

Site correlations, capacitance, and polarizability from protein protonisation fluctuations

Anže Božič

Department of Theoretical Physics, Jožef Stefan Institute, SI-1000 Ljubljana, Slovenia

Rudolf Podgornik

School of Physical Sciences and Kavli Institute for Theoretical Sciences,

University of Chinese Academy of Sciences, Beijing 100049,

China and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China and

*Wenzhou Institute of the University of Chinese Academy of Sciences, Wenzhou, Zhejiang 325000, China**

(Dated: September 5, 2021)

We generalize the Kirkwood-Shumaker theory of protonisation fluctuation for an anisotropic distribution of dissociable charges on a globular protein. The fluctuations of the total charge and the total dipole moment, in contrast to their average values, depend on the same proton occupancy correlator, thus exhibiting a similar dependence also on the solution pH. This has important consequences for the Kirkwood-Shumaker interaction and its dependence on the bathing solution conditions.

I. INTRODUCTION

Electrostatic interactions are ubiquitous in proteinaceous systems and make an essential contribution to protein structure, folding, binding, and condensation [1]. Proteins differ in a fundamental way from many other charged colloidal systems because their charge is not fixed, but depends through the dissociation equilibrium of constitutive AAs on local molecular geometry, dielectric inhomogeneities, vicinal charges, and solution composition. As soon as any of these parameters is changed, the protein charge responds in a process referred to as *charge regulation* (CR), through which the same protein acquires a different charge. CR effects are significant and have been shown to qualitatively change the phase diagram of protein solutions, including a liquid-liquid phase separation based on CR [2], while self-assembled protein structures appear to exhibit asymmetric constructs stabilized due to global charge redistribution [3].

Understanding the charging mechanism of dissociable AAs has a long history, starting with Linderstrøm-Lang of the Carlsberg Laboratory [4] and followed by the early general studies of the acid-base equilibria in polyelectrolytes by Lifson [5] and Marcus [6]. Protonisation/deprotonisation dissociation equilibria were introduced into protein electrostatics in the seminal works of Kirkwood and Shumaker [7, 8] and Tanford and Kirkwood [9], and later developed further through detailed simulation models of protein behaviour in biomolecular solutions [3, 10, 11]. The connection between the protonisation/deprotonisation dissociation equilibria and Poisson-Boltzmann electrostatics was developed by Ninham and Parsegian [12], who introduced the CR mechanism as a self-consistent relationship between the local

electrostatic potential and the dissociated state of chargeable surface groups described by the Langmuir isotherm, which has been recently extended to any adsorption isotherm [13]. CR in the context of proteins [14] and its development was authoritatively reviewed by Borkovec *et al.* [15].

An important consequence of the CR mechanism is the fact that the charging state of a protein is an annealed variable and can therefore exhibit thermal fluctuations, a theoretical prediction [7] borne out by experiments [16]. These thermal protonisation/deprotonisation fluctuations engender fluctuations of both the protein charge as well as the protein dipole moment that in turn lead to capacitance and polarizability of the protein. The charge and the dipole moment are then together at the origin of the fluctuation interactions usually referred to as the Kirkwood-Shumaker (KS) forces [8]. They have been quantified in detailed simulations of protein-protein interactions by Lund and Jönsson [10, 14] and Barroso da Silva and Dias [17]. Upgrading the original KS approach, the fluctuation interactions were cast into a more appropriate theoretical framework of monopole and dipole Casimir-type thermal fluctuation interactions [18, 19] or within the zero frequency Lifshitz term in the general theory of van der Waals interactions [20, 21].

Apart from their effect on the interactions, calculations of the protonisation fluctuation contribution to protein capacitance and polarizability are scarce, which is why we embark in this work upon a more detailed analysis of their properties. Specifically, based on previously introduced methodology [22, 23] we will investigate CR effects of the protein net charge and net dipole moment as well as the corresponding charge fluctuations, quantified by the capacitance response function, and the dipole fluctuations, quantified by the dielectric response function. Our work generalizes some aspects of the seminal approach of Kirkwood and Shumaker [7] and takes explicitly into account the electrostatic interactions on the weak-coupling level in the calculations of the response

* Also affiliated with Department of Physics, Faculty of Mathematics and Physics, University of Ljubljana, SI-1000 Ljubljana, Slovenia; podgornikrudolf@ucas.ac.cn

functions.

II. CHARGE REGULATION OF A CHARGED DIELECTRIC SPHERE

We will treat globular proteins as impermeable dielectric spheres with radius R , dielectric constant ε_∞ , and carrying N charges q_k , $k = 1, \dots, N$, located at positions \mathbf{r}_k , $|\mathbf{r}_k| = R$. This gives rise to a surface charge density on the protein,

$$\sigma(\Omega) = \frac{1}{4\pi R^2} \sum_k q_k^\pm \frac{\lambda_k}{\sinh \lambda_k} \exp(\lambda_k \cos \gamma_k), \quad (1)$$

where γ_k is the great circle distance between Ω and Ω_k . We have associated with each charge a finite size λ_k through a normal distribution on the sphere [24]; when $\lambda \rightarrow \infty$, this reduces to the Dirac delta function. For simplicity, we will assume that $\lambda_k = \lambda \forall k$, and we set $\lambda = 100$, since the exact choice of this parameter does not significantly influence the results [23]. In Eq. (1), we have already made a distinction between the charges based on whether they acquire a positive charge, $q_k^+ = e\eta_k$, or negative charge, $q_k^- = e(\eta_k - 1)$; $\eta_k \in [0, 1]$ determines the protonation/deprotonation state of each charge. The surface charge density can be expanded in terms of spherical harmonics $Y_{lm}(\Omega)$, with the expansion coefficients given by $\sigma(lm) = 4\pi \sum_k q_k g_l(\lambda_k) Y_{lm}^*(\Omega_k)$ [23], where we have defined $g_l(\lambda) = \lambda i_l(\lambda)/\sinh \lambda$ and $i_l(x) = \sqrt{\pi/2x} I_{l+1/2}(x)$ are the modified spherical Bessel functions of the first kind.

A. Free energy of the system

In the linearized Debye-Hückel (DH) approximation, the electrostatic free energy of a charged, impermeable dielectric sphere is [23]

$$F_{ES}[\sigma(lm)] = \frac{1}{32\pi^2 \varepsilon_w \varepsilon_0 R} \sum_l C(l, \kappa R) \sum_m |\sigma(lm)|^2, \quad (2)$$

where $\kappa = \sqrt{2en_0/\varepsilon_w \varepsilon_0 k_B T}$ is the inverse Debye screening length, n_0 is the monovalent salt concentration, $\bar{\varepsilon} = \varepsilon_\infty/\varepsilon_w$, and $\varepsilon_w = 80$ is the dielectric constant of water. In Eq. (2), $C(l, x)$ is the multipole coupling function as described in the Supplementary Information (SI), see Fig. S1, and it is important to note that it exhibits no exponential screening, so that the different multipoles in the electrostatic energy are coupled algebraically.

The electrostatic energy in Eq. (2) differs from and simplifies the well-known Kirkwood expression [25–27] in two ways: (i) the charges in the protein interior are not dissociated and the position of the charges is thus confined to the surface of the dielectric discontinuity; and (ii) the charges are not point-like but have a finite angular size described through a normal distribution on the sphere,

quantified by the parameter λ [24]. In addition, Kirkwood and Shumaker [7] consider the electrostatic interactions either only through the electrostatic self-energy or neglect them altogether.

The total free energy of the system is then composed of the electrostatic (ES) and charge regulation (CR) components and can be cast into the form [23]

$$\begin{aligned} \beta F[q_k] &= \beta F_{ES}[q_k] + \beta F_{CR}[q_k] \\ &= \beta F_{ES}[\eta_k] + \sum_k \beta \alpha_k \eta_k \\ &\quad + \sum_k [\eta_k \ln \eta_k + (1 - \eta_k) \ln(1 - \eta_k)], \end{aligned} \quad (3)$$

where $\beta \alpha_i = \ln 10(\text{pH} - \text{pK}_a^{(i)})$ and pK_a are the AA dissociation constants. For the CR part of the free energy we have chosen the form corresponding to the Langmuir-Davies dissociation isotherm (related to the Henderson-Hasselbalch isotherm) that is typically used to model the CR process in proteins [28]. Other forms, such as the Frumkin-Fowler-Guggenheim isotherm, would correspond to more detailed models of the CR process [29]. The thermodynamic equilibrium solutions for $\bar{\eta}_k$ and the corresponding charges \bar{q}_k are then obtained from an implicit solution of

$$\bar{\eta}_i = \left[1 + \exp \left(\ln 10(\text{pH} - \text{pK}_a^{(i)}) + \beta \frac{\partial F_{ES}[\bar{\eta}_k]}{\partial \bar{\eta}_i} \right) \right]^{-1}, \quad (4)$$

and the equilibrium free energy is thus given by

$$\begin{aligned} \beta F[\bar{q}_k] &= \beta F_{ES}[\eta_k] - \sum_i \frac{\partial \beta F_{ES}[\eta_k]}{\partial \bar{\eta}_i} \bar{\eta}_i \\ &\quad - \sum_i \ln \left(1 + \exp \left(-\beta \alpha_i - \beta \frac{\partial F_{ES}[\eta_k]}{\partial \bar{\eta}_i} \right) \right). \end{aligned} \quad (5)$$

In this way, we can express the total free energy of the system through its electrostatic part and its derivatives, and the form of Eq. (5) is universal for any form of F_{ES} .

B. Linearized CR approximation

Since the electrostatic free energy is based on the DH approximation, valid for small electrostatic potentials and thus necessarily for small electrostatic interactions, we can expand Eq. (4) in terms of the electrostatic contribution to obtain [23]

$$\bar{q}_i = q_i^{(0)} - c_i^{(0)} \frac{\partial F_{ES}[\bar{\eta}_k]}{\partial e \bar{\eta}_i}, \quad (6)$$

where the bare charge $q_i^{(0)}$ and the bare capacitance $c_i^{(0)}$ of the i -th moiety are defined in the SI. In this context, we use the term “bare” to denote that the electrostatic contribution to these quantities has not been taken into

account. From Eq. (2), we obtain the derivative of the free energy with respect to the chargeable moieties:

$$\beta e_0 \frac{\partial F_{ES}[\eta_k]}{\partial \eta_i} = \sum_k \xi_{ik} q_k, \quad (7)$$

with the electrostatic coupling matrix defined as

$$\xi_{ik} = \frac{\ell_B}{R} \sum_l (2l+1) g_l(\lambda_i)^2 C(l, \kappa R) P_l(\cos \gamma_{ik}). \quad (8)$$

Here, $\ell_B = \beta e_0^2 / 4\pi \epsilon_w \epsilon_0$ is the Bjerrum length in water and $\cos \gamma_{kt} = \cos \vartheta_k \cos \vartheta_t + \sin \vartheta_k \sin \vartheta_t \cos(\varphi_k - \varphi_t)$ is the cosine of the spherical distance on the sphere. Note that the magnitude of ξ_{ik} is determined by the ratio of $\kappa \ell_B$ and κR . The charges in the linearized CR approximation that follow from Eq. (6) are then obtained by solving a linear system of equations $\sum_k (\delta_{ik} + c_i^{(0)} \xi_{ik}) \bar{q}_k = q_i^{(0)}$ [23], giving the charges \bar{q}_k as a function of the capacitance and the electrostatic coupling matrix.

III. PROTEIN PROTONISATION FLUCTUATIONS

A. Fluctuations around the charge regulated thermal average

Fluctuations around the thermodynamic equilibrium solution [Eq. (3)] are obtained as

$$\beta F[\bar{q}_k + \delta q_k] = \beta F[\bar{q}_k] + \frac{1}{2} \frac{\partial^2 (\beta F[\bar{q}_l])}{\partial \bar{q}_k \partial \bar{q}_t} \delta q_k \delta q_t + \dots, \quad (9)$$

where $\delta q_i = q_i - \bar{q}_i$ is the deviation from the mean-field solution. We introduce the *proton occupancy correlator* $\langle \delta q_k \delta q_t \rangle$ as the average over fluctuations around the equilibrium value as

$$\langle \delta q_k \delta q_t \rangle = \prod_k \int d\delta q_k (\delta q_m \delta q_n) \times \exp \left(-\frac{1}{2} \sum_{k,t} \frac{\partial^2 (\beta F[\bar{q}_l])}{\partial \bar{q}_k \partial \bar{q}_t} \delta q_k \delta q_t \right). \quad (10)$$

To determine the proton occupancy correlator matrix $\langle \delta q_k \delta q_t \rangle$, we note that

$$\langle \delta q_k \delta q_t \rangle = \frac{1}{2} k_B T \left(\frac{\partial^2 F}{\partial \bar{q}_k \partial \bar{q}_t} \right)^{-1} = \frac{e^2}{2} \left(\xi_{kt} + \frac{e^2 \delta_{kt}}{|\bar{q}_k|(e - |\bar{q}_k|)} \right)^{-1}. \quad (11)$$

where we have used $\partial_q = (1/e) \partial_\eta$. In practice, we use a special case of the Woodbury matrix identity to determine the inverse [30]. The proton occupancy correlator matrix describes the correlations between the protonisation sites on the protein. From Eq. (11), it follows that proton occupancy fluctuations are anisotropic due to the electrostatic interaction coupling matrix ξ_{kp} [Eq. (8)].

B. Charge and dipole moment fluctuations

Charge and dipole moment fluctuations around the thermal equilibrium values for a charge distribution on a sphere can be defined as $\delta q = R^2 \int d\Omega \delta \sigma(\Omega)$ and $\delta \mathbf{p} = R^2 \int d\Omega \delta \sigma(\Omega) \mathbf{r}_\Omega$, respectively, where $\delta \sigma(\Omega) = \sigma(\Omega) - \bar{\sigma}(\Omega)$ and we have taken into account that the charges are confined to the surface of a sphere [31]. To calculate the fluctuations in the total charge of the protein around the mean-field solution, we have

$$\delta q = \frac{1}{4\pi} \int d\Omega \sum_{l,m} \delta \sigma(lm) Y_{lm}(\Omega) = \sum_k \delta q_k g_0(\lambda_k) \quad (12)$$

and as $g_0(\lambda) = 1$, it follows that

$$\langle \delta q^2 \rangle = \sum_{k,t} \langle \delta q_k \delta q_t \rangle = \frac{e^2}{2} \sum_{k,t} \left(\xi_{kt} + \frac{e^2 \delta_{kt}}{|\bar{q}_k|(e - |\bar{q}_k|)} \right)^{-1}. \quad (13)$$

Similarly, the integrals for the fluctuations in the dipole moment vanish for all $l \neq 1$, and we obtain

$$\langle \delta p^2 \rangle = R^2 \frac{(\lambda \coth \lambda - 1)^2}{\lambda^2} \sum_{k,t} \cos \gamma_{kt} \langle \delta q_k \delta q_t \rangle. \quad (14)$$

In the limit where we have no electrostatic contribution to CR, we have $\xi_{ik} = 0$ and thus

$$\langle \delta p^2 \rangle = \frac{e^2 R^2}{2} \frac{(\lambda \coth \lambda - 1)^2}{\lambda^2} \sum_k \bar{\eta}_k (1 - \bar{\eta}_k), \quad (15)$$

which is the result obtained by Kirkwood and Shumaker [7]. Our theory is clearly a generalization of the latter in the sense that it takes into account the electrostatic energy of the fluctuating states. Note again that the derivations presented here are general, and F_{ES} can be substituted with any form of electrostatic free energy of a given system.

From the definitions of $\langle \delta q^2 \rangle$ and $\langle \delta p^2 \rangle$ it transpires that the pH dependence in both cases is given solely by the correlator $\langle \delta q_k \delta q_t \rangle$. The conclusion would therefore be that both have a similar dependence on pH, the only difference being in the summation over the arclength $\cos \gamma_{kt}$ in Eq. (14). For a vanishing electrostatic coupling (vanishing ξ_{kt}), the charge and dipole fluctuations become proportional in the lowest order,

$$\langle \delta p^2 \rangle = R^2 \langle \delta q^2 \rangle, \quad (16)$$

where we have assumed for simplicity a Dirac-like distribution on a sphere, $\lambda \rightarrow \infty$. The electrostatic part ξ_{kt} is thus the one that underpins the difference between the pH dependence of $\langle q^2 \rangle$ and $\langle p^2 \rangle$. The stronger the electrostatic coupling in Eq. (11), the more the fluctuations deviate from locality and exhibit a collective behavior which acts to screen out the details in this dependence.

C. Capacitance and polarizability due to protonisation fluctuations

Since thermodynamic fluctuations in an annealed thermodynamic variable predict the response of the system, it follows that charge and dipole moment fluctuations are generally connected with capacitance and polarizability response functions, respectively. The Johnson-Nyquist fluctuation-dissipation formula, as an example of the general linear response theory, can be cast as [10]

$$\langle \delta q^2 \rangle = -k_B T \left. \frac{\partial \langle q \rangle}{\partial \phi} \right|_{\phi=0} = k_B T \mathcal{C}, \quad (17)$$

where ϕ is the electrostatic potential. Equation (17) provides the connection between the differential capacitance response function \mathcal{C} and charge fluctuations [28]. It can be derived from thermal fluctuations of the electrostatic potential by the Einstein formula [32]. From the linear approximation of Eq. (4) the differential capacitance response function can be rewritten as $\mathcal{C} = \frac{\beta e_0}{\ln 10} \frac{\partial \langle q \rangle}{\partial \text{pH}}$.

Protonisation fluctuations in addition instigate dipole moment fluctuations, which in turn lead to a correction of the dielectric constant of the protein ε_p through a variant of the Kirkwood formula [33, 34], derived by considering protonation-induced dipole fluctuation of a spherical cavity of radius R with an isotropic dielectric response ε_∞ [35]. The relation between dipole fluctuations and the excess polarizability response function α is then

$$\langle \delta p^2 \rangle = k_B T \left. \frac{\partial \langle \mathbf{p} \rangle}{\partial \mathbf{E}_0} \right|_{\mathbf{E}_0=0} = k_B T \alpha, \quad (18)$$

where \mathbf{E}_0 is the external electrostatic field. Taking into account the difference between the internal field \mathbf{E} and the external field, we can write

$$\frac{\partial \langle \mathbf{p} \rangle}{\partial \mathbf{E}_0} = \frac{\partial \langle \mathbf{p} \rangle}{\partial \mathbf{E}} \frac{\partial \mathbf{E}}{\partial \mathbf{E}_0} = 4\pi R^3 \varepsilon_0 \frac{(\varepsilon_p - \varepsilon_\infty) \varepsilon_w}{2\varepsilon_w + \varepsilon_\infty}. \quad (19)$$

Here, ε_∞ refers to all high frequency contributions—such as electronic and vibrational—*excluding* the protonisation fluctuations. As a caveat, *configurational fluctuations* of the average dipole moment are not included at this level.

D. Fluctuation interactions

The response functions of two proteins, separated by a fixed distance and coupled via a screened electrostatic potential, imply a fluctuation-driven interaction in the standard second order perturbation theory. The dominant contribution comes from monopole charge fluctuations and the subdominant term from the dipole charge fluctuations. For two proteins located at $\mathbf{r}_{1,2}$, the excess electrostatic free energy as a functional of the local

electrostatic potential configuration can be written as

$$\mathcal{H}[\phi(\mathbf{r})] = \mathcal{H}_0[\phi(\mathbf{r})] + \int_V d^3\mathbf{r} \tilde{\rho}(\mathbf{r}) \phi(\mathbf{r}) + \frac{1}{2} \iint_V d^3\mathbf{r} d^3\mathbf{r}' \phi(\mathbf{r}) \tilde{\mathcal{C}}(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}'), \quad (20)$$

where $\mathcal{H}_0[\phi(\mathbf{r})]$ is the electrostatic free energy of the bathing ionic solution that is taken in its approximate DH form, with external charges composed of a charge monopole and dipole

$$\tilde{\rho}(\mathbf{r}) = \sum_{i=1,2} (e_i + \mathbf{p}_i \cdot \nabla) \delta(\mathbf{r} - \mathbf{r}_i) = \sum_{i=1,2} \bar{q}_i \delta(\mathbf{r} - \mathbf{r}_i). \quad (21)$$

The generalized response function can similarly be written as composed of the capacitance $\mathcal{C}_i = \mathcal{C} = \beta \langle \delta q^2 \rangle$ and the polarizability $\alpha_i = \alpha = \beta \langle \delta p^2 \rangle$ response functions

$$\tilde{\mathcal{C}}(\mathbf{r}, \mathbf{r}') = \sum_{i=1,2} (\mathcal{C}_i + \nabla \alpha_i \nabla') \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_i). \quad (22)$$

Writing the partition function of this system in the form of a Hubbard-Stratonovich field functional integral, one remains to the lowest (Gaussian) order in the electrostatic potential fluctuations with free energy of the form

$$\beta \mathcal{F}_2[\bar{q}_i, \mathbf{r}_i] = \bar{q}_1 \bar{q}_2 V_0(\mathbf{r}_1, \mathbf{r}_2) - \frac{1}{2} (\tilde{\mathcal{C}}_1 \tilde{\mathcal{C}}_2 + \bar{q}_1^2 \tilde{\mathcal{C}}_2 + \bar{q}_2^2 \tilde{\mathcal{C}}_1) V_0^2(\mathbf{r}_1, \mathbf{r}_2) + \dots \quad (23)$$

where we have introduced $\tilde{\mathcal{C}}_i = \mathcal{C}_i + \nabla \alpha_i \nabla'$. Here, $V_0(\mathbf{r}, \mathbf{r}')$ is the effective electrostatic interaction between two Coulomb charges, equal either to the bare Coulomb interaction for two proteins in a pure solvent, $V_0(\mathbf{r}, \mathbf{r}') = 1/4\pi\varepsilon_w\varepsilon_0|\mathbf{r} - \mathbf{r}'|$, or to the DH potential for two proteins in a monovalent electrolyte, $V_0(\mathbf{r}, \mathbf{r}') = e^{-\kappa|\mathbf{r} - \mathbf{r}'|}/4\pi\varepsilon_w\varepsilon_0|\mathbf{r} - \mathbf{r}'|$. For $\alpha_i \rightarrow 0$, the result of Eq. (23) coincides with the KS interaction as derived by Lund and Jönsson [10]. It also resembles the recently calculated charge fluctuation Casimir interaction between dielectric spheres [21], but without the CR model.

When the average charge and dipole moment of two identical proteins vanish and in the limit of large screening, $\kappa|\mathbf{r}_1 - \mathbf{r}_2| \gg 1$, we have in the leading order in the separation distance

$$\beta \mathcal{F}_2[\mathbf{r}_i] \simeq -\frac{\beta^2}{2} [\langle \delta q^2 \rangle + \langle \delta p^2 \rangle \kappa^2]^2 V_0^2(\mathbf{r}_1, \mathbf{r}_2) + \dots \quad (24)$$

In this case, the effective strength of the KS interactions is additive in CR-generated capacitance *and* polarizability of the protein, mediated crucially by the screening length. This finding is distinct from both the original [25] and subsequent treatments of the KS interactions [10].

IV. EXAMPLES: GLOBULAR PROTEINS

We explore the results of our work on four well-studied globular proteins: bovine β -lactoglobulin (PDB:3NPO),

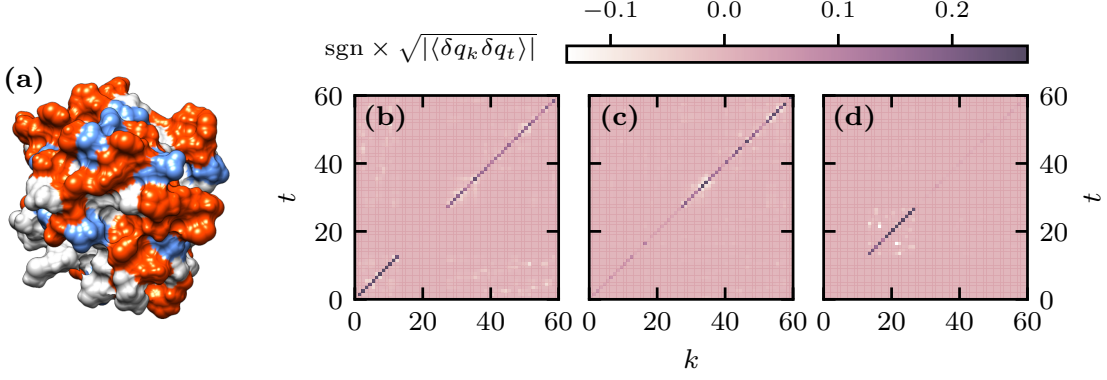


Figure 1. **(a)** An example of a globular protein, hisactophilin (PDB:1HCD), consisting of $N = 60$ charged AAs with different pK_a values. Coloured in red and in blue are the AAs which can become positively or negatively charged, respectively. The charge distribution of the protein can be mapped onto its circumscribed sphere with radius $R = 2.44$ nm. **(b)–(d)** Elements of the proton occupancy correlator matrix $\langle \delta q_k \delta q_t \rangle$ of hisactophilin at pH = 4 (a), 7 (b), and 10 (c). System parameters are $\varepsilon_\infty = 4$ and $n_0 = 100$ mM. For easier discernibility, we show the square root of the elements of the matrix, and the sign of the original matrix elements is preserved. Charge indices k, t refer to the dissociable AAs of the protein.

hen egg-white lysozyme (PDB:3WUN), hisactophilin (PDB:1HCD), and calmodulin (PDB:3CLN). In extracting their charge distributions (see the example in Fig. 1a), we follow previously established methods [22, 23]; details of the procedure are given in the SI. Using the four proteins as examples, we analyze their average charge and the corresponding charge fluctuations, their average dipole moment and the dipole moment fluctuations, their polarizability and the associated dielectric constant of the protein, and the strength of the generalized KS charge fluctuation interactions. In what follows, the charges q_i in the presence of CR are calculated from Eq. (6) and charges $q_i^{(0)}$ in the absence of CR are obtained from Eq. (S3) in the SI.

A. Charge and dipole moment fluctuations

Figure 2a shows how the total charge on hisactophilin changes as a function of pH, both in presence and in absence of CR. At the same time, Fig. 2a shows the changes in the square root of charge fluctuations as a function of pH. We observe two things: charge fluctuations are clearly maximal when the total charge changes the most with pH, and are minimal when the change in the total charge plateaus. This is, of course, a direct consequence of the theoretical derivation. However, the scale of the charge fluctuations is always fairly small compared to the overall charge on the proteins.

The changes in the dipole moment of hisactophilin as a function of pH are shown in Fig. 2b, again both in presence and in absence of CR. Figure 2b also shows the changes in the square root of dipole moment fluctuations as a function of pH. Here, the fluctuations are often maximal when the dipole moment is minimal and vice versa, although this does not seem to be a general rule. An-

other observation, stemming directly from our theoretical calculations, is that the pH dependence of the dipole moment fluctuations is essentially the same as the dependence of total charge fluctuations, which can be seen from comparing panels (a) and (b) of Fig. 2. This is a simple consequence of the fact that the total charge as well as the total dipole moment fluctuations depend on the same proton occupancy correlator $\langle \delta q_k \delta q_t \rangle$ (see the examples in Fig. 1 and Figs. S2 and S3 in the SI), which displays positive values for the diagonal terms (self energy) and negative values (anticorrelation) for off-diagonal terms.

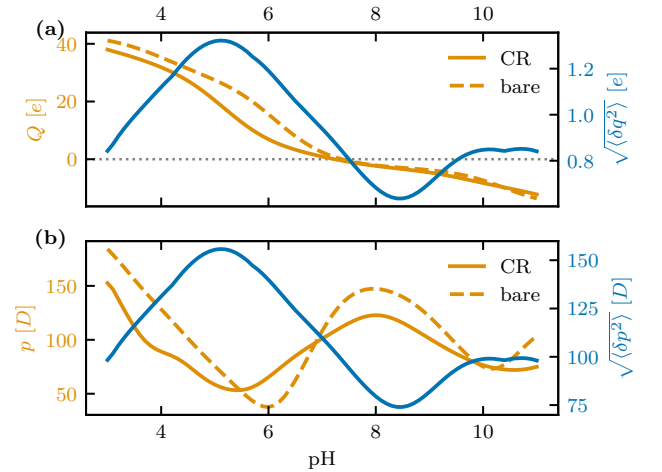


Figure 2. **(a)** Total charge Q and **(b)** total dipole moment p in presence (full line) and absence (dashed line) of CR as a function of pH (left y-axes). Right y-axes of panels (a) and (b) show the square root of charge fluctuations $\sqrt{\langle \delta q^2 \rangle}$ and dipole fluctuations $\sqrt{\langle \delta p^2 \rangle}$, respectively. Shown for hisactophilin (PDB:1HCD) with $\varepsilon_\infty = 4$ and $n_0 = 100$ mM.

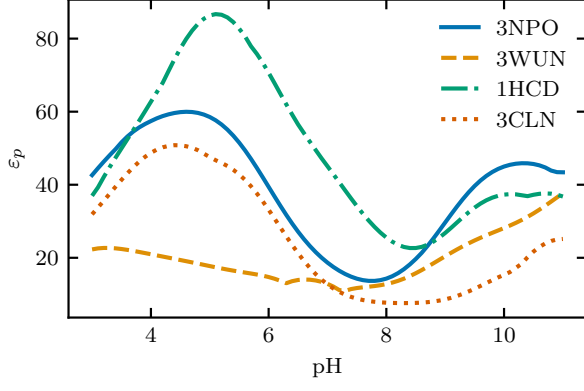


Figure 3. Correction to the protein dielectric constant ε_p due to dipole fluctuations for all four proteins studied. System parameters are $\varepsilon_\infty = 4$ and $n_0 = 100$ mM.

The intermittency in the diagonal components is due to full neutralization of some charged groups at those solution conditions. These observations are characteristic of all four proteins studied, which can be seen from Figs. S4 and S5 in the SI.

External parameters—namely, salt concentration n_0 and protein background dielectric constant ε_∞ —influence the fluctuations to various degrees. In general, changes in either of the two do not modify the pH dependence of the fluctuations to a significant extent, as can be seen from Figs. S6 and S7 in the SI on the examples of β -lactoglobulin and lysozyme, respectively. Changes in salt concentration tend to lead to somewhat larger changes in the magnitude of the fluctuations compared to changes in protein dielectric constant (Fig. S6), although the difference depends on the protein in question (Fig. S7). Moreover, while charge fluctuations depend only on the dimensionless parameter κR (through the electrostatic coupling matrix ξ_{ik}), dipole fluctuations are in comparison more sensitive to the choice of the protein radius, as they feature an explicit quadratic dependence on R in the lowest order [Eq. (16)].

B. Capacitance, polarizability, and thermal fluctuation interactions

Equations (18) and (19) allow us to determine the correction to the dielectric constant of the proteins due to the dipole component of the protonisation fluctuations. Namely, we have that

$$\varepsilon_p = \varepsilon_\infty + \frac{\beta \langle \delta p^2 \rangle}{4\pi \varepsilon_0 R^3} \frac{2\varepsilon_w + \varepsilon_\infty}{\varepsilon_w}. \quad (25)$$

This correction is shown in Fig. 3 for all four proteins studied here and for the background protein dielectric constant $\varepsilon_\infty = 4$ and salt concentration $n_0 = 100$ mM. The correction to the dielectric constant is very large and

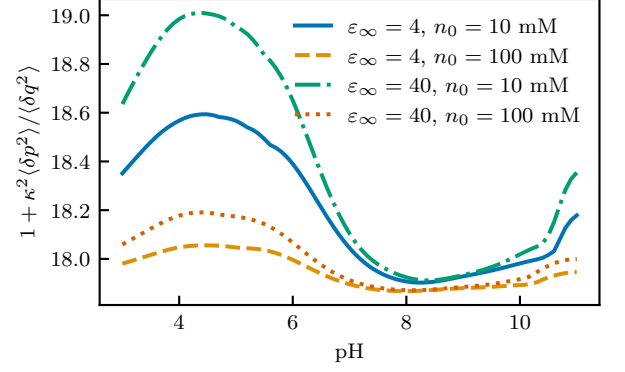


Figure 4. Relative contributions of charge fluctuations and dipole fluctuations to the fluctuation interactions between two identical proteins [Eq. (24)]. Shown on the example of calmodulin (PDB:3CLN) for different values of protein dielectric constant ε_∞ and salt concentration n_0 .

exceeds $\varepsilon_p \gtrsim 10$ for most of the pH range studied. This trend persists even at higher values of ε_∞ or lower values of n_0 (Fig. S8 in the SI).

Capacitance and polarizability of the protein also additively combine in the dominant term in the separation dependence of the fluctuation interaction between two identical proteins, which differs from the original KS form. The relative contribution of the two is further weighted by the screening length of the bathing solution. Figure 4 shows the scaled strength of the fluctuation interactions $1 + \kappa^2 \langle \delta p^2 \rangle / \langle \delta q^2 \rangle$ [see Eq. (24)] on the example of calmodulin for different values of ε_∞ and n_0 . Regardless of these two parameters, polarizability of the protein appears to be the dominant contribution to the interaction. The relative contributions of charge and dipole fluctuations are modified by changes in ε_p and n_0 , somewhat more so by the latter. And while the relative contributions of polarizability and capacitance to the fluctuations differ from protein to protein (Fig. S9 in the SI), the former is always dominant.

V. CONCLUSIONS

Based upon our previous work [23], we generalized the KS theory of protonisation fluctuation [7] by calculating the electrostatic and CR free energy contributions due to *anisotropic distribution* of dissociable charges on a sphere circumscribed to a protein. This allowed us to implement and solve for the weak-coupling CR boundary condition as well as quadratic fluctuations around it. We show that fluctuations in both the total charge and the total dipole moment depend on the same proton occupancy correlator $\langle \delta q_k \delta q_t \rangle$. As a result, both the capacitance (proportional to the charge fluctuations) as well as the polarizability (proportional to the dipole fluctuations) of the protein depend in a similar fashion on

the solution pH through the proton occupancy correlator. This is surprising, as these two lowest multipoles, i.e., monopole and dipole, describe two linearly independent terms in the spherical harmonic expansion of the electrostatic interaction energy; in fact, it is known that their equilibrium values, \bar{q} and \bar{p} , show vastly different pH behavior [22]. We also show that the strength of the KS fluctuation interaction in the asymptotic regime of protein-protein separation [10] is modified and depends additively on the capacitance and polarizability of the two interacting proteins. As a consequence, the strength of the KS interaction acquires a strong pH and salt de-

pendence.

ACKNOWLEDGMENTS

A.B. acknowledges funding from the Slovenian Research Agency ARRS (Research Core Funding No. P1-0055). R.P. acknowledges the support of the University of Chinese Academy of Sciences and the funding from the NSFC under grant No. 12034019.

-
- [1] H.-X. Zhou and X. Pang, Electrostatic interactions in protein structure, folding, binding, and condensation, *Chem. Rev.* **118**, 1691 (2018).
 - [2] B. Zheng, Y. Avni, D. Andelman, and R. Podgornik, Phase separation of polyelectrolytes: The effect of charge regulation, *J. Phys. Chem. B* **125**, 7863 (2021).
 - [3] T. Curk and E. Luijten, Charge regulation effects in nanoparticle self-assembly, *Phys. Rev. Lett.* **126**, 138003 (2021).
 - [4] J. Schellman and C. Schellman, Kaj Linderstrom-Lang (1896-1959), in *Selected Topics in the History of Biochemistry Personal Recollections. VI*, Comprehensive Biochemistry, Vol. 41, edited by G. Semenza and R. Jaenicke (Elsevier, 2000) pp. 45–89.
 - [5] S. Lifson, Potentiometric titration, association phenomena, and interaction of neighboring groups in polyelectrolytes, *J. Chem. Phys.* **26**, 727 (1957).
 - [6] R. Marcus, Calculation of thermodynamic properties of polyelectrolytes, *J. Chem. Phys.* **23**, 1057 (1955).
 - [7] J. G. Kirkwood and J. B. Shumaker, The influence of dipole moment fluctuations on the dielectric increment of proteins in solution, *Proc. Natl. Acad. Sci. USA* **38**, 855 (1952).
 - [8] J. G. Kirkwood and J. B. Shumaker, Forces between protein molecules in solution arising from fluctuations in proton charge and configuration, *Proc. Natl. Acad. Sci. USA* **38**, 863 (1952).
 - [9] C. Tanford and J. G. Kirkwood, Theory of protein titration curves. I. General equations for impenetrable spheres, *J. Am. Chem. Soc.* **79**, 5333 (1957).
 - [10] M. Lund and B. Jönsson, Charge regulation in biomolecular solution, *Q. Rev. Biophys.* **46**, 265 (2013).
 - [11] R. Lunkad, A. Murmiliuk, Z. Tošner, M. Štěpánek, and P. Košovan, Role of pKa in charge regulation and conformation of various peptide sequences, *Polymers* **13** (2021).
 - [12] B. Ninham and V. Parsegian, Electrostatic potential between surfaces bearing ionizable groups in ionic equilibrium with physiologic saline solution, *J. Theor. Biol.* **31**, 405 (1971).
 - [13] R. Podgornik, General theory of charge regulation and surface differential capacitance, *J. Chem. Phys.* **149**, 104701 (2018).
 - [14] M. Lund and B. Jönsson, On the charge regulation of proteins, *Biochemistry* **44**, 5722 (2005).
 - [15] M. Borkovec, B. Jönsson, and G. J. M. Koper, Ionization processes and proton binding in polyprotic systems: Small molecules, proteins, interfaces, and polyelectrolytes, in *Surface and Colloid Science*, edited by E. Matijević (Springer US, Boston, MA, 2001) pp. 99–339.
 - [16] S. Takashima, A study of proton fluctuation in protein. Experimental study of the Kirkwood-Shumaker theory, *J. Phys. Chem.* **69**, 2281 (1965).
 - [17] F. L. Barroso da Silva and L. G. Dias, Development of constant-pH simulation methods in implicit solvent and applications in biomolecular systems, *Biophys. Rev.* **9**, 699 (2017).
 - [18] N. Adžić and R. Podgornik, Charge regulation in ionic solutions: Thermal fluctuations and Kirkwood-Schumaker interactions, *Phys. Rev. E* **91**, 022715 (2015).
 - [19] N. Adžić and R. Podgornik, Field-theoretic description of charge regulation interaction, *Eur. Phys. J. E* **37**, 49 (2014).
 - [20] J. Mahanty and B. W. Ninham, *Dispersion forces* (London Academic Press, 1976).
 - [21] R. O. Nunes, B. Spreng, R. de Melo e Souza, G.-L. Ingold, P. A. Maia Neto, and F. S. S. Rosa, The Casimir interaction between spheres immersed in electrolytes, *Universe* **7** (2021).
 - [22] A. Božič and R. Podgornik, pH dependence of charge multipole moments in proteins, *Biophys. J.* **113**, 1454 (2017).
 - [23] A. Božič and R. Podgornik, Anomalous multipole expansion: Charge regulation of patchy inhomogeneously charged spherical particles, *J. Chem. Phys.* **149**, 163307 (2018).
 - [24] A. Božič, From discrete to continuous description of spherical surface charge distributions, *Soft matter* **14**, 1149 (2018).
 - [25] J. G. Kirkwood, Theory of solutions of molecules containing widely separated charges with special application to zwitterions, *J. Chem. Phys.* **2**, 351 (1934).
 - [26] C. Tanford and J. G. Kirkwood, Theory of protein titration curves. I. General equations for impenetrable spheres, *J. Am. Chem. Soc.* **79**, 5333 (1957).
 - [27] F. L. Da Silva, B. Jönsson, and R. Penfold, A critical investigation of the Tanford-Kirkwood scheme by means of monte carlo simulations, *Protein Sci.* **10**, 1415 (2001).
 - [28] Y. Avni, D. Andelman, and R. Podgornik, Charge regulation with fixed and mobile charged macromolecules, *Curr. Op. Electrochem.* **13**, 70 (2019).

- [29] A. Majee, M. Bier, R. Blossey, and R. Podgornik, Charge symmetry broken complex coacervation, *Phys. Rev. Res.* **2**, 043417 (2020).
- [30] H. V. Henderson and S. R. Searle, On deriving the inverse of a sum of matrices, *SIAM Review* **23**, 53 (1981).
- [31] A. Božič and R. Podgornik, Symmetry effects in electrostatic interactions between two arbitrarily charged spherical shells in the Debye-Hückel approximation, *J. Chem. Phys.* **138**, 074902 (2013).
- [32] A. Einstein, Über die Gültigkeitsgrenze des Satzes vom thermodynamischen Gleichgewicht und über die Möglichkeit einer neuen Bestimmung der Elementarquanten, *Ann. Phys.* **22**, 569–572 (1907).
- [33] T. Simonson and C. L. Brooks, Charge screening and the dielectric constant of proteins: insights from molecular dynamics, *J. Am. Chem. Soc.* **118**, 8452 (1996).
- [34] H. Fröhlich, *Theory of dielectrics* (Clarendon Press, Oxford, UK, 1949).
- [35] F. Stillinger, Low frequency dielectric properties of liquid and solid water, in *The Liquid state of matter: Fluids, simple and complex*, Studies in statistical mechanics, Vol. 8, edited by J. L. E.W. Montroll (Elsevier, 1982) pp. 341–431.