Dielectric Interpretation of Specificity of Ion Pairing in Water

ARTICLE *in* JOURNAL OF PHYSICAL CHEMISTRY LETTERS · DECEMBER 2009

Impact Factor: 7.46 · DOI: 10.1021/jz900151f

5 AUTHORS, INCLUDING:

Robert Vácha
Masaryk University
47 PUBLICATIONS **1,948** CITATIONS

SEE PROFILE

Implications of a high dielectric constant in proteins

Mikael Lund
Institute of Organic Chemistry and Biochemistry
The Academy of Sciences of the Czech Republic
Flemingovo nam. 2,
CZ-16610 Prague 6, CZECH REPUBLIC
*Phone: +420774311693 / Fax: +420220410320**

Bo Jönsson
Theoretical Chemistry, Chemical Center, Lund University
POB 124, S-22100 Lund, SWEDEN

Clifford E. Woodward
School of Chemistry, University College, University of New South Wales
Australian Defence Force Academy, Canberra ACT 2600, AUSTRALIA
(Dated: April 25, 2007)

Solvation of protein surface charges plays an important role for the protonation states of titratable surface groups and is routinely incorporated in low dielectric protein models using surface accessible areas. For many-body protein simulations, however, such dielectric boundary methods are rarely tractable and a greater level of simplification is desirable. In this work, we scrutinize how charges on a high dielectric surface are affected by the non-polar interior core of the protein. A simple dielectric model, which models the interior as a low dielectric sphere, combined with Monte Carlo simulations, show that for small, hydrophilic proteins the effect of the low dielectric interior is largely negligible and that the protein (and solution) can be approximated with a uniform high dielectric constant equal to that of the solvent. This is verified by estimates of titration curves and acidity constants for four different proteins (BPTI, calbindin D_{9k}, ribonuclease A, and turkey ovomucoid third domain) that all correlate well with experimental data. Furthermore, the high dielectric approximation follows as a natural consequence of the multipole expansion of the potential due to imbedded protein charges in the presence of the low dielectric core region.

PACS numbers:

Keywords: Protein electrostatics, Monte Carlo Simulation, Coarse Graining, Dissociation Constant, Charge Capacitance, Uniform dielectric.

Introduction

In the past century protein ionization has caught the attention of many researchers and a wealth of experimental data is available, especially potentiometric titration curves and NMR determined pK_a values. Theoreticians have not neglected the subject either and numerous protein solution models have been proposed for the calculation of pK_a values. These range from simple, spherical models as that of Tanford and Kirkwood [1–3] (TK), numeric solution of the Poisson-Boltzmann equation (PBE) and generalized Born model (GBM) calculations[4–9] as well as explicit solvent models[10, 11].

In the original TK model [1–3] titratable charged groups were assumed to be buried within a spherical protein with a dielectric constant much less than that of the surrounding (aqueous) solvent. Modern approaches utilizing electrostatic continuum models and the Poisson-Boltzmann theory, can be viewed as generalizations of the original TK model, which take into account the detailed structure of the protein surface. However, even within the context of these more "refined" approaches, two important questions arise: (i) *What is the dielectric constant of the protein?* and (ii) *Where is the dielectric interface between protein and solvent to be located?*

With respect to the first question, we note that the interior of a protein may have a higher than expected dielectric constant due to polarization contributions from the configurational freedom of polar side chains [12, 13] as well as proton fluctuations in titratable groups [14–17]. In addition to this, one may expect some intrusion by solvent into the protein interior. Many values for the interior dielectric response of proteins have been used in reported calculations, ranging from 2 to 40, approximately. Our purpose here is not to review this work but merely to note that there is little agreement about this quantity and, in macroscopic models, it may ultimately be best treated as an adjustable parameter[18] matched to experimental results.

With respect to the second question, we observe that the majority of charges in globular proteins are located "close" to the protein solution interface (Fig. 1). These surface groups are thus very accessible, and attractive,

*Electronic address: mlund@mac.com

to the surrounding solvent, as evidenced by the fact that proteins dissolve in water through solvation of polar surface groups and that water readily exchanges protons with protein side chains. This suggests that the position of the interface between the low dielectric region of the protein and the high dielectric aqueous solvent is expected to be buried within the protein surface. A similar conjecture led Tanford to state some thirty years ago [19] *"...that the Kirkwood model is not appropriate for proteins because the ionic groups often extend into the surrounding solvent sufficiently far so that the medium between them is pure solvent and the effect of the large dielectric cavity several Ångstroms away is minimal."* Thus the proper implementation of models containing complex shaped regions of different polarizability, is fraught with some uncertainty, quite apart from the intensive computational effort required for their numerical solution. It is therefore useful to investigate simpler models that display reasonable predictive qualities without the numerical cost. In connection with this, it is pertinent to note, that a number of researchers[6, 20–24] have obtained good agreement with experimental titration results using a *homogeneous* dielectric constant model (HDM) where the internal dielectric constant of the protein is set to that of the surrounding solvent. Despite its predictive success, the HDM is often considered as a limiting case with little physical validity. On the other hand, more recent attempts to obtain accurate numerical predictions have spawned a range of models proposing the use of geometry dependant dielectric constants and protein charge rescaling with questionable physical basis. These types of theories amount to *empirical* fits to experimental data based on an approximate functional form for the (free) energy. Attempts to physically justify these models rely on their apparent support of the assertion that electrostatic interactions will be screened less in the protein core, compared with at the surface.

In this paper we will investigate a simple protein model, wherein the expected hydrophobic core of the protein is modelled as a low dielectric sphere, imbedded within the (convoluted) surface of the protein (Fig. 2). In the region between the sphere and the protein surface, we shall assume that the dielectric constant is equal to that of the solvent. This is in recognition of the ability of the solvent to permeate this region and solvate charged residues and, further, that the residues themselves will often increase the local polarizability of this region. This model contains the essential physical features described above, while being simple enough to allow simple analytical expressions for the electrostatic potentials. In particular, the use of the spherical core allows the use of a simple multipole expansion [1, 2]. We shall use this model to calculate theoretical protein titration curves and make a comparison with a range of experimental results. Though crude, this model should give us some insight into the role played by the low dielectric interior in determining the titration behaviour. It is clear that the low dielectric interior of a typical protein will likely be non-spherical. Nevertheless, the model does suggest a suitable generalization with a sounder physical basis than current simplified approaches.

