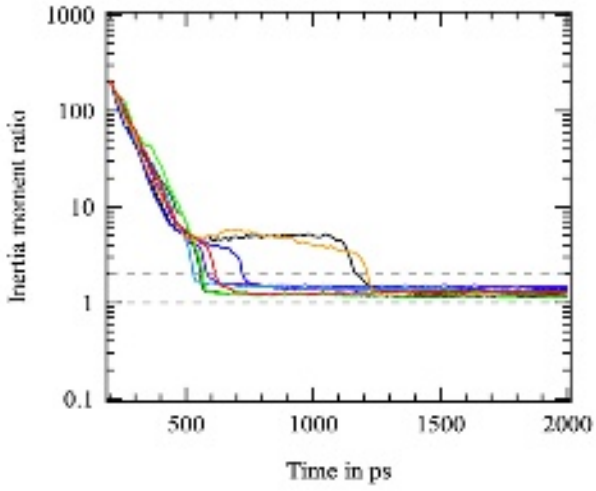
(81) Varghese, A.; Vemparala, S.; Rajesh, R. Phase Transitions of a Single Polyelectrolyte in a Poor Solvent with Explicit Counterions. *The Journal of Chemical Physics* **2011**, *135*, 154902.

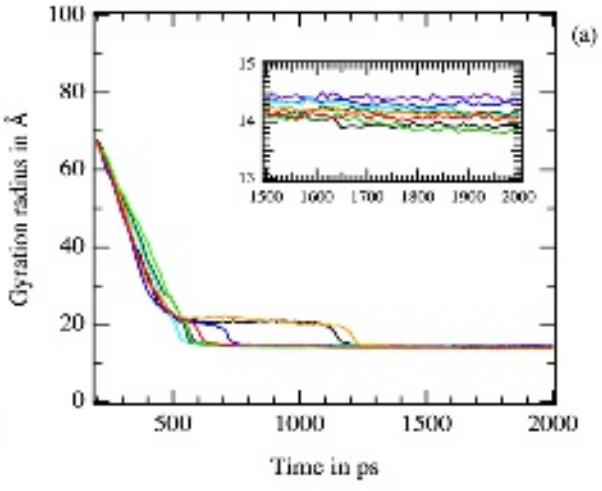
(82) Coles, J. P.; Masella, M. The Fast Multipole Method and Point Dipole Moment Polarizable Force Fields. *The Journal of Chemical Physics* **2015**, *142*, 024109.

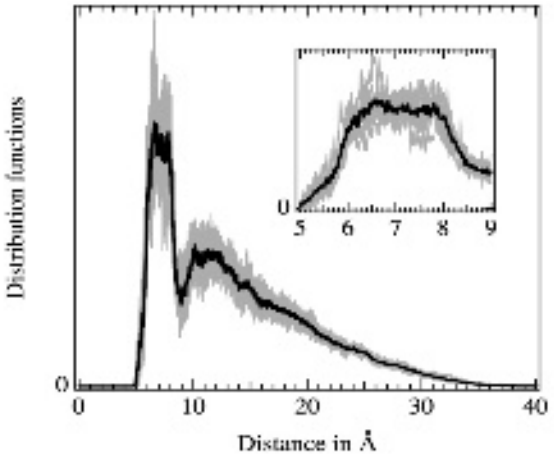

(a)
HPE local charges
1
2
3
$\vec{E}_3$
$\vec{E}_2$
α
$\vec{E}_1$
$\vec{E}_{tot}=\vec{E}_1+\vec{E}_2+\vec{E}_3$
HPE main chain
(b)
$\vec{E}_{tot}=\vec{E}_A+\vec{E}_B=\vec{0}$
α
$\vec{E}_B$
$\vec{E}_A$
HPE chain A
HPE chain B

