Counterion Condensation in Short Cationic Peptides: Limiting Mobilities Beyond the Onsager-Fuoss Theory

Erik Wernersson,[†] Jan Heyda,[†] Anna Kubíčková,[‡] Tomáš Křížek,[‡] Pavel Coufal,[‡] and Pavel Jungwirth[†]

†: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and Center for Biomolecules and Complex Molecular Systems, Flemingovo nám. 2, 16610 Prague 6, Czech Republic

‡: Charles University in Prague, Faculty of Science, Department of Analytical Chemistry, Albertov 2030, 12840 Prague 2, Czech Republic

Corresponding Author: Erik Wernersson

E-mail: erik.wernersson@uochb.cas.cz

Alt. E-mail: erik.g.wernersson@gmail.com

Fax (office): +420 220 410 320

List of Abbreviations:

OF (Onsager-Fuoss), DH (Debye-Hückel), PDS (propane-1,3-disulfonate), MD (molecular dynamics), MC (Monte Carlo)

Keywords:

Background electrolyte effects, Counterion condensation, Ion-pairing, Limiting mobility, Molecular modeling.

Word Count:

6940

Abstract

We investigated the effect of the background electrolyte anions on the electrophoretic mobilities of the cationic amino acids arginine and lysine and the polycationic peptides tetraarginine, tetralysine, nonaarginine, and nonalysine. Background electrolytes composed of sodium chloride, sodium propane-1,3-disulfonate, and sodium sulfate were used. For the amino acids, determination of the limiting mobility by extrapolation, using the Onsager-Fuoss theory expression, yielded consistent estimates. For the peptides, however, the estimates of the limiting mobilities were found to spuriously depend on the background electrolyte salt. This paradox was resolved using molecular modeling. Simulations, on all-atom as well as coarse grained levels, show that significant counterion condensation, an effect not accounted for in Onsager-Fuoss theory, occurs for the tetra- and nonapeptides, even for low background electrolyte concentrations. Including this effect in the quantitative estimation of the background electrolyte effect on mobility removed the discrepancy between the estimated limiting mobilities in different salts. The counterion condensation was found to be mainly due to electrostatic interactions, with specific ion effects playing a secondary role. Therefore, the conclusions are likely to be generalizable to other analytes with a similar density of charged groups and Onsager-Fuoss theory is expected to fail in a predicable way for such analytes.

1 Introduction

Ion pairing agents can be used to enhance selectivity in capillary electrophoresis (CE).[1] A class of compounds that has been used successfully as such agents is represented by organic sulfonates. Hexanesulfonic acid, under conditions where it is partly deprotonated, has been shown to enhance separation of polylysines of different lengths.[2] Conversely, the separation of sulfonate *analytes* can be enhanced by addition of aliphatic compounds containing two or more ammonium groups.[3-5] The selectivity was influenced by the match between the number and spacing of ammonium groups and analyte sulfonate groups. The binding of heparin, a sulfated glycosaminoglycan, to proteins has also been studied by CE together with molecular modeling.[6] The results suggest that the selectivity of the binding is partly determined by the match in the distributions of charged groups between the protein and the heparin ligands. Ion-pairing between trilysine and sulfated glycopeptides have been ingeniously exploited in CE-MS to improve the peptide fragmentation properties and to allow the MS to be run in cationic mode.[7]

In modeling of electrophoretic migration, the effect of the ions of the background electrolyte (BGE) in capillary electrophoresis experiments is typically taken into account using some version of the Onsager-Fuoss (OF) theory.[8] This theory is, in turn, based on the Debye-Hückel (DH) approximation for the structure of the "ionic atmosphere" of each ion.[9] Elaborations of this basic framework has recently been used to estimate analyte properties, such as protonation state, size and shape.[10-12] However, in situations where the electrostatic interaction energy is large compared to the thermal energy (k_BT), which is a common situation in aqueous solutions containing ionic species of valency greater than one, DH theory is invalid even for modest concentrations. In this regime, quantitative prediction of electrophoretic mobility is difficult to accomplish. Non-electrostatic interactions, absent in DH theory, may also play a large role in determining the degree of ion-pairing. This is illustrated by the fact that that pentafluoroprpropanoate has been used to separate peptides with substitutions that did not affect the total charge.[13] A semi-quantitative treatment in terms of ion-pairing between the analyte and the BGE ions does, however, capture the general features of the experimental trends. Ion-pairing due to

electrostatic interactions can, as an approximation, be described as a chemical equilibrium between free and paired ions in the spirit of Bjerrum theory.[14-18] As shown in our previous publication,[19] this concept can to a large degree account for the qualitative features of the ionic strength dependence of the mobility of cationic peptides in sodium sulfate BGE. Ion-pairing due to ion-specific interactions was also detected, but was found to be of secondary importance. Here, we use careful molecular modeling to make the analysis of electrophoretic data in terms of these concepts quantitative.

In this work we investigate the ion-pairing between cationic amino acids and peptides and the propane-1,3-disulfonate (PDS) ion ($[SO_3(CH_2)_3SO_3]^{2-}$), using CE and molecular simulations. The results are compared to those for sodium sulfate and sodium chloride and analyzed in a consistent way. The charged groups in PDS are chemically similar to sulfate, but carry only one negative charge each. In the OF theory, only the total charge and the limiting mobility of the ions matter, and PDS and sulfate should have a similar influence on electrophoretic mobility. When ion-pairing is taken into account, the situation becomes more complex. In particular, it is not clear from the outset whether the ion pairing properties of PDS resemble those of a monovalent or a divalent ion more closely.

We use two complementary modeling approaches to interpret our experimental data: Molecular dynamics (MD) simulations of an atomistically detailed model and Monte Carlo (MC) simulations of a coarse grained, 'primitive', model that contains only electrostatic and excluded volume interactions. The former approach in principle contains, within the accuracy of the molecular model, all intermolecular interactions and can, therefore, give information about the specificity of ionic interactions. However, all-atom simulations have to be performed at elevated salt concentrations because the number of water molecules required per ion would be unfeasibly large for concentrations in the experimentally accessible range. The coarse grained model is not subject to this limitation, and thus permits a more direct comparison to experiments. Any specificity in the interaction between charged groups is absent in the coarse grained model by design. While this can be seen a deficiency for some purposes, this feature can be exploited in conjuncture with all-atom simulations to estimate the relative importance of electrostatic and specific interactions.

