Electrostatics in macromolecular solutions

Jönsson, Bo; Lund, Mikael; Barroso da Silva, Fernando L.

Published in:
Food Colloids: Self-Assembly and Material Science

DOI:
10.1039/9781847557698-00127

2007

October 31, 2006

Electrostatics in Macromolecular Solution

Bo Jönsson, Mikael Lund and Fernando L. Barroso daSilva

Theoretical Chemistry, Chemical Center, POB 124, S-221 00 Lund, SWEDEN

Abstract

An overview of the interaction between charged macromolecules in aqueous solution is presented. The starting point is the dielectric continuum model and the Debye-Hückel equation. The usefulness of the simple theory is emphasized in particular for biological macromolecules, whose net charge or surface charge density often is low. With more highly charged macromolecules or aggregates it may be necessary to go beyond the simple Debye-Hückel theory and invoke the non-linear Poisson-Boltzmann equation or even to approach an exact solution using Monte Carlo simulations or similar techniques. The latter approach becomes indispensable when studying systems with divalent or multivalent (counter)-ions. The long range character of the electrostatic interactions means that charged systems of varying geometry - spheres, planes, cylinders... - often have many properties in common. Another consequence is that the detailed charge distribution on a macromolecule is less important. Many biological macromolecules contain titratable groups, which means that the net charge will vary as a consequence of solution conditions. This gives an extra attractive contribution to the interaction between two macromolecules, which might be particularly important close to their respective isoelectric points. The treatment of flexible polyelectrolytes/polyampholytes requires some extra efforts in order to handle the increasingly complex geometry. A theoretical consequence is that the number of parameters - chain length, charge density, polydispersity etc - prohibits the presentation of a simple unified picture. An additional experimental, and theoretical, difficulty in this context is the slow approach towards equilibrium, in particular with high molecular weight polymers. A few generic situations where polyelectrolytes can act both as stabilizers and coagulants can, however, be demonstrated using simulation techniques.

Introduction - The Dielectric Continuum Model

An aqueous solution containing biological molecules can in a general sense be described as an electrolyte solution. That is, it contains simple ions such as Na^+, K^+, Cl^- etc., but it can also include macromolecules with a net charge significantly different from unity. DNA, proteins and polysaccharides are important examples of natural origin but different synthetic additives can also be described as charged macromolecules, sometimes collectively refered to as *polyelectrolytes*. It is our intention to discuss the interaction/stability of biological polyelectrolytes in a few generic situations, some of which hopefully are of interest for a food chemist.

Despite the progress in computer technology and numerical algorithms during the last decades, it is still not feasible to treat a general solution of charged macromolecules in an atomistic model. This becomes especially clear when we are trying to calculate the interaction between macromolecules and how the interaction can be modulated by other charged species.

The alternative at hand is to use the dielectric continuum model, crudely refered to as *the Primitive Model*. The solvent is then described as a structureless medium solely characterized by its relative dielectric permittivity, ϵ_r. This simplification facilitates both the theoretical treatment and the conceptual understanding of electrostatic interactions in solution. In contrast to its name, it is a very sophisticated approximation, which allows an almost quantitative description of widely different phenomena such as sea water and cement paste! In the Primitive Model we treat all charged species as charged hard spheres and the interaction, between two charges i and j separated a distance r, can be formally described as,

$$u(r) = \frac{Z_i Z_j e^2}{4\pi\epsilon_0\epsilon_r r} \qquad r > d_{hc} \tag{1}$$

$$u(r) = \infty \qquad r < d_{hc} \tag{2}$$

where Z_i is the ion valency, e the elementary charge, ϵ_0 the dielectric permittivity of vacuum and d_{hc} is the hard sphere diameter of the ion. For simplicity, we will in this communication mostly assume it to be the same for all ionic species and equal to 4 Å.

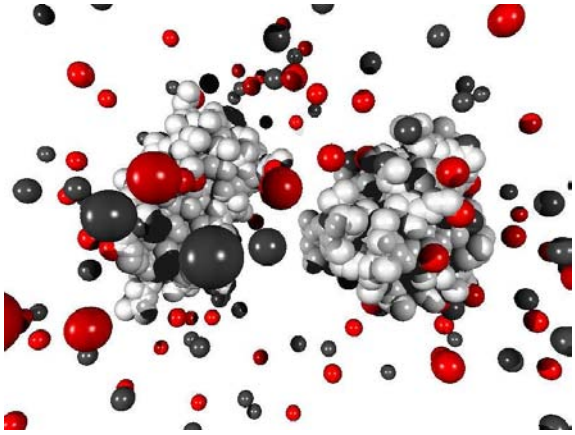

Figure 1: Snapshot from a MC simulation of two proteins. The black and grey spheres illustrate mobile cations and anions, while amino acids are depicted as a white spheres, clustered to form the two proteins. In a simulations, the proteins are displaced along a line and rotated independently. Ions are displaced in all three directions and the whole system is enclosed in sphere of appropriate radius.

These charges can be the small mobile ions in a salt solution, but they can also be the charged groups on a protein or some other macromolecule. The model is schematically depicted in Figure 1 with two macromolecules in a salt solution. We will solve this model exactly using Monte Carlo (MC) simulations or in an approximate way with either the Poisson-Boltzmann (PB) equation or its linearized version, the Debye-Hückel (DH) equation. For an introduction to the DH theory, the reader is recommended to consult the excellent textbook of Hill [1]. Engström and Wennerström [2] has solved the PB equation for a charged surface with neutralising counterions and their paper is a good starting point on this subject. Monte Carlo and other simulations are well described in the textbooks by Allen and Tildesley [3] and by Frenkel and Smit [4]. MC simulations allow us to emphasize where the simple theory is applicable and where a more accurate treatment is needed. The simulations also give an oppurtunity to clarify certain physical mechanisms, providing a deeper understanding of the system at hand.

The paper is arranged as follows:

- A simple electrolyte solution.
- A charged macromolecule in a salt solution.

- The interaction between two charged macromoleules.
- The addition of polyelectrolytes/polyampholytes.
- Attraction due to charge regulation.
- Protein polyelectrolyte complexes.

A Simple Electrolyte Solution

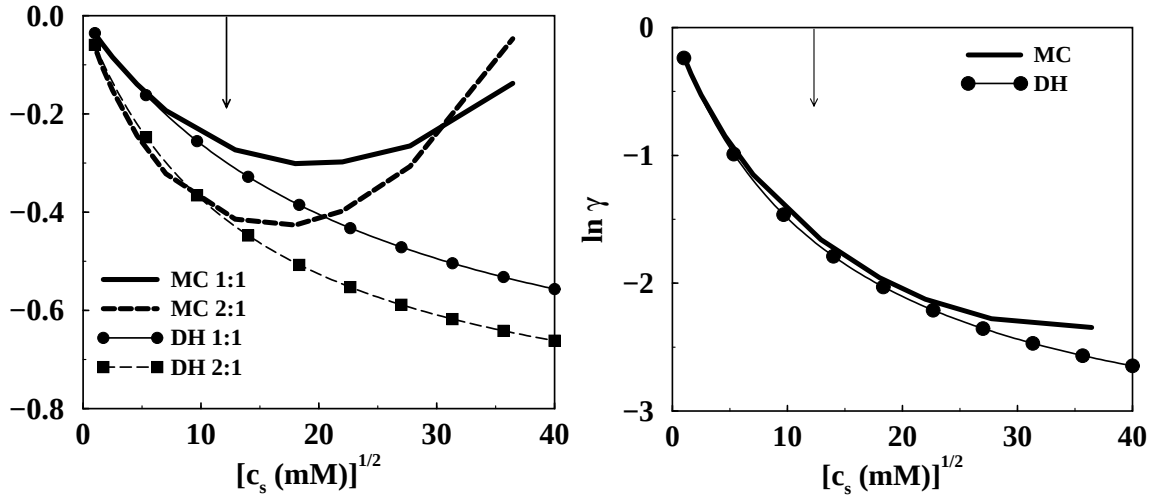

Figure 2: Individual activity factors from MC simulations and from the DH theory with the hard core diameter equal to 4 Å. a) Monovalent ion in a 1:1 and 2:1 salt and b) Divalent ion in a 2:1 salt. The arrows indicate physiological salt condition.