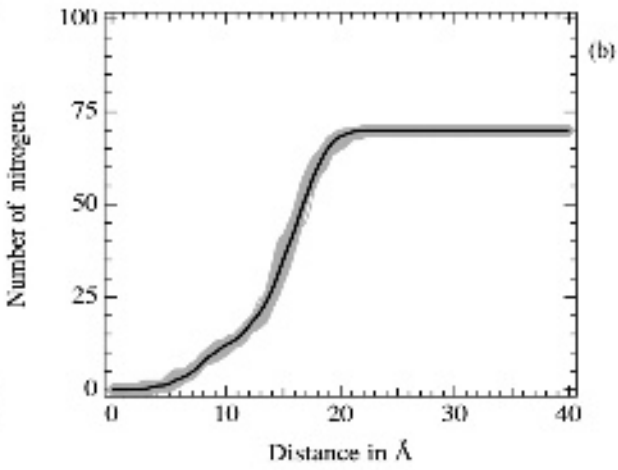

(a)
23 Å
(b)
301 Å
25 Å
(c)
110 Å
35 Å

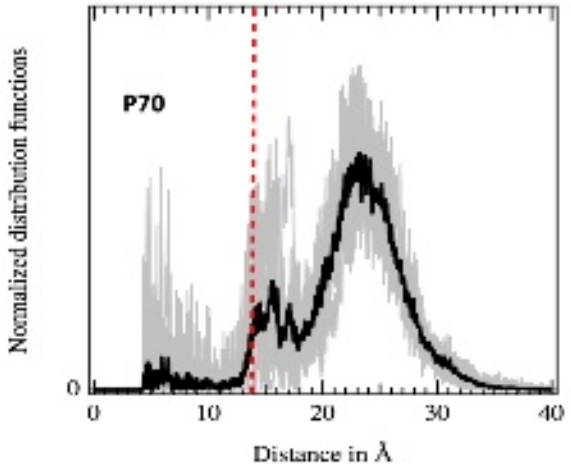

A
Explicit water
PPP
= 1 water molecule
PPP
= 8 water molecules
PPP
= larger water volumes
HPE
First level
Second level
Higher order levels
B
A
B
C
D
E
F