2 Materials and Methods

2.1 Chemicals

Sodium chloride (p.a.), sodium sulfate (p.a.) and thiourea (p.a.) were purchased from Lachema (Brno, Czech Republic). L-Arginine (99.5%) was provided by Fluka (Buchs, Switzerland), and D,L-lysine (98%) was obtained from Sigma (St. Louis, MO). Tetraarginine (trifluoroacetate salt) and tetralysine (acetate salt) were purchased from Bachem (Bubendorf, Switzerland). Nonaarginine was acquired from American Peptide Company (Sunnyvale, USA) and nonalysine was purchased from Biomatik (Cambridge, Canada). Methanol (HPLC-grade) was purchased from Merck (Darmstadt, Germany) and 1,3-propanedisulfonic acid disodium salt, sodium PDS, was obtained from Sigma (St. Louis, MO). Background electrolytes and samples were prepared using deionized water produced by a Milli-Q system (Millipore, Billerica, MA).

2.2 Instrumentation

A 7100 CE capillary electrophoresis system (Agilent Technologies, Waldbronn, Germany) was used for all experiments. A 75-µm-i.d. (325-µm-o.d.) fused-silica capillary coated with hydrophilic polymer (capillary electrophoresis polymer, CEP) was purchased from Agilent Technologies (Waldbronn, Germany) and cut to 80.0 cm total length (71.5 cm to the detection window). Prior to the first use, the capillary was flushed for 20 min with methanol and for 10 min with deionized water using a pressure of 100 kPa. The diode array UV detector was operated at a wavelength of 200 nm. The temperature was kept at 25 C by air cooling.

2.3 Mobility measurements

A procedure derived from the electro-osmotic flow (EOF) mobility measurement method described in ref [20] was employed to determine electrophoretic mobilities of the amino acids, tetrapeptides, and nonapeptides. As the aim of this study was to

investigate the influence of the BGE anion, it was essential to keep the BGE composition as simple as possible. Thus aqueous solutions of sodium chloride, sodium sulfate and sodium PDS were used as the BGE without any buffering component. Although a considerable pH gradient was created in the initially neutral BGE in the capillary during the measurement, the effect of this variations in pH was found negligible in our previous work. [19] We showed that the pH changes do not affect the protonation state of the peptides by comparison of mobilities obtained in buffered and non-buffered BGEs. Under such conditions, it was, however, found necessary to use a capillary with deactivated inner walls. Before each run, the capillary was flushed for 5 min with methanol, 5 min with deionized water, and 2 min with BGE using a pressure of 100 kPa. A zone of 2 mmol/L (mM) aqueous solution of the studied analyte with 0.1 g/L thiourea was injected into the capillary using a pressure of 5 kPa for 3 s. The zone was then pushed into the capillary by the application of 5 kPa of pressure for 2 min to avoid loss of the thiourea zone in the case of reverse EOF. Subsequently, a voltage of +10 kV was applied for 10 min. During this period, the thiourea zone was mobilized only by EOF, whereas the motion of the analyte zone resulted from the combination of EOF and electrophoretic mobility. Then, a zone of aqueous solution with a concentration of 0.1 g/L thiourea was injected using a pressure of 5 kPa for 3 s. Finally, all of the zones in the capillary were mobilized by the application of 5 kPa of pressure, and the UV detector recorded the zones passing through the detection window. At the beginning of this pressure mobilization, the registration of the UV signal was started. The electrophoretic mobility, μ, of each analyte was calculated according to

$$\mu = \frac{(t_2 - t_1) l_d}{t_3 t_U} \frac{l_c}{U} \quad (1)$$

where t_1 is the time needed for the analyte zone to be pushed through the detection window; t_2 and t_3 are the corresponding times for the first and second thiourea zones, respectively. l_c is the total capillary length and l_d is the length to the detection window. U is the voltage applied and t_U is the duration of voltage application. Examples of electropherograms are shown in the SI.

The mobilities of analytes were measured for ionic strengths of 10, 20, 50, 70, 100, 120, and 150 mM in sodium PDS, sodium chloride, and sodium sulfate BGE.

Finally, the mobilities of tetrapeptides in mixed aqueous salt solutions of sodium PDS and sodium sulfate were measured. The ionic strength of the BGE was maintained at 50 mM while the sodium sulfate fraction was increased from 0 to 1. All measurements were carried out in triplicate. Because the capillary was thermostatted and the applied voltage was relatively low, no effect of Joule heating on mobilities was considered. To further justify this approach, the dependence of the electric current in the capillary on the voltage applied was investigated. Even for the highest ionic strength (i.e. 150 mM), this dependence was linear within the range of 0-10 kV for all three BGEs.

2.4 All-Atom Molecular Dynamics Simulations

The solutes considered are cationic argininium and lysinium amino acids and the polycationic oligopeptides tetralysinium, nonalysinium, tetraarginium, and nonaarginium. Below, these are referred to as arginine, lysine, tetralysine, nonalysine, tetraarginine and nonaarginine, respectively, with the understanding that the side chains are fully protonated. In contrast to our earlier study,[19] the C- and N-terminals were not protected. The C-terminal was modeled as deprotonated and the N-terminal was modeled as protonated. This is intended to represent the dominant protonation state in pH neutral solution. Simulations were carried out for each of the solutes in sodium PDS and sodium chloride solution BGE. For the free amino acids, the simulated system contained one amino acid molecule and 575 water molecules with 3 PDS (or 6 chloride) and 5 sodium ions. For the tetrapeptides, the BGE composition was 2300 water molecules, 12 PDS ions (or 24 chloride ions), and 20 sodium ions, and for the nonapeptides, it was 5185 water molecules and 27 PDS (or 54 chloride ions), and 45 sodium ions. This corresponds to a molal salt concentrations of 0.24 m for sodium PDS and 0.48 m for sodium chloride in all cases. Note that the concentration of anions is higher than this, 0.28 and 0.56 m, respectively, due to the solute counterions.

For the amino acids and peptides the parm99 force field was used.[21] The partial charges for the free amino acids were taken from gas-phase Hartree-Fock calculations with the 6-31G* basis set according to the Merz-Kollman RESP procedure using the

Gaussian program.[22-24] For the PDS ions, we represented the aliphatic chain using the GAFF force field[25, 26] and used the same bonding and van der Waals parameters for oxygen and sulfur as in the sulfate model from ref [27]. The partial charges were determined in the same way as for the amino acids. In contrast to our earlier study,[19] the force fields used in these simulations were non-polarizable. The reason why polarizable simulations were found to be necessary in the previous work is that polarization effects are essential to describe the bulk properties of anions with a high charge density, such as sulfate.[28] As the two charged groups in PDS are well-separated, that situation does not arise for this ion. For this reason, the extra computational cost associated with polarizable force fields could be avoided. For sodium and chloride ions, the Smith-Dang parameters[29] were used, and for water, we used the SPC/E water model.[30]

The simulation box was approximately cubic. The temperature was kept at 300 K using the Berendsen thermostat and the pressure was kept at 1 atm (101325 Pa) using an analogous barostat.[31] The cut-off for short range interactions was 0.9 nm in all cases. Long-range electrostatic interactions were handled using particle mesh Ewald summation,[32] with a grid spacing of approximately 0.1 nm. At the start of the simulations, the oligopeptides were in a stretched out conformation with the side chains extending in the plane of the backbone in an alternating (all trans) configuration. After 0.5 ns of equilibration the trajectories were propagated for at least 20 ns for the nonapeptides, 50 ns for the tetrapeptides, and 100 ns for the free amino acids. This was sufficient to converge the ion distribution around the solutes.[33] The simulations were carried out using the AMBER 10 program package.[34]

2.5 Coarse Grained Monte Carlo Simulations

A coarse grained, 'primitive', model of the peptide and ions, simple enough to allow calculations for the low BGE concentrations of the experiments, was also considered. This model is intended to be parsimonious, with only electrostatic and excluded volume interactions taken into account. The latter were modeled as additive hard-sphere repulsions, i.e., the interaction energy is infinite if the distance between two

particles is smaller than the sum of their radii and zero otherwise. Only charged groups were modeled explicitly; non-charged groups and their various conformations were taken into accounts implicitly as constraints on the spatial distribution of charged groups, as explained below. The solvent was modeled as a dielectric continuum. The strength of the electrostatic interaction between two unit charges can be quantified in terms of the Bjerrum length,

$$l_B = \frac{e_0^2}{4\pi k_B T \epsilon \epsilon_0}, \tag{2}$$

where e_0 is the unit charge, ε_0 is the permittivity of vacuum, k_B is Boltzmann's constant, ε is the relative permittivity and T is the temperature. Coulomb's law for the interaction between two ions separated by a distance r can be expressed in terms of l_B as

$$\frac{u_{ij}^{El}(r)}{k_B T} = \frac{z_i z_j e_0^2}{4\pi k_B T \epsilon \epsilon_0 r} = \frac{z_i z_j l_B}{r}, \tag{3}$$

where z_i and z_j are the valencies of ions i and j, respectively. As the temperature enters the Hamiltonian via the electrostatic interactions only, all parameters of the model can be rendered dimensionless by expressing distances in units of l_B. The model parameters were chosen assuming l_B = 0.71 nm, appropriate for water at room temperature.[35] Similar primitive models have a long history of application to electrolyte solutions, colloidal dispersions, and electrical double layers,[36] and have recently been applied to biological ion channels.[37] In particular, a model of this type forms the physical basis of the DH theory of electrolyte solutions.[9] Restricting the model to this level of sophistication thus allows the performance of the mathematical approximations inherent in DH theory, and therefore the foundation of OF theory, to be tested. A type of coarse grained model on a similar level of complexity have recently been applied to predict the electrophoretic mobility of various peptide solutes.[38-40] In that model, the hydrodynamic interactions between ions was treated in great detail, allowing mobility predictions without adjustable parameters. The electrostatic interactions and BGE ion distributuion, however, were treated on the DH level.

The sodium, chloride, and sulfate ions, the charged terminal and side chain groups of the peptides, and the PDS sulfonate groups are all modeled as hard spheres with a

radius of 0.2 nm and the appropriate charge. This size falls roughly in the middle of the range of effective sizes typical for simple salts and is roughly consistent with the contact distances seen in the all-atom simulations.[41] The effective radius is a measure of the total repulsion between ions in aqueous environment, and therefore does not in general coincide with either the hydrodynamic ion radius or the radius in crystals. No attempt was made to distinguish between ions of different species, including the charged groups of lysine and arginine. The model is thus free from specific interactions in the usual sense of the word. In the case of species where there are several charged groups, the 'connectedness' of these groups is simulated by constraining them to move on the surface of a sphere, referred to as the confining sphere below. The confining sphere is penetrable to all other species, it has no other role in the simulation than to restrict the relative positions of charged groups belonging to the same molecule.

PDS is modeled as two anionic sites on a confining sphere of radius 0.27 nm, closely corresponding to the typical distance between each of the two sulfonate sulfur atoms and the carbon atom at the 2-position of the aliphatic chain in the all-atom simulations. The number of positive sites for an amino acid *n*-mer is *n*+1 and the number of negative sites is always 1. The radius of the confining sphere was taken as 0.35, 0.72, and 0.93 nm for the free amino acids, tetramers and nonamers, respectively. These numbers were selected on the basis of the typical distance between a central atom in each peptide and the ammonium nitrogen or the guanidinium carbon of side chain groups in the all-atom simulations. The central atom was taken as the α-carbon of the fifth residue for the nona-peptides, the amide nitrogen of the third residue for the tetra-peptides, and the β-carbon of the amino acids. The radii adopted for the confining spheres in the coarse grained models are within 0.03 nm of the distances within which half of the side chain groups can be found, on average, in each of the corresponding all-atom simulations.

The ion distribution functions around the coarse grained model peptides were calculated using the canonical (NVT ensemble) Metropolis MC scheme,[42] for five low ionic strengths of 10, 50, 100, and 150 mM, as well as for the higher concentrations of the corresponding all-atom simulations. The simulated electrolyte was confined in a spherical volume centered on the peptide. For sodium sulfate and

sodium PDS, the simulated system contained approximately 300 anions and 600 cations and for sodium chloride, the system contained 600 ion pairs. Counterions were added to the system to neutralize the peptide charge. The concentration was adjusted by changing the system volume. In addition, calculations were carried out for mixtures of sodium PDS and sodium sulfate with sulfate fractions of 0.2, 0.4, 0.6 and 0.8 at a constant ionic strength of 50 mM. Ions were initially distributed randomly in the simulation box. After a hundred thousand MC cycles of equilibration, data were collected over at least 5 million MC cycles. The simulations were carried out using small FORTRAN program that was written specifically for this purpose and compiled using the Intel FORTRAN compiler, version 11.0. The source code is included as SI. The MC method provides a solution for equilibrium distribution functions that, for a given interaction model, is exact in the sense that it does not contain systematic errors, only a statistical error that can be made arbitrarily small by averaging over a sufficient number of MC cycles. See the SI for a visualization of an example configuration.

3 Results and Discussion

3.1 Experimental Results

The mobilities of mono-, tetra-, and nonaarginine and mono-, tetra-, and nonalysine in sodium PDS, sodium chloride, and sodium sulfate BGE are shown in Figure 1. For the free amino acids and tetrapeptides, results for sodium sulfate and sodium chloride have already been reported in ref [19]. These systems are considered here for comparison with the PDS results. In all cases, the mobilities in sodium PDS are intermediate between those in sodium chloride and sodium sulfate. For free arginine the mobility in sodium PDS is closer to that in sodium chloride, while for lysine the PDS results are half-way between those for sodium chloride and sodium sulfate. For the tetra- and nonapeptides, the mobilities in sodium PDS are closer to those in sodium sulfate and have a similar concentration dependence. The difference between the mobilities in sodium PDS and in sodium sulfate are larger for the oligoarginines than for the oligolysines. This is likely to be related to specific interactions between

sulfate and the arginine guanidinium group, which favors ion pairing.[43, 44]

We attempted to fit the experimental data with the expression for the mobilities from the OF theory[8]

$$\mu_a(I) = \mu_a^0 - \left(B_1\mu_a^0 H_a + B_2\right)|z_a|\frac{\sqrt{I}}{1+Ba\sqrt{I}} \quad (4)$$

where I is the ionic strength and μ^0 is the limiting mobility, with indexes a referring to analyte. a is the 'ion size' parameter from DH theory. B, B_1 and B_2 are independent of the analyte properties, with $B \approx 3.286$ nm^{-1}, $B_1 \approx 0.7817$ M$^{-½}$, and $B_2 \approx 31.38\times10^{-9}$ m^2 V^{-1} s^{-1} M$^{-½}$ under the conditions of the mobility measurements. H_a was calculated according to the procedure for ternary systems in ref [8] (see sections 4.7 to 4.10). As H_a depends on μ^0, as well as the limiting mobilities of the BGE ions, a self-consistent solution had to be found by iteration. Limiting mobilities of BGE ions were taken from ref [35]. As we could not find the mobility of PDS in the literature, the mobility of glutarate (pentanedioate) was used as a proxy.

No sensible fit to the data could be obtained using the "limiting law" version of eq 4, i.e., with a = 0, using only μ^0 as a fitting parameter. Especially for the peptides, such fitting attempts give a too steep decrease in mobility with concentration to accommodate the experimental trends. Therefore, both μ^0 and a were used as fitting parameters. For the amino acids a reasonable fit could be obtained for physically reasonable values of a, in the range of several tenths of an nm (see SI for the fitted parameters). For the peptides, superficially plausible fits could be obtained, but only for a in the order of a few to several nm, which is unreasonably large. The mobilities of the amino acids in different BGE extrapolate to similar values at I=0, which is the behavior expected in light of OF theory. As can be seen in Figure 1, this is not the case for the tetra- and nonapeptides, for which the fitted limiting mobilities are consistently smaller in sodium sulfate and sodium PDS than in sodium chloride. Thus, the fits of eq 4 to the experimental data have no obvious physical meaning. OF theory, therefore, appears unsuitable for the purpose of extrapolating the current experimental data to infinite dilution, leading us to believe that even the lowest concentration considered here is outside the theory's range of validity. We were, however, able to estimate the limiting mobility using the predictions of the coarse grained model, as discussed below.

To further explore the difference in migration behavior between arginine and lysine in the presence of sulfate and PDS, we carried out mobility measurements in mixed solutions where the amounts of sulfate and PDS were varied at constant ionic strength. The results are shown in Figure 2. There is a decrease in mobility of both tetraarginine and tetralysine as PDS is replaced by sulfate, in accord with the results for pure solutions. The difference in mobility between arginine and lysine peptides increases with increasing sulfate fraction, as the mobility of tetraarginine decreases more quickly than that of tetralysine. The sulfate concentration dependence is near-linear for both peptides.

3.2 Simulation Results and Discussion

Due to the failure of OF theory to furnish consistent estimates of the limiting mobilities of the peptides, we turned to more detailed models to find an explanation to the experimental trends. In Figure 3 the electrical screening of the peptides and amino acids is shown in terms of the quantity

$$Q(s) = Z_p + \sum_j z_j c_j \int_0^s r^2 \, dr h_{aj}(r), \tag{5}$$

where Z_p is the total valency of the peptide or amino acid and $h_{ij} = g_{ij} - 1$ is the pair correlation function between species i and j. $c_j h_{aj}(r)$ is thus the deviation in local concentration from the bulk concentration of species j at a distance r from the analyte. $Q(s)$ may be thought of as the total number of unit charges within the radial distance s. $Q(s)$ for the sodium chloride solutions approaches zero monotonically in all cases, including the coarse grained model. In the sodium PDS solutions, on the other hand, the counter-charge overshoots the peptide charge in all cases, leading to negative values of $Q(s)$ for some s. This indicates a super-equivalent adsorption of PDS counterions. The counterion enrichment is more pronounced for arginine peptides than for lysine ones, and is the weakest (barely visible in the figure), for the coarse grained model. The all-atom simulations thus indicate that there is some degree of specific attraction between PDS and the peptides, which is stronger for oligoarginine than oligolysine. As the simulations are carried out at a concentration higher than the experiments, no fully quantitative comparison can be made. The fact that the measured salt dependence of the mobilities is similar between arginine and lysine

peptides suggest that the specific interactions are of secondary importance in the concentration range of the experiment.

In Figure 4, $Q(s)$ from the coarse grained model is shown for some concentrations in the experimental range. For the coarse grained amino acid in model sodium chloride solution, the agreement with DH theory is good. For the tetra- and nonapeptides, there are visible deviations, especially for model sodium sulfate and sodium PDS BGE. Notably, the slopes of the large-s "tails" of $Q(s)$ are similar to those predicted by DH theory, though the magnitude is much smaller. This result is closely reminiscent of the behavior of charged surfaces, which can be assigned an effective charge on the basis of the long-rang asymptotic behavior of the ion distribution in the electric double layer.[45] Qualitatively, the screening in these cases can thus be described as counterion condensation on the model peptides, which is visible as the steep decline in $Q(s)$ around the radius of the confining sphere. This behavior is reminiscent of the counterion condensation seen in polyelectrolytes.[46] This is hardly surprising as the peptides *are* polyelectrolytes, albeit shorter than what is commonly suggested by the use of that term. The counterion condensation is also commensurable with the predictions from a modified version Bjerrum theory presented in ref [19]. In contrast to Bjerrum theory, however, the current simulations give information that is sufficiently detailed and accurate to make quantitative predictions about the electrophoretic mobility.

Under the assumption that both the analyte and the ions can be treated as point particles with respect to the hydrodynamic interactions, the electrophoretic mobility can be calculated from the ionic correlation functions.[47] If relaxation effects and hydrodynamic interactions between ions are neglected, the mobility of an analyte at infinite dilution is given by

$$\mu_a = \mu_a^0 + \frac{2e_0}{3\eta} \sum_j z_j c_j \int_0^\infty r\, dr h_{aj}(r), \tag{6}$$

where μ_a^0 is the mobility at infinite dilution of salt, c_j is the concentration of species j, e_0 is the unit charge, and η is the solvent viscosity, taken from ref [35]. The sum is over all BGE ion species. This expression follows from eqs. 4.14 to 4.16 of ref [47] if the contributions to the Onsager coefficient from hydrodynamic interactions between ions and relaxation effects, which are intractable given the information available, are

omitted. Note the difference between eq 6 and OF theory in this respect: In the latter, the ion distributions are calculated within the drastic approximations of DH theory, but relaxation effects are included in a consistent way. The second term in eq 6 can be expressed in terms or the electrostatic potential at the center of mass of the solute due to the ionic atmosphere and would then take on a form analogous, but not equivalent, to the Hückel expresion for electrophoretic mobility, which is commonly used in colloid science.

The limiting mobilities, as calculated from the experimental mobilities and from the coarse grained simulations via eq 6, for the various systems are shown in Figure 5. The average of the estimates of μ^0 are 26.7×10^{-9}, 57.5×10^{-9}, and 97.7×10^{-9} $m^2\ V^{-1}\ s^{-1}$ for arginine, tetraarginine, and nonaarginine, respectively, and 28.9×10^{-9}, 60.7×10^{-9}, and 102.1×10^{-9} $m^2\ V^{-1}\ s^{-1}$ for lysine, tetralysine, and nonalysine. The resulting estimates are consistent within 20% in all cases, typically better than that. This level of consistency is remarkable, considering the simplicity of the coarse grained model and the magnitude of the correction to the measured mobility given by eq 6. This is not to say, however, that the residual variation is insignificant. Indeed, it does contain all relaxation effects and effects of ion-specific interactions.

The consistency of the estimates of the limiting mobilities in Figure 5 suggests that the effect of salt on the electrophoretic mobility is mainly caused by counterion condensation due to electrostatic interactions, and that both ion-specific interactions and relaxation effects play a secondary role in the experimental concentration range. For the nonapeptides, the second term in eq 6, that contains the effect of the model electrolyte, including condensed counterions, accounts for as much as about 90 % of the magnitude of the estimated limiting mobility in some cases. For the amino acids, the effect is much smaller in all cases and roughly consistent with OF theory. We emphasize that, within its range of validity, OF theory is preferable to procedure behind Figure 5, as OF theory contains relaxation effects. Neglecting such effects is likely justified for the peptides as a first approximation, owing to the large influence of ion-peptide association on the mobility. For the approach adopted here to be generally applicable for accurate determination of limiting mobilities, however, it would be necessary to find a feasible way to evaluate the full expression from ref [47]. Refinements of the analyte model may also be required to improve the quality of

the estimates of the limiting mobility, but as the long-range Coulomb interaction is the driving force for counterion condensation, the sensitivity to other details should be moderate.

Importantly, the predicted limiting mobilities implicitly constitute an estimate of the diffusion coefficient, through the Einstein relation, which can be measured using other techniques than electrophoresis. As such, diffusion coefficient measurements would likely be less sensitive to ion association than measurements of electrophoretic mobility, and they would provide a suitable test of the presents estimates of the limiting mobilities. These estimates are roughly consistent with the mobility of a sphere with the same size as the confining spheres in the coarse grained model, and are thus of a physically reasonable magnitude.

Turning to the issue of the specificity of ion-sidechain interactions, the coarse grained model calculation for the mixed PDS-sulfate system predict, via eq 6 an almost perfectly linear drop of approximately 5×10^{-9} m^2 V^{-1} s^{-1} in mobility as the sulfate fraction is varied from 0 to 1, see SI. This is in excellent agreement with the experimental results for tetraarginine in Figure 2, but the variation is smaller for tetralysine. This indicates that there is a larger difference in the interaction of PDS and sulfate with arginine than with lysine. Presumably, this reflects the specific interaction between sulfate and the arginine guanidinium side chain groups. Note that the systematic differences between sodium PDS and sodium sulfate seen in Figures 1 and 2 are unlikely to be due to relaxation effects. The larger, and therefore more slowly diffusing, PDS ion should, all else being equal, give rise to a greater slowing down of the analyte due to relaxation effects than the smaller sulfate ions. This intuitive conclusion is supported by numerical calculation of H_a within the OF framework (see eq 4).

4. Concluding Remarks

We carried out capillary electrophoresis experiments to measure the mobility of cationic amino acids, tetrapeptides, and nonapeptides in sodium chloride, sodium sulfate, and sodium propane-1,3-disulfonate BGE. Extrapolations to infinite dilution using OF theory provided a consistent estimate of the limiting mobilities of the amino

acids, but failed to do so for the peptides. Molecular modeling, both on all-atom and coarse grained levels, suggests that this failure of OF theory is due to counterion condensation. The measured mobilities thus pertain to peptide-counterion aggregates, which gives rise to a BGE dependence that cannot be captured by the OF theory. Comparison with a coarse grained model, that contains only electrostatic and excluded volume interactions, allowed us to estimate the limiting mobility in a way that removed most of the BGE dependence. This indicates that this behavior is predominantly due to electrostatic interactions. All-atom simulations, for technical reasons performed at elevated salt concentration, suggest that arginine peptides interact more strongly with PDS than lysine peptides, but this effect appears to be of secondary importance in the concentration range of the experiments. Taken together, the result suggest that the linear treatment of electrostatic interactions inherent in the OF theory causes it to fail unexpectedly for analytes with a high density of charged groups, even at low BGE concentration. This observation may be rationalized by the fact that such analytes *locally* violate the condition for the applicability of OF theory that "the total concentration of ions must be low", as quoted verbatim from ref [8].

Acknowledgments

A.K., T.K. and P.C. would like to thank grant number SVV 2011-263204 from the Grant Agency of the Charles University and Research Project MSM0021620857 of the Ministry of Education, Youth and Sports for funding. P.J. acknowledges the Czech Science Foundation (Grant 203/08/0114), the Ministry of Education (Grant LC 512), and the Academy of Sciences (Praemium Academie) for support. J.H. thanks the IMPRS in Dresden for support.

Conflict of Interest Statement

The authors declare no conflict of interest.

5. References

[1] Shelton, C. M., Koch, J. T., Desai, N., Wheeler, J. F., *J. Chromatography A* 1997, *792*, 455–462.

[2] Mikšík, I., Charvátová, J., Eckhardt, A., Cserháti, T., Forgács, E., Deyl, Z., *J. Chromatogr. B.* 2004, *800*, 161–167.

[3] Okada, T., *Anal. Chem.* 1996, *68*, 1158–1163.

[4] Takayanagi, T., Wada, E., Motomizu, S., *Analyst* 1997, *122*, 57–62.

[5] Takayanagi, T., Wada, E., Motomizu, S., *Analyst* 1997, *122*, 1387–1391.

[6] Kayitmazer, A. B., Quinn, B., Kimura, K., Ryan, G. L., Tate, A. J., Pink, D. A., Dubin, P. L., *Biomacromolecules* 2010, *11*, 3325–3331.

[7] Imami, K., Ishihama, Y., Terabe, S., *J. Chromatogr. A* 2008, *1194*, 237–242.

[8] Onsager, L., Fuoss, R. M., *J. Phys. Chem.* 1932, *36*, 2689–2778.

[9] Debye, P., Hückel, E., *Physik Z.*1923, *24*, 185–206.

[10] Piaggio, M. V., Peirotti, M. B., Deiber, J. A., *Electrophoresis* 2006, *27, 4631*–4647.

[11] Piaggio, M. V., Peirotti, M. B., Deiber, J. A., *J. Sep. Sci.* 2008, *31*, 548–554.

[12] Piaggio, M. V., Peirotti, M. B., Deiber, J. A., *J. Sep. Sci.* 2010, *33*, 2423–2429.

[13] Popa, T. V., Mant, C. T., Hodges, R. S., *Electrophoresis* 2007, *28*, 2181–2190.

[14] Bjerrum., N., *Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd.* 1926, *7*, 1.

[15] Justice, J.-C., Justice, M.-C., *Faraday Discuss. Chem. Soc.* 1977, *64*, 265–273.

[16] Koval, D., Kašička, V., Jiráček, J., Collinsová, M., *Electrophoresis* 2003, *24*, 774–781.

[17] Koval, D., Kašička, V., Zuskova, I., *Electrophoresis* 2005, *26*, 3221–3231.

[18] Allison, S. A., Pei, H., Baek, S., Brown, J., Lee, M. Y., Nguyen, V., Twahir, U. T., Wu, H., *Electrophoresis* 2010, *31*, 920–932.

[19] Wernersson, E., Heyda, J., Kubíčková, A., Křížek, T., Coufal, P., Jungwirth, P., *J. Phys. Chem. B* 2010, *114*, 11934–11941.

[20] Williams, B. A., Vigh, G., *Anal. Chem.* 1996, *68*, 1174–1180.

[21] Wang, J. M., Cieplak, P., Kollman, P. A., *J. Comput. Chem.* 2000, *21*, 1049–1074.

[22] Singh, U. C., Kollman, P. A., *J. Comp. Chem.* 1984, *5*, 129–145.

[23] Besler, B. H., Merz, K. M., Kollman, P. A., *J. Comp. Chem.* 1990, *11*,431–439.

[24] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J.,

Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J., Gaussian 09, revission a.02, 2009.

[25] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., Case, D. A., *J. Comput. Chem.* 2004, *25*, 1157–1174.

[26] Wang, J., Wang, W., Kollman, P. A., Case, D. A., *J. Mol. Graphics Modell.* 2006, *27*, 247–260.

[27] Cannon, W. R., Pettitt, B. M., McCammon, J. A., *J. Phys. Chem.* 1994, *98*, 6225–6230.

[28] Wernersson, E., Jungwirth, P., *J. Chem. Theory Comput.* 2010, *6*, 3233–3240.

[29] Smith, D. E., Dang, L. X., *J. Chem. Phys.* 1994, *100*, 3757–3766.

[30] Berendsen, H. J. C., Grigera, J. R., Straasma. T. P., *J. Phys. Chem.* 1987, *91*, *6269*–6271.

[31] Berendsen, H. J. C., Postma, P. J. M., DiNola, A., Haak, J. R., *J. Chem. Phys.* 1984, *81*, 3684–3690.

[32] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., Pedersen, L. G., *J. Chem. Phys.* 1995, *103*, 8577–8593.

[33] Vondrášek, J., Mason, P. E., Heyda, J., Collins, K. D., Jungwirth, P., *J. Phys. Chem. B Lett.* 2009, *113*, 9041–9045.

[34] Case, D. A., Darden, T. A., Cheatham, III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K.F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin,

M.G., Sagui, C., Babin, V., Kollman, P.A., Amber 10, 2008.

[35] Lide, D. R. (Ed.), *Handbook of Chemistry and Physics, 90th ed.* CRC Press, Boca Ranton, FL, 2010.

[36] Vlachy, V., *Annu. Rev. Phys. Chem.* 1999, *50*, 145–165.

[37] Boda, D., Giri, J., Henderson, D., Eisenberg, B., Gillespie, D., *J. Chem. Phys.* 2011, *134*, 055102.

[38] Allison, S. A., Pei, H. X., Allen, M., Brown, J., Kim, C. I., Zhen, Y., *J. Sep. Sci.* 2010, *33*, 2439-2446.

[39] Allison, S. A., Pei, H. X., Twahir, U., Wu, H. F., Cottet, H., *J. Sep. Sci*. 2010, 33, 2430-2438.

[40] Pei, H. X., Allison, S., *J. Chromatogr. A* 2009, *1216*, 1908-1916.

[41] Abbas, Z., Ahlberg, E., Nordholm, S., *J. Phys. Chem. B* 2009, *113*, 5905–5916.

[42] Frenkel, D., Smit. B., *Understanding Molecular Simulation*. Academic Press, San Diego, CA, 2002.

[43] Best, M. D., Tobey, S. L., Anslyn, E. V., *Coord. Chem. Rev.* 2003, *240*, 3–15.

[44] Mason, P.E., Dempsey, C. E., Nielson, G. W., Brady, J. W., *J. Phys. Chem. B* 2005, *109*, 24185–24196.

[45] Ulander, J., Greberg, H., Kjellander, R., *J. Chem. Phys.* 2001, *115*, 7144–7160.

[46] Manning, G. S., *Acc. Chem. Res.* 1979, *12*, 443–449.

[47] Friedman, H. L., Altenberger, A. R. *J. Chem. Phys.* 1983, *78*, 4162–4173.

Figure Captions

Figure 1. Measured electrophoretic mobilities in sodium chloride (green symbols), sodium propane-1,3-disulfonate (red symbols), and sodium sulfate (blue symbols) background electrolyte for the compounds indicated as a function of ionic strength. The curves are fitting functions according to eq 4, see text and SI.

Figure 2. Effective electrophoretic mobility of arginine (red symbols) and lysine (blue symbols) in mixed solutions of sodium propane-1,3-disulfonate and sodium sulfate for varying fraction of sulfate. The total ionic strength is 50 mM. The lines correspond to linear fits to the data, but this form has no special physical significance; their purpose is to guide the eye.

Figure 3. Screening function $Q(s)$, see eq 5, for the all-atom simulations of arginine and lysine mono-, tetra-, and nonamers, as well as the coarse grained (CG) models of those species, for sodium chloride (green curves) and sodium propane-1,3-disulfonate (red curves) background electrolyte at approximately 500 and 250 mM concentration, respectively. See sections 2.4 and 2.5.

Figure 4. Screening function $Q(s)$, see eq 5, for the coarse grained models of amino acid mono-, tetra-, and nonamers at low background electrolyte concentrations, 10 (blue), 50 (red), 150 mM (green). Dashed curves correspond to the prediction from Debye-Hückel theory. The vertical, thin lines show the radius of the confining spheres.

Figure 5. Limiting mobilities as estimated from the experimental data and the coarse grained model through eq 6. Triangle, rhombus and circle symbols are for mono-, tetra-, and nonapeptides, respectively, and green, red, and blue color is for sodium chloride, sodium propane-1,3-disulfonate, and sodium sulfate background electrolytes, respectively. The horizontal lines correspond to the limiting mobilities of perfect spheres of the same size as the confining spheres in the coarse grained model as estimated from the Stokes-Einstein relation. The error bars for the contribution of the experimental error to the total error of the estimate are smaller then the size of the symbols and are therefore omitted.

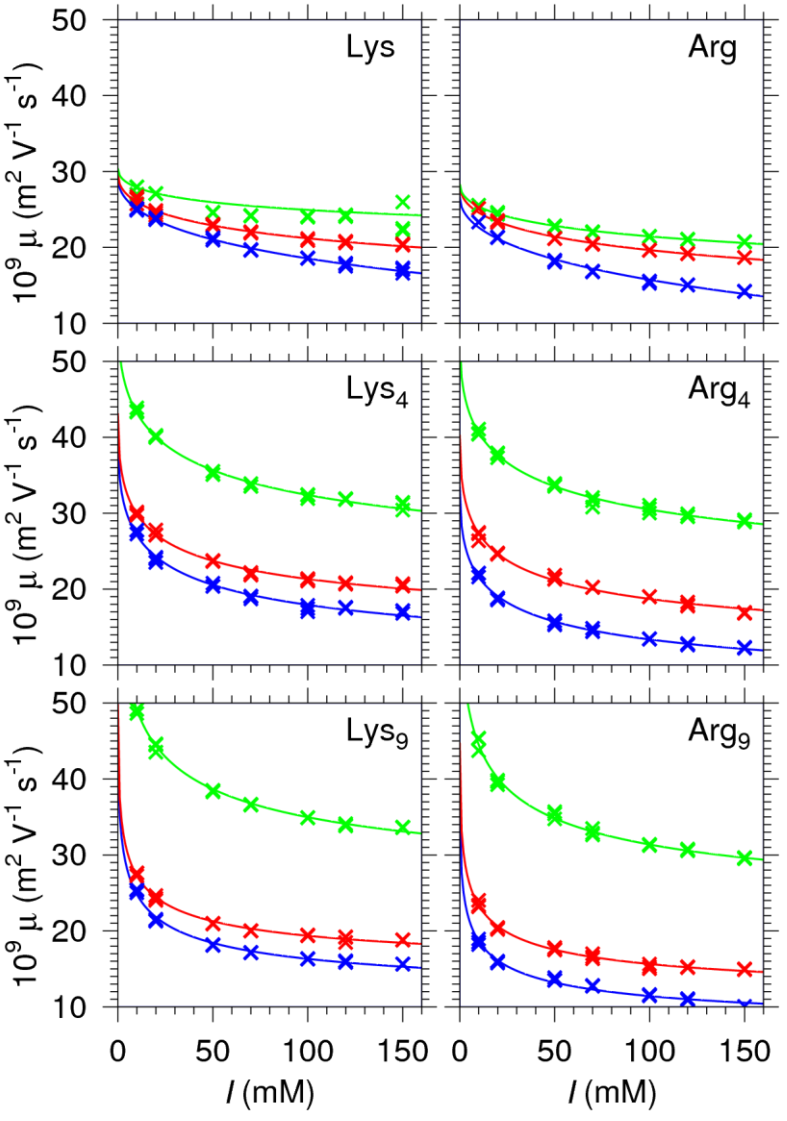

Lys
Arg
Lys4
Arg4
Lys9
Arg9
10^9 μ (m^2 V^-1 s^-1)
I (mM)
10
20
30
40
50
0
50
100
150

Figure 1:

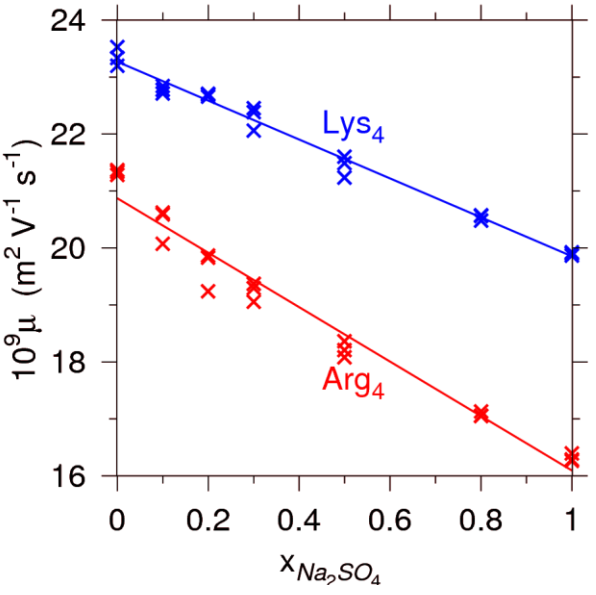
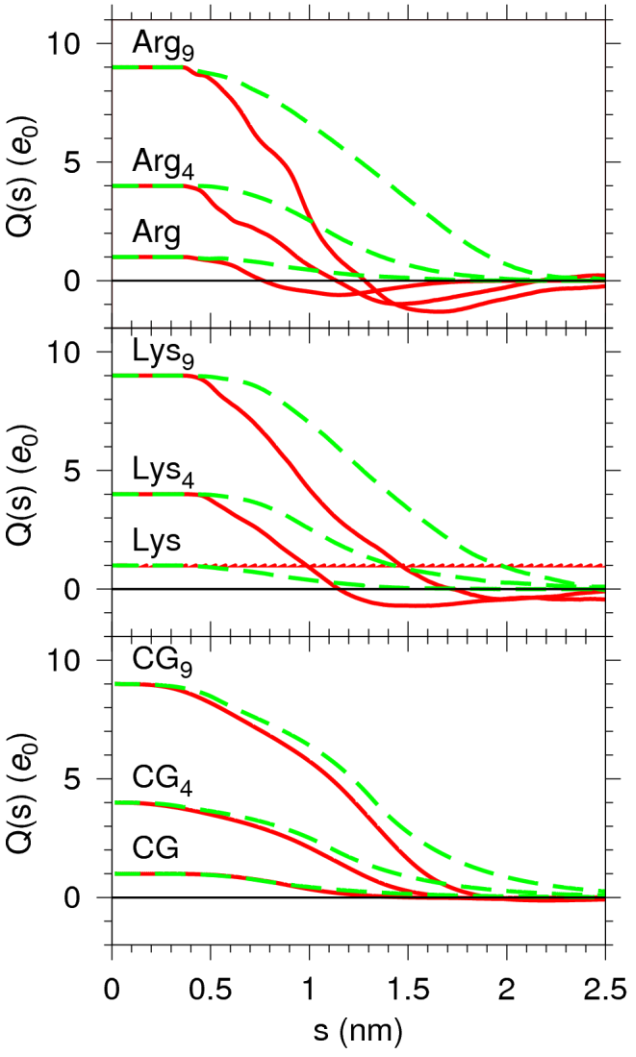


Figure 2:


Arg9
Arg4
Arg
Lys9
Lys4
Lys
CG9
CG4
CG
Q(s) (e0)
10
5
0
0
0.5
1
1.5
2
2.5
s (nm)
Figure 3:

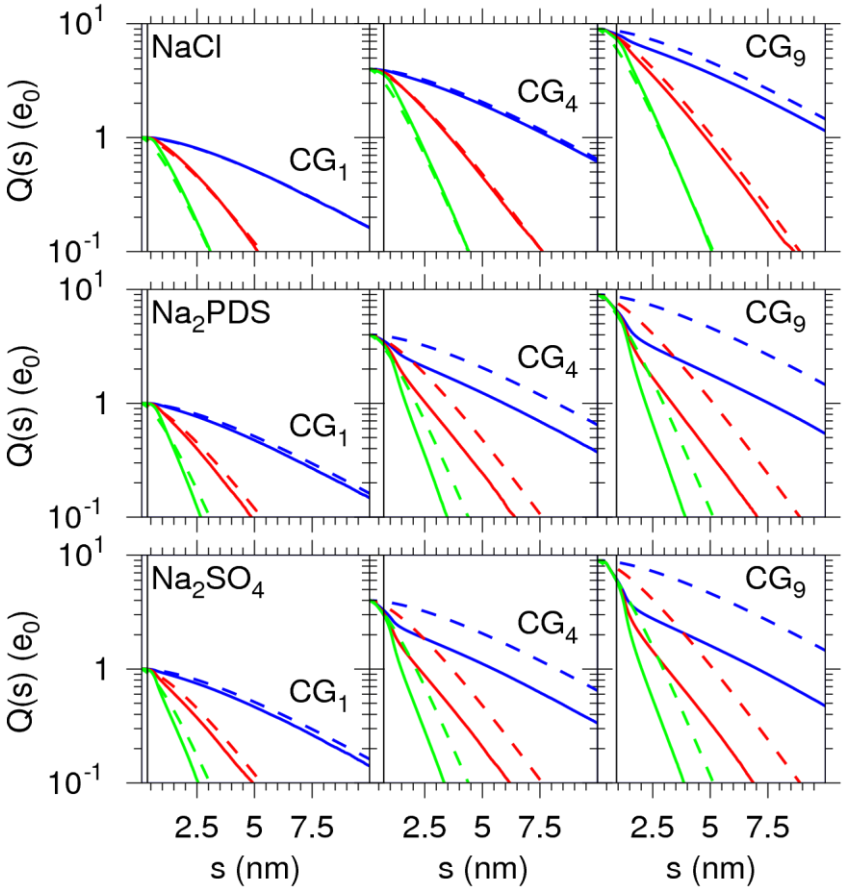


Figure 4:

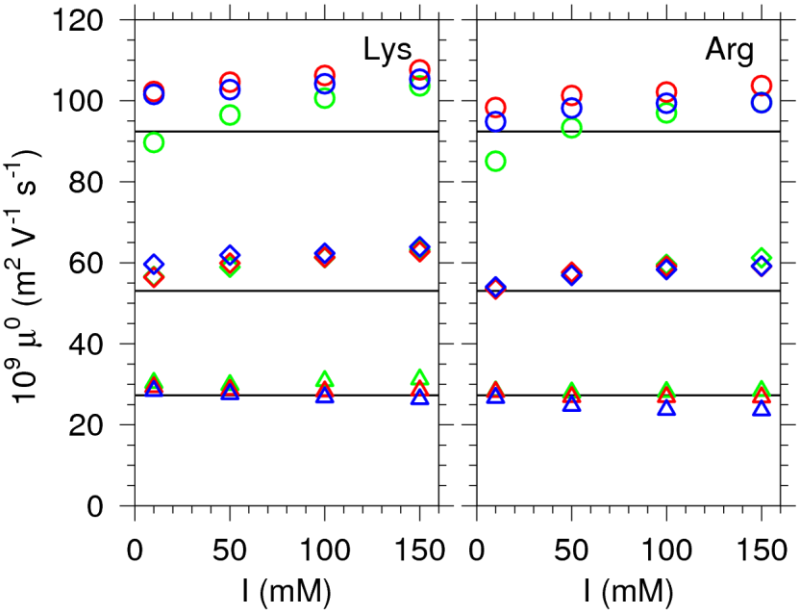


Figure 5: