PAPER

Role of entropy in determining the phase behavior of protein solutions induced by multivalent ions†

Anil Kumar Sahoo, [ID] *[abc] Frank Schreiber, [ID] [d] Roland R. Netz [ID] [ce] and
Prabal K. Maiti [ID] *[a]

Recent experiments have reported lower critical solution temperature (LCST) phase behavior of aqueous solutions of proteins induced by multivalent ions, where the solution phase separates upon heating. This phenomenon is linked to complex hydration effects that result in a net entropy gain upon phase separation. To decipher the underlying molecular mechanism, we use all-atom molecular dynamics simulations along with the two-phase thermodynamic method for entropy calculation. Based on simulations of a single BSA protein in various salt solutions (NaCl, CaCl$_2$, MgCl$_2$, and YCl$_3$) at temperatures (T) ranging 283–323 K, we find that the cation–protein binding affinity *increases* with T, reflecting its thermodynamic driving force to be entropic in origin. We show that in the cation binding process, many tightly bound water molecules from the solvation shells of a cation and the protein are released to the bulk, resulting in entropy gain. To rationalize the LCST behavior, we calculate the ζ-potential that shows *charge inversion* of the protein for solutions containing multivalent ions. The ζ-potential increases with T. Performing simulations of two BSA proteins, we demonstrate that the protein–protein binding is mediated by multiple cation bridges and involves similar dehydration effects that cause a large entropy gain which more than compensates for rotational and translational entropy losses of the proteins. Thus, the LCST behavior is entropy-driven, but the associated solvation effects are markedly different from hydrophobic hydration. Our findings have direct implications for tuning the phase behavior of biological and soft-matter systems, *e.g.*, protein condensation and crystallization.

[a] *Center for Condensed Matter Theory, Department of Physics, Indian Institute of Science, Bangalore-560012, India. E-mail: 201992kumarsahoo@gmail.com, maiti@iisc.ac.in*

[b] *Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam, Germany*

[c] *Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

[d] *Institute for Applied Physics, University of Tübingen, 72076 Tübingen, Germany*

[e] *Department of Physics, Indian Institute of Science, Bangalore-560012, India*

† Electronic supplementary information (ESI) available: Distribution of anions around the protein; electrostatic potential map and cation/anion number density map; cation binding/unbinding kinetics; time series of the number of protein-bound water molecules; radial distribution functions for oxygen of water around cations; entropy contribution of water in the second solvation shell of cations; comparison of entropy obtained from the 2PT method and the temperature derivative of free energy; the T-dependence of binding energy; T and salt concentration dependence of the dielectric constant; the T-dependence of entropy and energy of ion binding excluding the contribution due to dehydration of the second solvation shell of the ion; note on the T-dependence of free energy; theory of the 2PT method; note on volume calculation; derivation of the expression for ζ-potential (eqn (9)); table for comparison between computed and experimental values for structural parameters and entropy of ion hydration; and table for water released from protein and cation surfaces in a binding process. See DOI: 10.1039/d1sm00730k

Introduction

Ions play an important role in many biophysical processes, *e.g.*, allosteric regulation, enzymatic activity, DNA condensation, and protein solubility and crystallization. Starting from the pioneering works by Hofmeister, there has been immense progress made to better understand ion–protein interactions.[1,2] In recent years, due to various applications in biology, medicine and physics, there is increasing interest to tune and control the phase behavior of protein solutions using multivalent ions.[3] Diverse phenomena induced by multivalent ions have been realized in experiments. These include: (i) reentrant condensation of proteins in bulk solution[4] as well as reentrant surface-adsorption of proteins[5] by varying the concentration of Y^{3+} or other trivalent cations, (ii) pathway-controlled protein crystallization,[6] (iii) clustering,[7] (iv) liquid–liquid phase separation,[7,8] and (v) lower critical solution temperature (LCST) phase behavior.[9] Although many aspects regarding ion–protein interactions have been qualitatively understood, a fundamental and quantitative understanding is required for further developments in this field.

Of particular interest is the LCST phase behavior for a solution of bovine serum albumin (BSA) proteins in the presence of Y^{3+} ions.[9] At low temperatures, the proteins remain well

dispersed in solution, whereas upon increasing temperature up to 300 K, the proteins attract each other, and the solution separates into protein-rich and protein-poor phases. We note that aggregation of proteins can also be caused by thermal denaturation, but in the experiments Matsarskaia et al.[9] stayed well below the protein denaturation temperature and observed LCST behavior only for solutions containing trivalent ions.[10] This precludes denaturation as a mechanism and suggests that the LCST behavior is related to ion-mediated protein aggregation.

It has been suggested that the LCST behavior is due to the combination of effects associated with the solvation of the protein and the multivalent ions, and that entropy is the driving force.[9] However, the molecular mechanism of the LCST behavior has not been quantitatively identified. A quantitative understanding of the thermodynamics of this process requires an accurate estimation of various entropy contributions associated with the ion–protein complex formation and the subsequent ion-mediated protein–protein aggregation. The total entropy change includes entropy costs due to (i) hindrance in the translation of a protein-bound ion, (ii) restrictions on the translational and rotational motions of proteins, (iii) hydration/dehydration of the protein and ions, and (iv) conformational changes of the protein. The latter is mainly important for metalloregulatory allosteric proteins. Quantifying all these entropy contributions in experiments remains a daunting task, even with the present-day techniques that provide residue-level dynamic information.[11] In this regard, molecular simulations[12,13] along with accurate and robust entropy calculation techniques provide an alternative and reliable approach.

To understand the mechanistic details and the thermodynamic driving force for the intriguing phenomena related to ion-mediated protein–protein interactions, we have performed large-scale molecular dynamics (MD) simulations of a single and two BSA proteins in chloride salt solutions of Y^{3+} and several other cations found in physiological conditions, such as Na^+, Ca^{2+}, and Mg^{2+} in the temperature range of 283–323 K. The simulation details are presented in the Methods section. A snapshot of the initial configuration of the simulated single-protein system is shown in Fig. 1A. We investigate the specific nature of ion–protein interactions and quantify the free energy, various entropy contributions as well as electrostatics of the system. Our study reveals crucial solvation/desolvation phenomena giving rise to an entropic driving force for ion–protein binding, in contrast to common expectations. From simulations of the systems involving two BSA proteins, it is found that Y^{3+} ions link the two proteins to form a dimer. Hence, the process of ion-mediated protein–protein binding is argued to be entropy-driven, as a large number of tightly bound water molecules are released from the proteins and the mediating cations' surfaces to the bulk solution.

Results

BSA protein–ion interaction and ion binding kinetics

To investigate the nature of ion–protein interactions, we calculate the number distribution of ions $N(r)$ along the protein's
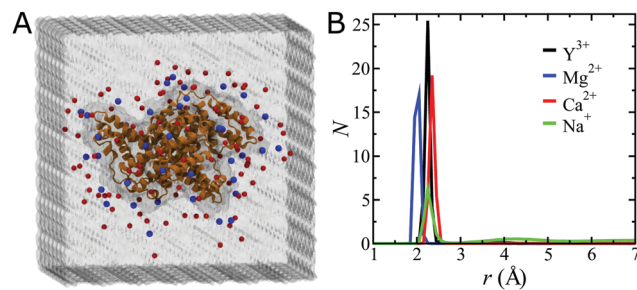

Fig. 1 (A) Snapshot for the starting simulation box (of size 13.2 × 13.2 × 13.1 nm³) containing a single BSA protein in 30 mM YCl_3 solution. The protein (orange) is shown in cartoon representation. Y^{3+} (blue) and Cl^- (red) ions are represented as VDW spheres. Water molecules are represented as a continuum in semi-transparent mode for clarity. Note that the system was also simulated in other salt solutions ($MgCl_2$, $CaCl_2$, and $NaCl$) of the same ionic strength as in the case of YCl_3. (B) For each type of cation, the total number of cations N found within shells of width dr = 0.1 Å, present at a shortest distance r from the protein surface, is plotted as a function of r for the simulation performed at 303 K. $N(r)$ for each cation type is averaged over the last 100 ns of the simulation time. Note that $N(r)$ is enhanced near the protein surface, representing the strong affinity of cations for the negatively charged protein (−16 e at pH 7).

surface-normal direction. $N(r)$ for the cations are shown in Fig. 1B, while $N(r)$ for Cl^- ion for the different ionic solutions are plotted in Fig. S1 in the ESI.† We find that cations are mostly present near the protein, with the relative propensity of binding showing the following trend: monovalent < divalent < trivalent. These cations predominantly pair with the negatively charged carboxylate groups of aspartate and glutamate surface residues of the protein. Interestingly, even in NaCl solution, Cl^- ions are found to be largely present near the protein, and the number of Cl^- ions present near the protein decreases in the following order: YCl_3 > $MgCl_2$ ≈ $CaCl_2$ > NaCl (Fig. S1 in the ESI†). This suggests that Cl^- ions interact with the $-NH_3^+$ group of the protein surface residues, and also interact, via ion-pair formation, with the cations present in the vicinity of the protein.

A protein surface is, however, far from uniform and if some extended patches are present on its surface, strong affinity of multivalent ions is expected even if the net charge of the protein is small or even of the opposite sign.[14] We indeed find a positively charged patch and a few extended negatively charged patches from the electrostatic potential map for BSA (Fig. S2A in the ESI†). We find higher density of cations (anions) near negatively (positively) charged patches even for monovalent ions (Fig. S2B and C in the ESI†).

To check how tightly the cations are bound to the protein, we monitor their binding/unbinding kinetics. An ion is defined as bound if it is within a cutoff distance r_c from any atom of the protein, otherwise the ion is unbound or free. From the $N(r)$ plot in Fig. 1B, r_c's for the different cations are chosen as 2.8 Å (Na^+), 2.7 Å (Ca^{2+}), 2.3 Å (Mg^{2+}), and 2.5 Å (Y^{3+}). We find intermittent binding/unbinding events for both Na^+ and Ca^{2+} ions (Fig. S3 in the ESI†). While the binding/unbinding events for Na^+ ions are frequent, prolonged bindings are observed for Ca^{2+} ions. For these two cation types, the binding time, i.e., the duration for which an ion remains bound once it comes within
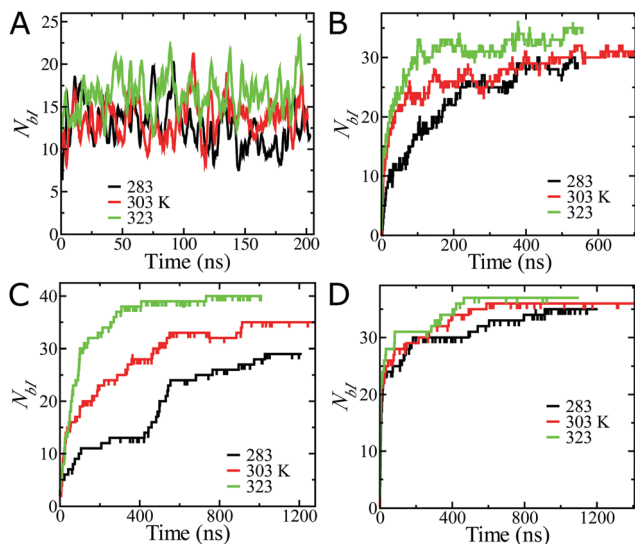
Fig. 2 Time series of the total number of protein-bound cations (N_{bI}) at several temperatures for Na^+ (A), Ca^{2+} (B), Mg^{2+} (C) and Y^{3+} (D) ions. For the multivalent cations, N_{bI} increases upon increasing temperature.

distance of r_c from the protein, is broadly distributed, owing to stochastic effects and the surface heterogeneity of the protein. In contrast, only one unbinding event is observed for Mg^{2+} within 1.27 μs, whereas no unbinding of Y^{3+} is seen within 1.45 μs (Fig. S3 in the ESI†). As the water escape time in the first solvation shell of Mg^{2+} is ~1.5 μs,[15] it presumably requires very long simulations (100 μs to 10 ms) to obtain sufficient unbinding statistics for Mg^{2+} and Y^{3+} ions. Performing such long, all-atom simulations for our system is out of reach of our computational capabilities.

For each cation type, the total number of protein-bound cations, N_{bI}, is plotted as a function of the simulation time at three different temperatures in Fig. 2. No ion is bound to the protein at the beginning of a simulation, and N_{bI} gradually increases with the simulation time. N_{bI} eventually reaches a saturation value, at a time required for equilibration of the ion distribution around the protein. This ion equilibration time differs for each cation, which can be rationalized by considering the ion–water exchange kinetics that strongly depends on the cation's charge and size.[15] Counterintuitively, we find from Fig. 2 that N_{bI} increases with increasing temperature. This effect is prominent for all the cations, except Na^+. In contrast, the number of protein-bound water, i.e., the total number of water molecules present within 3 Å from the protein surface decreases with the increase in temperature as expected (Fig. S4 in the ESI†). Although an increase in the binding affinity of any two objects by raising the temperature is not new—e.g., hydrophobic interaction strength increases with temperature,[16] it is surprising to be observed in a system involving strong electrostatic interactions and can be rationalized by the temperature dependence of dielectric and hydration effects.[17,18] For a quantitative understanding of this, we calculate various thermodynamic quantities such as the free energy, enthalpy, and various entropy contributions as discussed below.

Thermodynamics of cation binding to the protein

The free energy of a cation binding to the protein, ΔG_b, for temperatures in the range of 283–323 K is shown in Fig. 3A (see Methods for the calculation details). For each cation type, ΔG_b is always negative, and its magnitude increases with the increase in temperature. $|\Delta G_b|$ follows the trend: $Na^+ < Ca^{2+} \approx Mg^{2+} < Y^{3+}$. By changing temperature from 283 K to 323 K, we see the highest change in ΔG_b for Y^{3+} (-1.21 kcal mol^{-1}), whereas the least change is observed for Na^+ binding (-0.52 kcal mol^{-1}). The changes in ΔG_b for Ca^{2+} and Mg^{2+} ions are -0.71 and -1.03 kcal mol^{-1}, respectively.

The increase in binding affinity of the cations with solely increasing temperature (Fig. 3A) cannot be explained by considering the energy of binding, for purely thermodynamic reasons, as described in the ESI,† Section 1. Further, it should be noted that since the dielectric constant of water decreases as $\sim T^{-3/2}$, any electrostatic interaction in water is predominantly entropic in nature.[17,18] Therefore, entropy must be playing a dominant role here.

The binding free energy for an ion is given by

$$\Delta G_b(T) = \Delta E_b(T) - T\Delta S_b(T), \quad (1)$$

where ΔE_b and ΔS_b are the energy and entropy of binding, respectively and T is the temperature. For the calculation of ΔS_b, one needs to correctly account for "hydration effects" associated with the ion binding process, such as partial desolvation of both the protein and ion. The radial distribution functions for water molecules around a cation, both free in solution and bound to the protein surface, clearly show partial dehydration of the first and second solvation shells (SS's) of each cation (Fig. S5 in the ESI†). ΔS_b in eqn (1) consists of three terms—the loss in entropy of a protein-bound ion ($\Delta S_{P,I}$), the gain in entropy due to release of tightly-bound water molecules from the first and second SS's of the ion ($\Delta S_{I,W}$), and the gain in entropy of water molecules released to the bulk due to desolvation of the protein surface residue where the ion binds ($\Delta S_{P,W}$). Together, it can be written as

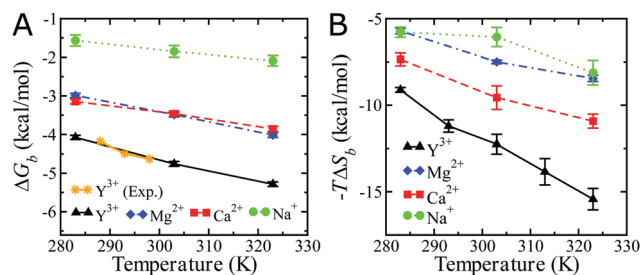$$\Delta S_b = \Delta S_{P,I} - \Delta S_{I,W} - \Delta S_{P,W}. \quad (2)$$

Fig. 3 Temperature dependence of (A) the free energy, ΔG_b, and (B) the total entropy contribution, $-T\Delta S_b$, for each cation binding to the protein. Error bars represent the standard deviation. The different lines are for guiding the eye. The experimental binding free energies for Y^{3+} at different temperatures shown in (A) are taken from ref. 9.
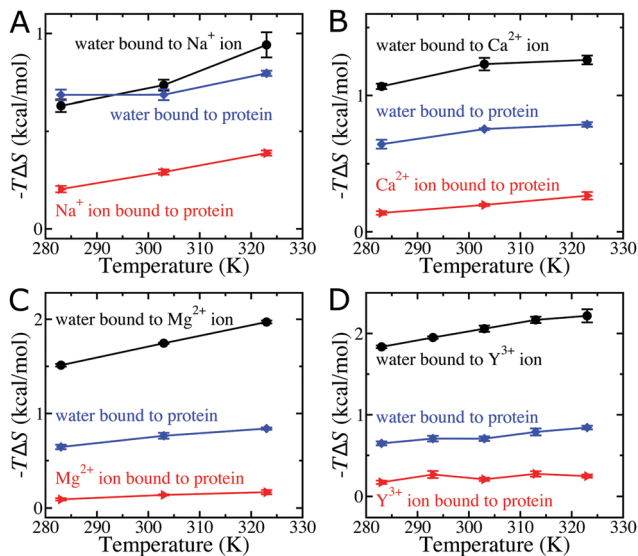
Fig. 4 Entropy contributions $\Delta S_{P,I}$ (red), $\Delta S_{I,W}$ (black), $\Delta S_{P,W}$ (blue) from eqn (2) as a function of temperature for the protein in NaCl (A), CaCl$_2$ (B), MgCl$_2$ (C), and YCl$_3$ (D) solutions. Error bars represent the standard deviation. The lines are for guiding the eye.

We have used the two-phase thermodynamic (2PT) method[19–21] to calculate all the terms on the right hand side of eqn (2). The theory of the 2PT method is described in the ESI,† Section 2 and calculation details are given in the Methods section. We first validate the 2PT method for ionic solutions by reproducing from the simulation data the experimental ion hydration entropy in bulk (ΔS_{hyd}) for the different ion types (see Table S1 in the ESI†). Then, we proceed, using 2PT, with calculations of the entropy differences for a protein-bound ion, a protein-bound water, and a water in the first SS of the cation as shown in Fig. 4, as well as the entropy difference of a water in the second SS of the cation as shown in Fig. S6 in the ESI.† Note that for calculations of the various entropy contributions shown in Fig. 4 and Fig. S6 (ESI†), the reference values are taken as the respective absolute entropies in the bulk water. $\Delta S_{I,W}$ in eqn (2) is then calculated by multiplying the per water entropy differences with the corresponding numbers of water molecules released in the partial dehydration of both the first and second SS's of the cation (values given in Table S2 in the ESI†), and adding both terms. Similarly, $\Delta S_{P,W}$ in eqn (2) is evaluated by multiplying the per water entropy difference with the number of water molecules released in the partial dehydration of the protein surface residue (values given in Table S2 in the ESI†). From Fig. 4, we see that the entropy loss of a protein-bound cation is more than compensated by the entropy gain of water molecules released to bulk by the partial dehydration of both the cation and protein. The cation desolvation entropy contributes the highest to the thermodynamics of protein–ion binding for all the multivalent cations, whereas both the protein and ion desolvation entropies contribute equally for Na$^+$ binding.

The total entropy contribution ($-T\Delta S_b$ in eqn (1)) for each cation plotted in Fig. 3B is always negative, and it decreases (becomes more negative) with increasing temperature. We have also estimated $-T\Delta S_b$ from the temperature dependence of ΔG_b using the thermodynamic relation $\Delta S_b = -\partial \Delta G_b / \partial T$, and we find that the temperature dependence trend is the same as obtained from the 2PT method, though the values obtained from both methods match only semi-quantitatively (see Fig. S7 in the ESI,† for comparison). For each cation, $-T\Delta S_b$ is more negative than the binding free energy ΔG_b throughout the temperature range studied in this work (Fig. 3). Therefore, the process of a cation binding to the protein is entropy driven. The above observations, in particular, explain the enhancement of the protein-binding affinity of a multivalent cation with increasing temperature (Fig. 2). The total entropy contribution as shown in Fig. 3B is the highest for Y^{3+} across the whole temperature range, followed by Ca^{2+} > Mg^{2+} ≈ Na$^+$—representing a delicate dependency of entropy on the charge and size of a cation. Note that though the entropy contribution due to a water molecule released from Mg^{2+} is more than that for Na$^+$ and Ca^{2+} ions (Fig. 4), the altered trend in ΔS_b for Mg^{2+} in Fig. 3B is rationalized by the lower number of water molecules released in the process of a Mg^{2+} ion binding, compared to that for Na$^+$ and Ca^{2+} bindings.

The large (and negative) value of the entropy contribution, $-T\Delta S_b$, must be partially compensated by a positive binding energy ΔE_b to result in a small (and negative) value of the binding free energy ΔG_b. ΔE_b, calculated by using the thermodynamic relation given in eqn (1), is plotted as a function of temperature in Fig. S8 in the ESI.† ΔE_b is positive throughout the temperature range, in agreement with the experiment,[9] but is comparable to the magnitude of $-T\Delta S_b$. The increase in ΔE_b with temperature (Fig. S8, ESI†) can be rationalized by the enhancement in the electrostatic interaction strength due to the decrease in the water dielectric constant ε, as explained below.[17] The electrostatic free energy $\Delta G \propto \varepsilon^{-1}$ and $\varepsilon \propto T^{-\alpha}$, thus $\Delta G = -CT^\alpha$. Here C is a constant and the negative sign is due to $\Delta G < 0$ in our case. The entropy follows as $\Delta S = -\partial \Delta G / \partial T = C\alpha T^{\alpha-1}$. The internal energy results as $\Delta E = \Delta G + T\Delta S = -CT^\alpha + C\alpha T^\alpha = C(\alpha - 1)T^\alpha$. As long as the exponent $\alpha > 1$, ΔE is always positive and increases as T^α. For pure water $\alpha > 1$ at all temperatures (Fig. S9 in the ESI†). Although α slightly decreases with the addition of salt (*viz.* 1 M NaCl solution in Fig. S9, ESI†), α is significantly greater than 1 for the temperature regime investigated in our simulations, which explains the observed increase in ΔE_b with increasing temperature.

The temperature-dependent increase in ΔE_b follows the trend: Y^{3+} > Ca^{2+} > Na$^+$ ≈ Mg^{2+} (Fig. S8 in the ESI†). By changing temperature from 283 K to 323 K, the change in ΔE_b for Na$^+$, Ca^{2+}, Mg^{2+}, and Y^{3+} is found to be 1.82, 2.86, 1.71, and 5.11 kcal mol^{-1}, respectively.

The large value of ΔE_b can be understood by considering the energetic penalties associated with the desolvation of both the protein and cation. For example, ΔE_b for Y^{3+} ion at 300 K decreases from 7.50 to 3.71 kcal mol^{-1} if we exclude the contribution due to the dehydration of the second SS of Y^{3+} (Fig. S10B in the ESI†). Fig. S10 (ESI†) also highlights that the effect of the second SS is significant for the accurate description of solvation thermodynamics of cations, and cannot be neglected even for monovalent ions, *e.g.*, Na$^+$.

Preferential interaction coefficients

The interaction of ions with proteins, whether these are enriched or depleted from the protein surface, can be quantified by experimentally measuring the preferential interaction coefficient Γ_{23}. The thermodynamic definition of Γ_{23} is the change in chemical potential of the protein due to the addition of ions; it can also be expressed as the change in ion concentration to maintain constant chemical potential when a protein is added to the solution:[22]

$$\Gamma_{23} = -\left(\frac{\partial \mu_2}{\partial \mu_3}\right)_{m_2,T,P} = -\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_3,T,P}, \qquad (3)$$

where μ is the chemical potential, m is the molal concentration, and the subscripts 1, 2, and 3 stand for water, protein, and ion, respectively. Record $et~al.$,[23] based on the molal concentration definition, developed a two-domain molecular model for the estimation of Γ_{23} in terms of the difference in ion concentration in the local domain near the protein surface and the bulk solution as follows:

$$\Gamma_{23} = \left\langle N_3^{\text{local}} - N_1^{\text{local}}\left[\frac{N_3^{\text{bulk}}}{N_1^{\text{bulk}}}\right]\right\rangle, \qquad (4)$$

where N_i is the number of molecules of type i and $\langle \cdot \rangle$ represents the time average. For the calculation of Γ_{23} using eqn (4) a boundary or a distance cutoff needs to be chosen for defining the local and bulk domain, but the choice is arbitrary. Γ_{23} is instead estimated at each value of r, the distance from the protein surface, asuming that it is the boundary: $\Gamma_{23}(r) = \langle N_3(r) - N_1(r)[N_3^{\text{bulk}}/N_1^{\text{bulk}}] \rangle$. The distance r^* after which $\Gamma_{23}(r)$ becomes constant is defined as the actual boundary. In our simulations the total numbers of water molecules (N_1) and ions (N_3) are constant, thus the above expression for $\Gamma_{23}(r)$ is further simplified as:[24]

$$\Gamma_{23}(r) = \left\langle N_3(r) - N_1(r)\left[\frac{N_3 - N_3(r)}{N_1 - N_1(r)}\right]\right\rangle. \qquad (5)$$

In a salt solution, cations and anions are distributed around the protein. We obtain preferential interaction parameters for the cation $\Gamma_{2,+3}(r)$ and anion $\Gamma_{2,-3}(r)$ separately by using $N_{+/-3}(r)$ as the cation or anion distribution, respectively in eqn (5). $\Gamma_{2,+3}(r)$ and $\Gamma_{2,-3}(r)$ are shown for different salt solutions in Fig. 5. Experimentally, it is impossible, however, to separate the cationic and anionic contribution to the measured value of Γ_{23} for a salt solution. For a salt of monovalent cation and anion, the preferential interaction parameter is given by[23]

$$\Gamma_{23} = \frac{1}{2}(\Gamma_{2,+3} + \Gamma_{2,-3} - |Q_2|), \qquad (6)$$

where $|Q_2|$ is the protein's net charge that is subtracted from Γ_{23}, as Q_2 counterions (cations in case of BSA protein) are accumulated near the protein surface to neutralize its charge and do not contribute to the preferential interaction. For a salt of multivalent cation/anion, it is straight-forward to generalize eqn (6) by scaling $\Gamma_{2,+3}(r)$, $\Gamma_{2,-3}(r)$, and Q_2 with valency of the
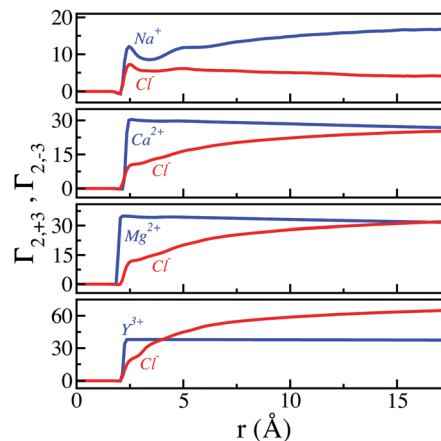

Fig. 5 Preferential interaction coefficients for cations $\Gamma_{2,+3}$ and anions $\Gamma_{2,-3}$ of the different salt solutions at 303 K as a function of the boundary distance cutoff r. $\Gamma_{2,+3}(r)$ reaches a saturation value rather quickly compared to $\Gamma_{2,-3}(r)$, for solutions containing multivalent cations.

anion z_-, valency of the cation z_+, and charge on the counterion z_+, respectively.

$$\Gamma_{23} = \frac{1}{2}\left(\frac{\Gamma_{2,+3}}{z_-} + \frac{\Gamma_{2,-3}}{z_+} - \frac{|Q_2|}{z_+}\right). \qquad (7)$$

For the BSA protein, using in eqn (7) $\Gamma_{2,+3}$ and $\Gamma_{2,-3}$ at $r^* = 17$ Å (by which all curves reach their respective saturation values as seen in Fig. 5), we obtain preferential interaction coefficients for different salts: NaCl ($\Gamma_{23} = 2.44$), CaCl$_2$ ($\Gamma_{23} = 15.67$), MgCl$_2$ ($\Gamma_{23} = 19.83$), and YCl$_3$ ($\Gamma_{23} = 26.87$). Positive values of Γ_{23} for all the different salt types reflect that these salt ions are attracted towards the protein surface, predominantly due to electrostatic interactions which tend to neutralize the protein charge. For salt containing multivalent ions, Γ_{23} is significantly larger than that for NaCl, which suggests that addition of trivalent ions in the protein solution affects the solution stability[25] and stabilizes protein dimer formation as seen in our simulations.

ζ-Potential of the protein and the protein–protein interaction

ζ-Potential measurements for a protein in an ionic solution report on charge compensation by the counterions and thus have direct implications for protein–protein association and the phase behavior of the solution. ζ-Potentials are defined by the electrophoretic mobility.[9,10,26] From the simulation data, we calculate the surface potential at one ionic diameter away from the protein surface (see Methods). Note that the surface potential typically serves as a good approximation for the ζ-potential for proteins and colloidal systems; however, the surface and ζ-potential values might differ significantly for extended surfaces with high surface charge densities.[27]

As shown in Fig. 6, the ζ-potential of the protein in the NaCl solution is negative at all temperatures, as expected based on the protein net charge of $-16~e$. In contrast, the ζ-potential is positive for all multivalent cation-chloride solutions at all temperatures, indicating sign reversal of the effective charge
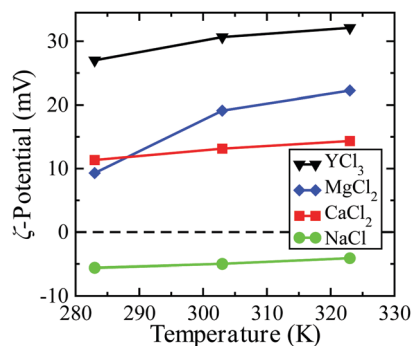
Fig. 6 Temperature dependence of the ζ-potential of the protein in various salt solutions. The lines are guide to the eye only.

of the protein (Fig. 6). This *charge inversion* phenomenon in the presence of multivalent cations is due to strong interactions of the cations predominantly with the COO^- groups of the protein's surface residues and can be rationalized by considering strong charge–charge correlations.[28] Note that similar charge reversal of proteins in the presence of trivalent cations has also been reported both in experiments[4,9,26] and simulations[13] as well as in a coarse-grained analytical model.[29] As shown in Fig. 6, with the increase in temperature, the ζ-potential of the protein increases for all cation types; the highest change is seen for $MgCl_2$, whereas the effect is minimal for NaCl. The ζ-potential of the protein at 283 K is higher in $CaCl_2$ solution than in $MgCl_2$ solution, and *vice versa* at 323 K. These observations are consistent with the trends for the temperature dependence of binding free energies of the different cations (Fig. 3A).

Protein–protein binding mediated by cation bridges

Protein aggregation seen in experiments[9] was hypothesized to be mediated by cation bridges.[6] To explicitly demonstrate the multivalent ion-mediated protein–protein binding, we have performed three independent simulations with different orientations of two BSA proteins in YCl_3 solution, as shown in Fig. 7A (left panel). In every simulation, we find that two BSA proteins, which are initially placed far apart, approach each other (see the timeseries of the total number of inter-protein residue–residue contacts in Fig. 7B) and eventually form a dimer mediated by 1–5 Y^{3+} ions (see snapshots in Fig. 7A [middle panel]). The Y^{3+} ion bridges remain stable over a 1 μs timescale, as evident from the time series plot for the number of bridging cations (Fig. 7C). A Y^{3+} ion bridge is stabilized by coordination of multiple caboxylate groups of each protein with the cation, as evident from snapshots in Fig. 7A (right panel). Note that even for the stable, Y^{3+} ion-bridged protein dimer complex, the relative orientation between the two proteins changes over time but very slowly (see the orientational autocorrelation function in Fig. 7D). This reveals the conformational flexibility of the protein dimer complex.

To compare monovalent and multivalent cations, we have also simulated the above three systems (shown in Fig. 7A) in NaCl solution, at the equivalent ionic strength as for YCl_3 solution. In sharp contrast to the case of YCl_3, we find that Na^+ ion bridges between two BSA proteins form transiently and remain stable only for 1–20 ns (Fig. 7E). These results

demonstrate the need for multivalent ions in protein cluster formations, in agreement with experiments.[4]

Discussion

The temperature behavior of the ζ-potential found in Fig. 6 is in qualitative agreement with the experiments.[9] As the ζ-potential is influenced by the number of surface-bound ions and the binding affinity of ions increases with temperature, the ζ-potential is expected to increase with temperature irrespective of the salt concentration of the solution. ζ-Potential values estimated from our simulations, however, are larger than that reported in the experiments[9] presumably because of the YCl_3 concentration difference. The ζ-potential increases with increasing YCl_3 concentration as found in experiments,[9] thus we expect the simulation and experimental results to match if the same salt concentration is used.

It should be noted that the YCl_3 concentration used in our simulations is 30 mM which is much higher than the 1 mM concentration used in the experiment. A direct comparison between all-atom simulations and the experiments[9] at low concentration of multivalent ions is rather difficult for the following reason. We consider a higher YCl_3 ion concentration in our simulations in order to obtain statistically converged results with sufficient number of ions. Obtaining well-converged results for proteins at low salt concentrations with enough number of ions would require significantly larger system sizes. Simulating such large systems is very demanding at the all-atom level, but it is feasible at a coarse-grained level as shown in a recent study.[13] However, the solvation effects, which are crucial for the accurate prediction of protein–ion binding thermodynamics, are not properly taken into account in such coarse-grained simulations.

The LCST phase behavior found in experiments[9] can be rationalized by the temperature dependence of the ζ-potential and the microscopic picture emerges from our simulations. For sufficiently low Y^{3+} concentration, at low temperatures due to the reduced binding affinity of counterions, the ζ-potential is expected to become negative (and large) and the proteins are expected to repel each other, keeping the solution stable. With increasing temperature, counterion binding affinity for the protein increases, and hence the ζ-potential increases and becomes positive at a sufficiently high temperature. In the temperature range (293–313 K) where the ζ-potential is small (−5 to +5 mV),[9] the proteins are predicted to attract each other, eventually causing the solution to phase separate into protein-rich and protein-poor phases.

The protein–protein binding at a low concentration of multivalent ions occurs *via* cation bridging, as shown in Fig. 7, as well as suggested from experiments.[6,9] A cation bridge formation—like the first step of a cation binding to the protein—requires desolvation of both the protein-bound cation and the surface residue of another protein that will bind to the cation. These processes involve the release of many tightly-bound water molecules to the bulk that results in a significant entropy gain, which contributes at least 10–15 kcal mol^{-1} (depending on the temperature) to the total free energy, as shown in Fig. 3B for
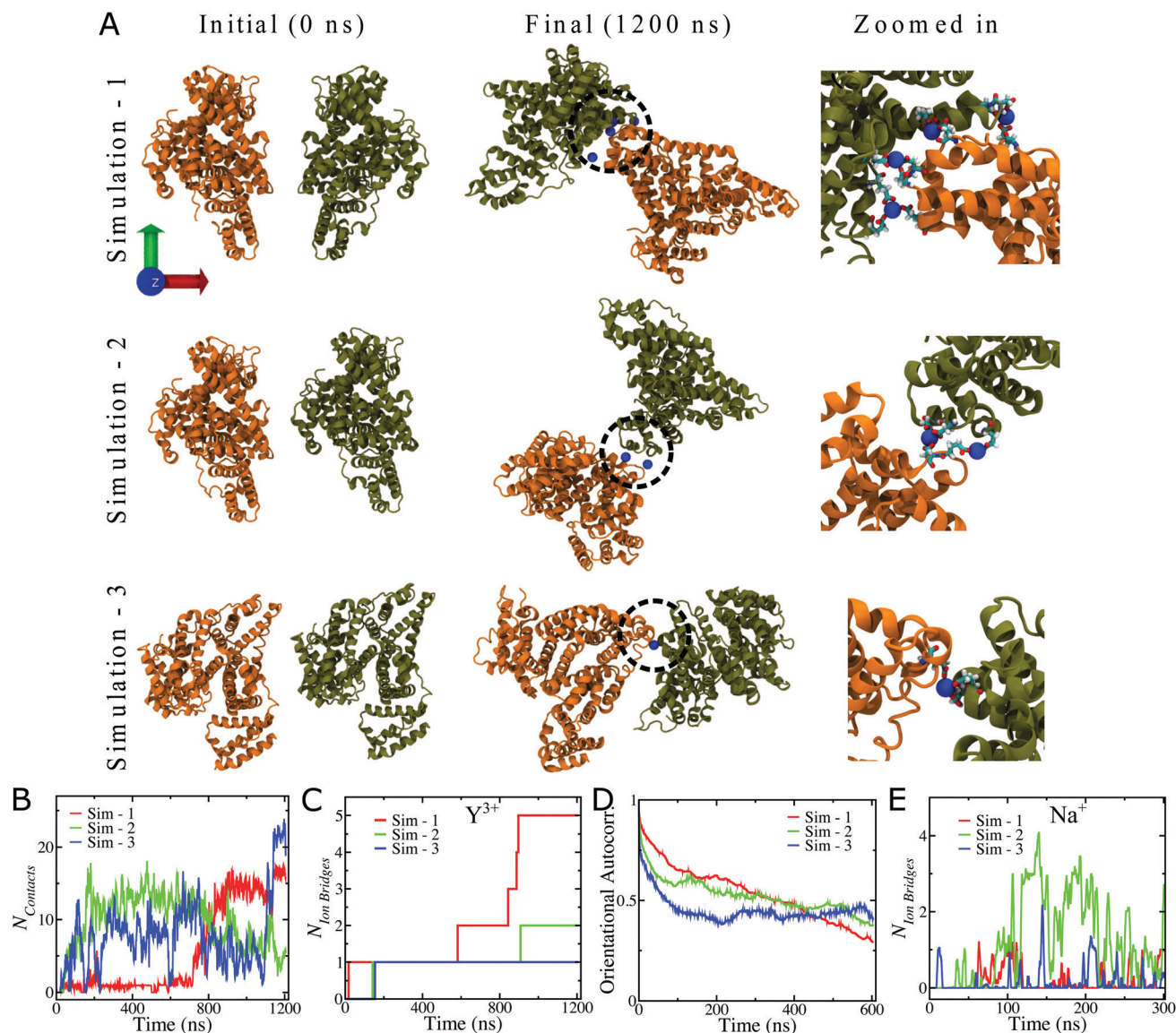
Fig. 7 Results for two BSA proteins (A–D) in 30 mM YCl$_3$ and (E) 180 mM Nacl solution at 303 K. (A) The initial and final configuration of two BSA proteins (represented in orange and tan). Bridging Y^{3+} ions in the final structure are shown as blue spheres, and amino acid residues of the two proteins involved in the formation of ion bridges are highlighted in the ball–stick representation (see the right panel for the zoomed-in version). Water and other ions are omitted for clarity. (B) Time series of the total number of inter protein residue–residue contacts. (C) Time series of the total number of Y^{3+} ion bridges that link the two proteins to form a dimer. (D) The average autocorrelation function for the relative orientation angles between the principal axes of two proteins as a function of time delay. (E) Time series of the total number of Na$^+$ ion bridges that form transiently between two BSA proteins in 180 mM NaCl solution at 303 K. All observables are defined in Methods.

a Y^{3+} ion binding. As multiple cation bridges are formed in a protein–protein binding (see Fig. 7, and also found in experiments[6]), the net entropy gain due to cation and protein desolvation more than compensates for the translational and rotational entropy losses of the proteins during protein–protein binding. Therefore, the LCST phase behavior[9] is entropy-driven.

Conclusions

In summary, by performing fully atomistic MD simulations of a BSA protein in different cation-chloride solutions (NaCl, CaCl$_2$,

MgCl$_2$, and YCl$_3$) and by calculating various entropy contributions, we demonstrate that multivalent cation binding to the protein is an entropy-driven phenomenon. The loss in entropy of a protein-bound cation is more than compensated by the entropy gain of water molecules due to the partial dehydration of both the cation and the cation-bound surface residue of the protein. A particularly interesting observation is the significant difference in the binding/unbinding kinetics of Ca^{2+} and Mg^{2+} ions (see Fig. S3, ESI†)—although having comparable binding free energies (Fig. 3A), which can be related to the recent finding that the ion–water exchange kinetics strongly depends on the size of a cation.[15] It will thus be interesting to investigate

in future simulation studies the universality of the ion size dependence of ion–protein binding kinetics and thermodynamics.

The ζ-potential calculation shows *charge inversion* of the protein in all solutions containing multivalent cations, but not in the monovalent NaCl solution (Fig. 6). The LCST phase behavior observed in the experiment[9] can be rationalized by considering the temperature-dependent increase in the ζ-potential of the protein and the associated *charge inversion* phenomenon. The protein–protein interaction involves: (i) the ion binding to the protein, and (ii) the subsequent protein–protein binding by cation bridging (Fig. 7). In both processes many tightly-bound water molecules are released to the bulk, which results in a thermodynamic driving force for the LCST behavior that is entropic in nature, in agreement with the experiment.[9]

This work shows that similarly to hydrophobic association, entropy plays a pivotal role in systems involving strong electrostatic interactions, revealing intriguing hydration and dielectric effects. Our results are important for the basic understanding of ion effects in soft matter and biology, and the insights gained here will be useful in studies of ion-mediated surface adsorption and crystallization of proteins. Moreover, molecular-level understanding of interactions of heavy metals—usually not found in healthy cells—with different biomolecules, as studied here, can provide insights for carcinogenicity and neurotoxicity induced by exposure to such environmental contaminants.

Methods

Model building and force field parameters

The initial structure of BSA protein was obtained from the crystal structure available in the protein data bank (PDB ID: 3V03). The charge or protonation state of each residue of the protein was chosen at neutral pH 7 depending on the residue's pK_a value, and the assigned charges were fixed over the simulation time. Note, however, that pK_a depends on the ionic strength (through the activity coefficients), and the reported pK_a values of amino acids are typically determined in a solution of high ionic strength.[30] In particular, the apparent pK_a values of carboxyl groups shift up slightly in the presence of multivalent cations at low salt concentrations. If the pH of the solution differs from 7 in an experiment due to the CO_2 content in air (which lowers the water pH down to 5.6[31]), this could make some of the acidic peptide groups less charged. But, the experiments described in Matsarskaia *et al.*[9] were performed in air and in ultrapure (MilliQ, 18.2 Mega Ohm) water which had previously been degassed under vacuum to eliminate the CO_2 contributions. Also, it is known from experiments that the addition of multivalent metal cations such as Al^{3+} and Fe^{3+} changes the pH of the solution due to hydrolysis of these cations, which can change the charge states of the protein surface residues. However, this effect is less significant for Y^{3+} ion.[10] Therefore, our assigned fixed charges of the protein residues at pH 7 for the different cations (Na^+, Ca^{2+}, Mg^{2+},

and Y^{3+}) is assumed to mimic the experimental conditions sufficiently well.

The ff14SB force field parameters[32] were used for the protein. The system was solvated with TIP3P[33] water model using the *xleap* module of the AMBER17 tools[34] in a way such that there exists at least 17 Å solvation shell in between the solute and simulation box wall. The final unit shell for simulation is a rectangular box of size $13.2 \times 13.2 \times 13.1\ nm^3$ that contains $\sim 200\,000$ atoms (Fig. 1A). The system was simulated in four different salt solutions, namely NaCl, $MgCl_2$, $CaCl_2$ and YCl_3. Depending on the ion type, an appropriate number of counterions were added to ensure the charge neutrality of the simulation box. To simulate the system at a specified salt concentration, enough numbers of counterions/coions, estimated from the mole fraction of counterions/coions and water, were further added to the system. For YCl_3, the system was simulated at 30 mM salt solution. For the other salts, the system was simulated at the equivalent ionic strength as in the case of YCl_3, *e.g.*, 180 mM for NaCl. Especially for multivalent ions, the electronic polarization effect contributes significantly to the total interaction energy of such an ion with another charged object. The recently developed Li/Merz ion parameters[35–37] with 12-6-4 Lennard-Jones (LJ)-type nonbonded interaction terms take care of the electronic polarization effect and have been shown to well reproduce the experimental measurables, such as the ion–oxygen (of water) distance, the ion–water coordination number, and the hydration free energy of mono- and multi-valent ions. We have provided in Table S1 in the ESI,† the structural parameters and entropy of ion hydration for the different ions calculated from our simulation data, which quantitatively match with the corresponding experimental values. Therefore, we used Li/Merz ion parameters for an accurate modeling of the ion–water and ion–protein interactions.

MD simulation details

All the simulations were performed using the PMEMD module of the AMBER14 package.[38] Periodic boundary condition was used for all the simulations. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm[39] that allowed the use of a time step of 2 fs for the integration of Newton's equation of motion. The temperature of the system was maintained using a Langevin thermostat[40] with the collision frequency of 5.0 ps^{-1}. Berendsen weak coupling method[41] was used to apply a pressure of 1 atm with isotropic position scaling with a pressure relaxation time constant of 2.0 ps. Particle mesh Ewald[42] sum was used to compute long-range electrostatic interactions with a real space cutoff of 10 Å. van der Waals and direct electrostatic interactions were truncated at the cutoff. The direct sum non-bonded list was extended to cutoff + "nonbond skin" (10 + 2 Å).

The solvated systems with harmonic restraints (force constant of 500 kcal mol^{-1} $Å^{-2}$) on the position of each atom of the protein were first subjected to 2000 steps of steepest descent energy minimization, followed by 1000 steps of conjugate gradient minimization to remove bad contacts present in the initially built systems. The restraints on the protein atoms were

sequentially decreased to zero during further 4000 steps of energy minimization. The energy minimized systems were then slowly heated from 10 K to the desired temperature in many steps during the first 210 ps of MD simulation. During this time, the solute particles were restrained to their initial positions using harmonic restraints with a force constant of 20 kcal mol^{-1} Å^{-2}. The first 2 ns of equilibration simulations were performed in the *NPT* ensemble to attain the proper water density. Simulations were then switched to the *NVT* ensemble for further production runs of 200–1450 ns, depending on cation types.

Data analysis

All the analyses were carried out by using home-written codes and/or the AMBER17 tools.[34] Images were rendered using the Visual Molecular Dynamics software.[43]

The free energy of ion binding, ΔG_b, was calculated using the expression

$$\Delta G_b = -k_B T \ln(C_{bI}/C_{fI}), \qquad (8)$$

where k_B is the Boltzmann constant, and C_{bI} and C_{fI} are the concentration of bound and free ions, respectively. The expressions for calculations of the concentrations are $C_{bI} = N_{bI}/V_s$ and $C_{fI} = N_{fI}/V_f$, where V_s is the volume of the shell around the protein surface where ions are considered as bound, N_{fI} (= total number of ions $-N_{bI}$) is the number of free ions, and V_f is the free volume available for ions. The volumes were calculated following the protocol described in ref. 44, by using the Gromacs program *gmx sasa*.[45] Further details on the volume calculation are provided in the ESI,† Section 3. The last 200 ns data for each ion type (the last 150 ns for Na$^+$) was taken for the calculation of ΔG_b, whereas the rest of the data served for the equilibration.

The reported entropy contributions in Fig. 4 and Fig. S6 (ESI†) were obtained by calculating absolute molar entropies for free and protein-bound ions, free and protein-bound water molecules, and water molecules in the first and second SS's of the cation by using the 2PT method.[19–21] To generate trajectories for 2PT calculations, simulations were restarted after 100–500 ns, depending on the ion type. 3 short (40 ps) *NVT* trajectories for each system at each temperature were generated with coordinates and velocities saved every 4 fs. To calculate the 2PT entropy for bound ions, we performed the analysis for all bound ions and got the average entropy per ion, similarly for bound water. In Fig. 4 and Fig. S6 (ESI†), each point and the corresponding error bar are the average and standard deviation of 3 different simulations, respectively.

The surface or ζ-potential was obtained by calculating as a function of r (the distance from the center of mass of the protein) the electrostatic potential profile, $\phi(r)$, for the system as follows.[46] $\phi(r)$ was calculated by solving the Poisson equation, *i.e.*, by carrying out a double integration of the charge density profile, $\rho(r)$, obtained from our MD simulation by using the following expression.

$$\phi(R) - \phi(r) = -\frac{1}{\varepsilon}\int_r^R dr_1 \frac{1}{r_1^2}\int_0^{r_1} dr_2 r_2^2 \rho(r_2), \qquad (9)$$

where R (= 65 Å) is the radius of the inscribed sphere within the rectangular MD simulation box and ε is the dielectric permittivity of water. A derivation of eqn (9) is given in the ESI,† Section 4. At a temperature T, ε was calculated from the Bjerrum length, λ_B, of water (= 7 Å) by using the relation: $\varepsilon = e^2/4\pi\lambda_B k_B T$, where e is the elementary charge. It should, however, be noted that for the temperature behavior of $\varepsilon \propto T^{-\alpha}$, the exponent α in experiments is close to $^3/_2$ but we take it to be 1 which is close to what is seen in simulations for a rather similar water model.[18] So the above approximation for $\varepsilon(T)$ is deemed to be good for our purpose. Finally, the ζ-potential was obtained as $\zeta = \phi(R) - \phi(R_h + 2r_c)$. Here, the hydrodynamic radius of the protein, R_h, was taken to be 36 Å,[47] and r_c's for the different cations (Fig. 1B) were taken to be 2.8 Å (Na$^+$), 2.7 Å (Ca^{2+}), 2.3 Å (Mg^{2+}) and 2.5 Å (Y^{3+}).

$N_{Contacts}$ shown in Fig. 7B is defined as the total number of inter-protein amino acid residue–residue contacts, and such a contact is counted if at least one pair of atoms from residues of two different proteins are within 3 Å.

$N_{IonBridges}$ shown in Fig. 7C is defined as the total number of ions bridging two different proteins, and an ion bridge is counted if an ion is present within 3 Å from both the proteins' surfaces.

The average orientational autocorrelation function shown in Fig. 7D is defined as

$$C(t) = \frac{1}{3}\sum_{i=1}^{3}\frac{\langle\cos\theta_i(0)\cdot\cos\theta_i(t)\rangle}{\langle\cos\theta_i(0)\cdot\cos\theta_i(0)\rangle}, \qquad (10)$$

with $\cos\theta_i(t) = \hat{e}_i^A\cdot\hat{e}_i^B$, where $\hat{e}_i^A$ and $\hat{e}_i^B$ are the unit vectors along the principal axes of proteins A and B, respectively, and the angular bracket represents the average over time origins.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 P. Jungwirth and P. S. Cremer, *Nat. Chem.*, 2014, **6**, 261–263.
2 H. I. Okur, J. Hladilkova, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer and P. Jungwirth, *J. Phys. Chem. B*, 2017, **121**, 1997–2014.
3 O. Matsarskaia, F. Roosen-Runge and F. Schreiber, *ChemPhysChem*, 2020, **21**, 1742.

4   F. Zhang, M. Skoda, R. Jacobs, S. Zorn, R. A. Martin, C. Martin, G. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *Phys. Rev. Lett.*, 2008, **101**, 148101.

5   M. R. Fries, D. Stopper, M. K. Braun, A. Hinderhofer, F. Zhang, R. M. Jacobs, M. W. Skoda, H. Hansen-Goos, R. Roth and F. Schreiber, *Phys. Rev. Lett.*, 2017, **119**, 228001.

6   F. Zhang, G. Zocher, A. Sauter, T. Stehle and F. Schreiber, *J. Appl. Crystallogr.*, 2011, **44**, 755–762.

7   F. Zhang, F. Roosen-Runge, A. Sauter, R. Roth, M. W. Skoda, R. M. Jacobs, M. Sztucki and F. Schreiber, *Faraday Discuss.*, 2012, **159**, 313–325.

8   F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. Skoda, R. M. Jacobs, M. Stzucki and F. Schreiber, *Soft Matter*, 2012, **8**, 1313–1316.

9   O. Matsarskaia, M. K. Braun, F. Roosen-Runge, M. Wolf, F. Zhang, R. Roth and F. Schreiber, *J. Phys. Chem. B*, 2016, **120**, 7731–7736.

10 F. Roosen-Runge, B. S. Heck, F. Zhang, O. Kohlbacher and F. Schreiber, *J. Phys. Chem. B*, 2013, **117**, 5777–5787.

11 D. A. Capdevila, K. A. Edmonds, G. C. Campanello, H. Wu, G. Gonzalez-Gutierrez and D. P. Giedroc, *J. Am. Chem. Soc.*, 2018, **140**, 9108–9119.

12 B. Hess and N. F. van der Vegt, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 13296–13300.

13 C. Pasquier, M. Vazdar, J. Forsman, P. Jungwirth and M. Lund, *J. Phys. Chem. B*, 2017, **121**, 3000–3006.

14 C. Yigit, J. Heyda and J. Dzubiella, *J. Chem. Phys.*, 2015, **143**, 08B606_1.

15 Y. Lee, D. Thirumalai and C. Hyeon, *J. Am. Chem. Soc.*, 2017, **139**, 12334–12337.

16 D. Chandler, *Nature*, 2005, **437**, 640.

17 J. N. Israelachvili, *Intermolecular and Surface Forces*, Academic press, 2015.

18 F. Sedlmeier and R. R. Netz, *J. Chem. Phys.*, 2013, **138**, 03B609.

19 S.-T. Lin, M. Blanco and W. A. Goddard III, *J. Chem. Phys.*, 2003, **119**, 11792–11805.

20 S.-T. Lin, P. K. Maiti and W. A. Goddard III, *J. Phys. Chem. B*, 2010, **114**, 8191–8198.

21 T. A. Pascal, S.-T. Lin and W. A. Goddard III, *Phys. Chem. Chem. Phys.*, 2011, **13**, 169–181.

22 V. Pierce, M. Kang, M. Aburi, S. Weerasinghe and P. E. Smith, *Cell Biochem. Biophys.*, 2008, **50**, 1–22.

23 M. T. Record Jr and C. F. Anderson, *Biophys. J.*, 1995, **68**, 786–794.

24 D. Shukla, C. Shinde and B. L. Trout, *J. Phys. Chem. B*, 2009, **113**, 12546–12554.

25 F. Zhang, S. Weggler, M. J. Ziller, L. Ianeselli, B. S. Heck, A. Hildebrandt, O. Kohlbacher, M. W. Skoda, R. M. Jacobs and F. Schreiber, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 3450–3457.

26 A. Kubíčková, T. Křížek, P. Coufal, M. Vazdar, E. Wernersson, J. Heyda and P. Jungwirth, *Phys. Rev. Lett.*, 2012, **108**, 186101.

27 Y. Uematsu, R. R. Netz and D. J. Bonthuis, *Langmuir*, 2018, **34**, 9097–9113.

28 A. Y. Grosberg, T. Nguyen and B. Shklovskii, *Rev. Mod. Phys.*, 2002, **74**, 329–345.

29 F. Roosen-Runge, F. Zhang, F. Schreiber and R. Roth, *Sci. Rep.*, 2014, **4**, 1–5.

30 J. Reijenga, A. van Hoof, A. van Loon and B. Teunissen, *Anal. Chem. Insights*, 2013, **8**, ACI-S12304.

31 Y. Uematsu, D. J. Bonthuis and R. R. Netz, *J. Phys. Chem. Lett.*, 2017, **9**, 189–193.

32 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

33 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

34 D. A. Case, D. Cerutti, I. T. E. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. S. W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, D. York and P. Kollman, *AMBER, 2017*, University of California, San Francisco, 2017.

35 P. Li and K. M. Merz Jr, *J. Chem. Theory Comput.*, 2013, **10**, 289–297.

36 P. Li, L. F. Song and K. M. Merz Jr, *J. Phys. Chem. B*, 2014, **119**, 883–895.

37 P. Li, L. F. Song and K. M. Merz Jr, *J. Chem. Theory Comput.*, 2015, **11**, 1645–1657.

38 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.

39 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

40 W. F. Van Gunsteren and H. J. C. Berendsen, *Mol. Simul.*, 1988, **1**, 173–185.

41 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

42 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

43 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

44 O. Becconi, E. Ahlstrand, A. Salis and R. Friedman, *Isr. J. Chem.*, 2017, **57**, 403–412.

45 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19–25.

46 P. K. Maiti and R. Messina, *Macromolecules*, 2008, **41**, 5002–5006.

47 Y. Li, G. Yang and Z. Mei, *Acta. Pharm. Sin. B*, 2012, **2**, 53–59.