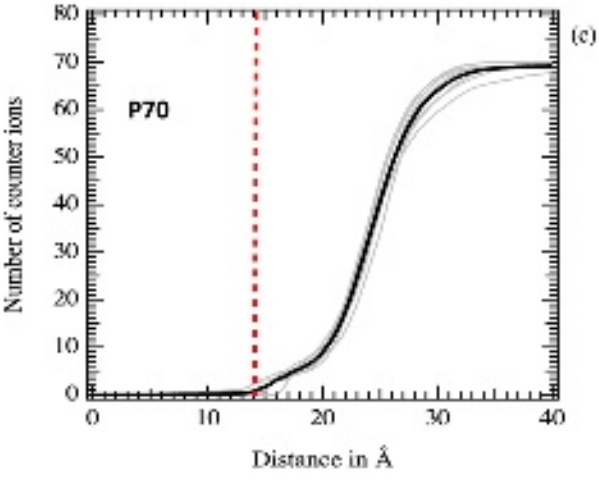

Inertia moment ratio
Time in ps

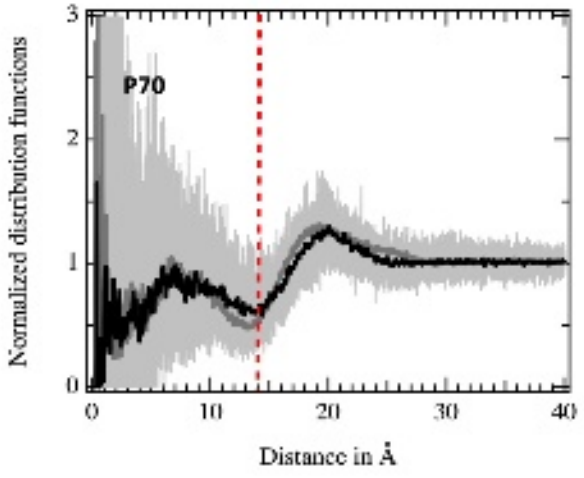

(a)
Gyration radius in Å
Time in ps

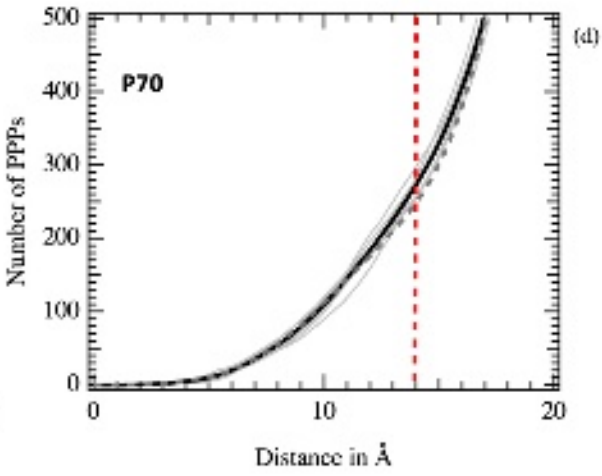

Distribution functions
Distance in Å

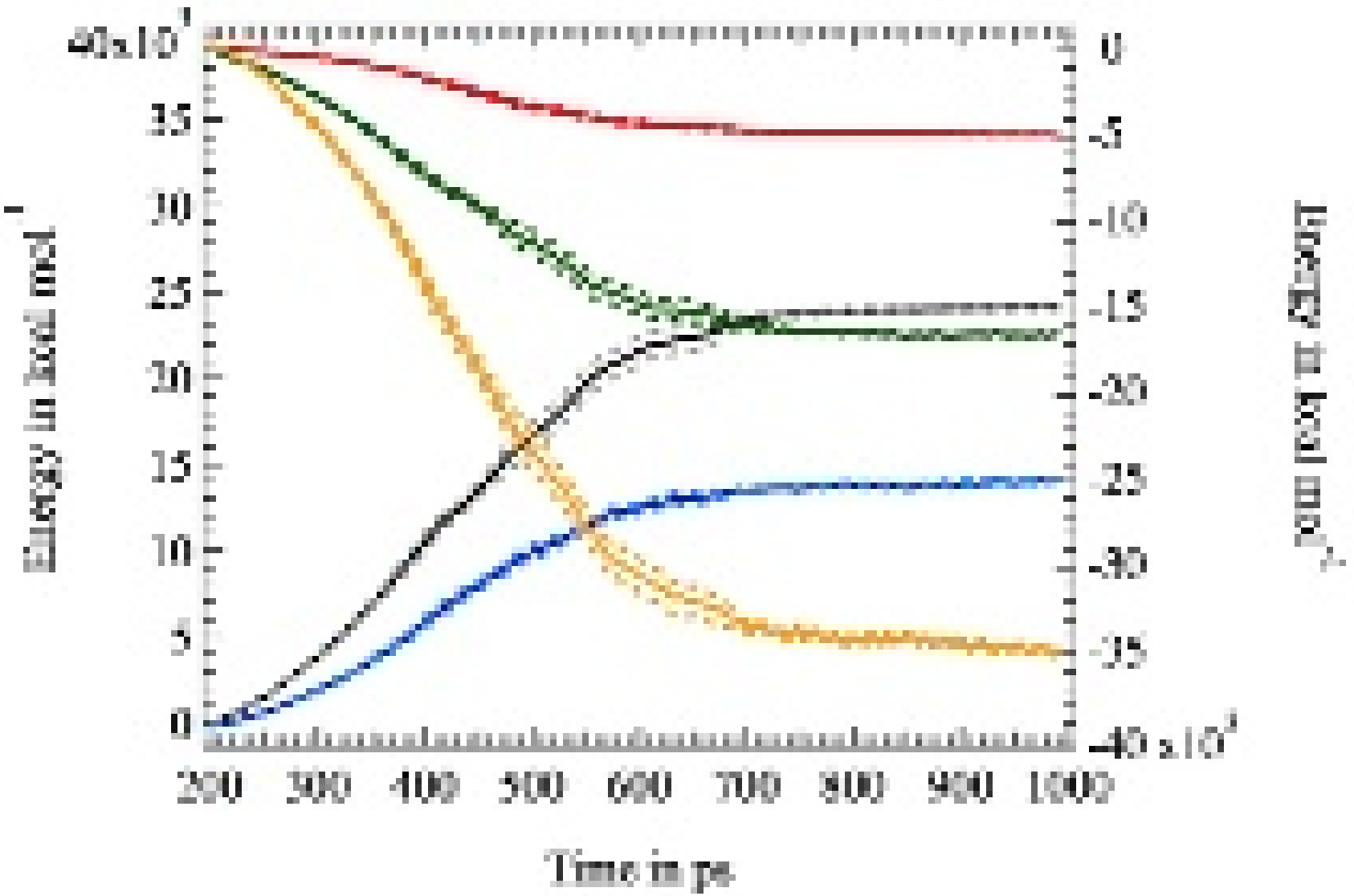

(b)
Number of nitrogens
Distance in Å

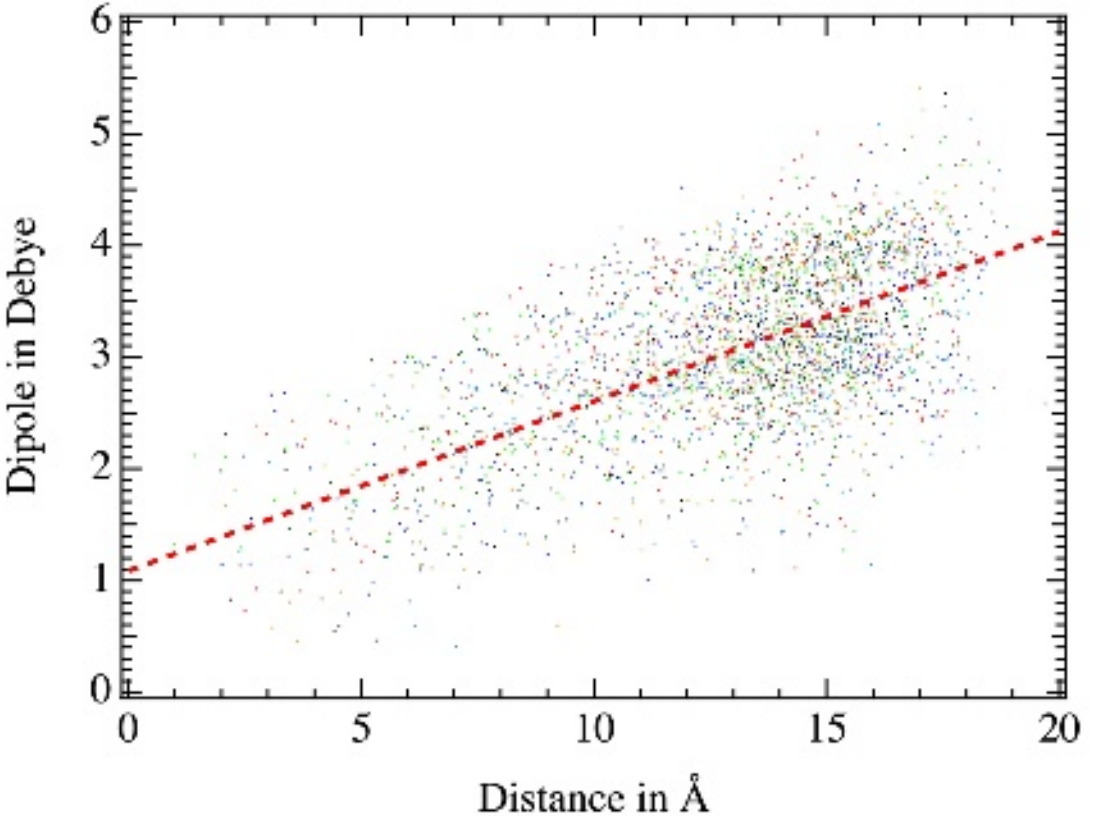

P70
Normalized distribution functions
Distance in Å

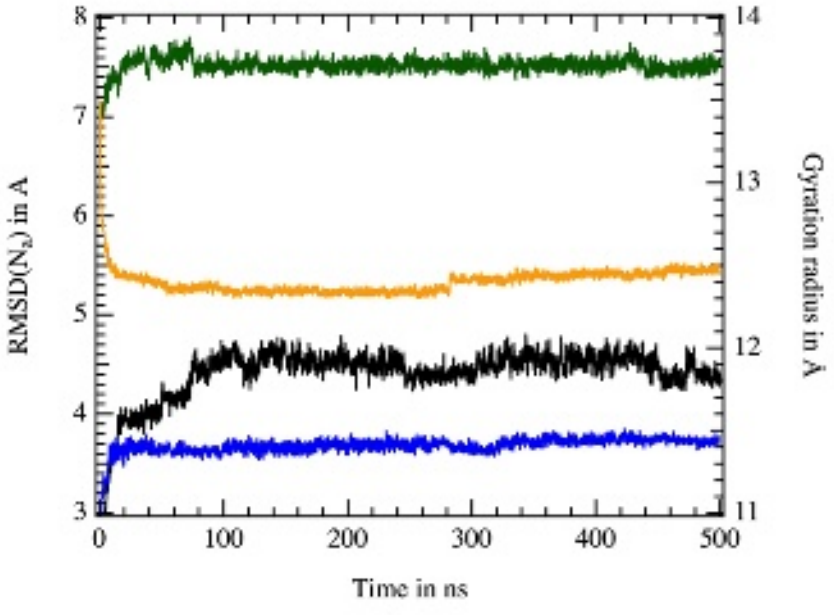

(c)
P70
Number of counter ions
Distance in Å

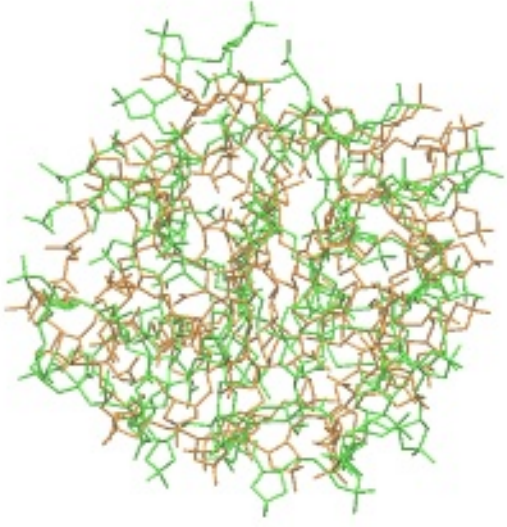

P70
Normalized distribution functions
Distance in Å

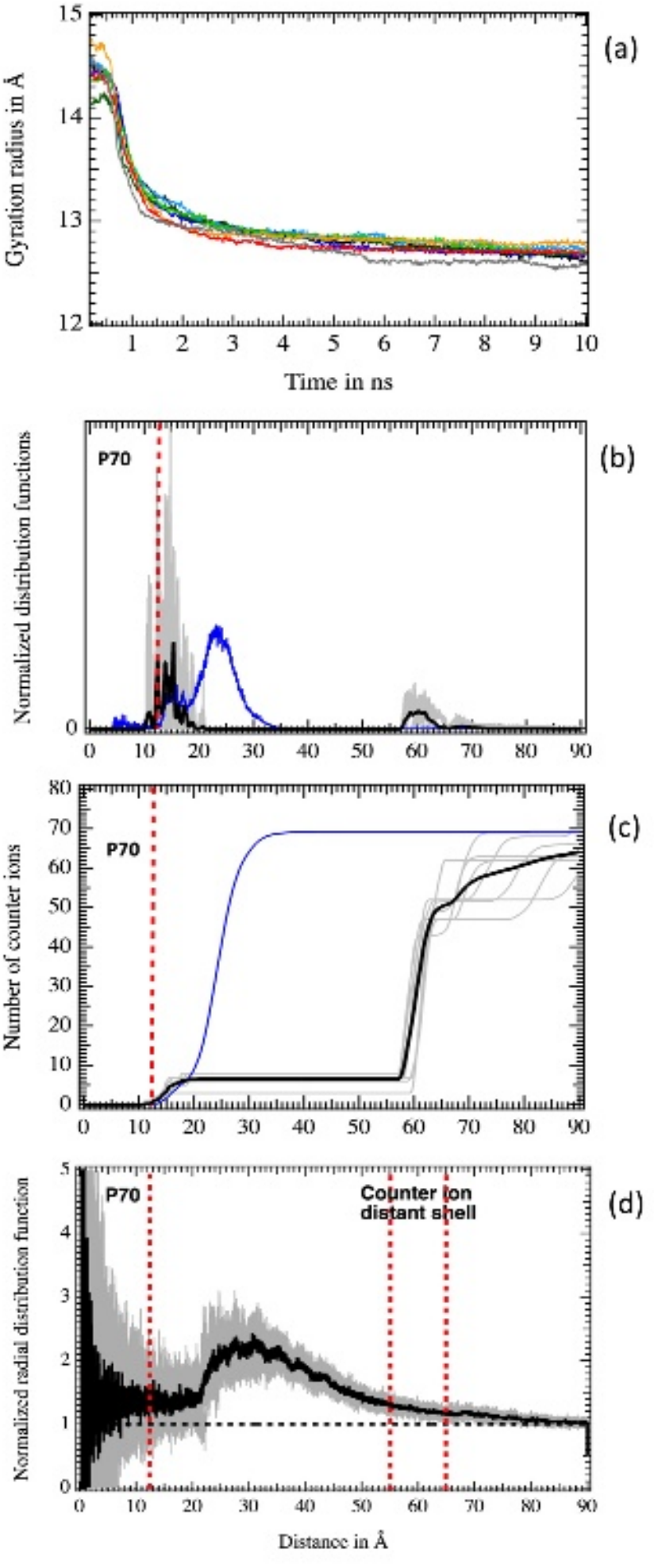

(d)
P70
Number of PPPs
Distance in Å