An important property in an electrolyte solution is the activity factor, γ, or excess chemical potential, μ_{ex}, which is a part of the total chemical potential, μ,

$$\mu = \mu_0 + kT \ln c + kT \ln \gamma = \mu_0 + \mu_{id} + \mu_{ex} \tag{3}$$

μ_0 is an uninteresting reference chemical potential and c is the concentration. It is straightforward to calculate γ in a Monte Carlo simulation, but we can also obtain it from the Debye-Hückel approximation,

$$kT \ln \gamma^{DH} = -\frac{Z^2 e^2 \kappa}{8\pi\epsilon_0\epsilon_r(1+\kappa d_{hc})} \tag{4}$$

The important quantity in eq.(4) is the inverse screening length, κ,

$$\kappa^2 = \frac{e^2}{\epsilon_0\epsilon_r kT}\sum_i c_i z_i^2 \tag{5}$$

which is proportinal to the ionic strength. Figure 2 shows how γ varies as a function of salt concentration for two different salts. The accuracy of the simple DH theory is surprisingly good and the main discrepancy comes from the too approximate treatment of the excluded volume effect, *i.e.* the hard core interaction. A knowledge of γ allows us to calculate a number of interesting quantities. For example, we can calculate the dissolution of carbon dioxide in the ocean. The high salt content of the oceans increases the solubility of CO_2, which is apparent from the equilibrium relations,

$$H_2CO_3 \rightarrow HCO_3^- + H^+ \qquad K_1 = \frac{c_{HCO3}c_H}{c_{H2CO3}}\frac{\gamma_{HCO3}\gamma_H}{\gamma_{H2CO3}} = K_1^S\frac{\gamma_{HCO3}\gamma_H}{\gamma_{H2CO3}} \tag{6}$$

$$\mathrm{HCO_3^-} \rightarrow \mathrm{CO_3^{2-}} + \mathrm{H^+} \qquad K_2 = \frac{c_{CO3}c_H}{c_{HCO3}}\frac{\gamma_{CO3}\gamma_H}{\gamma_{HCO3}} = K_2^S\frac{\gamma_{CO3}\gamma_H}{\gamma_{HCO3}} \tag{7}$$

Note that *thermodynamic* equilibrium constants, K_1 and K_2, are true constants in contrasts to the *stoichiometric* ones, K_1^S and K_2^S. Table 1 presents experimental and simulated activity factors for some salts relevant for sea water. The departure from ideality ($\gamma = 1$) is non-negligible and as a consequence the dissolution of CO_2 in sea water is significantly larger than in fresh water. The excellent agreement between measured and simulated activity factors in Table 1 gives a strong support for the Primitive Model.

Salt	γ_{Exp}	γ_{Sim}
Na_2SO_4	0.37	0.37
K_2SO_4	0.35	0.36
NaCl	0.67	0.67
KCl	0.66	0.66
$CaSO_4$	0.14	0.15

Table 1: Experimental [5, 6] and simulated [7] mean activity factors in sea water at 298 K. The salinity is 3.5 %.

A Charged Macromolecule in a Salt Solution

We can use the activity factors in order to study how the binding of a charged ligand to a charged macromolecule is affected by addition of salt or changes in pH - a change in pH means that the net charge of both ligand and macromolecule can vary. The changes will affect the electrostatic interactions and are almost quantitatively captured by the activity factors. The simplest approach would then be to treat the macromolecule as a charged spherical object and directly apply eq.(4). Let us take the calcium binding to the small chelator 5,5'-Br_2BAPTA as an example [8],

$$Ch + Ca^{2+} \rightarrow ChCa \qquad K = \frac{c_{ChCa}}{c_{Ch}c_{Ca}}\frac{\gamma_{ChCa}}{\gamma_{Ch}\gamma_{Ca}} = K_s\frac{\gamma_{ChCa}}{\gamma_{Ch}\gamma_{Ca}} \tag{8}$$

Since K is a true constant we can write a relation between the stoichiometric binding constants at two different salt concentrations as,

$$K_s^I\frac{\gamma_{ChCa}^I}{\gamma_{Ch}^I\gamma_{Ca}^I} = K_s^{II}\frac{\gamma_{ChCa}^{II}}{\gamma_{Ch}^{II}\gamma_{Ca}^{II}} \tag{9}$$

The charge of the chelator is $-4e$ at neutral pH and it is assumed to have a radius of 7 Å. When calcium is bound to the chelator it is simply modeled by a reduction of the chelator charge from $-4e$ to $-2e$. This simple model captures the salt dependence from 1 mM to 1 M salt. Table 2 shows how the stoichiometric binding constant, K^s, varies with salt concentration. Both simulated and DH results are in excellent agreement with experiment.

A quantitatively more correct alternative is to use the so-called Tanford-Kirkwood (TK) model [9]. The TK model takes the detailed charge distribution into account and solves the electrostatic problem using a variant of the DH approximation. The final result is the free energy for the macromolecule in a salt solution. For not too highly charged macromolecule this is usually a very efficient and reliable approach and the relevant equations are easily evaluated numerically. Figure 3 shows how the calcium binding constant to the small protein calbindin D_{9k} varies with salt concentration [10]. Both simulated and TK results are based on the detailed charge distribution of the protein with the calbindin structure obtained from an x-ray study [11].

c_s (mM)	ΔpK_s^{Exp}	ΔpK_s^{Sim}	ΔpK_s^{DH}
2	0.00	0.00	0.00
10	0.26	0.32	0.32
25	0.64	0.60	0.59
50	0.89	0.85	0.84
100	1.20	1.12	1.11
300	1.58	1.58	1.59
500	1.77	1.79	1.81
1000	1.97	2.05	2.09

Table 2: Shift in the stoichiometric calcium binding constant for the chelator BAPTA. 2 mM salt has been taken as a reference point and the shifts are calculated relative this value.

The agreement between the two theoretical approaches is excellent and so is the comparison with experimental results.

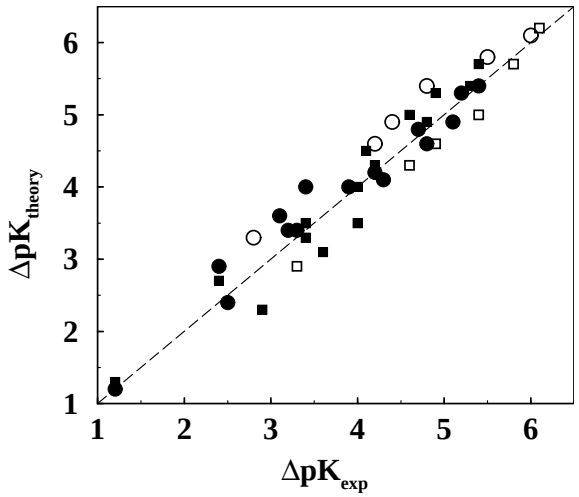

Figure 3: A comparison of experimental and theoretical binding constant shifts for the calcium binding protein calbindin D_{9k}. The electrostatic interactions have been modified by adding salt in the range 2-150 mM and by mutating (neutralizing) charge residues in the protein [10]. The symbols represent different mutations (charge neutralization of acidic residues) and different salt concentrations. Spheres are simulated data and squares are calculated using the TK approach. Filled symbols describe the addition of KCl and open symbols the addition of K_2SO_4. The dashed line corresponds to perfect agreement. The shifts are calculated relative to the native protein at 2 mM salt concentration.

It is interesting to investigate the limitations of the TK approach and one should expect deviations from the simulated values for a really highly charged protein. This is indeed the case and Figure 4 reveals a typical behaviour for the binding of a charged ligand to an oppositely charged macromolecule or particle. That is, when the charge reaches a certain niveau, then the electrostatic response is no longer linear but it approaches an asymptotic value. This means that the binding becomes "saturated" and, for example, a further increase of negatively charged residues in a protein does not lead to an increased binding of calcium.

The electrostatic model in colloid chemistry has always been one with a uniform dielectric permittivity for the whole system, typically chosen to be equal to that of water. In the calculations reported above we have followed this tradition. Obviously, the dielectric permittivity of a protein is different from that of bulk water, but we do not know its exact value and to be more formal, it is not a well-defined quantity. We also note that charged species prefer the high dielectric region - ions dissolve in water and not in oil! Another way to express it is to say that the electric field lines remain in the aqueous phase, hence a small body of low dielectric material

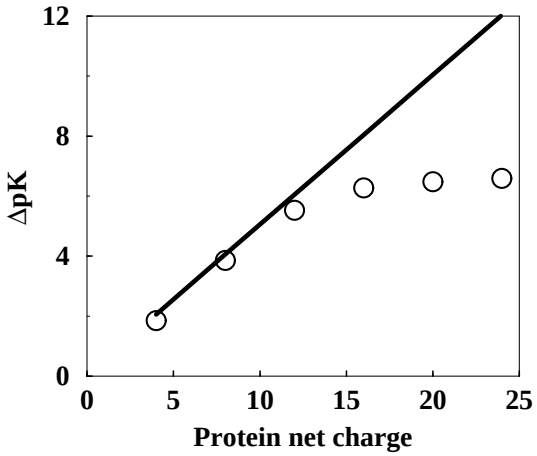

Figure 4: Binding constant shifts as a function of protein net charge - comparison of DH (line) and MC simulations (symbols). The protein is modelled as a sphere of radius is 14 Å with two binding sites close to the surface. The shift refers to a change in salt concentration from 1 to 500 mM. The protein concentration is $20\mu M$ and the binding process involves two divalent ions.

has only a marginal effect on the electrostatic interactions. These conclusions are supported by a wealth of experimental results on colloidal systems.

In biophysics, the opposite paradigm prevails and the low dielectric interior of a protein is usually assumed to be the clue to many properties of biochemical interest. The electrostatic approach is based on the PB or DH equation. A technical feature with the "low dielectric" assumption is that the calculations contain a divergence, which can cause numerical problems. Or, it can be used as a "fitting parameter". The divergence in electrostatic calculations invoking a low dielectric region is apparent in many applications. One very clear such example is the determination of apparent pK_a's in the protein calbindin - see Table 3, which have been determined experimentally by Kesvatera *et al.* [12] and theoretically by Spassov and Bashford [13] using a low dielectric response for the protein. Juffer and Vogel [14] have extended the Debye-Hückel calculations of Spassov and Bashford and allowed for a high dielectric response from the protein. The paper by Kesvatera also contains results from MC simulations using a uniform dielectric response equal to that of water. Obviously the calculations using a low dielectric interior containing charged groups are unable to describe the electrostatic interactions in calbindin and the results are unphysical.

Amino Acid	Exp.	Theory-Spassov	Theory-Juffer	Theory-Kesvatera
Glu-27	6.5	**21.8**	5.2	4.7
Asp-54	3.6	**16.9**	4.8	4.4
Asp-58	4.4	**9.1**	4.8	4.8
Glu-60	6.2	**13.2**	5.6	6.0
Glu-65	5.4	**12.7**	4.6	5.0
Rms.	-	**10.6**	0.93	0.92

Table 3: The apparent pK_a of titrating acidic groups in calbindin D_{9k}. Experimental and various theoretical results. The "low dielectric" results of Spassov and Bashford have been highlighted. Both Spassov-Bashford and Juffer-Vogel have used the DH approximation, but the latter authors have assumed a uniformly high dielectric permittivity in the same way as Kesvatera *et al.*. The rms deviations are given in units of pK_a.

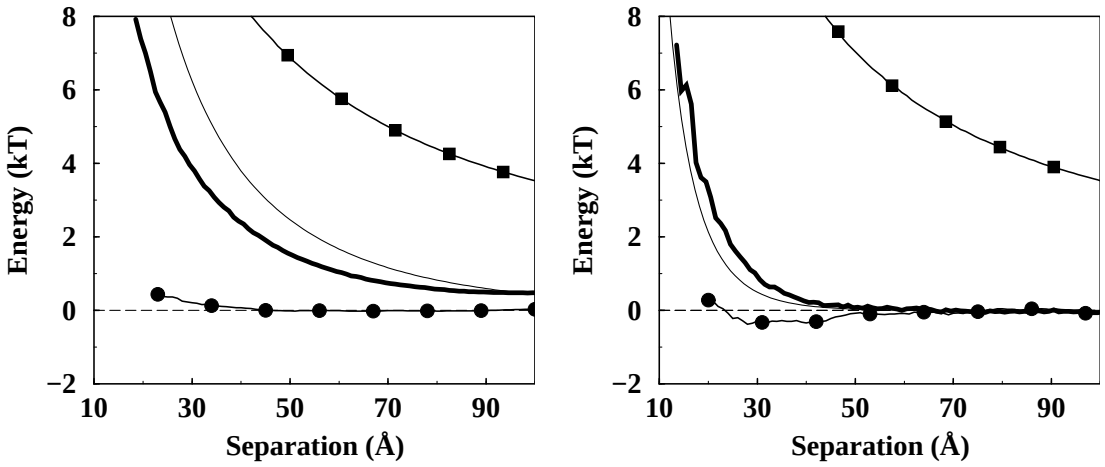

Figure 5: The interaction between two smMLCK peptides at two different salt concentrations; left=4 mM and right=100 mM of a monovalent salt. The smMLCK peptide consists of 15 amino acids and its net charge is +7e. Solid fat lines show the simulated free energy of interaction, while thin solid line is from the screened Coulomb interaction, eq.(10). The thin line with filled circles is the simulated total energy of interaction and the line marked with filled squares is the electrostatic interaction between the charges on the two peptides only.

The Electrostatic Interaction Between Two Proteins

The interaction of two peptides

Calmodulin binds to myosin light chain kinase (MLCK) via a small peptide rich in basic residues. Calmodulin and the peptide forms a complex, which has been isolated and crystallized. We have taken the peptide, smooth muscle MLCK (= smMLCK), from this complex and studied the interaction between a pair. The net charge of smMLCK at neutral pH is close to $+7e$ and the two peptides repel each other, see Figure 5a, that is the free energy of interaction is positive. The unscreened direct electrostatic interaction between the peptides is of course strongly repulsive, but the total electrostatic energy, including the background electrolyte, is essentially zero or sligthly attractive for all separations. Thus, the repulsion between the equally charged peptides is totally dominated by the entropy - the entropy of salt and counterions. An increase in salt concentration from 4 to 100 mM does not change this picture - Figure 5b.

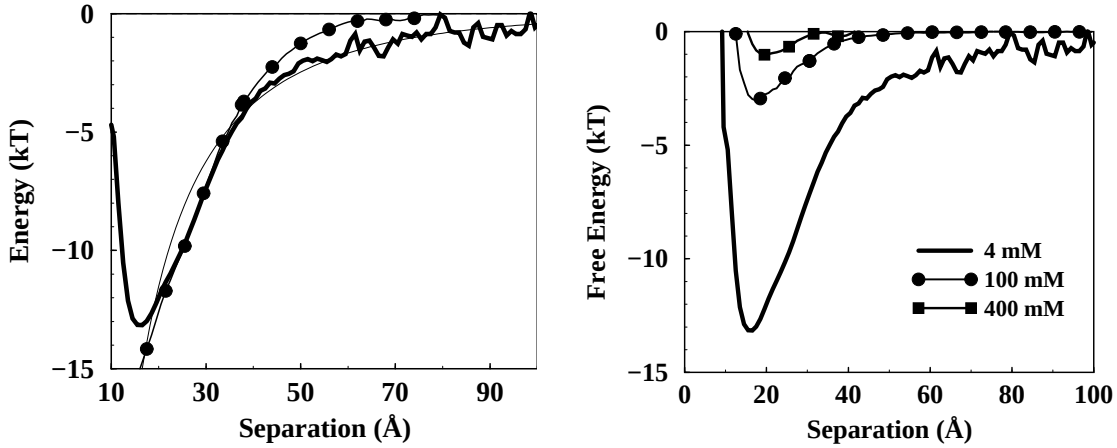

Figure 6: The interaction between an smMLCK peptide and a fragment of calmodulin. The net charge of smMLCK is $+7e$ and the calmodulin fragment has a charge of $-8e$. Left: The salt concentration is 4 mM. The solid fat line is the simulated free energy of interaction, while the thin line is the corresponding screened Coulomb interaction. The thin line with filled circles is the simulated total energy of interaction. Right: The effect of added salt on the free energy of interaction.

A different picture emerges for the interaction of two oppositely charged peptides. Figure 6

shows the free energy of interaction between smMLCK and a peptide section from calmodulin, comprising Glu45-Glu67 with a net charge of $-8e$. The interaction free energy is strongly attractive and so is the total energy. Thus, the attraction is energy driven and the entropy change is in this case only marginal.

The interaction between two charged macromolecules in a salt solution is screened by salt particles and one can derive an expression for their free energy of interaction, $A(r)$, based on the DH approximation,

$$A(r)/kT = l_B Z_1 Z_2 \frac{\exp(-\kappa r)}{r} \tag{10}$$

where we for convenience have introduced the Bjerrum length, $l_B = e^2/4\pi\epsilon_0\epsilon_r kT$. Note that $A(r)$ is a free energy. Figure 5 shows that the screened Coulomb potential is a good approximation and it is semi-quantitatively correct at both salt concentrations.

The results presented here for these peptides is generic and is found in many cases with charged macromolecules or particles. The geometry is not crucial and the same qualitative behaviour is found for both interacting planes and interacting spheres. The screened Coulomb potential captures the change in free energy when the two macromolecules approach each other. It is, however, questionable to partition the screened Coulomb interaction into energy and entropy terms. More elaborate forms of the screened Coulomb potential can be derived [15], where the macromolecular size is taken into account. The comparison in this section has been limited to a uni-uni valent electrolyte and to situations where κ^{-1} is of the same order or larger than the macromolecular dimension. To extend the use of the screened Coulomb potential to multivalent electrolytes usually leads to qualitatively incorrect results - see next section.

The effect of multivalent ions

Above we have shown how the simple theory, the screened Coulomb potential, is capable of an almost quantitative description of the interaction between two charged proteins. This good agreement is limited to systems containing only monovalent counterions. There is a qualitative difference between the interaction of two charged macromolecules in the presence of monovalent and in the presence of multivalent counterions. In the latter case the mean field approximation behind the DH equation breaks down and one has to rely on simulations or more accurate theories like the hypernetted chain equation [16, 17]. The deviation from the mean field description due to *ion-ion correlations* has such a physical origin that the effect should be independent of the particular geometry of the charged aggregates. Clearly there are quantitative differences between cylindrical, spherical or irregularly shaped or flexible charged colloidal species, but the basic mechanism operates in the same way. The importance of ion-ion correlations can be seen from Fig.7, where the free energy of interaction for two charged spherical aggregates has been calculated from an MC simulation. For monovalent counterions there is a monotonic repulsion in accordance with the screened Coulomb equation, eq.(10), but with multivalent counterions or a solvent with a low dielectric permittivity, the entropic double layer repulsion decreases and eventually the correlation term starts to dominate. This phenomenon can be seen as a balance between entropy and energy. For two weakly or moderately charged macromolecules with monovalent counterions, the dominant contribution to the free energy of interaction comes, as we have seen in Figure 5, from a reduction in entropy when the two counterion clouds start to overlap. The energy of interaction is always attractive and is only weakly dependent on the counterion valency. The important difference between a system with monovalent or divalent counterions, is the reduced entropy of the latter due to a lower number density of counterions. Thus, any change that reduces the entropy and/or increases the electrostatic interactions will eventually lead to a net attractive interaction.

This is for a model system with spheres with net charges, but the same mechanism is

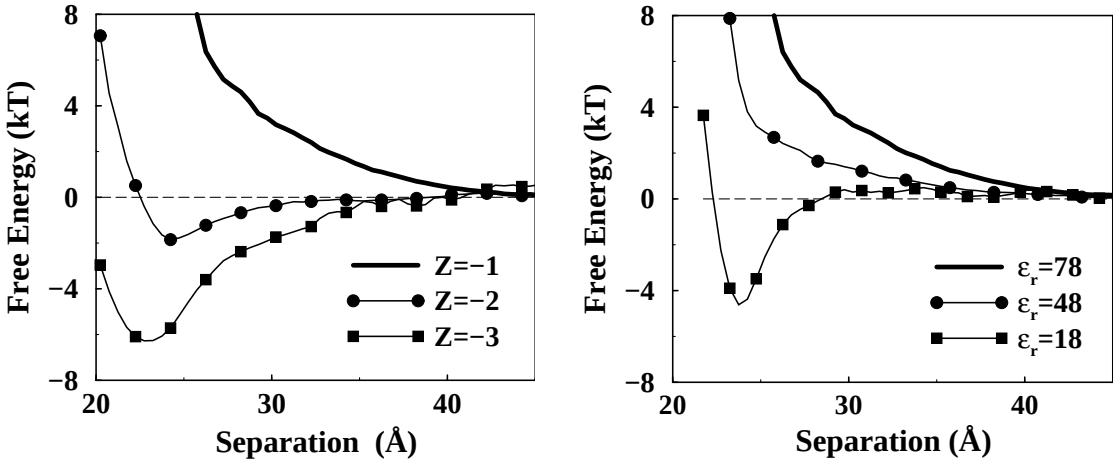

Figure 7: a) The free energy of interaction between two spherical aggregates of radius 10 Å and net charge 24. The system contains no salt but only counterions of different valency. The dielectric permittivity is 78 and the temperature 298 K. b)The same as in a) with monovalent counterions and variation of the relative dielectric permittivity.

also operating between two protein molecules with discrete charge distributions and irregular form[18] and between two DNA molecules [19].

The effect of titrating groups

All proteins and many other macromolecules contain ionizable residues whose ionization status depends on the interaction with other molecules. This means that the electrostatic interaction between two proteins, besides the interaction between their average charges, also will contain terms originating from induced charges. These interactions can be formalized in a statistical mechanical perturbation approach [20, 21] and a protein is characterized not only by its average net charge, but also by its *capacitance*. The induction interaction is important for the interaction of an approximately neutral protein with another charged macromolecule. The protein capacitance is a function of the number of titrating residues and will display maxima close to the pK_a's of the titrating amino acids. In this section we will derive a formal expression for the capacitance. Consider the macromolecules A and B, described by two set of charges $[\mathbf{r}_i, z_i]$ and $[\mathbf{r}_j, z_j]$, respectively. Their mass centra are separated by $\mathbf{R}$, which means that the distance between two charges i and j is given by $r_{ij} = |\mathbf{R} + \mathbf{r}_j - \mathbf{r}_i|$. The average net charge of the distributions need not be zero, that is $\langle Z_A \rangle \neq 0$, where $\langle Z_A \rangle = \langle \sum z_i \rangle$. The free energy of interaction can be written as,

$$A(R)/kT = -\ln \langle \exp(-U(R)/kT) \rangle_0 \approx \langle U(R)/kT \rangle_0$$
$$-\frac{1}{2}\left\langle (U(R)/kT)^2 \right\rangle_0 + \frac{1}{2}[\langle U(R)/kT \rangle_0 + \frac{1}{2}\left\langle (U(R)/kT)^2 \right\rangle_0]^2 \qquad (11)$$

where $U(R)$ is the interaction between the two charge distributions and $\langle ... \rangle_0$ denotes an average over the unperturbed system, which in the present case is the single isolated protein in solution. The interaction energy is simply the direct Coulomb interaction between the two charge distributions,

$$U(R)/kT = \sum_i \sum_j \frac{l_B z_i z_j}{r_{ij}} \qquad (12)$$

We can make a Taylor series expansion of U, assuming that $R >> r_i$. This expansion will include ion-ion interaction, ion-dipole interaction, dipole-dipole interaction etc. It will also include charge-induced charge and induced charge-induced charge interactions. Thus, we can

write an approximation to the free energy including all terms of order up to $1/R^2$. Note that the ion-dipole interaction disappears in first order and that the first non-vanishing dipole term, $-l_B^2 Z^2 \mu^2/6R^4$ is of order $1/R^4$.

$$A(R)/kT \approx \frac{l_B \langle Z_A \rangle \langle Z_B \rangle}{R} - \frac{l_B^2}{2R^2}(\left\langle Z_A^2 \right\rangle - \langle Z_A \rangle^2)(\left\langle Z_B^2 \right\rangle - \langle Z_B \rangle^2)$$
$$-\frac{l_B^2}{2R^2}((\left\langle Z_A^2 \right\rangle - \langle Z_A \rangle^2) \langle Z_B \rangle^2 + (\left\langle Z_B^2 \right\rangle - \langle Z_B \rangle^2) \langle Z_A \rangle^2) \tag{13}$$

The first term is the direct Coulomb term and the following term is the *induced charge-induced charge* and the last terms are the *charge-induced charge* interactions. Note also that $\langle Z^2 \rangle \neq \langle Z \rangle^2$. If the molecules are identical, that is $\langle Z_A \rangle = \langle Z_B \rangle = \langle Z \rangle$, then the expression simplifies to,

$$A(R)/kT \approx -\frac{l_B \langle Z \rangle^2}{R} - \frac{l_B^2}{2R^2}(\left\langle Z^2 \right\rangle - \langle Z \rangle^2)^2 - \frac{l_B^2}{R^2}(\left\langle Z^2 \right\rangle - \langle Z \rangle^2) \langle Z \rangle^2 \tag{14}$$

and if $pH = pI$, then $< Z >= 0$ and the induced charge-induced charge interaction becomes the leading term,

$$A(R) \approx -\frac{l_B^2 \langle Z^2 \rangle^2}{2R^2} \tag{15}$$

The above equations show that the fluctuating charge of a protein or macromolecule may under certain circumstances contribute significantly to the net interaction. We can define a "charge polarizability" or a capacitance, C, as

$$C = \left\langle Z^2 \right\rangle - \langle Z \rangle^2 \tag{16}$$

With this definition of the capacitance, Eq.(13) can be rewritten in a more compact form,

$$A(R)/kT \approx \frac{l_B \langle Z_A \rangle \langle Z_B \rangle}{R} - \frac{l_B^2}{2R^2}(C_A C_B + C_A \langle Z_B \rangle^2 + C_B \langle Z_A \rangle^2) \tag{17}$$

We can use general electrostatic equations and relate the capacitance to the charge induced by a potential $\Delta\Phi$,

$$Z_{ind} = \frac{C\Delta\Phi}{kT} \tag{18}$$

The capacitance, C, can also be derived from the experimental titration curve. For a single titrating acid the ionization degree, α, can be found in any elementary physical chemistry textbook,

$$\log K = -pH + \log \frac{\alpha}{1-\alpha} \tag{19}$$

Taking the derivative of α wrt to pH gives,

$$\frac{d\alpha}{dpH} = \alpha(1-\alpha) = C \ln 10 \tag{20}$$

where in the second step we have identified the capacitance defined in Eq.(16). We can obtain an approximate value for the capacitance in a protein assuming that there is no interaction between the titrating sites. A protein contains several titrating groups like aspartic and glutamic acid, histidine etc., each with an ideal pK value. Denoting different titrating groups with γ and their number with n_γ, then the total capacitance can be approximated with,

$$C^{ideal} = \frac{1}{\ln 10} \sum_\gamma n_\gamma \frac{10^{pH-pK_\gamma}}{(1+10^{pH-pK_\gamma})^2} \tag{21}$$

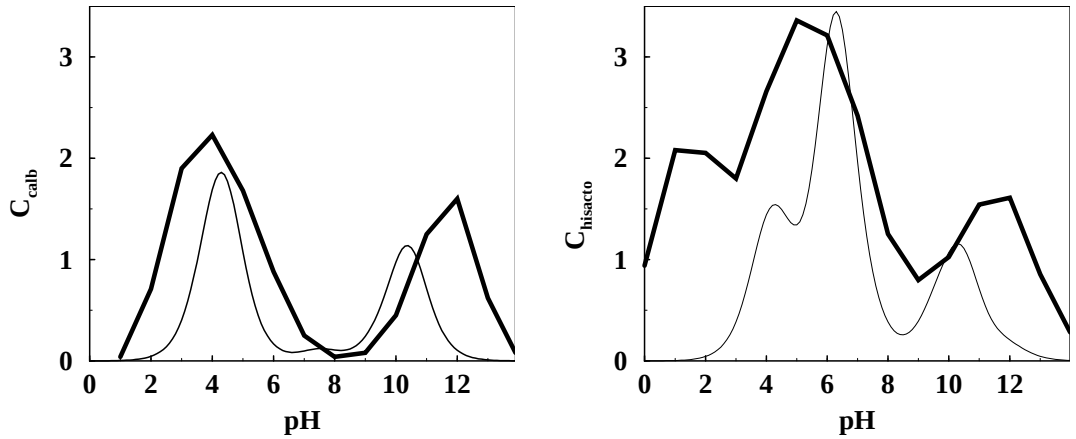

Figure 8: Left: The capacitance for calbindin D_{9k} as a function of pH. The thick solid curve is from a MC simulation of the atomistic model , while the thin solid line is the ideal capacitance calculated from Eq.(21). pI for calbindin is approximately 4.2. Right: The capacitance for hisactophilin as a function of pH with symbols as before. pI for hisactophilin is 7.3

We have calculated the capacitance for a number of proteins with different characteristics in terms of number and type of residues. A MC simulation has to be performed at each pH at given salt and protein concentrations. Unless otherwise stated we have used a salt concentration of 70 mM and a protein concentration of 0.7 mM. Figure 8a shows the capacitance for calbindin. The main difference from the ideal capacitance curve is a strong broadening of two peaks corresponding to the response from acidic and basic residues, respectively. If the protein has a significant net charge, the true curve will also shift away from the ideal one, as is seen for calbindin at high pH.

The protein hisactophilin is of the same size as calbindin, but it has a slightly different capacitance curve, see Figure 8b. The protein contains 31 histidine residues, which is reflected in a large maximum for $C_{hisacto}$ at $pH \approx 5$. The downward shift of the maximum is due to the high positive charge of hisactophilin at low pH. The net charge is +28 at $pH = 3$ and +23 at $pH = 4$. The isoelectric point found from the simulations is $pI = 7.3$, which is in good agreement with experimental estimates.

The electrostatic interaction between two proteins will be dominated by the direct Coulomb interaction provided that the net charge, Z, is sufficiently different from zero. The induced interactions will only play an important role at pH values close to the isoelectric point of one of the proteins - this can be seen from Eq.(17). Figure 9a shows the free energy of interaction between the two proteins calbindin and lysozyme at $pH = 4$, which is close to the isoelectric point for calbindin. At contact there is a significant difference in interaction energy between a model with fixed charges compared to a situation where the proteins are free to adjust their charges.

The difference in free energy between the two models is mainly due to the interaction between the induced charge in calbindin and the permanent charge in lysozyme. This is a typical result and significant effects from charge regulation can be expected when one of the interacting proteins has a large net charge and the other a large capacitance. Following Eq.(17) we can approximate the difference as,

$$(A_{reg}(R) - A_{fix}(R))/kT = \Delta A(R)/kT = -\frac{l_B^2}{2R^2}(C_{calb}C_{lys} + C_{lys}Z_{calb}^2 + C_{calb}Z_{lys}^2) \qquad (22)$$

and Figure 10 shows an almost perfect agreement between the simulated free energy difference and the calculated one according to Eq.(22).

An interesting result is that despite that both calbindin and lysozyme are positively charged at $pH = 4$, there is still an attractive electrostatic interaction between the two. Such an

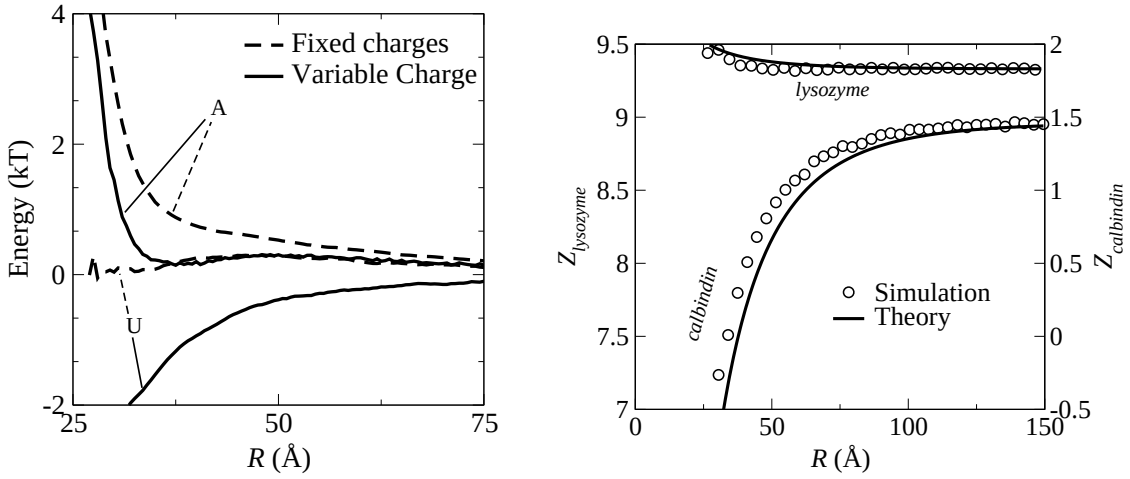

Figure 9: a) The energy amd free energy of interaction between calbindin and lysozyme at $pH = 4$ for a protein model with fixed charges (dashed lines) and one with charge regulation (solid lines). The amino acid model is used and the salt concentration is b) The variation of net charge of calbindin (solid line) and lysozyme (dashed line) as a function of their separation. The simulations are based on the amino acid model. $pH = 4$ and salt concentration is 5 mM.

attraction could of course be due to charge-dipole and/or dipole-dipole interactions, but they do not seem to be important in the present case: the main contribution to the interaction free energy comes from the induced charges. This is further demonstrated in Figure 9b, where one can follow how the net charge of calbindin goes from ≈ 1.4 at infinite separation to ≈ -0.5 at contact between calbindin and lysozyme. We will come back to this issue when discussing protein polyelectrolyte complexation.

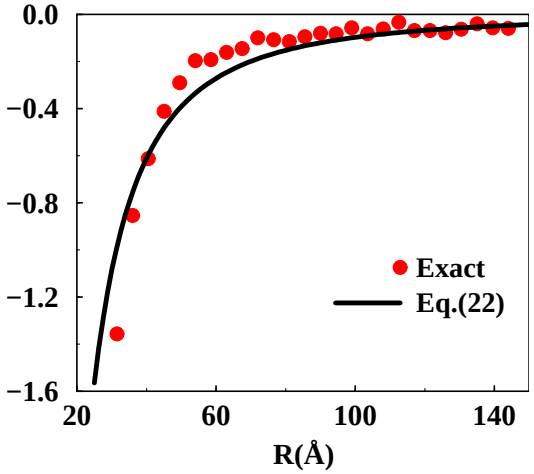

Figure 10: The difference in free energy of interaction between calbindin and lysozyme at $pH = 4$ for a protein model w regulation and one with fixed charges. R is the separation between the mass centra of the two proteins. Symbols denote the simulated difference (see Figure 9) and the solid line is obtained from Eq.(22) with $Z_{calb} = 1.16$, $C_{calb} = 2.23$, $Z_{lys} = 10.2$ and $C_{lys} = 0.88$.

Bridging attraction with polyelectrolytes

Adsorption of a polyelectrolyte to an aggregate is a necessary, but not sufficient condition, in order to attain a modulation of the free energy. It actually has to adsorb to both aggregates in order to form *bridges*, see Figure 11, that can lead to attractive interactions. For highly charged polyelectrolytes and oppositely charged macromolecules, bridge formation is usually a very effective way of destabilization. From simulations and mean field theories, we know that

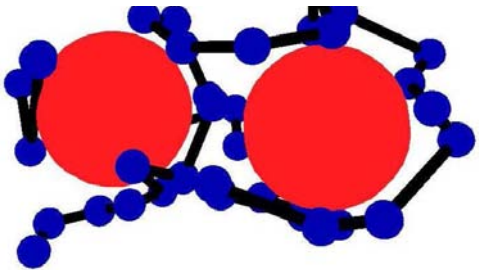

Figure 11: Snapshot from a MC simulation of system containing two charged macromolecules and an oppositely charged polyelectrolyte.

the attraction is rather short ranged and that it typically only extends over distances of the order of the monomer-monomer separation [22, 23, 24, 25]. Figure 12a shows what happens if a polyelectrolyte salt is added to a solution of two charged macromolecules. The double layer repulsion is replaced with a short range attraction with a minimum at a surface-to-surface separation of approximately a monomer-monomer distance.

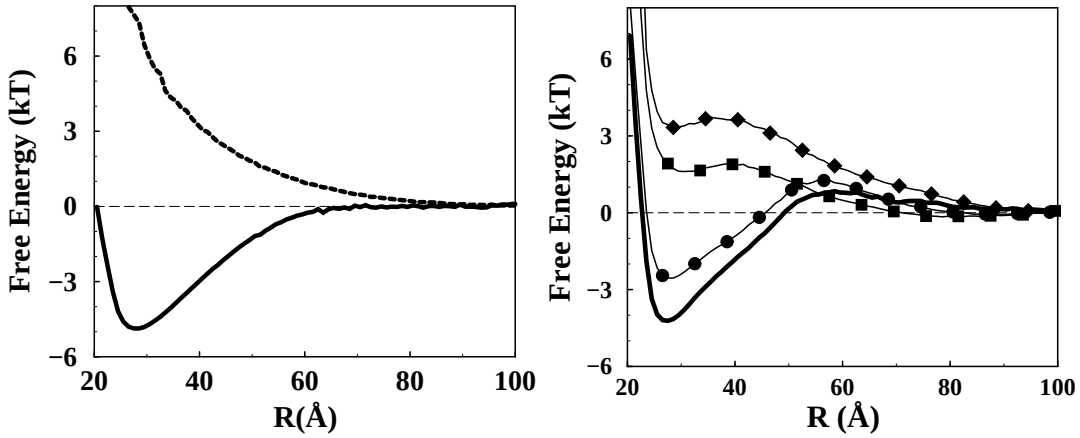

Figure 12: a) The free energy of interaction between two charged spheres as a function of separation, in the presence of a polyelectrolyte salt (solid line) and in the presence of a 1:1 salt (dashed line). The charge of the aggregates is 10e and the radius is 10 Å. The freely jointed polyelectrolyte chain contains 10 charged monomers separated a distance of 6 Å. b) The free energy of interaction between two negatively charged spheres in the presence of a single neutral polyampholyte chain with 40 monomers. The charge topology has been varied and the following notation is used: di-block (solid line with no symbols), tri-block (+10,-20,+10) (circles), tetra-block (+10,-10,+10,-10) (squares) and "reversed" tri-block (-10,+20,-10) (diamonds). Each macromolecule has a charge of +20e and the radius is 10 Å.

A polyelectrolyte adsorbs readily to an oppositely charged macromolecule and in the presence of several charged spheres it becomes of course entropically favourable for the chain to adsorb to more than one sphere. This can only be accomplished at short separations, since the chain tries to avoid placing charges far from the charged aggregates, where the potential is high. Thus, a weakly charged chain, *i.e.* a chain with large separation between the charged monomers, will lead to a more long ranged but weaker attraction. In general, one finds that highly charged systems give rise to fewer, but stronger "bridges", and there will be an optimal choice of polyelectrolyte structure for the attraction between the colloids.

The interaction between charged macromolecules is, from an electrostatic point of view, rather insensitive to the addition of neutral *random* polyampholytes. It is only with block-polyampholytes that the normal double layer repulsion can be decreased in the same way as with oppositely charged polyelectrolytes. The oppositely charged block acts in the same way as an oppositely charged polyelectrolyte. The only complication or constraint is that the equally charged blocks should avoid the aggregates. If the polyampholyte has a net charge, then it behaves qualitatively as a weakly charged polyelectrolyte. A mixing of positively and negatively

charged monomers allows a tailoring of the range and magnitude of the attraction. Figure 12b shows the free energy of interaction between two charged macromolecules with different types of polyampholytes. A naive picture of a tri-block between two adsorbing macromolecules, which seems to be true for neutral block-copolymers, is one where the two ends of the PA chain adsorb to one aggregate each and "pull" them together. Such a structure is quite common in a simulation, but it does not lead to a significant "pulling" force due to the weak force constant of a long segment of negatively charged monomers. Another way to express this is that the free energy gain of adsorbing a PA chain is approximately distance independent for a tri-block of the type (-10,+20,-10). Figure 13 is a snapshot from the simulation and demonstrates this conformation.

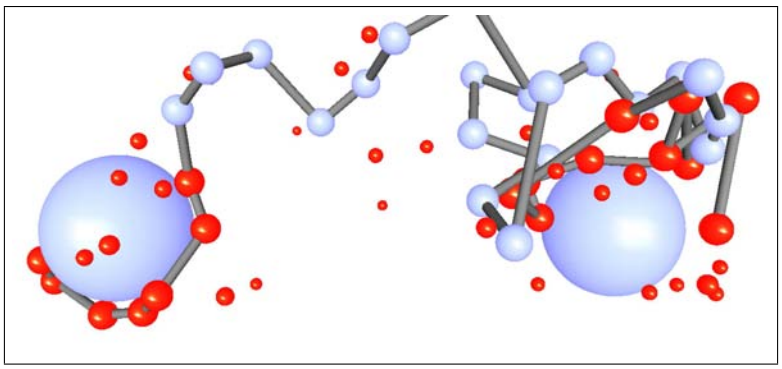

Figure 13: A tri-block, (+10,-20,+10), adsorbing to two negatively charged, $Z = -20$, macroions. Counterions and positively charged monomers are shown in grey and negatively charged monomers in black.

Protein polyelectrolyte complexation

The complexation of polyelectrolytes and proteins is extensively used in pharmaceutics, foods and cosmetics. [26, 27, 28, 29, 30, 31, 32, 33] The subject has been addressed by a number of authors exploring it from experimental measurements [32, 33, 34, 35] to theoretical modeling [36, 37, 38]. The strength of interaction is to a large extent regulated by electrostatic interactions, governed by key parameters such as pH and salt concentration.

A particularly interesting observation [33, 36, 39] is the apparently paradoxical formation of soluble complexes at conditions where the net charges of the protein and the polyelectrolyte have the same sign. Experimental studies of Dubin, Kruif and co-workers [33, 36, 39] have demonstrated this special feature of the polymer/protein complexation. The term complexation "on the wrong side" has been used, meaning that a polyanion forms a complex with a protein at a pH above the isoelectric point of the protein. The molecular interpretation of such studies has focused on the assumption of "charged patches" on the protein surface [33, 40, 34, 37].

A formal way to describe the interaction between oppositely charged patches on two macromolecules is in terms of a multipole expansion. That is, for two neutral protein molecules the leading terms would then be dipole-dipole, dipole-quadrupole, etc. Other electrostatic properties of the protein, however, may be more important and Kirkwood and Shumaker [20] demonstrated theoretically already in 1952 that fluctuations of residue charges in two proteins can result in an attractive force. Recently, we have taken up this idea and used MC simulations and a charge regulation theory in order to explain protein-protein and protein-polyelectrolyte association in a purely electrostatic model [21, 41]. A charge regulation mechanism has also been suggested by Biesheuvel and Cohen-Stuart [42].

We can use simulated capacitances and dipole moments in order to analytically calculate the ion-induced charge and ion-dipole contributions to the interaction free energy according

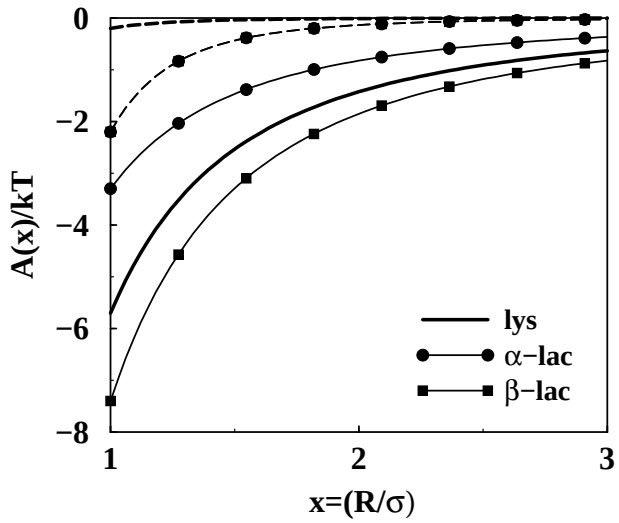

Figure 14: The contribution to the free energy of interaction from the charge - induced charge term (solid lines) and the ion-dipole term (dashed lines). Lines without symbols describe lysozyme, filled circles refer to α-lactalbumin and filled squares refer to β-lactoglobulin, respectively. The free energies are calculated from eq.(17) using simulated capacitances and dipole moments. Note that the ion-dipole terms for α-lactalbumin and β-lactoglobulin coincide.

to eq.(17). The results indicate that the regulation term is by far the most important term for lysozyme, while for α-lactalbumin and β-lactoglobulin the two terms are of comparable magnitude. The curves in Figure 14 should of course be regarded as qualitative and not quantitative. However, they still give, as will be seen below, a correct picture of the behaviour of the three proteins. The contact separation has been defined as the protein radius plus the polyelectrolyte radius, $R_p + R_{pe}$. The latter has been chosen as half the end-to-end separation of the corresponding neutral ideal polymer. Both the protein and polyelectrolyte radii are approximate, but even with a rather generous variation of these values the general picture of Figure 14 will remain the same. The regulation term decays slower than the ion-dipole term, which means that it will gain in relative importance at larger separation, see Figure 14. This means that even if the two terms are comparable at contact, the regulation term can still dominate the contribution to, for example, the second virial coefficient.

We have performed four different simulations for each protein-polyelectrolyte complex:

- A: the "neutral" protein, that is all charges have been set to zero.
- B: the protein with fixed charges at each amino acid residue.
- C: the protein with an ideal dipole at its center of mass.
- D: the protein with titrating amino acid residues.

The first set of simulations (A) describes only the shape of the protein and the free energy of interaction is of course everywhere repulsive. The second set of simulations (B) uses fixed fractional charges on all residues, which has been determined in a separate simulation of the isolated protein at the appropriate pH. In the next set (C), the charge distribution of the protein is replaced by an ideal dipole. In the fourth and final set (D) the amino acids are allowed to titrate and this simulation contains all electrostatic contributions including the ion-induced charge term. The difference between set B and C describes the importance of higher order electrostatic moments, quadrupole, octupole etc. in the protein, while a comparison of sets C and D reveals the effect of the regulation mechanism.

The calculated free energy of interaction, $A(R)$, for the three proteins at their respective pI all show a clear minimum, see Figure 15. The relative strength of the minima are in qualitative

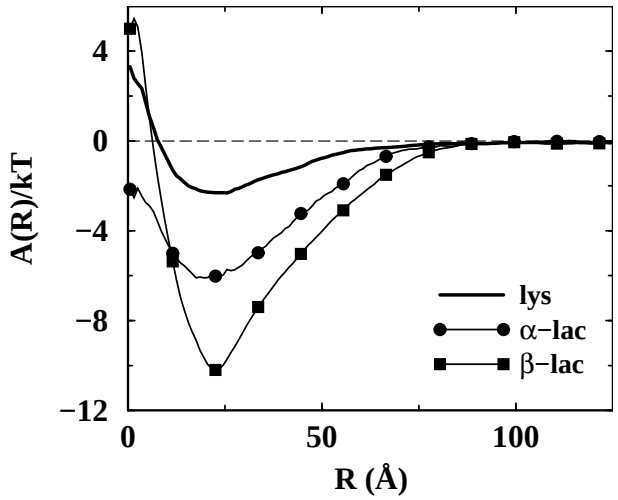

Figure 15: The free energy of interaction between the centers of mass of the protein and the polyelectrolyte at low salt concentration obtained from MC simulations with Model D. The curves have been calculated at the respective isoelectric points for lysozyme (no symbols), α-lactalbumin (filled circles) and β-lactoglobulin (filled squares).

agreement with perturbation calculations, *cf.* Figure 14, while the actual numbers are approximately half the values predicted by second order perturbation theory. The minima appear at roughly the same separation despite the fact that β-lactoglobulin is more than twice as big as the two others. This can be explained by the elongated form of the former, which also results in a more long ranged attraction. The separation R can approach zero, which corresponds to a situation where the polyelectrolyte wraps around the protein. Note, however, that $A(0)$ is repulsive indicating that the "wrapping" of the chain around the proteins is an entropically unfavourable structure.

The attractive minimum in the protein-polyelectrolyte complex is reduced upon addition of salt [37] and we can use the minima of $A(R)$ in Fig.15 in order to estimate the critical ionic strength. Assuming that the salt screening can be described by simple Debye-Hückel theory and that the complex can be defined as dissolved when the interaction is less than kT, we get the following relation,

$$\exp(-2\kappa R_{min})|A(R_{min})| \leq kT \tag{23}$$

The factor of two in the exponent comes from the fact that the second order terms dominate the interaction. Following this recipe we find that approximately 10 and 20 mM salt is sufficient to dissociate the α-lactalbumin and β-lactoglobulin polymer complexes, respectively.

Thus, we have shown that a polyanion can form a complex with a neutral protein molecule. Next, we will make a numerically more rigorous partitioning of contributions to the free energy of interaction shown in Fig. 15. The minimum for lysozyme is solely due to charge regulation, Fig. 16a. If the charge distribution on lysozyme is considered fixed, then the polyanion-lysozyme interaction is essentially everywhere repulsive. Replacing the detailed charge distribution with an ideal dipole at the mass center has a small effect on the free energy. This means that the ion-dipole interaction gives a very small attractive contribution, while the effect from higher order moments is negligible.

As shown in Fig. 16b, the polyanion interacts more strongly with α-lactalbumin than with lysozyme. For α-lactalbumin the regulation term increases the depth of the minimum from approximately 4 to 6 kT. An interesting effect is that the dipolar protein shows a stronger interaction than the protein with a detailed but fixed charge distribution. This means that the ion-quadrupole interactions etc. add repulsive contributions to the interaction.

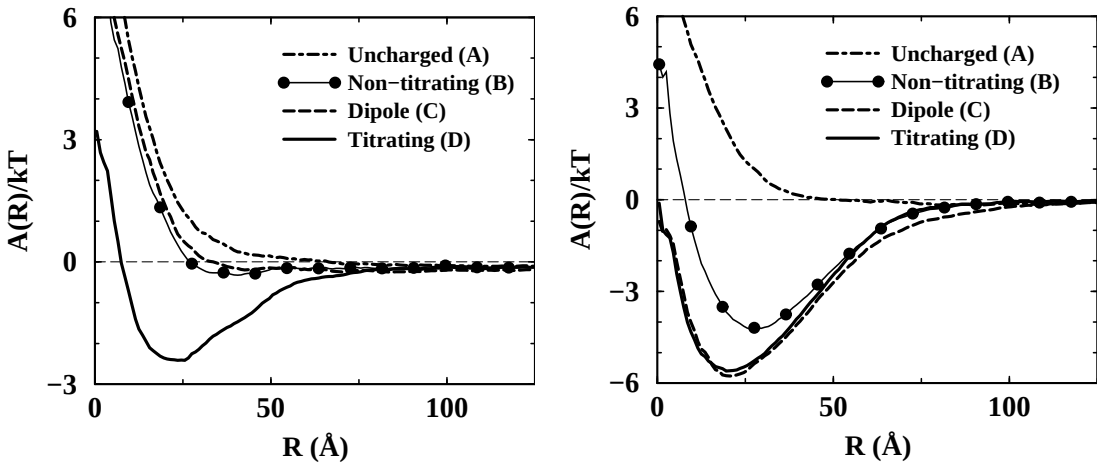

Figure 16: The free energy of interaction between the centers of mass of lysozyme and the polyanion. The free energies have been calculated at pI and the four curves correspond to the different cases mentioned in the text. a) Lysozyme and the polyanion and b) α-lactalbumin and the polyanion.

Conclusion

We have demonstrated a few generic situations where electrostatic interactions between charged macromolecules seem to play an important role. With Monte Carlo simulations we can obtain the exact answer within the given interaction model, which allows us to test the validity of approximate theories. Many biochemical systems are comparatively weakly charged, in contrast to many inorganic systems, and simple theories based on the Debye-Hückel approximation give accurate answers. The long range character of the Coulomb interaction usually means that the geometry and detailed distribution of the charged groups are less important for the interaction of two charged macromolecules.

References

[1] T. L. Hill, *An Introduction to Statistical Thermododynamics* (Dover Publications Inc., New York, 1986).

[2] S. Engström and H. Wennerström, J. Phys. Chem. **82**, 2711 (1978).

[3] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).

[4] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 1996).

[5] R. F. Platford and T. Dafoe, J. Mar. Res. **23**, 63 (1965).

[6] R. F. Platford, J. Mar. Res. **23**, 55 (1965).

[7] M. Lund, B. Jönsson, and T. Pedersen, Mar. Chem. **80**, 95 (2003).

[8] L. A. Svensson, E. Thulin, and S. Forsén, J. Mol. Biol. **223**, 601 (1992).

[9] C. Tanford and J. G. Kirkwood, J. Am. Chem. Soc. **79**, 5333 (1957).

[10] B. Svensson, B. Jönsson, C. E. Woodward, and S. Linse, Biochemistry **30**, 5209 (1991).

[11] D. M. E. Szebenyi and K. Moffat, J. Biol. Chem. **261**, 8761 (1986).

[12] T. Kesvatera, B. Jönsson, E. Thulin, and S. Linse, Proteins **45**, 129 (2001).

[13] V. Spassov and D. Bashford, Protein Sci. **7**, 2012 (1998).

[14] A. H. Juffer and H. J. Vogel, Proteins **41**, 554 (2000).

[15] B. Beresford-Smith. *Some aspects of strongly interacting colloidal interactions*. PhD thesis, Australian National University, Canberra, (1985).

[16] R. Kjellander and S. Marčelja, Chem. Phys. Letters **112**, 49 (1984).

[17] R. Kjellander and S. Marčelja, J. Chem. Phys. **82**, 2122 (1985).

[18] M. Lund and B. Jönsson, Biophys. J. **85**, 2940 (2003).

[19] S. M. Mel'nikov, M. O. Khan, B. Lindman, and B. Jönsson, J. Am. Chem. Soc. **121**, 1130 (1999).

[20] J. G. Kirkwood and J. B. Shumaker, Chemistry **38**, 863 (1952).

[21] M. Lund and B. Jönsson, Biochemistry **44**, 5722 (2005).

[22] C. E. Woodward, B. Jönsson, and T. Åkesson, J. Chem. Phys. **89**, 5145 (1988).

[23] R. Podgornik, J. Phys. Chem. **95**, 5249 (1991).

[24] J. Ennis, L. Sjöström, T. Åkesson, and B. Jönsson, Langmuir **16**, 7116 (2000).

[25] L. Sjöström and T. Åkesson, J. Coll. Interface Sci. **181**, 645 (1996).

[26] C. Schmitt, C. Sanchez, S. Desobry-Banon, and J. Hardy, Crit. Rev. Food Sci. Nutr. **38**, 689 (1998).

[27] J. L. Doublier, C. Garnier, D. Renard, and C. Sanchez, Curr. Opin. Colloid. Interface Sci. **5**, 202 (2000).

[28] S. Zancong and S. Mitragotri, Pharm. Res. **19**, 391 (2002).

[29] G. Jiang, B. H. Woo, F. Kangb, Jagdish Singhb, and Patrick P. DeLuca, J. Controlled Release **79**, 137 (2002).

[30] M. Simon, M. Wittmar, U. Bakowsky, and T. Kissel, Bioconjugate Chem. **15**, 841 (2004).

[31] J. A. Hubbell, Science **300**, 595 (2003).

[32] M. Girard, S. L. Turgeon, and S. F. Gauthier, J. Agric. Food Chem. **51**, 6043 (2003).

[33] C. G. de Kruif, F. Weinbreck, and R. de Vries, Curr. Opin. Colloid Interface Sci. **9**, 340 (2004).

[34] E. Seyrek, P. L. Dubin, C. Tribet, and E. A. Gamble, Biomacromolecules **4**, 273 (2003).

[35] R. Hallberg and P. L. Dubin, J. Phys. Chem. B **102**, 8629 (1998).

[36] K. R. Grymonpré, B. A. Staggemeier, P. L. Dubin, and K. W. Mattison, Biomacromolecules **2**, 422 (2001).

[37] R. de Vries, J. Chem. Phys. **120**, 3475 (2004).

[38] F. Carlsson, P. Linse, and M. Malmsten, J. Phys. Chem. B **105**, 9040 (2001).

[39] R. de Vries, F. Weinbreck, and C. G. deKruif, J. Chem. Phys. **118**, 4649 (2003).

[40] T. Hattori, R. Hallberg, and P. L. Dubin, Langmuir **16**, 9738 (2000).

[41] F. L. B. da Silva, M. Lund, B. Jönsson, and T. Åkesson, J. Phys. Chem. B **110**, 4459 (2006).

[42] M.A. Cohen Stuart P.M. Biesheuvel, Langmuir **20**, 2785 (2004).