Model and theory

X-ray or nuclear magnetic resonance determined protein structures are represented by either a collection of spheres (diameter, σ=4 Å) each representing an atom or else spheres representing amino acid residues located at the residue center-of-mass, with a diameter, σ_R, determined from the residue molecular weight, M_w, according to the formula,

$$\sigma_R/2 = \left(\frac{3M_w}{4\pi\rho N_{Av}}\right)^{1/3} \qquad \rho = 1 \text{ g/cm}^3. \tag{1}$$

The latter (coarse grained) model reduces the number of particles from the order of thousands to hundreds and is useful when studying several protein molecules [25–27]. We use a rigid model for the protein without permanent partial charges. That is, the only charges on the protein arise from the presence or absence of acidic protons at the titrateable sites.

The low dielectric interior of the protein is modelled as a sphere with dielectric constant ϵ_p and a radius a about the protein centre of mass. The surface of the sphere lies within the protein surface. The solvent is treated as a continuum with dielectric constant, ϵ_s, and mobile counter ions and salt are modelled as monovalent charged hard spheres with radius σ_s (=4 Å). While it is assumed that water can penetrate the outer protein surface to some extent, and contribute to the dielectric constant in the region between the inner sphere and surface, we assume that solvated ions are too bulky to penetrate into the protein.

The configurations of the mobile ions are sampled in the canonical ensemble distribution using the Metropolis Monte Carlo (MC) algorithm[28]. The total energy is calculated as a sum of interactions between pairs of species, which include electrostatic and hard sphere parts. For interactions between mobile ions and sites within the protein, the hard sphere function takes into account exclusion by the protein surface. In addition to the pair interactions, there is a self energy contribution, $u^{self}(r_i)$ due to the interaction between ions and the dielectric regions (as described in more detail below).

$$U = \sum_{i,j}\left(u^{el}(r_{ij}) + u^{hs}(r_{ij})\right) + \sum_i u^{self}(r_i) \qquad i \neq j \tag{2}$$

$$u^{hs}(r_{ij}) = \begin{cases} 0 & r_{ij} > (\sigma_i + \sigma_j)/2 \\ \infty & \text{otherwise} \end{cases} \tag{3}$$

The Electrostatic Pair Interaction

The electrostatic part of the pair potential, u^{el}, is found by identifying the appropriate boundary conditions and solving the Poisson and Laplace equations for a spherical discontinuity, yielding[29, 30] the following infinite order multipole expansions,

$$u^{el}(r_{ij}) = \begin{cases} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_s}\left(\frac{1}{r_{ij}} + \sum_{n=1}^{\infty}\frac{r_j^n}{r_i^{n+1}}\Gamma_1\right) & r_i > a > r_j \\ \frac{q_i q_j}{4\pi\epsilon_0\epsilon_s}\left(\frac{1}{r_{ij}} + \sum_{n=1}^{\infty}\frac{a^{2n+1}}{(r_i r_j)^{n+1}}\Gamma_1\right) & r_i > a \wedge r_j > a \\ \frac{q_i q_j}{4\pi\epsilon_0\epsilon_p}\left(\frac{1}{r_{ij}} + \sum_{n=0}^{\infty}\frac{(r_i r_j)^n}{a^{2n+1}}\Gamma_2\right) & r_i < a \wedge r_j < a \end{cases} \quad (4)$$

with $\Gamma_1 = \frac{(\epsilon_s-\epsilon_p)P_n(\cos\theta)}{\epsilon_p+\epsilon_s(n+1)/n}$ and $\Gamma_2 = \frac{(\epsilon_p-\epsilon_s)P_n(\cos\theta)}{\epsilon_s-\epsilon_p n/(n+1)}$. $P_n(\cos\theta)$ is the Legendre polynomial evaluated for the angle, θ between $\mathbf{r}_i$ and $\mathbf{r}_j$.

The general form of the pair interaction, as described by Eq. 4 is quite revealing. If at least one of the charges is in the high dielectric region, the direct Coulombic interaction (the first terms in the first two cases in Eq. 4) is the same as for two charges immersed in the (infinite) solvent alone. The higher order terms ($n \geq 1$) correspond to interaction between one of the charges with multipoles induced at the spherical dielectric boundary by the other charge. These multipole forces are of shorter range than the direct Coulombic interaction. It is worth noting, that the multipolar nature of the interactions is not dependent on the spherical geometry assumed for the core region. A more complex shape would also give rise to pair interactions with higher-order multipolar terms, albeit with different "polarisabilities", reflecting the departure from the spherical shape (see below).

Ignoring these higher-order terms corresponds to the HDM, discussed earlier. Thus, the approximations inherent in the HDM can be justified on a physical basis, if the short-range multipolar contributions can be ignored relative to the long-ranged Coulombic term. Alternatively, any correction to the HDM must involve contributions, which reflect the multipolar interactions that have been neglected. Below, we show that the angular dependence of such contributions can lead to non-monotonic deviations from the HDM.

Interestingly, when both charged sites are in the low dielectric sphere, the leading pair term corresponds to the Coulomb interaction between two charges immersed in the *low* dielectric continuum. Thus, when dealing with systems with deeply buried titrating sites, a simple generalization of the HDM is to let the interactions between the deep charges be determined by the low dielectric constant. This type of "two-dielectric" model is reminiscent of theories that define geometry dependent values of the dielectric constant [31, 32]. In this case, however, the simple (spherical) model allows us to rigorously identify the nature of the energy terms that have been neglected. It should be pointed out that a similar analysis of this type of model was recently used in order to improve parameters for the GBM [33]. However, in that case, perfect conducting boundary (PCB) conditions were assumed at the protein surface in order to simplify expressions. This is inapplicable in our case, as the field due to charges within the protein cannot extend to the external mobile salt ions with PCB conditions. Furthermore, the multipolar character of the pair interactions between charges is not captured within the GBM.