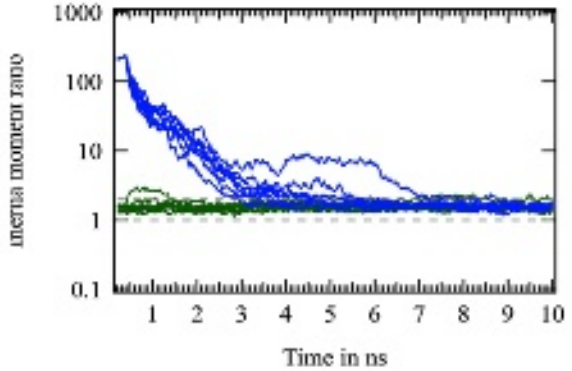

40x10^3
35
30
25
20
15
10
5
0
Energy in kcal mol^-1
0
-5
-10
-15
-20
-25
-30
-35
-40 x10^3
Energy in kcal mol^-1
200
300
400
500
600
700
800
900
1000
Time in ps

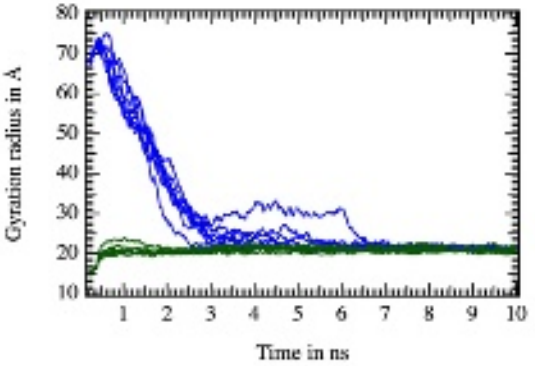

Dipole in Debye
Distance in Å
0
1
2
3
4
5
6
0
5
10
15
20


RMSD(N$_a$) in A
Gyration radius in Å
Time in ns
8
7
6
5
4
3
14
13
12
11
0
100
200
300
400
500

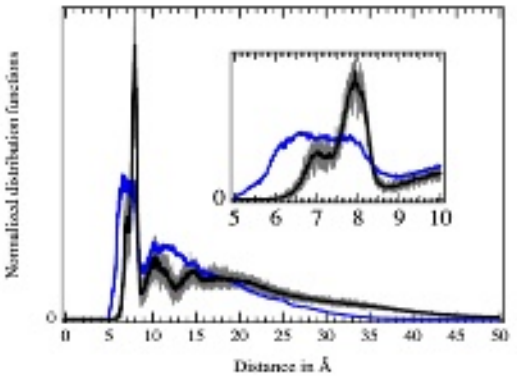


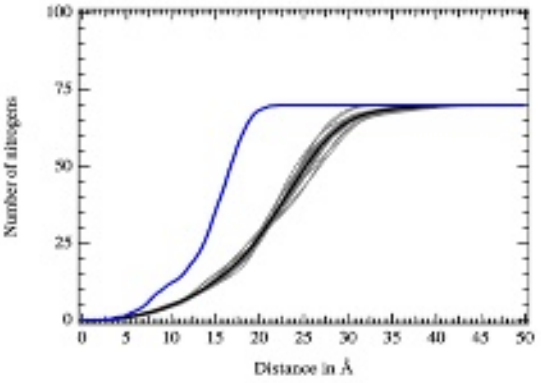

(a)
Gyration radius in Å
Time in ns
(b)
P70
Normalized distribution functions
(c)
P70
Number of counter ions
(d)
P70
Counter ion
distant shell
Normalized radial distribution function
Distance in Å


inertia moment ratio
1000
100
10
1
0.1
1 2 3 4 5 6 7 8 9 10
Time in ns

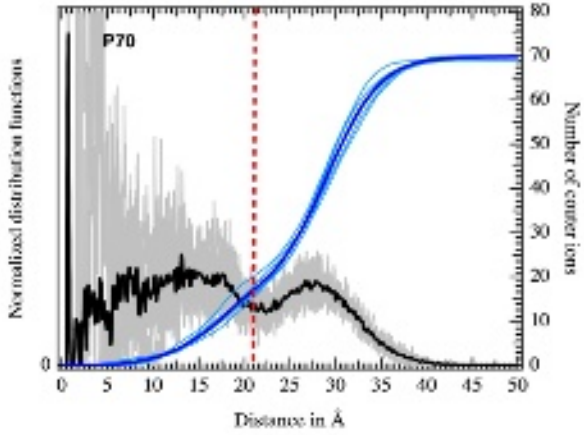

Gyration radius in Å
80
70
60
50
40
30
20
10
1 2 3 4 5 6 7 8 9 10
Time in ns

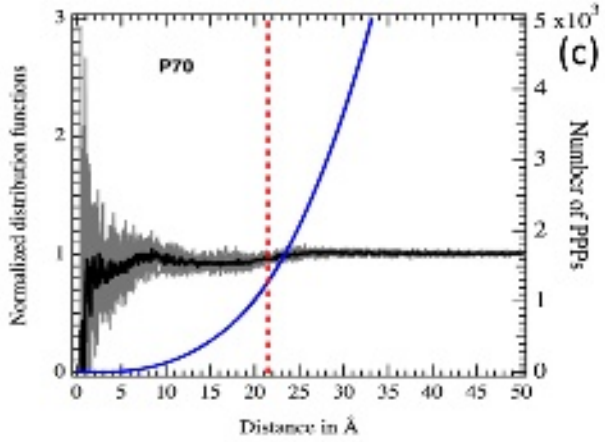
(a)


Normalized distribution functions
0
0 5 10 15 20 25 30 35 40 45 50
Distance in Å
0
5 6 7 8 9 10

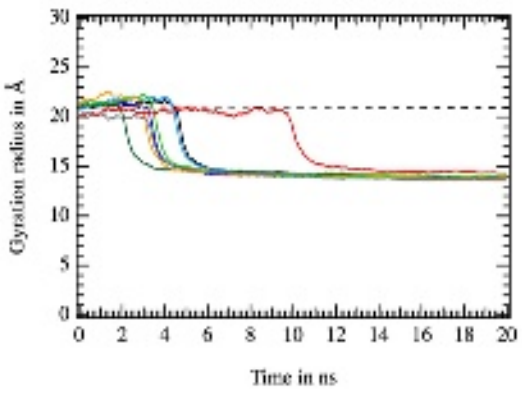

Number of atoms
100
75
50
25
0
0 5 10 15 20 25 30 35 40 45 50
Distance in Å

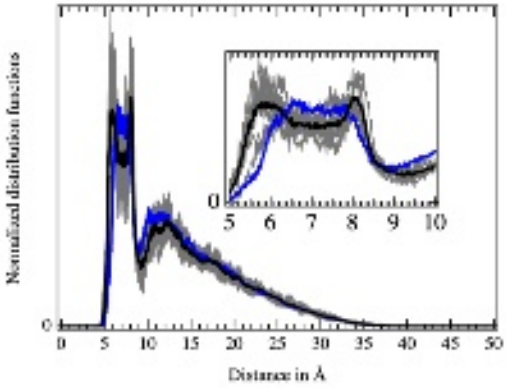
(b)


P70
Normalized distribution functions
Number of counter ions
0
80
70
60
50
40
30
20
10
0
0 5 10 15 20 25 30 35 40 45 50
Distance in Å

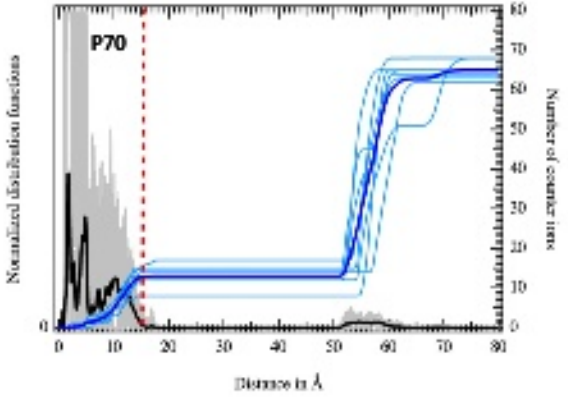

P70
Normalized distribution functions
Number of PPPs
3
2
1
0
5 x10-3
4
3
2
1
0
0 5 10 15 20 25 30 35 40 45 50
Distance in Å

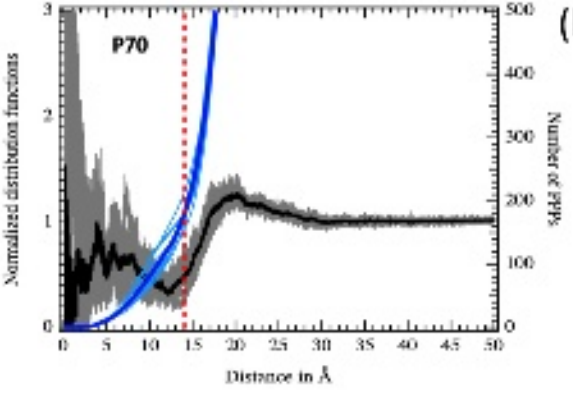
(c)


Gyration radius in Å
Time in ns

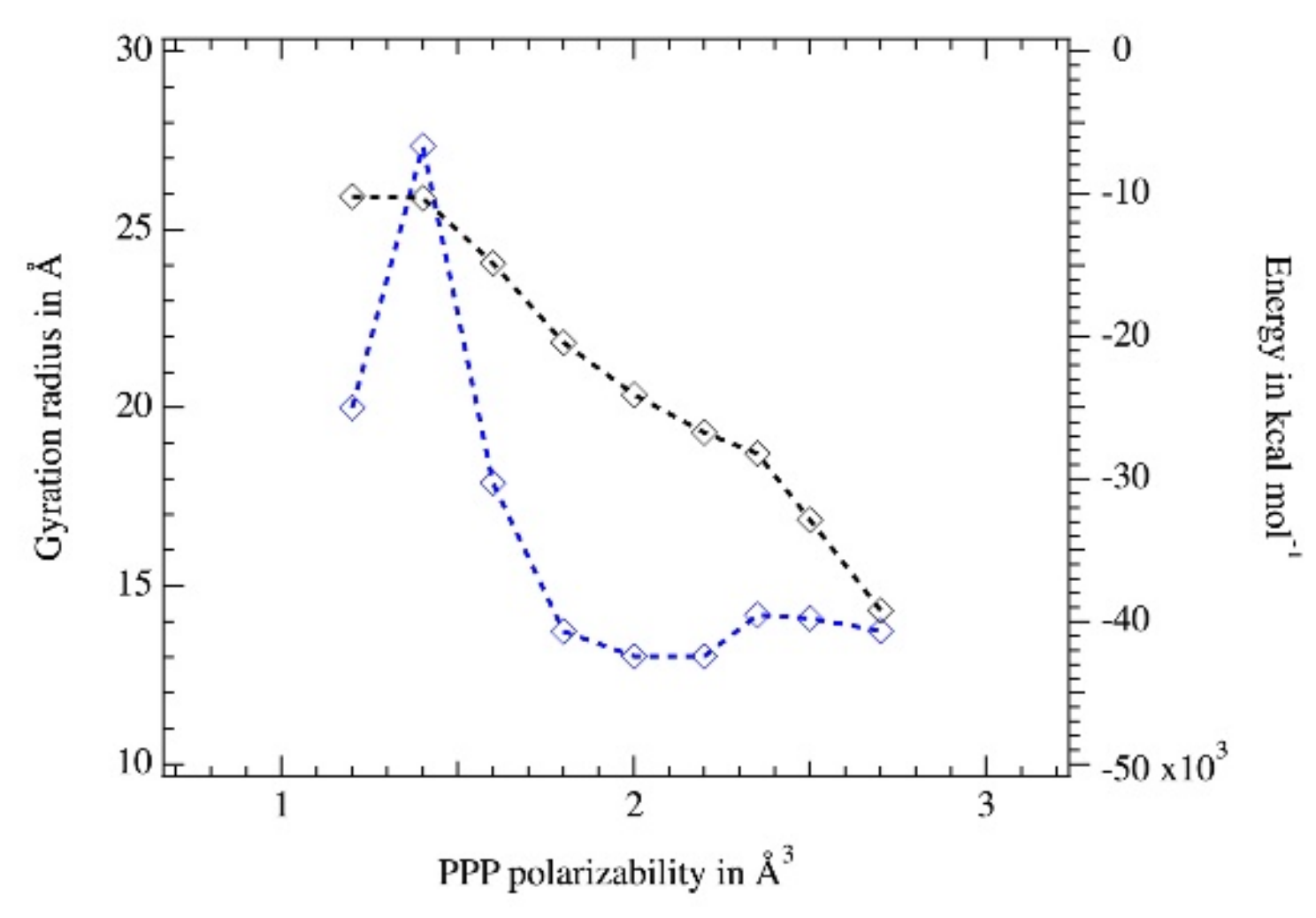

Normalized distribution functions
Distance in Å

(a)


P70
Normalized distribution functions
Number of contact lines
Distance in Å


P70
Normalized distribution functions
Number of PPIs
Distance in Å

(b)


30
25
20
15
10
Gyration radius in Å
1
2
3
PPP polarizability in Å^3
0
-10
-20
-30
-40
-50 x10^3
Energy in kcal mol^{-1}