

Supporting Information:

Role of Entropy in Determining the Phase Behavior of Protein Solutions Induced by Multivalent Ions

Anil Kumar Sahoo,^{*,†,‡,¶} Frank Schreiber,[§] Roland R. Netz,^{¶,||} and Prabal K.
Maiti^{*,†}

[†]*Center for Condensed Matter Theory, Department of Physics, Indian Institute of Science,
Bangalore-560012, India*

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

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

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

^{||}*Department of Physics, Indian Institute of Science, Bangalore-560012, India*

E-mail: 201992kumarsahoo@gmail.com(A.K.S.); maiti@iisc.ac.in(P.K.M)

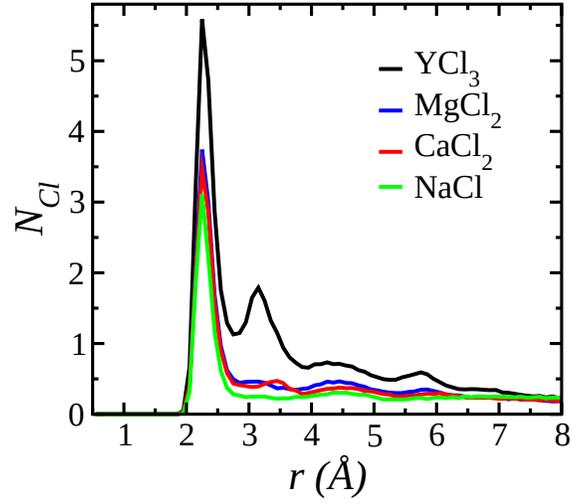


Figure S1: For each of the different ionic solutions, the total number of Cl^- ions (N_{Cl}) found within a shell of width $dr = 0.1 \text{ \AA}$, present at a shortest distance r from the protein surface, is shown as a function of r for the simulation performed at 303 K. $N_{Cl}(r)$ for each ionic solution type is averaged over the last 100 ns of the simulation time.

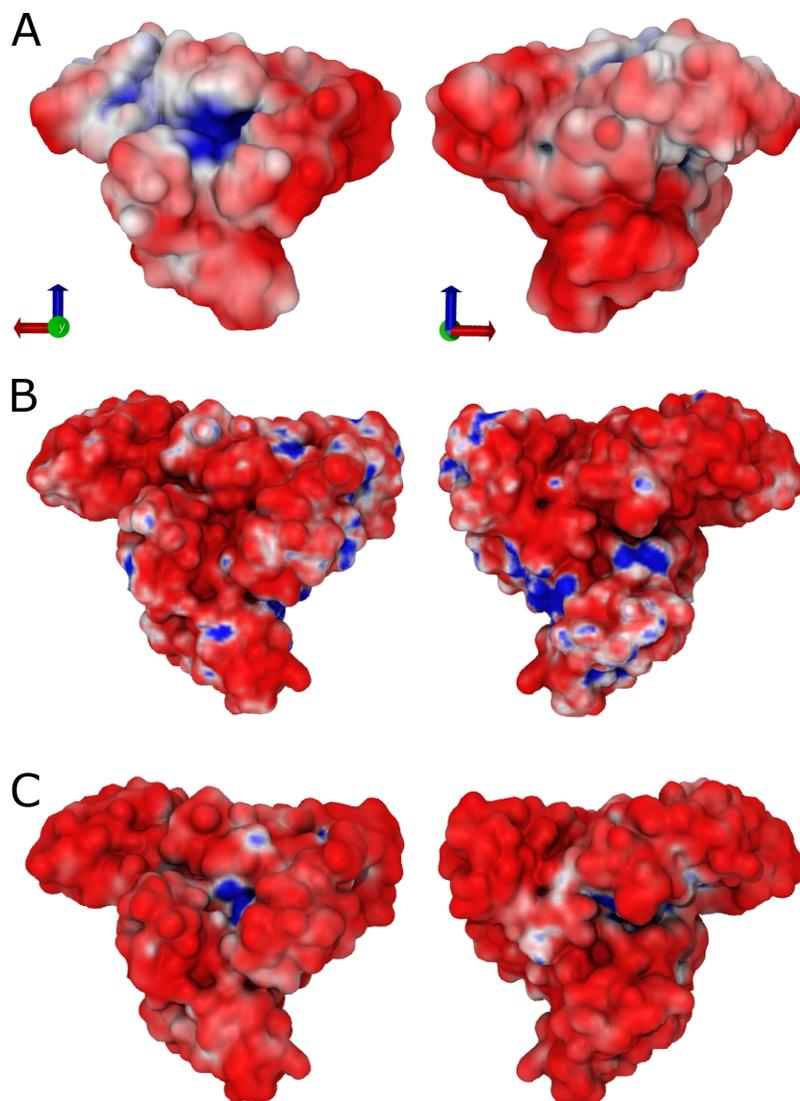


Figure S2: (A) Electrostatic potential map for BSA, left: front view, right: back view. Red represents negative potential, whereas blue represents positive potential. (B) Na⁺ and (C) Cl⁻ number density map obtained from ion distributions within a 5 Å shell from the protein surface sampled in the last 100 ns of the simulation at 303 K. Red represents lower density, whereas blue represents higher density.

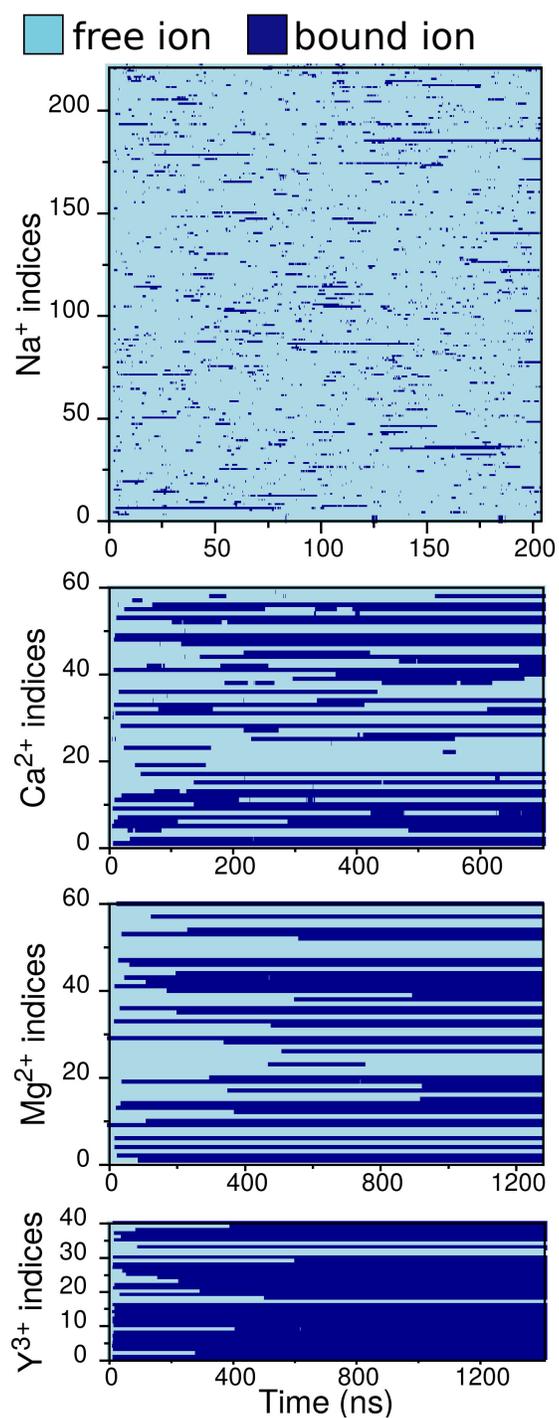


Figure S3: Time series showing events of cation binding and unbinding from the protein surface at 303 K for each Na^+ , Ca^{2+} , Mg^{2+} , and Y^{3+} . The length of a continuous dark-blue (light-blue) line represents the time for which a cation remains bound (unbound) to the protein surface.

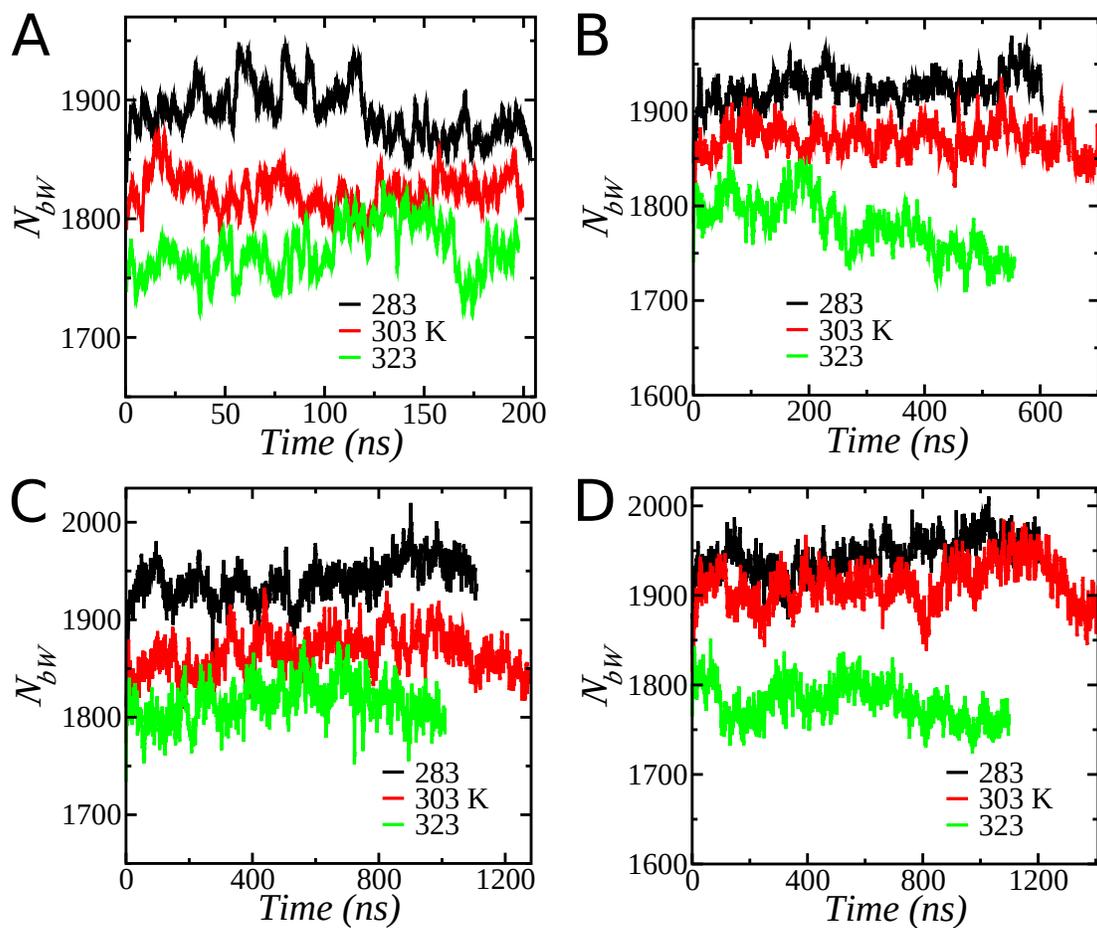


Figure S4: Time series of the number of protein-bound water molecules at different temperatures for NaCl (A), CaCl₂ (B), MgCl₂ (C), and YCl₃ (D) solutions.

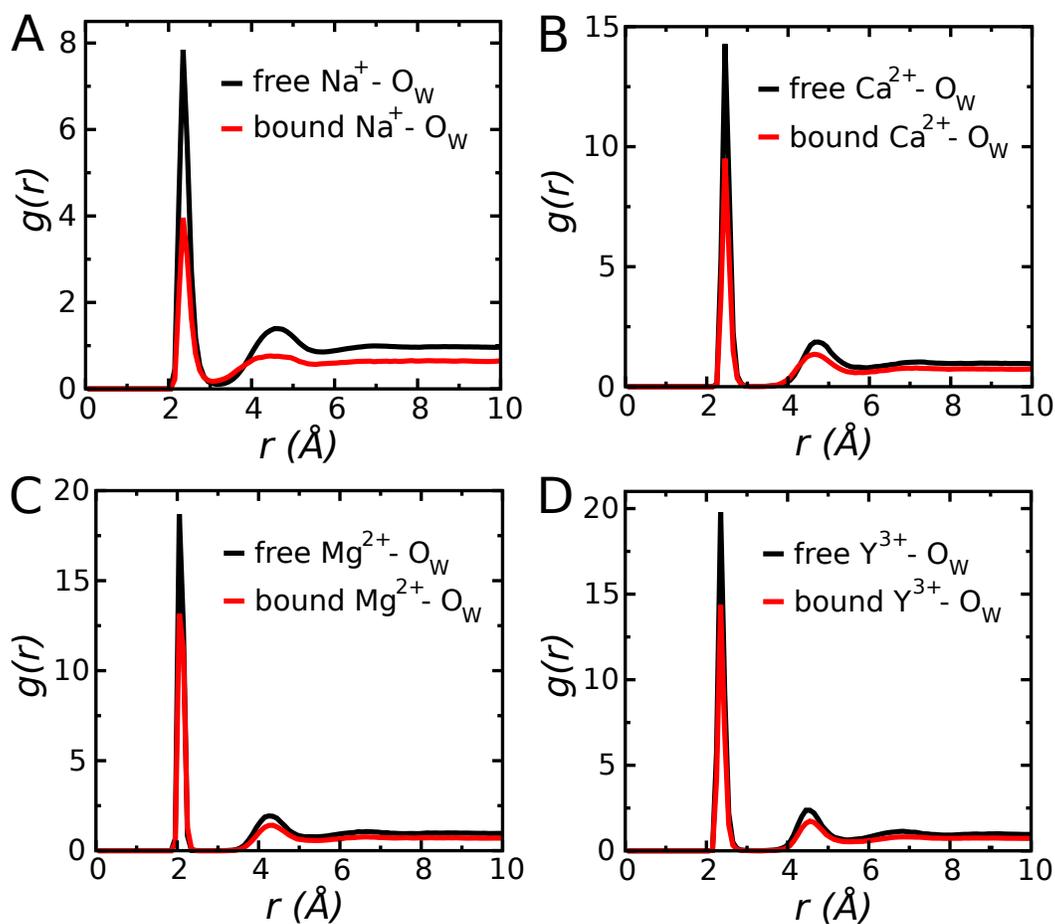


Figure S5: The radial distribution functions (RDF) at 303 K for oxygen atoms of water around Na^+ (A), Ca^{2+} (B), Mg^{2+} (C), and Y^{3+} (D) ions. In each case, the results for the cation free in solution and the cation bound to the protein surface are shown. Water molecules are released from both the first and second solvation shells of each cation, when the cation binds to the protein surface.

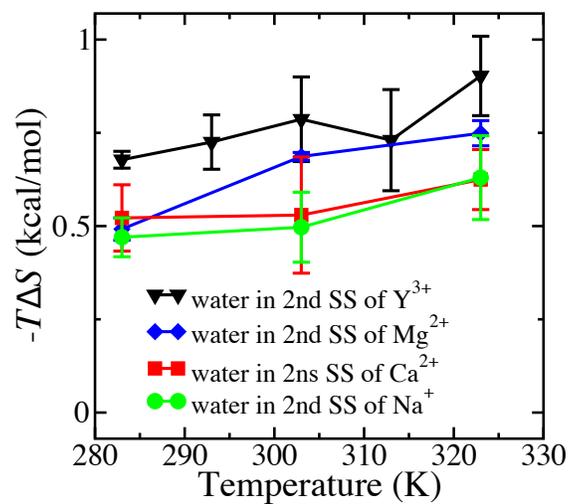


Figure S6: Temperature dependence of the difference in entropies of a water in the second solvation shell of a cation and a water in bulk. Error bars represent the standard deviation. The lines are for guiding the eye.

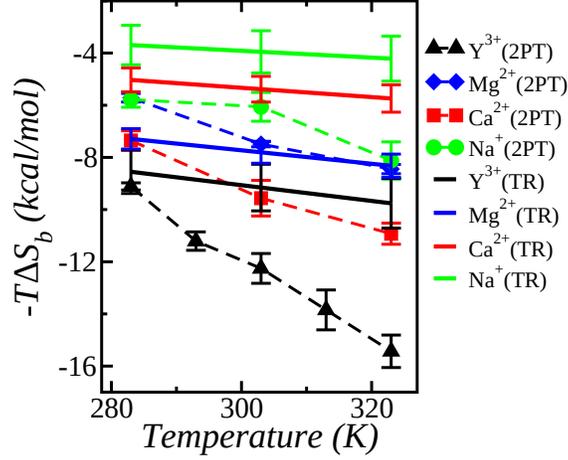


Figure S7: Comparison of the total entropy contribution ($-T\Delta S_b$) obtained from the 2PT method and that obtained from the temperature dependence of the ion-binding free energy ΔG_b (see Fig. 3A in the main text) using the thermodynamic relation (TR), $\Delta S_b = -\partial\Delta G_b/\partial T$. The error bar for the 2PT method represents the standard deviation of four independent calculations at each temperature, whereas the error bar for TR denotes the standard linear-regression error associated with the estimation of the slope of temperature versus binding energy data. Note that for each cation type, the entropy obtained from each method is negative and sufficiently large to drive the ion binding. Given the higher “true” error (compared to the computed error here) usually associated with the entropy obtained from TR, we deem the agreement between the widely different methods sufficient for our purposes.

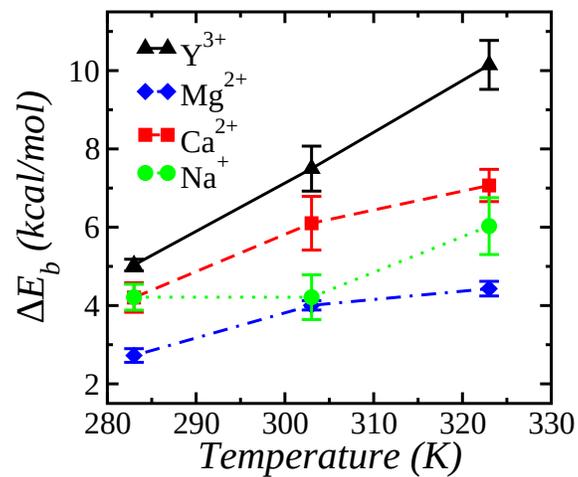


Figure S8: Temperature dependence of the protein–ion binding energy ΔE_b for the different cations. The different lines are for guiding the eye. The error bars are the propagation errors coming from the errors in ΔG_b and ΔS_b (see Eq. 1 in the main text).

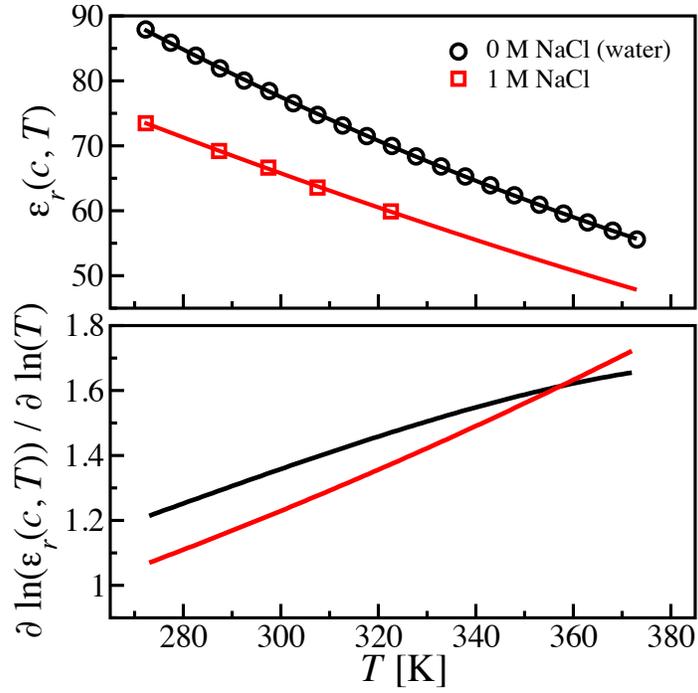


Figure S9: (top) Temperature dependence of the relative dielectric constant ϵ_r of NaCl solutions. Open circles are experimental data for pure water,^{S1} whereas open squares are experimental data for 1 M NaCl solution.^{S2} Curves represent polynomial fits to the data. (bottom) First logarithmic derivative of ϵ_r which determines the exponent α of the scaling law: $\epsilon_r(c, T) \propto T^{-\alpha(c)}$.

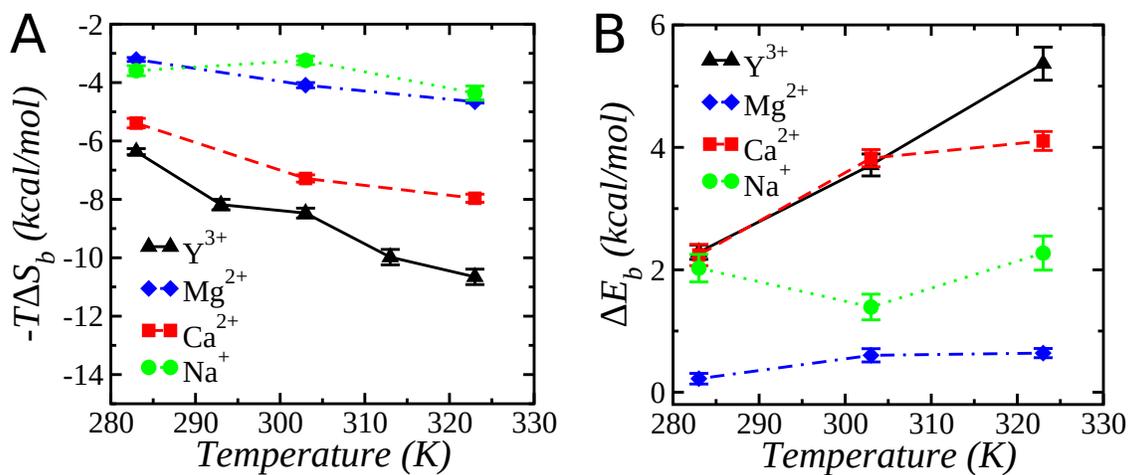


Figure S10: Temperature dependence of the total entropy contribution $-T\Delta S_b$ (A) and the total energy contribution ΔE_b (B) for the different cations, where the respective contributions due to the dehydration of the second solvation shell are not included. The different lines are for guiding the eye. For comparison where all the contributions are included, see Fig. 3B (in the main text) and Fig. S8 for $-T\Delta S_b$ and ΔE_b , respectively.

1. Temperature dependence of the free energy

The free energy is given by

$$F = U - TS = U(T) - TS(T) = F(T, V),$$

where U is the internal energy, S is the entropy, V is the volume, and T is the temperature.

Taking the derivative of F

$$\implies \left. \frac{\partial F(T, V)}{\partial T} \right|_V = -S(T) = U'(T) - S(T) - TS'(T).$$

In the above equation, we have used the relation: $U'(T) = C_V = TS'(T)$, where C_V is the specific heat at constant V . Therefore, the change of U with T does not enter the T -dependence of F .

2. Two-phase thermodynamic (2PT) method for entropy calculation

The 2PT method was developed by Lin *et al.*^{S3,S4} The central hypothesis of the 2PT method is that the density of states (DoS) of a fluid can be treated as a combination of gas and solid-like components. The DoS of a fluid has a zero-frequency diffusive mode $S(0)$, similar to a gas, and a maximum at some finite frequency followed by an exponential decay at higher frequencies, similar to a solid.

Lin *et al.*^{S3} showed that the thermodynamic properties can be estimated by treating the DoS of a fluid as a sum of solid-like ($S^s(\nu)$) and gas-like ($S^g(\nu)$) contributions. Thermodynamic quantities for a solid can be estimated by treating its vibrational modes as a system of noninteracting harmonic oscillators, as in the Debye model.^{S5} The gas part is described as a low-density hard-sphere fluid. The velocity autocorrelation function decays exponentially

for this model,^{S5} and hence the DoS can be calculated analytically. Thus, the calculation of entropy for solid and gas requires knowledge of the DoS.

The translational density of states $S(\nu)$ of a system is defined as the mass-weighted sum of atomic spectral densities $s_j^k(\nu)$

$$S(\nu) = \frac{2}{k_B T} \sum_{j=1}^N \sum_{k=1}^3 m_j s_j^k(\nu), \quad (\text{S1})$$

where m_j is the mass of the j th atom, k refers to the three Cartesian directions, and $s_j^k(\nu)$ is given by:

$$s_j^k(\nu) = \lim_{\tau \rightarrow \infty} \frac{\left| \int_{-\tau}^{\tau} \nu_j^k(t) e^{-i2\pi\nu t} dt \right|^2}{\int_{-\tau}^{\tau} dt} = \lim_{\tau \rightarrow \infty} \frac{\left| \int_{-\tau}^{\tau} \nu_j^k(t) e^{-i2\pi\nu t} dt \right|^2}{2\tau}. \quad (\text{S2})$$

Here, $\nu_j^k(t)$ is the k th component of the velocity of atom j . It can be shown that the atomic spectral density $s_j^k(\nu)$ can be obtained from the Fourier transform of the velocity auto-correlation function (VACF) $c_j^k(t)$ ^{S3}

$$s_j^k(\nu) = \lim_{\tau \rightarrow \infty} \int_{-\tau}^{\tau} c_j^k(t) e^{-i2\pi\nu t} dt, \quad (\text{S3})$$

where $c_j^k(t)$ is given by:

$$c_j^k(t) = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} \nu_j^k(t+t') \nu_j^k(t') dt'. \quad (\text{S4})$$

Thus, Eq. S1 can be rewritten as:

$$S(\nu) = \frac{2}{k_B T} \lim_{\tau \rightarrow \infty} \int_{-\tau}^{\tau} \sum_{j=1}^N \sum_{k=1}^3 m_j c_j^k(t) e^{-i2\pi\nu t} dt. \quad (\text{S5})$$

More generally, it can be written as:

$$S(\nu) = \frac{2}{k_B T} \lim_{\tau \rightarrow \infty} \int_{-\tau}^{\tau} C(t) e^{-i2\pi\nu t} dt. \quad (\text{S6})$$

In the above equation, $C(t)$ can be either the mass-weighted translational VACF determined from the center of mass velocity $V_i^{cm}(t)$ of the i th molecule,

$$C_T(t) = \sum_{i=1}^N \langle m_i V_i^{cm}(t) \cdot V_i^{cm}(0) \rangle \quad (\text{S7})$$

or the moment-of-inertia weighted angular velocity auto-correlation function

$$C_R(t) = \sum_{i=1}^3 \sum_{i=1}^N \langle I_{ij} \omega_{ij}(t) \omega_{ij}(0) \rangle, \quad (\text{S8})$$

where I_{ij} and ω_{ij} are the j th components of the moment of inertia tensor and the angular velocity of the i th molecule, respectively. One can obtain the translational or rotational DoS depending on the use of $C_T(t)$ or $C_R(t)$ in Eq. S6.

In the 2PT method, the DoS is decomposed into a gas-like diffusive component and a solid-like nondiffusive component, $S(\nu) = S^g(\nu) + S^s(\nu)$, using the fluidity factor f which is a measure of the fluidity of a system. f is estimated in terms of the dimensionless diffusivity Δ using the universal equation:^{S3}

$$2\Delta^{-9/2} f^{15/2} - 6\Delta^{-3} f^5 - \Delta^{-3/2} f^{7/2} + 6\Delta^{-3/2} f^{5/2} + 2f - 2 = 0. \quad (\text{S9})$$

The diffusivity Δ can be uniquely determined for a thermodynamic state of the system using the equation:

$$\Delta(T, \rho, m, S_0) = \frac{2S_0}{9N} \left(\frac{6}{\pi}\right)^{2/3} \left(\frac{\pi k_B T}{m}\right)^{1/2} \rho^{1/3}, \quad (\text{S10})$$

where $S_0 = S(0)$ is the zero-frequency component of the DoS function (translational or rotational). Knowing f from Eqs. S9 and S10, the gas-like diffusive component of the DoS

can be obtained using a hard-sphere diffusive model:

$$S^g(\nu) = \frac{S_0}{1 + \left[\frac{\pi S_0 \nu}{6fN} \right]^2}. \quad (\text{S11})$$

Lin *et al.*^{S3} used the gas–solid decomposition scheme only for the translational DoS of monoatomic fluids. In a later work, Lin *et al.*^{S4} showed that for polyatomic fluids, the rotational entropy can be computed if the decomposition scheme is used for the rotational DoS as well. Separate fluidity factors f s are determined for the translational and rotational DoS using the translational and rotational diffusivities in Eq. S9. Then, the gas-like component of entropy is calculated using Eq. S11 with S_0 being $S_{tran}(0)$ or $S_{rot}(0)$ for the translational and rotational cases, respectively. Once such decomposition of DoS is done, each thermodynamic quantity A_m can be computed from the solid-like and gas-like DoS functions with the corresponding weight functions as follows.

$$A_m = \beta^{-1} \left[\int_0^\infty d\nu S_m^g(\nu) W_{A,m}^g + \int_0^\infty d\nu S_m^s(\nu) W_{A,m}^s \right], \quad (\text{S12})$$

where m can be translational, rotational or vibrational. The weight functions are provided in Ref.^{S4} For the rigid TIP3P^{S6} water model used in our simulations, the contribution due to intra-molecular vibration is zero.

3. Definition of volume for the calculation of ion concentration

Following Ref.,^{S7} the volume of the shell surrounding the protein surface (V_s) where ions are considered as bound is defined as $V_s = V_{ps} - V_{prot}$. V_{prot} is the volume of the protein calculated by rolling a small probe sphere of radius 0.5 Å on the protein. V_{ps} is the volume occupied by the protein and the shell around the protein that contains the bound ions, and

V_{ps} is calculated by rolling around the protein surface a probe sphere of radius r_{ps} , which is equivalent to the contact distance of a bound ion from the protein surface. The volume available for free ions (V_f) is defined as $V_f = V_{box} - V_{ps}$, where V_{box} is the volume of the simulation box. Note that the calculation of V_{ps} using this method is *ad hoc* and influences the concentration of bound and free ions, and hence the value of ΔG_b (*cf.* Eq. 3 in the main text). We, therefore, varied r_{ps} to get an optimized probe radius, such that the calculated ΔG_b matches the experimental ΔG_b for Y^{3+} , as shown in the figure below.

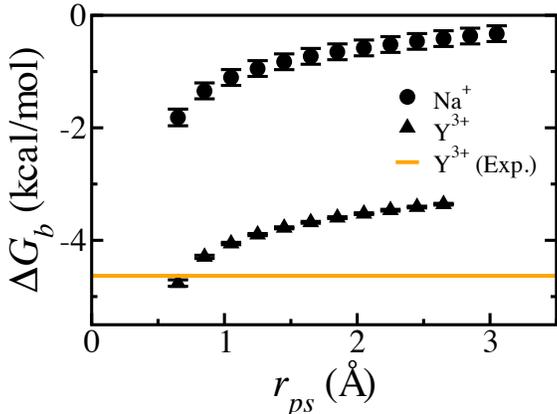


Figure S11: Dependence of the binding free energy ΔG_b on the probe radius r_{ps} used for defining the shell volume V_s . The experimental data for Y^{3+} is taken from Matsarskaia *et al.*^{S8}

4. The surface or ζ -potential calculation from the simulation data

For the estimation of the surface or ζ -potential of the protein in a salt solution, we consider Poisson's equation

$$\nabla^2 \phi = -\frac{\rho}{\epsilon}, \quad (\text{S13})$$

in *spherical polar* coordinates. Here, ϕ is the electrostatic potential, ρ is the charge density, and ϵ is the dielectric permittivity of water. For simplicity, we assume that the protein is a spherically isotropic object. To justify this, we calculate the asphericity, δ , of BSA protein

($\delta = 1 - 3[I_x I_y + I_y I_z + I_x I_z]/[I_x + I_y + I_z]^2$, where I_x , I_y , and I_z are the principal moments of inertia) and find δ to be very small ($= 0.0196 \pm 0.0016$). Now, ϕ and ρ are functions of only the radial distance r . Under these considerations, Eq. S13 becomes

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \phi(r) = -\frac{\rho(r)}{\epsilon} \quad (\text{S14})$$

$$\implies \frac{\partial}{\partial r} (r^2 E(r)) = \frac{\rho(r)}{\epsilon} r^2, \quad (\text{S15})$$

where the electric field $E = -\frac{\partial\phi(r)}{\partial r}$. Integrating both sides of the above equation from 0 to r_1

$$r_1^2 E(r_1) = \frac{1}{\epsilon} \int_0^{r_1} dr_2 \rho(r_2) r_2^2 \quad (\text{S16})$$

$$\implies -\frac{\partial\phi(r)}{\partial r} = \frac{1}{\epsilon} \frac{1}{r_1^2} \int_0^{r_1} dr_2 \rho(r_2) r_2^2. \quad (\text{S17})$$

To obtain the electrostatic potential profile $\phi(r)$, integrating both sides of the above equation from r to R (the maximum radial distance possible due to the finite size of the simulation box under *periodic boundary condition*)

$$\implies \phi(R) - \phi(r) = -\frac{1}{\epsilon} \int_r^R dr_1 \frac{1}{r_1^2} \int_0^{r_1} dr_2 \rho(r_2) r_2^2. \quad (\text{S18})$$

Note that this equation is the same as Eq. 9 in the main text. The surface or ζ -potential of the protein is defined as the electrostatic potential at one ionic diameter away from the protein surface. So, the ζ -potential is evaluated from the above equation using the charge density $\rho(r)$ obtained from the simulation as, $\zeta = \phi(R) - \phi(R_h + 2r_c)$. Here, R_h is the hydrodynamic radius of the protein obtained to be 36 Å from dynamic light scattering experiments,^{S9} and r_c is the effective radius of the counterion. The values of the parameters used to obtain the ζ -potential are provided in the Methods section in the main text.

Table S1: Structural parameters (such as the coordination number of the ion, N_{hyd} , and the ion–oxygen distance for water present in the ion’s 1st solvation shell, d_{I-O}^{1stSS} , and 2nd solvation shell d_{I-O}^{2ndSS}) and entropy (ΔS_{hyd}) of ion hydration obtained from simulations. The corresponding experimental values are given within the brackets. The calculated values for d_{I-O}^{1stSS} , d_{I-O}^{2ndSS} , and N_{hyd} are in quantitative agreement with the corresponding experimental values.^{S10,S11} The computed values of ΔS_{hyd} match well with the experimental values^{S12} for all the ions, except for Na⁺. The $\sim 50\%$ overestimation in the calculated ΔS_{hyd} for Na⁺ may be due to the inaccurate estimation of the entropy of a water molecule present in the 2nd SS of Na⁺. The lifetime of a water molecule in the 2nd SS of Na⁺ ion is 10–15 ps. Within this short time period, the velocity–velocity autocorrelation—which is needed to obtain the spectral density-of-states that serves as an input for the 2PT entropy calculation^{S3,S4}— is not well converged.

Ion	d_{I-O}^{1stSS} (Å)	d_{I-O}^{2ndSS} (Å)	N_{hyd}	ΔS_{hyd} (J mol ⁻¹ K ⁻¹)
Na ⁺	2.35±0.05 (2.34)	4.55±0.05 (—)	5.7 (5.6±0.3)	-168.08±20.94 (-111.2)
Ca ²⁺	2.45±0.05 (2.46)	4.65±0.05 (4.58)	8.0 (8)	-270.57±39.89 (-252.4)
Mg ²⁺	2.05±0.05 (2.09)	4.25±0.05 (4.35)	6.0 (6)	-316.09±2.72 (-331.2)
Y ³⁺	2.35±0.05 (2.37)	4.50±0.05 (4.40)	9.0 (8)	-460.77±29.96 (-482.5)
Cl ⁻	3.15±0.05 (3.14)	5.05±0.05 (4.99)	7.2 (7)	-71.56±9.43 (-75.7)

Table S2: In the process of a cation binding to the protein, the average numbers of water molecules released from the protein surface and the first and second solvation shells (SS) of the cation are given at each temperature for the different cations.

System	Temperature	Protein	1 st SS of cation	2 nd SS of cation
Protein in NaCl	283	4.37	2.47	4.65
	303	4.55	2.52	5.68
	323	4.39	2.82	5.96
Protein in CaCl ₂	283	4.74	2.32	3.76
	303	5.63	2.62	4.30
	323	5.34	3.17	4.74
Protein in MgCl ₂	283	2.68	1.04	5.08
	303	2.90	1.15	4.95
	323	2.79	1.26	5.06
Protein in YCl ₃	283	4.16	2.09	4.05
	293	5.07	2.49	4.17
	303	5.10	2.47	4.81
	313	5.61	2.69	5.29
	323	5.40	2.87	5.29

References

- (S1) Robinson, R. A.; Stokes, R. H. *Electrolyte solutions*; Courier Corporation, 2002.
- (S2) Giese, K.; Kaatze, U.; Pottel, R. Permittivity and dielectric and proton magnetic relaxation of aqueous solutions of the alkali halides. *J. Phys. Chem.* **1970**, *74*, 3718–3725.
- (S3) Lin, S.-T.; Blanco, M.; Goddard III, W. A. The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of Lennard-Jones fluids. *J. Chem. Phys.* **2003**, *119*, 11792–11805.
- (S4) Lin, S.-T.; Maiti, P. K.; Goddard III, W. A. Two-phase thermodynamic model for efficient and accurate absolute entropy of water from molecular dynamics simulations. *J. Phys. Chem. B* **2010**, *114*, 8191–8198.
- (S5) McQuarrie, D. A. *Statistical mechanics*; Harper and Row, 1976.
- (S6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (S7) Becconi, O.; Ahlstrand, E.; Salis, A.; Friedman, R. Protein-ion Interactions: Simulations of Bovine Serum Albumin in Physiological Solutions of NaCl, KCl and LiCl. *Isr. J. Chem.* **2017**, *57*, 403–412.
- (S8) Matsarskaia, O.; Braun, M. K.; Roosen-Runge, F.; Wolf, M.; Zhang, F.; Roth, R.; Schreiber, F. Cation-Induced Hydration Effects Cause Lower Critical Solution Temperature Behavior in Protein Solutions. *J. Phys. Chem. B* **2016**, *120*, 7731–7736.
- (S9) Li, Y.; Yang, G.; Mei, Z. Spectroscopic and Dynamic Light Scattering Studies of the Interaction Between Pterodonic Acid and Bovine Serum Albumin. *Acta. Pharm. Sin. B* **2012**, *2*, 53–59.

- (S10) Marcus, Y. Effect of ions on the structure of water: structure making and breaking. *Chem. Rev.* **2009**, *109*, 1346–1370.
- (S11) Jalilehvand, F.; Spångberg, D.; Lindqvist-Reis, P.; Hermansson, K.; Persson, I.; Sandström, M. Hydration of the calcium ion. An EXAFS, large-angle X-ray scattering, and molecular dynamics simulation study. *J. Am. Chem. Soc.* **2001**, *123*, 431–441.
- (S12) Marcus, Y. *Ions in water and biophysical implications: from chaos to cosmos*; Springer Science & Business Media, 2012.