Electrostatic Self-Energy

The electrostatic self-energy term is due to a charge interacting with the polarisable medium. The latter induces a reaction field with which the charge interacts. Thus, the self-energy can be obtained by the usual "charging" integration of the potentials, Eq.(4), for $i = j$. We note that the Coulombic term gives rise to an infinite contribution if we assume point charges. By, instead, allowing for the charge to be spread over the surface of a sphere of radius σ_B, the self-energy contribution that arises from the Coulomb term is just the classic Born expression for solvation. The Born energy, ΔG_B), estimates the free energy cost of transferring a charge, q, from a region with dielectric constant ϵ_s to another with with dielectric constant ϵ to be,

$$\Delta G_B = -\frac{q^2}{4\pi\epsilon_0}(1/\epsilon_s) - 1/\epsilon)/\sigma_B \quad (5)$$

The other component to the self-energy comes from the induced multipoles at the dielectric boundary. For a charge outside the sphere ($r > a$), this is given by,

$$u^{mult}(r) = \frac{1}{2}q\phi^{rf}(r) = \frac{q^2}{8\pi\epsilon_0\epsilon_s}\sum_{n=1}^{\infty}\frac{\Gamma_1}{r}\left(\frac{a}{r}\right)^{2n+1} \quad (6)$$

Titrations

Titration is accomplished by exchanging protons (+1 charges) between the sites in the protein and the solvent. The energy change associated with this process is given by[17, 34],

$$\Delta U = \Delta U^{el} \pm (\mathrm{pH} - \mathrm{pK}_a)kT\ln 10 \quad (7)$$

where kT is the thermal energy and (+) applies for protonation and (-) for deprotonation. The term (pH − pK_a)$kT\ln 10$ includes all non-electrostatic contributions to the free energy of titration, including bonding and (de)solvation components. These are obtained from titration data on small amino acid *model* compounds. In this work we used the following pK_a values: Ctr 3.8; Asp 4.0; Glu 4.4; His 6.3; Ntr 7.5; Tyr 9.6; Lys 10.4; Cys 10.8; Arg 12.0. This approach is often employed for protein titration calculations.

The quantity, ΔU^{el}, in Eq.(7) accounts for the electrostatic energy change that occurs when a charge is transferred from the appropriate model compound to the

corresponding site in the protein. In order to calculate, ΔU^{el}, one needs to properly model the electrostatic environment of protons in the model compound as well as that of the protein. In this work, we shall assume that the low dielectric region of the *model compounds* is small enough to be neglected. The validity of these assumptions will eventually be verified upon comparison with experiment.

Finally, we also calculated the excess chemical potentials, $\mu_{\rm ex}$ for ion pairs in the solution. These were obtained using the Widom particle insertion method[35]. In this case a cation/anion pair are simultaneously inserted at random positions in the cell and $\mu_{\rm ex}$ can be sampled according to

$$\mu_{\rm ex} = -kT \ln \left\langle e^{-\Delta U/kT} \right\rangle_0 , \qquad (8)$$

where ΔU is the total energy change for the process evaluated according to Eq. 2.

Results and Discussion

Electrostatics of the Model

We begin by considering the one- and two-body contributions to ΔU^{el}. We shall denote these as $\Delta U^{el,1}$ and $\Delta U^{el,2}$, respectively. In this work we will only consider titrating protons which are in the high dielectric region in both the protein. This means the Born contribution to the self-energy, Eq.(5), for titration is zero. Thus, for a proton removed from the model compound and placed in the protein at a distance r_0 from the center of the low dielectric sphere ($r_0 > a$),

$$\Delta U^{el,1} = u^{rf}(r_0) \qquad (9)$$

where u^{rf} is the reaction potential defined earlier, Eq.(6). We note that this implies that $\Delta U^{el,1}$, will generally be positive, i.e., it disfavours the transfer of charge to the protein, due to the presence of the low dielectric region. A charge placed near the dielectric interface will be subject to an unfavourable reaction potential. The magnitude of this effect is illustrated in Fig. 4, where we have used MC simulations to calculate the excess chemical potential, $\mu_{\rm ex}$, for an ion outside a low dielectric sphere. Here we have considered two different radii for the sphere (a = 7Å, 15Å) and also investigated the screening effect of the electrolyte solution at a typical (physiological) concentration of 150 mM. At the larger radius, $\Delta U^{el,1}$, is more positive for a given distance from the dielectric boundary. This is due to a greater amount of low dielectric material being placed closer to the ion. Furthermore, as the reaction field is relatively short ranged, we observe essentially no screening in the presence of salt. The effect of the low dielectric region on the, $\Delta U^{el,1}$, is not large. With reasonable choices for the radius of the low dielectric sphere and salt concentrations, $\mu_{\rm ex}$ is only a few tenths of a kT.

The two-body contribution, $\Delta U^{el,2}$ is a sum of the electrostatic pair interactions between the titrating proton and all the other charges present on the protein. As shown in Eq.(4), the pair interaction is mediated by the intervening dielectric material and any dielectric boundaries present. The presence of the boundaries gives rise to induced multipoles that can have some unexpected consequences. To illustrate this, we present in Fig 3, the interaction between two like charges close to the boundary in the protein. One may expect that the low dielectric region would have a tendency to enhance the electrostatic repulsion between charges, compared with a homogenous dielectric model, with a uniform dielectric constant equal to the solvent value. However, when the charges are only a few Ångstroms from the surface of the low dielectric sphere, there is little change in their interaction, Fig.3 (a). The distortion of electric field lines by the low dielectric region has minimal effect on the direct Coulomb interaction.

This notwithstanding, the effect of the sphere can be significant when the charges are closer to the boundary (on either side). Furthermore, the angular dependence of the induced multipolar interactions can bring about non-monotonic corrections to the HDM. This is best observed by looking at the difference between the ion pair interaction in the HDM and in the presence of the discontinuity, Fig.3 (b). For charges just outside the spherical surface, the pair interaction is more repulsive than the HDM at small angles, but then becomes slightly more attractive at large angles. The effect is more pronounced the closer the charges are to the surface. This behaviour can be understood by from the orientation of the multipoles induced at the dielectric boundary. Consider, for example, just the dipole induced by one charge, which interacts with the second charge. When this second charge is brought into the vicinity of the first it is repelled by the induced dipole at small angles and attracted by the dipole at large angles (greater than 90). Though truncation at the dipole term is an approximation, especially at close separations, the addition of the higher order induced multipoles does not alter the qualitative picture.

It is interesting to look at the effect of placing both charges into the low dielectric sphere, though we should point out that this scenario is never realised in the calculations on proteins that we present later. Here we expect the HDM to underestimate the repulsion between the charges. Indeed, the first term in the pair potential expansion, Eq.(4), is a Coulomb interaction screened by the *internal* dielectric constant, ϵ_p, rather than the solvent dielectric constant (as assumed in the HDM). However, when the charges are close to each other and only slightly below the surface ($a = r_i + 0.1$Å), it turns out that the HDM overestimates the repulsion. This counter-intuitive result is again caused by the angular dependence of the induced multipoles at the dielectric boundary. Looking at the first-order term again, the dipole induced by a given charge is now attractive to a second charge at small angles. This is because the greatest polarization

response is in the high dielectric solvent, external to the low dielectric sphere in which the charges are placed. The HDM will underestimate the repulsion at larger angles, when the total interaction is small in any case. When the charges are even further below the surface of the low dielectric sphere ($a = r_i + 2$Å), the effect of the induced multipoles is weaker and the interaction is much more repulsive than in the HDM, especially at small angles. These results demonstrate that the exact pair interaction can be quite sensitive to the position of the dielectric discontinuity, especially when the charges are close to the boundary [36]. In addition, the correction to the simple HDM has a complex angular dependence from the multipoles induced at the dielectric boundary. In general, simplified approaches such as the generalized Born theory do not include these type of multipolar corrections.

As indicated earlier, the simple spherical model we have proposed thus far, may be too simplistic to apply with confidence to real proteins, as the low dielectric interior region of the protein is unlikely to be spherical. On the other hand, the general form of the interaction expressions, Eq.(4), will be maintained for non-spherical geometries, but with generalized polarizabilities reflecting the asymmetry of the boundaries. For example, truncating at the induced dipole level, one would obtain for Eq.(4),

$$u^{el}(r_{ij}) \approx \begin{cases} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_s}\left(\frac{1}{r_{ij}} + \frac{r_j}{r_i^2}A_1 P_1(cos\theta)\right) & r_i > a > r_j \\ \frac{q_i q_j}{4\pi\epsilon_0\epsilon_s}\left(\frac{1}{r_{ij}} + \frac{a^3}{(r_i r_j)^2}A_2 P_1(cos\theta)\right) & r_i > a \wedge r_j > a \\ \frac{q_i q_j}{4\pi\epsilon_0\epsilon_p}\left(\frac{1}{r_{ij}} + \frac{(r_i r_j)}{a^3}A_3 P_1(cos\theta)\right) & r_i < a \wedge r_j < a \end{cases} \quad (10)$$

with A_n generalized polarizabilities. A simple theory for protein titration suggests itself by letting the A_n be determined empirically. They would then reflect both the shape and the internal polarizability of the internal region of the protein under investigation. This type of perturbation approach would be viable, provided the interactions are dominated by the zeroth order terms (Coulombic interactions) in the potential expressions above. As noted earlier, for the case where charges reside in the high dielectric region, the zeroth order approximation corresponds to the HDM.

We will not investigate this semi-empirical approach in this paper. Instead, we will use the simple spherical model, to ascertain the relative importance of induced multipolar corrections to the actual titration behaviour of four different proteins: calbindin, bovine pancreas trypsin inhibitor (BPTI), ribonuclease A, and turkey ovomucoid (third domain).

Monte Carlo Simulations of Protein Titration

We used MC simulations to calculate the titration curve for ribonuclease A. The results given in Fig. 5 used the HDM, i.e., we assumed a uniform dielectric constant throughout the solution and the protein ($\epsilon_p = \epsilon_s$=80). Surprisingly, good agreement is found between the computed and experimentally measured titration curves, despite the absence of a low dielectric protein interior. The simulated results remained equally good, even in the presence of a low dielectric spherical region, whose diameter was chosen so as not overlap any charged groups. The calculated results differ significantly from the ideal or "null" curve, which does not include electrostatic interactions at all. This is an important observation and hints that the underlying physical mechanisms captured by the HDM are largely accurate and that the low dielectric interior can be treated as a perturbation.

For BPT1 and turkey ovomucoid we were able to obtain experimentally determined individual pK_a values, providing us with a much more detailed test for the simulations. The results are presented in Tables I,II, where we also compare with theoretical predictions based on the Poisson-Boltzmann equation (PBE)[5] and a modified Tanford-Kirkwood (MTK) approach[37]. Again, we have used the HDM in our simulations. In both cases the model agrees well with experimental data. Furthermore, the simulation results are in overall better agreement with the experiments than the alternative models (see also Khandogin and Brooks [38] work on ovomucoid), which both assume low dielectric interiors for the protein. This is surprising, given that the simulations use the simple HDM, without multipolar corrections. The better performance of the simulations may be as a result of the "exact" treatment given to ion-ion correlations, which is not the case in the mean-field theories. Clearly, parameterized multipole corrections will cause our simulated results to give even better agreement with experiment. It is worth noting that the MTK calculations are significantly improved by introducing either a higher dielectric constant (reported up to 20), or letting the flexible side chains relax in the electric field. As is discussed by Havranek and Harbury[37] and others [18, 39], buried polar groups may require a more subtle description.

As a final example, Fig. 6 shows experimental [24, 40] and calculated titration curves for calbindin D_{9k}. In this case, our simulations did incorporate a low dielectric spherical interior in the protein. The radius of this sphere was chosen as large as possible, while still keeping the surface protein charges external to it. The agreement with the experimental data is good, however, the theory predicts that the protein is more difficult to charge than is actually the case. We repeated the simulations with no low dielectric sphere (HDM), however, no notable deviations from our previous results was obtained. It is possible that the introduction of multipolar corrections may give rise to better predictions. However, it is more likely that in a real protein, some degree of charge-charge correlation takes place in the flexible side-chains so as to lower the electrostatic free energy. Such a mechanism is, of course, absent in a rigid protein model.

Finally, it is interesting to consider the effect of varying the intrinsic acid constants for the titrating groups. This can be conveniently estimated using the *protein charge*

capacitance[17] which is a molecular measure of how the protonation state responds to an external electric potential, ϕ_{ext}:

$$C = \langle Z^2 \rangle - \langle Z \rangle^2 = -\frac{\partial Z}{\beta e \partial \phi_{ext}} = -\frac{1}{\ln 10}\frac{\partial Z}{\partial \mathrm{pH}} \qquad (11)$$

where Z is the total protein charge number. As evident from Eq. 7 shifting the pK_a values is equivalent to applying an external potential. At pH 4 the capacitance for calbindin is approximately 2 (from the derivative of the titration curve) and shifting the intrinsic pK_a values by 0.4 units will cause the protein to bind two more protons ($\Delta Z \approx -C \ln 10 \Delta pK_a$). At a pH where C is small, much larger shifts are required to significantly influence the titration curve. Generally, electrostatic interactions tend to lower and broaden the capacitance peaks, thus making the model calculations less sensitive to the choice of intrinsic parameters. For example, the ideal capacitance for calbindin at pH 4 is more than twice as large as in the real protein. As for BPTI and ovomucoid third domain, the root-mean-square value (rms) was found to be 0.4-0.5 pK units (Table I and II) indicating that better agreement with experiment may be obtained by slightly changing the intrinsic pK_a values. At the moment we do not offer any controlled way of effecting this, we merely note this as an operational possibility.

Conclusions

We have investigated a simple spherical model for proteins which accounts for a low dielectric interior. A possible generalization of the model is proposed with empirically adjustable polarizabilities to account for the general shape and polarization of the interior region. It is argued that maintaining the functional form of the multipolar corrections, suggested by the spherical solution, is desirable, given the non-monotonic nature of the corrections to the zeroeth-order Coulombic terms. That is, the simple spherical model highlights some counterintuitive behaviours, for example, the interaction between a pair of charges is not always enhanced by placing them in the vicinity of the low dielectric core. This is due to the orientation of the induced multipoles, relative to the charges.

We compared simulations using the zeroeth-order model with titration experiments on four different proteins, where titrating groups were close to the protein-solvent interface (and assumed external to the low dielectric region). Under these conditions the zeroeth-order model corresponds to the HDM. We showed that the HDM did remarkably well when compared with the experiments and, on this basis, we concluded that the multipolar corrections, due to the low dielectric interior would be small. Indeed, the HDM performed on average better than other approaches, which assumed a uniform low dielectric interior bounded by the protein-solvent interface. Such theories tend to obtain reasonable fits to experiment if the internal dielectric constant is chosen as an average between a typical hydrocarbon value ($\epsilon \approx 4$) and a much higher value reflecting greater protein polarizability and solvent penetration at the protein surface. Typically, an optimal *average dielectric constant* of around 20 gives a good fit to experiments. Our model is very different to these, as it maintains a type of "onion ring" approach to ϵ, which takes into account the expected inhomogeneity of the protein dielectric constant. For this reason we think our model has a sounder physical basis.

This notwithstanding, we found that the HDM is able to accurately predict both site and overall titration behaviour for a range of small proteins. This is invaluable for many-body protein simulations (protein aggregation etc.) where dielectric boundaries pose a difficult computational problem. Further, since salt particles are treated explicitly, the model is capable of describing multivalent ions and/or high salt concentrations. Thus, the straightforward strategy of assuming a uniform dielectric model seems to provide an accurate and simple model for protein electrostatics.

Acknowledgments

We wish to thank Professor Bengt Jönsson, Department of Biophysical Chemistry, Lund University for useful discussions and, for financial support: The Solander Program, The Swedish Foundation for Strategic Research and The Research School in Pharmaceutical Sciences, Sweden.

[1] J. G. Kirkwood, J. Chem. Phys. **2**, 351 (1934).
[2] C. Tanford and J. G. Kirkwood, J. Am. Chem. Soc. **79**, 5333 (1957).
[3] F. L. B. da Silva, B. Jönsson, and R. Penfold, Protein Science **10**, 1415 (2001).
[4] J. Warwicker and H. C. Watson, J. Mol. Biol. **157**, 671 (1982), URL `http://dx.doi.org/10.1016/0022-2836(82)90505-8`.
[5] J. Antosiewicz, J. M. Briggs, A. H. Elcock, M. K. Gibson, and J. A. McGammon, J. Comp. Chem. **17**, 1633 (1996).
[6] E. Demchuk and C. Wade, J. Phys. Chem. **100**, 17373 (1996).
[7] E. R. Georgescu, E. G. Alexov, and M. R. Gunner, Biophys J. **83**, 1731 (2002).
[8] J. Warwicker, Protein Science **13**, 2793 (2004).
[9] M. Feig and C. L. B. III, Current Opinion in Structural Biology **14**, 217 (2004).
[10] T. Simonson, J. Carlsson, and D. A. Case, J. Am. Chem.

Soc. **126**, 4167 (2004).
[11] C. Tan, L. Yang, and R. Luo, J. Phys. Chem. **110**, 18680 (2006).
[12] P. E. Smith, R. M. Brunne, A. E. Mark, and W. F. van Gunsteren, J. Phys. Chem. **97**, 2009 (1993).
[13] E. G. Alexov and M. R. Gunner, Biophys J. **5**, 2075 (1997).
[14] J. G. Kirkwood and J. B. Shumaker, Proceedings of the National Academy of Sciences **38**, 863 (1952).
[15] J. G. Kirkwood and J. B. Shumaker, PNAS **38**, 855 (1952), http://www.pnas.org/cgi/reprint/38/10/855.pdf, URL http://www.pnas.org.
[16] F. Bordi, C. Cametti, and G. Paradossi, Journal of Physical Chemistry **95**, 4883 (1991), ISSN 0022-3654, URL http://pubs3.acs.org/acs/journals/doilookup?in_doi=10.1021/j100165a051.
[17] M. Lund and B. Jönsson, Biochemistry **44**, 5722 (2005), URL http://dx.doi.org/10.1021/bi047630o.
[18] C. N. Schutz and A. Warshel, Proteins: Structure, Function, and Genetics **44**, 400 (2001).
[19] C. Tanford and R. Roxby, Biochemistry **11**, 2192 (1972).
[20] A. Warshel, S. T. Russel, and A. K. Churg, Proc. Natl. Acad. Sci. USA **81**, 4785 (1984).
[21] T. Kesvatera, B. Jönsson, E. Thulin, and S. Linse, Biochemistry **33**, 14170 (1994).
[22] T. Kesvatera, B. Jönsson, E. Thulin, and S. Linse, J. Mol. Biol. **259**, 828 (1996).
[23] R. Penfold, J. Warwicker, and B. Jönsson, J. Phys. Chem. B **102**, 8599 (1998).
[24] T. Kesvatera, B. Jönsson, E. Thulin, and S. Linse, Proteins **37**, 106 (1999).
[25] F. Carlsson, M. Malmsten, and P. Linse, J. Phys. Chem. **105**, 12189 (2001).
[26] M. Lund and B. Jönsson, Biophys. J. **85**, 2940 (2003), URL http://www.biophysj.org/cgi/content/abstract/85/5/2940.
[27] S. McGuffee and A. Elcock, Journal of the American Chemical Society **128**, 12098 (2006), URL http://dx.doi.org/10.1021/ja0614058.
[28] N. A. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[29] C. J. F. Böttcher, *Theory of Electric Polarization* (Elsevier, Amsterdam, 1973).
[30] C. E. Woodward and B. R. Svensson, J. Phys. Chem. **95**, 7471 (1991).
[31] M. S. Wisz and H. W. Hellinga, Proteins: Structure, Function, and Genetics **51**, 360 (2003), URL http://dx.doi.org/10.1002/prot.10332.
[32] V. Teixeira, C. Cunha, M. Machuqueiro, A. Oliveira, B. Victor, C. Soares, and A. Baptista, Journal of Physical Chemistry B **109**, 14691 (2005), ISSN 1520-6106, URL http://pubs3.acs.org/acs/journals/doilookup?in_doi=10.1021/jp052259f.
[33] T. Grycuk, The Journal of Chemical Physics **119**, 4817 (2003), URL http://link.aip.org/link/?JCP/119/4817/1.
[34] C. Labbez and B. Jönsson, Submitted (2007).
[35] B. Widom, J. Chem. Phys. **39**, 2808 (1963).
[36] M. Scarsi, J. Apostolakis, and A. Caflisch, J. Phys. Chem. **101**, 8098 (1997).
[37] J. J. Havranek and P. B. Harbury, Proc. Natl. Acad. Sci. **96**, 11145 (1999).
[38] J. Khandogin and C. L. r. Brooks, Biophys J **89**, 141 (2005), ISSN 0006-3495 (Print).
[39] C. A. Fitch, D. A. Karp, K. K. Lee, W. E. Stites, E. E. Lattman, and B. G.-M. E., Biophys J. **82**, 3289 (2002).
[40] T. Kesvatera, B. Jönsson, E. Thulin, and S. Linse, Proteins **45**, 129 (2001).
[41] M. Marquart, J. Walter, J. Deisenhofer, W. Bode, and R. Huber, Acta Crystallogr., Sect. B **39**, 480 (1983).
[42] R. Richarz and K. Wüthrich, Biochemistry **17**, 2263 (1978).
[43] K. Wuthrich and G. Wagner, J. Mol. Biol. **130**, 1 (1979).
[44] W. Bode, A. Z. Wei, R. Huber, E. Meyer, J. Travis, and S. Neumann, EMBO **5**, 2453 (1986).
[45] W. Shaller and A. D. Robertson, Biochemistry **34**, 4714 (1995).
[46] C. Tanford and J. D. Hauenstein, J. Am. Chem. Soc. **78**, 5287 (1956).
[47] D. D. Leonidas, R. Shapiro, L. Irons, N. Russo, and K. R. Acharya, Biochemistry **36**, 5578 (1997).
[48] D. M. E. Szebenyi and K. Moffat, J. Biol. Chem. **261**, 8761 (1986).

TABLE I: Measured (∼25-50 mM salt) and calculated pK_a values for BPTI. The MC (35 mM salt) and PB (150 mM salt) results are based on the crystal structure[41] (4PTI). The MC calculation was performed using an all atom description and a uniform dielectric response ($\epsilon_p = \epsilon_s = 80$).

	Ideal	PB[5]	MC	Exp.[42, 43]
Asp3	4.0	3.3	2.8	3.0
Asp50	4.0	2.5	2.5	3.4
Glu7	4.4	4.7	2.9	3.7
Glu49	4.4	3.5	3.8	3.8
Lys15	10.4	10.4	10.3	10.6
Lys26	10.4	10.4	10.7	10.6
Lys41	10.4	10.3	11.1	10.8
Lys46	10.4	10.0	10.4	10.6
Ctr	3.8	3.8	2.7	2.9
Ntr	7.5	7.2	7.3	8.1
rms	0.6	0.7	0.5	

TABLE II: Measured and calculated pK_a values for turkey ovomucoid third domain at 10 mM 1:1 salt concentrations. The MC and MTK calculations are both based on the crystal structure[44] (1PPF). In the simulations, the amino acid model was applied.

	Ideal	MTK[37]		MC	Exp.[45]
		ϵ_p=4, relaxed	ϵ_p=20	ϵ_p=80	
Asp7	4.0	2.1	3.0	3.2	< 2.6
Glu10	4.4	4.0	3.5	3.9	4.1
Glu19	4.4	3.1	2.7	3.4	3.2
Asp27	4.0	2.9	3.7	2.7	< 2.3
Glu43	4.4	5.6	4.7	4.1	4.7
Ctr56	3.8	2.6	3.2	2.5	< 2.5
rms	1.2	0.5	0.7	0.4	

Figure 1. Distances (Å) to the center-of-charge of: titrateable sites (circles) and the protein surface closest

to that site (line). Plotted for calbindin (left) and ribonuclease (right) using an all atom description based on the crystal structures.

Figure 2. Illustration of a protein model, where all charges (salt and protein) are located in a high-dielectric region (ϵ_s) and the non-polar interior is described by a spherical, charge-less low dielectric cavity (ϵ_p) of radius a.

Figure 3. Top: Effect of the non-polar protein interior (ϵ_p) on the electrostatic pair interaction between two positive charges situated in a polar region (ϵ_s) a few Ångstroms away from the dielectric interface. Calculated using Eq. 4 with $r_i = r_j = 15$ Å and $a = 13$ Å. Bottom: Plotted as the difference with various values of a.

Figure 4. MC simulation of the decay of the excess chemical potential of a univalent ion outside a low dielectric, neutral sphere (ϵ_p=2, ϵ_s=80) of different radii, a. The energy offsets are arbitrary.

Figure 5. Measured[46] and calculated titration curve for ribonuclease A at 10 mM salt. The MC results are based on the crystal structure 1AFU[47].

Figure 6. Titration curve for calbindin obtained from intrinsic pK_a-values (ideal), measurements[24, 40], and MC simulations using both the HDM and assuming the protein core to be a low dielectric spherical cavity of radius $a = 8$ Å and dielectric constant, ϵ_p=2. The simulations were performed using the full atomic structure (PDB entry 3ICB[48]) and with a ionic strength of 5 mM matching the experimental conditions.

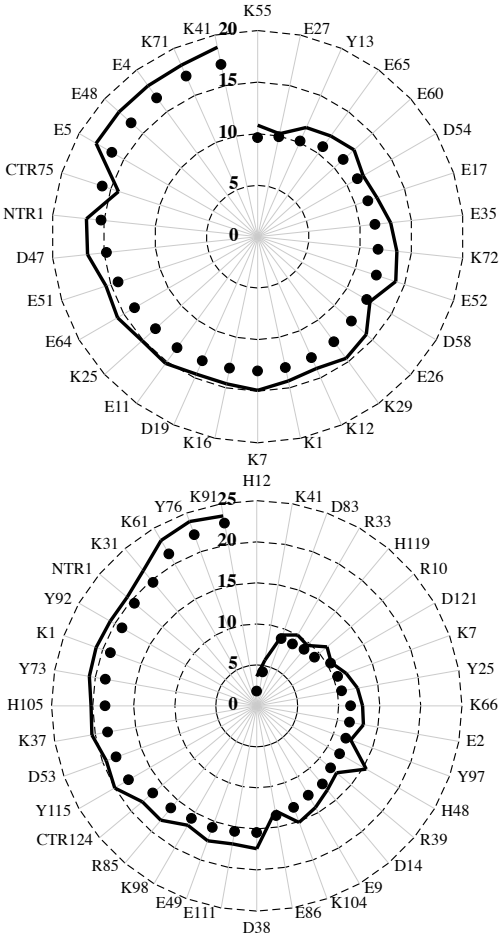

K55
E27
Y13
E65
E60
D54
E17
E35
K72
E52
D58
E26
K29
K12
K1
K7
D19
E11
K25
E64
E51
D47
NTR1
CTR75
E5
E48
E4
K71
K41
20
15
10
5
0
H12
K91
Y76
K61
K31
NTR1
Y92
K1
Y73
H105
K37
D53
Y115
CTR124
R85
K98
E49
E111
D38
E86
K104
E9
D14
R39
H48
Y97
E2
K66
Y25
K7
D121
R10
H119
R33
D83
K41
25
20
15
10
5
0

FIG. 1:

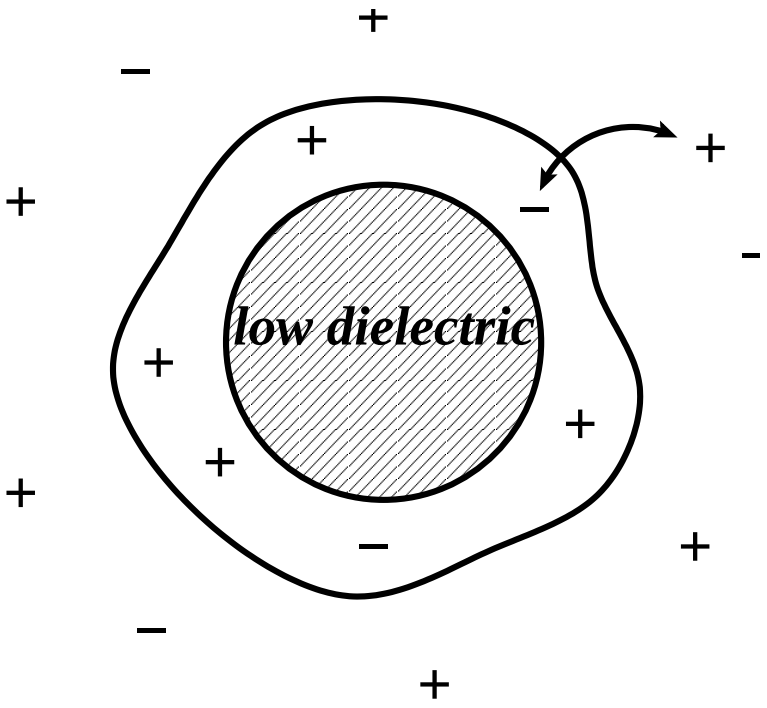
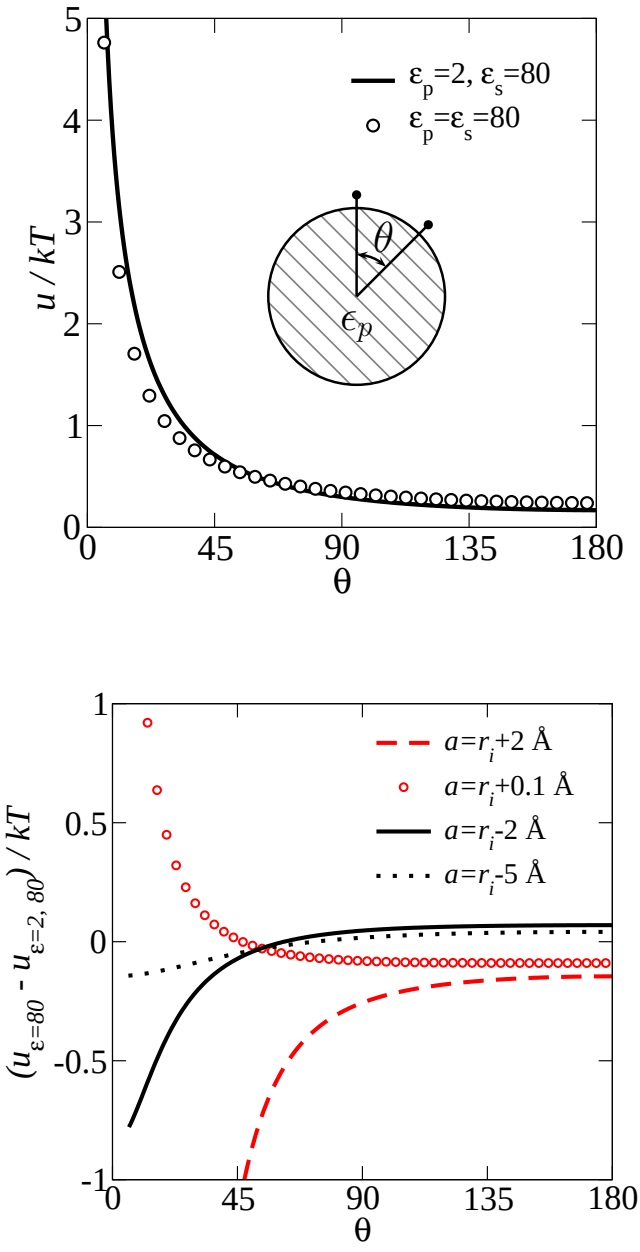

low dielectric

FIG. 2:


ε_p=2, ε_s=80
ε_p=ε_s=80
θ
ε_p
u / kT
θ
0
45
90
135
180
0
1
2
3
4
5

a=r_i+2 Å
a=r_i+0.1 Å
a=r_i-2 Å
a=r_i-5 Å
(u_ε=80 - u_ε=2, 80) / kT
θ
0
45
90
135
180
-1
-0.5
0
0.5
1

FIG. 3:

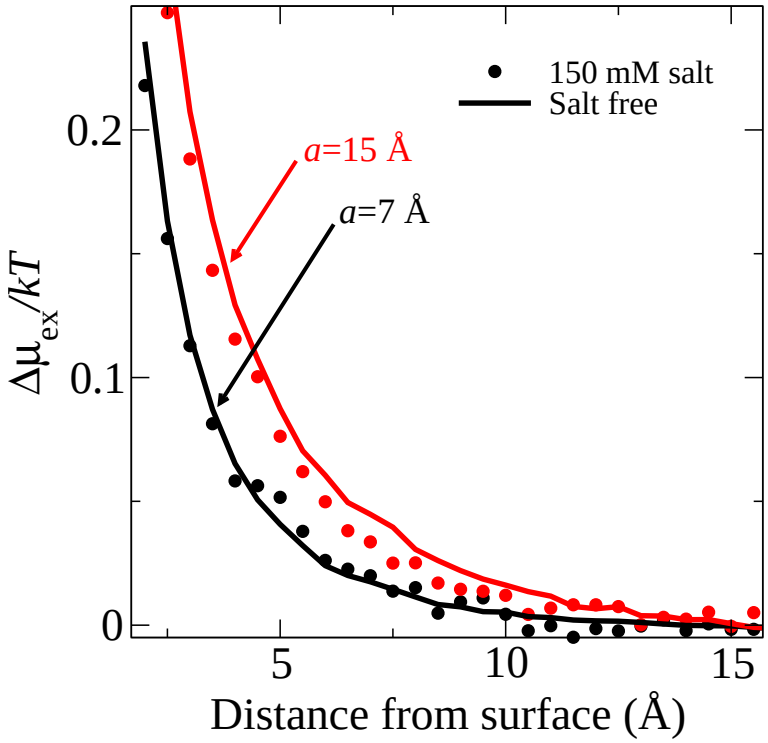

150 mM salt
Salt free
a=15 Å
a=7 Å
Δμex/kT
0.2
0.1
0
5
10
15
Distance from surface (Å)

FIG. 4:

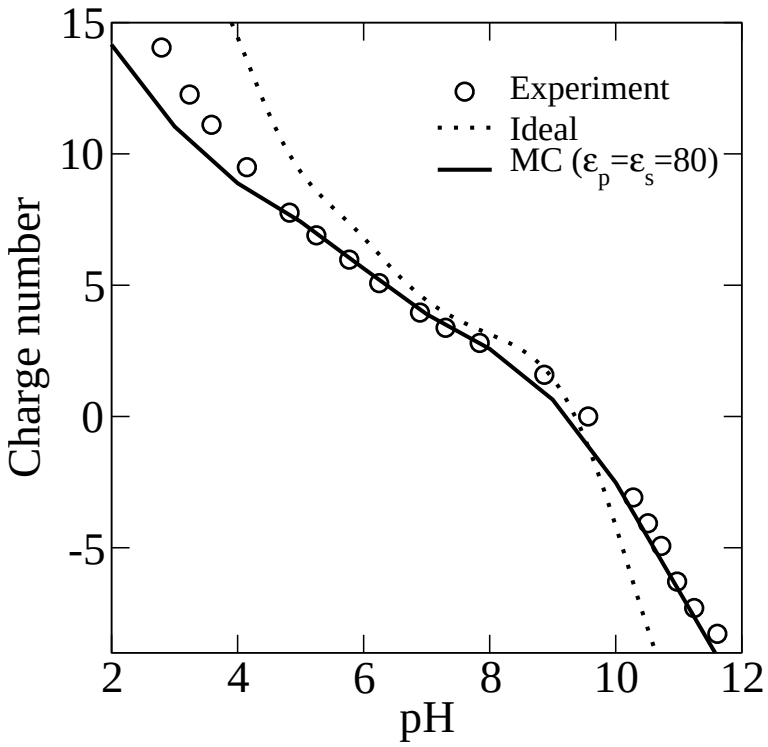

Experiment
Ideal
MC (εp=εs=80)
Charge number
15
10
5
0
-5
2
4
6
8
10
12
pH

FIG. 5:

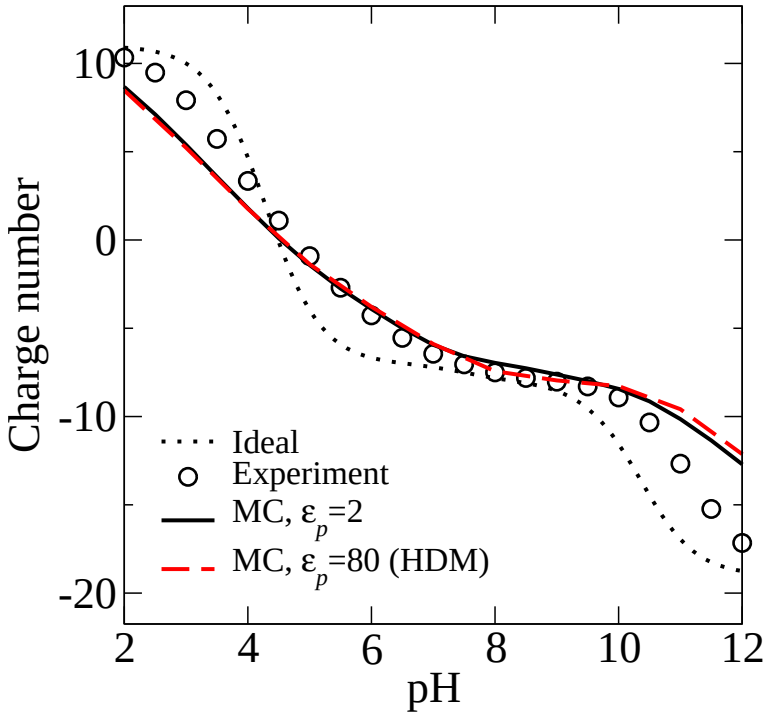

Charge number
10
0
-10
-20
2
4
6
8
10
12
pH
Ideal
Experiment
MC, ε$_p$=2
MC, ε$_p$=80 (HDM)

FIG. 6: