Spring 4-4-2011

Theoretical Modeling of Oligopeptides through Capillary Electrophoresis and Tarnsport Studies

Umar T. Twahir
Georgia State University, utwahir1@student.gsu.edu

THEORETICAL MODELING OF OLIGOPEPTIDES THROUGH CAPILLARY ELECTROPHORESIS AND TARNSPORT STUDIES

by

UMAR TARIQ TWAHIR

Under the Direction of Dr. Stuart Allison

ABSTRACT

Within this study, the focus will be on oligoglycines. Numerous studies pertaining to the mobility and conformations of oligoglycines have been completed, as this is a driving force for the study. The oligopeptide is modeled using a "coarse-grained" model created in the Allison lab at Georgia State University [Xin,Y.,et. al, J. Phys. Chem. B 2006, 110, 1038-1045], which will be briefly explained within this paper. Oligoglycines will be studied in a few different systems, as the overall charge on the peptide and system will affect its mobility. The conclusion drawn is that the peptide adopts three different conformations based on the temperature of the system and length of the peptide; random conformation at high temperatures, and compact conformations at low temperature. Oligoglycines of length three to five amino acids adopts a cyclic conformation at low temperatures. [Allison, S., et al., J. Sep. Sci. 2010, 33, 2430- 2438.]

INDEX WORDS: Electrophoretic effect, Relaxation effect, Complex formation, Electrophoretic mobility, Electrophoretic mobility of peptides

THEORETICAL MODELING OF OLIGOPEPTIDES THROUGH CAPILLARY ELECTROPHORESIS AND TARNSPORT STUDIES

by

UMAR TARIQ TWAHIR

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Masters of Science

In the College of Arts and Sciences

Georgia State University

2011

THEORETICAL MODELING OF OLIGOPEPTIDES THROUGH CAPILLARY ELECTROPHORESIS AND TARNSPORT STUDIES

by

UMAR TARIQ TWAHIR

Committee Chair: Dr. Stuart Allison

Committee: Dr. Donald Hamelberg

Dr. Kathryn Grant

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Science

Georgia State University

May 2011

ACKNOWLEDGEMENTS

Throughout my time at Georgia State University there have been many interactions with faculty, staff, and peers that have led to achieving all that I have, and helped to me to reach my educational goals at this institution. A select few have left a lasting impression, including but not limited to, Dr. Stuart Allison, who has been a mentor, and great inspiration in my educational career; As well as my upper class men and women in the lab, Dr. Hongxia Pei, and PhD Candidate Hengfu Wu. To them I say many thanks, and show great appreciation for their support and help along this journey.

Contributors to this work:

Stuart A. Allison

Hongxia Pei

Henfu Wu

Herve Cottet

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Chapter 1

OVERVIEW

of

Viscosity of Dilute Bead Model Arrays,

The Electrophoretic Mobility of Small Organic Ions,

Electrical Conductivity of Molecules,

&

Rotational Diffusion of DNA

Overview

The overall focus of my research was the understanding of the fundamentals behind molecule transport, including the factors that effect their movements in solutions and gels. Initially the determination of short-range viscosity interactions of alkanes in a benzene solvent and duplex DNA in an aqueous media were investigated. Viscosity can be defined as the fluid friction that occurs between solute and solvent. As molecules become larger the viscosity of the compound increases as interactions occur between solute and solvent. Temperature is also a factor which can increase or decrease viscosity. The solute concentration has a direct affect on the viscosity of the solvent. The specific viscosity of the solvent can then be directly related to the concentration of solute and the intrinsic viscosity [η]. Intrinsic viscosity is a solutes contribution to the viscosity of a solution. The interaction of solvent and solute can be broken down into three main categories, including the translational friction factor, rotational friction factor, and the volume factor. The model utilized is a coarse-grained solvent continuum bead model in which the solute is represented as a rigid shaped particle or as a rigid bead array that is immersed in a solvent that is represented as a continuum incompressible Newtonian fluid having a specific viscosity. The model used was successful in determining the intrinsic viscosity of duplex DNA of 100 up to 600 or more base pairs when modeled as a wormlike chain. The alkanes in bezene solution act as a model to show the benefits as well as weaknesses of the program. Overall the ability to account for the length dependence of viscosity can be achieved with the model by setting parameters for the system at hand. The theory of viscosity used allows for a representation of suspended macromolecules in dilute solutions that can account for solute-solvent interactions. It is

then shown using this approach that negative intrinsic viscosities are possible, and is used to account as mentioned before, for the length dependence that viscosity exhibits for certain systems. The work has been published, and more detail can be found on the research there [1].

The focus was then shifted away in the direction of studying the electrophoretic mobility of small organic ions, more so, the effects of ionic strength and complex formation on ion electrophoretic mobility. In order to understand the mobility of compounds when electrophoresis is used as a mode of separation, it was broken down into two pieces. The ionic strength and the formation of a complex were seen as two major factors that affect the electrophoretic mobility of a compound. As a current is applied to the solution a few things become apparent. We see the electrophoretic effect taking place causing backflow to occur, leading to the reduction of the mobility. It is also seen that the relaxation effect causes a reduction in the mobility of the compound. When the disturbances begin to occur between the counter-ion and the compound we see the mobility begin to decrease. Two models that were use to correct for the relaxation effect were the Onsager-Fouss-Pitts model, and the Overbeek-Booth-Wiersema model. Both allowed for the correction of the relaxation factor, however, analysis showed that the Onsager model only worked well for smaller and weakly charged particles. The Overbeek model worked over a wider range of compound sizes and charge, making it a better model for the purpose of addressing relaxation. Nevertheless, when the mobilities were compared to that of experimental data it was seen to still not correlate at either low or high slat concentration. It was found that the reason for this was due to a complex formation that occurs between the buffer and compound during separation. It was found

that there are four possible complexes that may form between the molecule of interest and buffer. Therefore the Henry-Yoon-Kim model was then used to address the complex formation that occurred. When conjoining the Henry and Overbeek model, a good approximation can be made for mobilities of compounds, as the relaxation effect and complex formation occur. Using the effective sphere model allowed for the assumption that the only two factors when model compounds that needs to be used were the size and shape. Therefore, isolating the hydrodynamic radius, and charge of the compound, then applying this to the Henry-Overbeek model, mobilities can be effectively identified for compounds. The Henry-Overbeek model corrected for the relaxation effect and complex formation allowing it possible to correlate experimental data with the effective sphere model. Theoretical data was compared to data found in experimentation on six compounds, salicylic acid, sorbic acid, citraconic acid, 4,5-dichlorophtalic acid, 5-sulfoisophtalic acid, and trimesic acid. The work has been published and more details can be found on the research there [2].

Using the concepts of molecule mobility, electrical conductivity of ionic solutions can also be determined. Electrical conductivity is defined as the ability of a material to conduct an electrical current. As in past research numerous time was spent on the studies of molecule transport and mobility. Using the basis of the Onsager-Fouss-Pitts and the Overbeek-Booth-Wiersema models, linking the underlying concepts, the conductivity of molecules in solution can be identified and modeled. On the basis of Ohms law a relationship between charged carriers and mobility of an electron (μ_e), and faraday's constant (F) can be linked to equate a value for conductivity (K) as shown in equation 1.

$$K = -F\, C_e\, \mu_e \quad (1)$$

Understanding the fact that the charged carriers are the significant factors in modeling conductivity of a molecule in solution, equation 1 can be rewritten to include the charge (z) and concentration (m) of the electrolyte solution as shown in equation 2 and 3.

$$K = +F \Sigma z_j m_j \mu_j \qquad (2)$$

$$\Lambda_j = F | \mu_j | \qquad (3)$$

As shown in equation 3 the conductivity (Λ) of a molecule can be simplified to multiplying the mobility of the ion (μ_j) by Faraday's constant (F), after accounting for the summation of carrier charge as shown in equation 2. Using the Onsager Fouss model parameters, the experimental mobilities of binary electrolytes can be attained and compared to experimentally found results to ensure model success. Important parameters necessary to understand include the charge of each ion in the molecule, radius of each molecule, size exclusion radius used as an adjustable parameter and the limiting equivalent conductance of the molecule. In order to expand the scope, theoretical approach for large ions was attempted. Using the full Poisson Boltzmann distribution, the ability to calculate the conductivity for larger ions can be accomplished. However, an issue that presents itself is that the Brownian motion of the ions are neglected. In order to include the Brownian motion correction, inclusion of the diffusion constant (D) becomes necessary. Solving for the apparent diffusion for each ion will include the Brownian motion by solving for each ion using equation 4 and 5.

$$(4)\ D_j^{app} = D_j + D_k$$

$$(5)\ a_j^{app} = (1/a_j + 1/a_k)^{-1}$$

The use of the sub terms j and k are used to signify that one ion represents the reference ion and the other, the ion of interest. The equation is solved twice in order to attain the

apparent diffusion for each ion, holding each fixed to solve for the other. The program calculates the mobilities (E) of each ion which then can be used to solve for the conductivity using equation 6.

$$(6)\ \Lambda = 120.920\ (|\ E_{+}\ | + |\ E_{-}\ |)\ cm^2 / ohm\ mole$$

Using this approach expands the theoretical model to a wider variety of larger and smaller electrolytes. However, what is seen when approaching more complex charged binary electrolytes, the ability to successfully model conductance is lost. In order to account for this, it was discovered that complexes form dipolar ions. In order to account for this it becomes necessary to account for the dissociation constant once in solution, which also means incorporating the activity coefficients for the ions. In order to solve for the conductance two adjustable parameters are used, the variation of both the size exclusion radius and the dissociation constant. Both parameters are adjusted until model conductivities fit the experimental conductivities. This work has been published and more details can be found on the research there [3].

Current research is focused on the understanding of the rotational diffusion of DNA in a gel. Many techniques have been employed in studying the behaviors of peptides, DNA, and other molecules. When a molecule, or as applied in this study DNA, is subjected to an electrical current as in electrophoresis, two motions occur, translational and a rotational motion. As studied before translational motion has been successfully modeled. In this study we explore the rotational motion in a gel, as modeling has never been conducted as of yet. When the DNA is subjected to electrophoresis, there is a motion along the X, Y and Z direction. The motion along the X, and Y are the same representing rotation along the short axis of the fragment. The rotation along the Z-axis

can be represented as a head to tail rotation. When placed in a gel the rotation of the fragment is greatly affected by the gel it is placed in. There are two affecting factors, long-range hydrodynamic interaction and direct interaction, both of which would reduce the rotational diffusion. The gel, agarose in this study is made up of fibers that are separated by a specific distance. Depending on the length of the fragment the rotation will be held to specific constraints. Experimental data was obtained from, Stellwagen C. N., Colloid and Surfaces, Physiochemical and Engineering Aspects 209, 2002, 107-122, and a model, effective medium model, was utilized to account for the long-range hydrodynamic interaction. [4,5,6] We'd now like to apply our model to a bead array of N beads rotating with angular velocity ω about its mass center in an effective medium (a gel). When a molecule is subjected to a rotational force, a constant torque is applied, in which is the force allowing an angular velocity to occur. The torque can then be related to the rotational friction along the y-axis.

$$T = f_{r\perp}\,\omega_y \qquad (7)$$

The torque, T, is equated to the product of the perpendicular friction, $f_{r\perp}$, and angular velocity ω_y.

Using the hydrodynamic continuum, the parallel friction can be solved, as the torque can be treated as tensor. Using this quantity, the friction can also be treated as tensor quantity, replacing the perpendicular friction with $\underline{f_R}$. Using the Einstein-Sutherland equation, the rotational diffusion can be solved for perpendicular and parallel motion.

$$D_{R\perp} = \frac{k_B T}{f_{R\perp}} \quad (8)$$

$$D_{R||} = \frac{k_B T}{f_{R||}} \quad (9)$$

$$\underline{D}_R = k_B T \, \underline{f}_R^{-1} \quad (10)$$

Experimental data can be attained about rotational diffusion, and modeling can define rotational friction factors. Equations 8 thru 10 allow a connection to experimental and theoretical data. The use of electric birefringence is used to study the rotational relaxation of the DNA in which the rotational diffusion can be extrapolated. [4]

$$\tau_R = \frac{1}{6D_{R\perp}} \quad (11)$$

As mentioned before, a gel reduces the rotational diffusion of the molecule being studied. The long-range hydrodynamic interaction is the first and most prevalent interaction. The second interaction that is now being studied is the direct interaction between the gel fibre's and the molecule. Both interactions become have higher effects in a dense gel, while in a dilute one only the long-range hydrodynamic interaction becomes pertinent. In modeling a strand of DNA, an important output quantity is a reduced friction tensor.

$$XR = \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{pmatrix} \quad (12)$$

The Λ represent the eigen values of the reduced frictional tensor. Each eigen value is solved by an equation related to that of the rotational diffusion.

$$D_{Ri} = \frac{k_B T}{f_{Ro}} \left(\frac{f_{R\phi}}{f_{Ri}} \right) = \frac{k_B T}{8\pi\eta a^3 N \Lambda_i} \quad (13)$$

$$\tau_1 = \frac{1}{6D_{R1}} = \frac{4\pi\eta a^3 N}{3k_B T}\Lambda_1 \qquad (14)$$

When using electric birefringence, the rotational relaxation time is obtained, as is necessary to relate it to data that can be obtained from modeling as shown in equation 14. It is then necessary to relate the gel being used to analyze the molecule, to the molecule in itself. The gel is related to the gel concentration, λ.

$$\frac{1}{\lambda^2} = -\frac{3\rho_g \omega_s r_g^{\,2}}{20M} * \left[\ln\left(\frac{M}{\rho_g \omega_s}\right) + 0.931\right] \qquad (15)$$

The gel fibre radius, r_g , density of the dry gel, ρ_g, ratio of hydrated gel to dry gel, ω_s, and the gel concentration is grams dry gel per milliliter of water, M, is shown in equation 15. Once all parameters were set, the gel fibre radius was set, as well as the persistent length. The persistent length is the length, which is traveled along the chain until the chain bends by approximately 90 degrees. For a DNA strand of 622 base pairs one persistence length is 500 Å, overall the 622 base pair chain has a persistent length of 4.2 P. However, varying the persistent length can allow for the conformation to be varied until a conformation that a theoretical value of rotational relaxation matches the experimental rotational relaxation time. The DNA chain was modeled using a bead modeling method as a wormlike chain of 88 beads. The program was successful in modeling the rotation of DNA of 622 base pairs. At a low gel concentration of $M \leq 0.010$ g/ml a good agreement between experimental and theoretical data is achieved. The persistent length was set to 65nm, and the gel fibre radius was set to 2.5nm. The persistence length was initially varied and later set to 65nm, and the gel fibre radius was varied at 1.52 nm, 2.0 nm, and 2.5 nm. The data was then plotted versus experimental data to see which gel fibre radius best fits the data as shown in Figure 1.1.

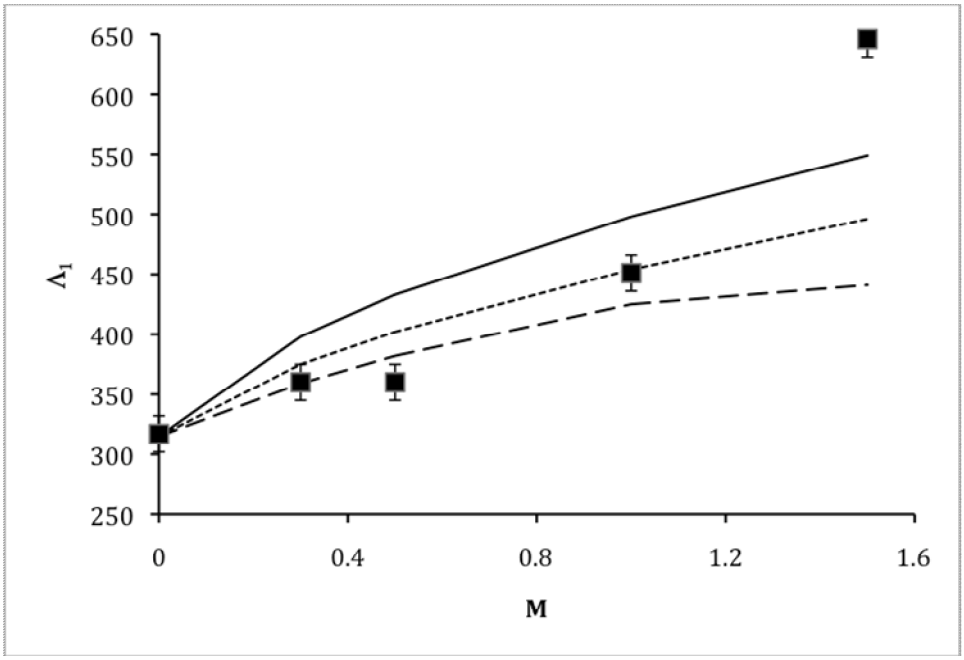

Figure 1.1: Frictional force eigen value versus gel fibre radius. Solid Line- 1.52nm, Small dashed line- 2.0nm, Large dashed line- 2.5nm, Squares- experimental data.

When gel concentration is increased other interactions are had not explained by the model. When the gel concentration becomes greater than 0.010 g/ml the gel fibre spacing (A) becomes comparable to the persistence length of the chain, causing a direct interaction between gel and molecule. It is suspected that direct interaction between the base pairs and gel fiber causing the rotational diffusion to be reduced. Wu, H., Twahir, U., Allison, S.A., Davis, A.N., Duodo, E., Kashani, B.B., Lee, Y.K., Pena, C., Witley, N., *Rotational Diffusion of Macromolecules and Nanoparticles Modeled as Non-Overlapping Bead Arrays in an effective Medium*, has been submitted for review, and pending response. Topics that have been discussed thus far are projects that have also been worked on, while completing the research for the project that will be discussed in detail within this thesis. All works discussed above have been published and can be found for

further information on any individual topic. The rest of work will now focus on the study completed on oligoglycines. The focus of this thesis will however utilize information found in previous study on the electrophoretic mobility of molecules, now expanding the scope to include oligoglycines as outlined in the abstract. The work has also been published as it is reflected in this piece [7].

Chapter 2

Oligoglycine Modeling Studies

Introduction 1

Over the last 20 years, free solution capillary electrophoresis, FSCE, and related techniques have emerged as major analytical tools in the separation of peptides as discussed in a number of reviews [8-14]. Peptide separations have been and continue to be a specific application of considerable interest [15-22]. Understanding how experimentally controlled variables influence the electrophoretic mobility, μ, is important in defining conditions that are optimal in a particular separation. Such variables include composition of background electrolyte, BGE, temperature, T, solvent, and pH. Two different general approaches have emerged in modeling μ of peptides. The first is semi-empirical and has its origins in the work of Offord [23], in which the mobility is written $\mu \cong c_1 Z/M^{\alpha}$[1-5, 9-11, 13] or $c_2 \log(Z+c_3)/M^{\alpha}$ where Z is the effective peptide charge, M is the molecular mass, c_1, c_2, c_3, and α are empirical constants. However, these models fail when applied to hydrophobic or highly charged peptides [20]. A number of more sophisticated semi-empirical computer models have been developed in which μ is correlated with additional variables (descriptors) in addition to Z and M [25-27]. The second approach is fundamental in nature and is grounded on continuum electro-hydrodynamic theory. It has its origins in the electrophoresis and conductivity of spherical ions [28-36]. This has been extended to other particles including long [30] and finite [38] rods, axi-symmetric ellipsoids [39-41], and rigid particles of arbitrary shape and charge distribution [42-44]. In order to deal with the problem of flexible structures in which the model particle passes through many conformational states, a bead modeling methodology, BMM, has been developed and applied to peptides [45-49]. A peptide composed of n amino acids is modeled as $2n$ beads and this shall be called the B model in

the present work. Independently and following a distinctly different strategy, coarse grained bead modeling has been applied to evenly charged synthetic polyelectrolytes in pure water where it has correctly predicted the variation in mobility with degree of polymerization [50].

We are confident that the "fundamental approach" discussed at the end of the previous paragraph is capable of accurately predicting μ of peptides in FSCE. A principle objective of the present work is to provide a convincing demonstration of this by examining the dependence of μ on temperature, ionic strength, and composition of BGE for oligoglycines under well defined conditions [19, 22]. An advantage of studying these peptides is that their charge state and pK_a values of the C- and N- terminal charge groups are now well known [51, 52]. Aside from assumptions regarding peptide secondary structure, this makes it possible to model μ *without adjustable parameters* and compare it directly with experiment. As shown in the present work, it is also possible to draw some fairly definite conclusions about peptide solution conformation versus temperature and how components of the BGE may influence peptide conformation and possibly form complexes with the peptide.

The B model is outlined in Section 2. Since the model and procedure for calculating mobilities (BMM) have been discussed at length in previous work [45-49], detail shall be kept to a minimum. However, recent refinements in the BMM are included in the Appendix. For weakly charged particles and peptides in particular [19, 22], it is possible to extrapolate mobility to the limit of zero salt concentration and zero peptide concentration, $\mu^{0\infty}$. The advantage of doing this is that dependence of mobility on BGE drops out and $\mu^{0\infty}$ can be related to hydrodynamic radius, A, and the translational

diffusion constant in the limit of zero concentration, D_T^0. In Section 3.1, we examine the temperature dependence of $\mu^{0\infty}$ for oligoglycines (n = 2 to 11) measured by Plasson and Cottet [22] and using the BMM, address what the mobility measurements are telling us about the solution conformation of oligoglycines as a function of size/length and temperature. These experiments were carried out at pH = 11.5 where the charge state of the oligoglycines is -1. In the low or mid pH ranges, however, peptide charge is not so certain. Section 3.2 discusses the theoretical and experimental charge state of the peptides under conditions of general pH. Mobility modeling in low pH phosphate and citrate buffers and high pH borate buffers is the subject of Section 3.3. Section 4 discusses the main conclusions of this work and explores possible alternative explanations for the temperature dependence of A for the smaller oligoglycines at low temperature.

Experimental 2

We employ a coarse grained model to represent a peptide that captures many of the gross structural characteristics of an all atom representation. This model has been described in detail previously [46-49] and only an outline shall be provided here. A peptide consisting of n amino acids is modeled as 2n non-overlapping beads. Two beads are assigned to each amino acid with the radius of one bead, the backbone bead, fixed at 0.19 nm, and the radius of the other bead, the side bead, variable. The value of 0.19 nm corresponds to one half of the alpha-carbon to alpha-carbon distance typical of peptides [53]. The radius of the side beads are parameterized on the basis of the translational diffusion constant of the corresponding amino acid [54, 55]. Due to the rigidity of the peptide bond unit, the conformation of a peptide chain can be accurately defined by a

succession of dihedral angles [56]. Peptide conformations can be produced by generating a series of dihedral angles and using rotation matrices [56] to place the beads at their proper positions. Assumptions about local secondary structure (helix, random coil, etc.) can be made which defines ranges over which the dihedral angles can vary [48]. Typically, 100 to several thousand different conformations are generated and various transport properties (translational diffusion constants, electrophoretic mobilities) can be calculated for each conformation and subsequently averaged. Details of the methodology are described in detail in previous work [46-49]. Figure 2.1 shows representative conformations of B model hexaglycine in three different assumed secondary structural motifs [48].

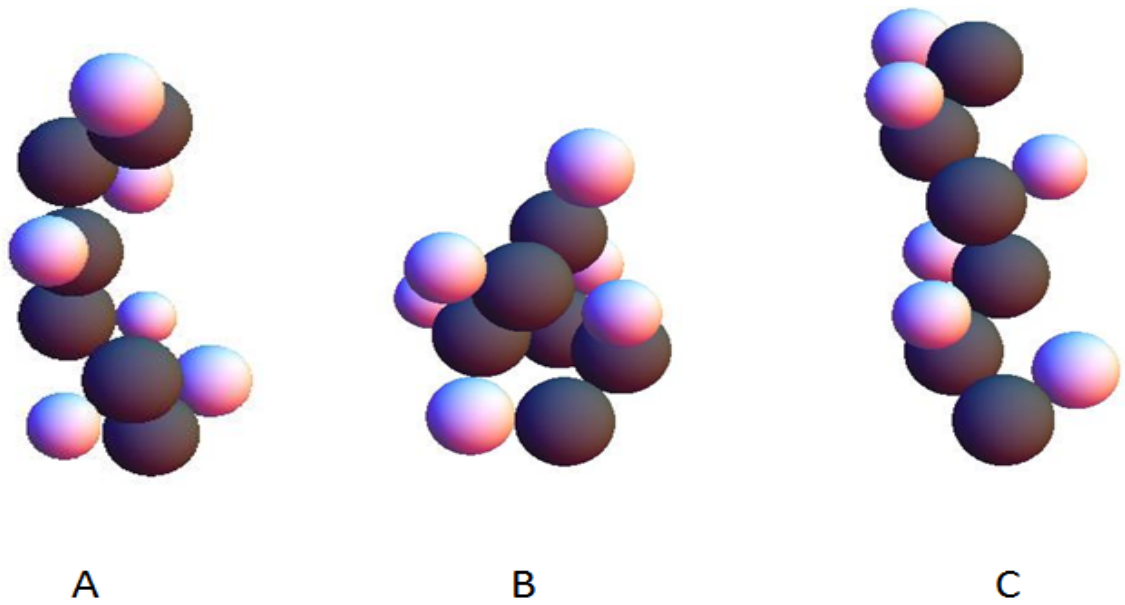

Figure 2.1 B models for Hexa-glycine in 3 different secondary structural motifs. (A) a typical random conformation, (B) an α-helix, (C) a P_{II}-helix. Dark and light grey spheres denote backbone and side beads, respectively.

It should be emphasized that the "random" structure shown in Fig. 1A is only one of an infinite number of possible conformations that range from compact to extended. The range of possible conformations for the α- and P_{II}-helices is much more restricted due to the limited allowed range of the dihedral angles [48].

One significant modification in the BMM algorithm has been made that ultimately yields the same model mobility as before, but in a more direct manner. Although this modification is important, it is undoubtedly not of interest to the general reader and is consequently placed in the Appendix. Eqs. (A9) and (A10) summarize the model mobility of a particular B model configuration.

Results 3

3.1 Hydrodynamic Radii of Oligoglycines

The hydrodynamic radius, A, of a peptide (or any particle) can be defined in terms of the translational diffusion constant in the limit of zero peptide concentration, D_T^0, in a solution of viscosity η_0 and temperature T

$$A = \frac{k_B T}{6\pi\eta_0 D_T^0} \qquad (1)$$

In Eq. (1), k_B is the Boltzmann constant. If the peptide bears charge eZ (e is the protonic charge), A can also be defined in terms of its electrophoretic mobility in the limit of zero peptide concentration and ionic strength, $\mu^{0\infty}$,

$$A = \frac{eZ}{6\pi\eta_0 \mu^{0\infty}} \qquad (2)$$

If A is in Angstroms, η_0 is in centipoise, and $\mu^{0\infty}$ is in 10^{-9} m^2 V^{-1} sec^{-1}, Eq. (2) reduces to

$$A = \frac{84.99Z}{\eta_0 \mu^{0\infty}} \qquad (3)$$

Values of $\mu^{0\infty}$ for oligoglycines (n = degree of polymerization = 2 to 11) for T ranging from 15 to 60 ºC have been reported by Plasson and Cottet [22]. Survay et al. [19] previously carried out similar studies for n = 2 to 6 at 25 ºC in a variety of buffers and their $\mu^{0\infty}$ and A values are consistent with those of reference 22. The experiments in reference 22 were carried out in triethylamine buffers of variable ionic strength at pH = 11.5 and $\mu^{0\infty}$ were extracted by a straightforward curve fitting procedure. At this pH, Z = -1. From Eq. (3), it is therefore straightforward to determine A versus n over this temperature range. (To be precise, mobilities in the limit of zero ionic strength, but finite peptide concentration, μ^{∞}, are reported in references 19 and 22. As discussed in Section 5, however the discrepancy between $\mu^{0\infty}$ and μ^{∞} is concluded to be negligible. In the present work, the "0" and "∞" in the exponent of μ refers to the limit of zero peptide concentration and zero ionic strength, respectively.)

The BMM approach has been outlined in Section 2 and is used here to determine A of oligoglycines. Recently, field gradient NMR was used to measure D_T^0 of glycine and diglycine (G_2) in order to parameterize more accurately the side bead radii of glycine. At 25 ºC in dilute aqueous buffer, D_T^0 for G_2 extrapolated to zero peptide concentration was found to equal 7.6 ± 0.1 x 10^{-10} m^2 s^{-1}, or equivalently A = 3.22 ± 0.05 Å [55]. From this study, side bead radii of 1.64 Å for glycine located at the C- or N- terminal position of a peptide or 1.43 Å for glycine located in a interior position were deduced. It should be emphasized, however, that these values were based on the assumption that G_2 adopted a "random" secondary structure in aqueous solution at 25 ºC [48]. In modeling the oligoglycines in the present work, assumptions are made about the secondary structure in advance which, in turn, determines the allowed ranges of dihedral, phi-psi, angles in

modeling [48]. Average D_T^0 and A values are derived from 100 to 1000 randomly generated conformations.

Plotted in Figure 3.1 are experimental A values (in Å) from reference 22 denoted by symbols (T = 288.15 K, filled squares; 298.15 K, x's; 311.15 K, unfilled triangles; 323.15 K, +'s; 333.15 K, unfilled squares).

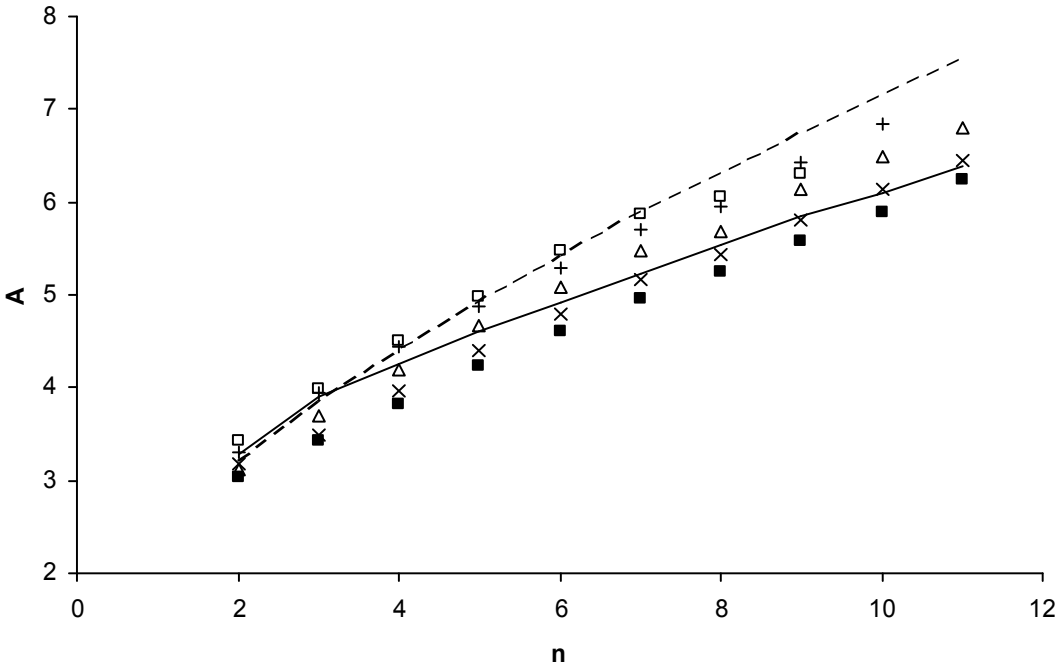

Figure 3.1: A versus n for oligoglycines. Values of A (hydrodynamic radii) are in Å. Experimental values come from reference 15 and are denoted by symbols (T = 288.15 K, filled squares; 298.15 K, x's; 311.15 K, unfilled triangles; 323.15 K, +'s; 333.15 K, unfilled squares). BMM modeling results are the solid line for an assumed α-helical secondary structure, and the dashed line for an assumed "random" secondary structure.

The behavior expected for peptides in α-helical and "random" secondary structures are denoted by the solid and dashed lines, respectively, and come from modeling. At high T, the experimental results conform to the "random" model for n < 8. For longer oligoglycines at high T, a slightly more compact conformation is indicated. However, a comparison of experimental and modeling A values indicate a compact average conformation at the lower temperatures of 15 and 25 ºC. For the longer peptides (n > 7), an α-helical secondary structure is fairly consistent with experiment. It is rather surprising to note that for small peptides in the n = 3 to 6 range, even a relatively compact α-helical structure cannot account for the small A values observed. It should also be emphasized that similar experimental results have been reported independently in a variety of buffers, but only at 25 ºC [19].

Consider G_4 where A from reference 22 equals 3.83, 3.96, 4.20, 4.44, and 4.51 Å at 15, 25, 38, 50, and 60 ºC, respectively. For a "random" model, which basically samples all possible conformations, the average A equals 4.42 Å, but the range is from 3.75 (compact) to 4.67 (extended) Å for particular conformations. Shown in Figure 3.2 are two particular G_4 conformations that yield A ≤ 3.85 Å.

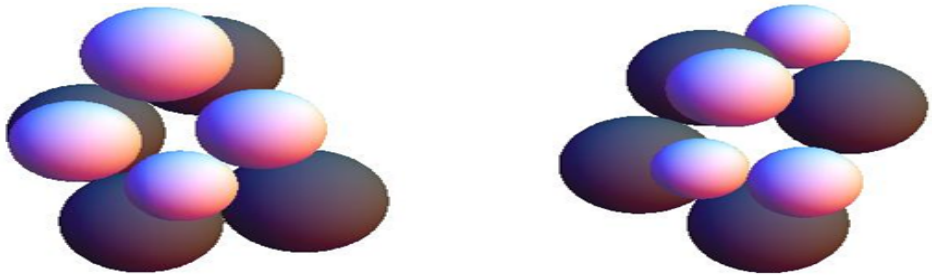

Figure 3.2: Compact B models for tetraglycine. The two cyclic structures shown have A values consistent with experiment at 15 °C. Dark and light grey spheres denote backbone and side beads, respectively.

What is notable about these compact conformations is that they are cyclic. Is it possible that the solution conformation of G_3 through G_6 in solution is cyclic? In the gas phase, there is strong experimental evidence (from IR spectroscopy) coupled with quantum mechanical modeling that this is indeed the case [57]. For G_4 and G_5, cyclic structures stabilized by intramolecular hydrogen bonding are the low energy forms. For G_3, linear and cyclic structures are low in energy. The experimental mobility measurements [19, 22] coupled with B-model modeling are consistent with cyclic structures for n = 3, 4, and 5 at low T, but random conformations at high T.

In the remainder of this paper, attention shall be turned to the mobility of oligoglycines in different BGE solutions of finite concentration. At low pH, the charge state of the C-terminal must be considered and Section 3.2 addresses that point. In terms of modeling, we shall consider two classes. In the "random" class, peptide conformations are generated assuming a "random" secondary structural motif, SSM. In the second "restricted" class, peptide conformations are also generated assuming a "random" SSM, but only those conformations with D_T^0 values falling above a certain value, D_T^*, are accepted. The value of D_T^* is chosen so that $\langle A \rangle$ matches the experimental value determined from $\mu^{0\infty}$ following Eq. (3). In the "restricted" class, conformations sampled will be much more compact than is typical in the "random" class.

3.2 Comparison of Model and Experimental Charges of Oligoglycines

For an oligoglycine of degree of polymerization n, the peptide charge is determined by the protonation state of the N-amino and C-carboxylic acid terminal groups. At low pH, the N terminal is fully protonated. Let pK_{a1} denote the *thermodynamic pK* of the C-terminal and let HX_nH^+ and HX_n denote the C-terminal protonated and deprotonated

forms bearing net charges of +1 and 0, respectively. In this notation, *X* denotes glycine and *n* is the degree of polymerization. The left hand *H* (of HX_nH^+) denotes the protonation of the N-terminal amino group and the right hand *H* the protonation of the C-terminal carboxylate group. We can write

$$K_{a1} = \frac{a(H^+)\,a(HX_n)}{a(HX_nH^+)} = \frac{10^{-pH}[HX_n]}{[HX_nH^+]}\frac{1}{\gamma_\pm} = \frac{K'_{a1}}{\gamma_\pm} \qquad (4)$$

$$K'_{a1} = \frac{10^{-pH}[HX_n]}{[HX_nH^+]} \qquad (5)$$

$$pK_{a1} = pK'_{a1} + \log_{10}\gamma_\pm \qquad (6)$$

In Eq. (4), *a* denotes activity, brackets denote molar concentrations, and $\gamma_\pm$ the (monovalent) activity coefficient of HX_nH^+. The latter term can be approximated using the extended Debye-Huckel law for a finite monovalent ion of radius $A_\pm$. (See reference [58] or standard Physical Chemistry textbooks.)

$$-\log_{10}\gamma_\pm = \frac{c_1\sqrt{I}}{1+c_2 A_\pm\sqrt{I}} \qquad (7)$$

$$c_1 = \frac{1.82\,x10^6}{(\varepsilon_r T)^{3/2}} \quad (M^{-1/2}) \qquad (8)$$

$$c_2 = \frac{502.8}{(\varepsilon_r T)^{1/2}} \quad (M^{-1/2}\,nm^{-1}) \qquad (9)$$

where *I* is the ionic strength (in M) and $A_\pm$ is in nm. It is customary to make the approximation of setting $A_\pm$ equal to 0.5 nm for peptides [51, 59]. It should be emphasized that the activity correction defined by Eq. (7) above is small under the aqueous low salt conditions usually encountered in FSCE. In water at 25 ºC (ϵ_r = 78.54), Eq. (4) becomes

$$-\log_{10}\gamma_{\pm} = \frac{0.508\sqrt{I}}{1+1.643\sqrt{I}} \qquad (10)$$

For the deprotonation of an amino group appropriate at high pH, a very similar analysis to that given above holds, except that the $+\log_{10}\gamma_{\pm}$ term in Eq. (3) is replaced with $-\log_{10}\gamma_{\pm}$. The *thermodynamic* K_{a1} is independent of the BGE, but K_{a1}', defined by Eq. (5), does depend on the BGE through the activity coefficient. However, the activity correction given by Eq. (6) that relates pK_{a1} to pK_{a1}', is usually small and amounts to a correction of 0.1 pH unit or less.

Plasson and Cottet have determined, among other things, the pK_{a1} of the smaller ($n < 10$) oligoglycines [51, 52]. Table 2 of reference 51 gives pK_{a1} values for oligoglycines of variable length in aqueous media from 15 to 60 °C. For oligoglycines, pK_{a1} increases from 3.10 to 3.31 as n increases from 2 to 10 at 25 °C. In other words, the C-terminal of G_2 is more acidic than the C-terminal of longer oligoglycines. This can be readily understood on physical grounds on the basis of the phenomenon of 'charge regulation' [44,60, 61].

In the above example of oligoglycines at low *pH*, the presence of a neighboring positive charge from the protonated amino group on the N-terminal serves to stabilize the deprotonated form of the C-terminal and this serves to make the C-terminal more acidic. However, as the length of the peptide increases, the average distance between the N and C terminals increase and this reduces the 'charge regulation' effect. Using thermodynamic free energy arguments, it is straightforward to account for the 'charge regulation' effect in modeling. Specifically, for the B model summarized in Section 2, we can write for the deprotonation of charge site k of a B model peptide in a particular conformation [48],

$$pK'_{ak}(B) \;=\; pK^{0}_{ak} \;\pm\; \log_{10}\gamma_{\pm} \;-\; \frac{0.434}{k_B T}\sum_{j\neq k} Z_j (C_{jk} + C_{kj}) \qquad (11)$$

In Eq. (11), pK_{ak}^{0} represents an intrinsic dissociation constant in the absence of charge regulation, the "+" case of the "$\log_{10}\gamma_{\pm}$" term is taken for deprotonation of neutral sites (such as the C-terminal and other carboxylic acid sites), the "-" case is taken for deprotonation of charged sites (such as the N-terminal and other protonated amino sites), the sum over j is over all other charge sites in the peptide of charge Z_j (in protonic units), and the C's are well defined mathematical quantities that depend on peptide conformation, bead sizes, and ionic strength. The C's are given by Eq. (A8) of reference [48] and shall not be reproduced here. What is new about Eq. (11) above is the inclusion of the activity coefficient correction term.

The only adjustable model parameters in the 'charge regulation' methodology summarized in Eq. (11) above are the intrinsic pK_{ak}^{0} and also assumptions about the secondary structure of the peptide. As discussed at the end of the previous subsection, two B-model classes are considered, "random" and "restricted". Since pK_{a1} (pK_a(C-term)) and pK_{a1}^{0} (which denotes the intrinsic pK_a of the C-terminal) are linearly related, it is straightforward to adjust the latter to obtain optimal fits between model and experiment. In modeling, we set T = 25 °C, pH = 2.0, I = 0.01 M that match reasonably well the conditions used in the low pH mobility measurements of reference 19. In the next subsection, model and experimental mobilities are compared under these conditions. Good agreement between experimental and model pK_{a1} values is achieved when pK_{a1}^{0} is set to 3.39 and 3.47 for "random" and "restricted" models, respectively. Figure 3.3 shows how pK_{a1} between experiment [51] and models compare versus n at 25 °C. Clearly, the "charge regulation" B model is able to account quite well (within about 0.05

pH unit) for the length dependence of pK_{a1} for oligoglycines. For the "restricted" model, the non-monotonic variation of pK_{a1} with n is reproducible.

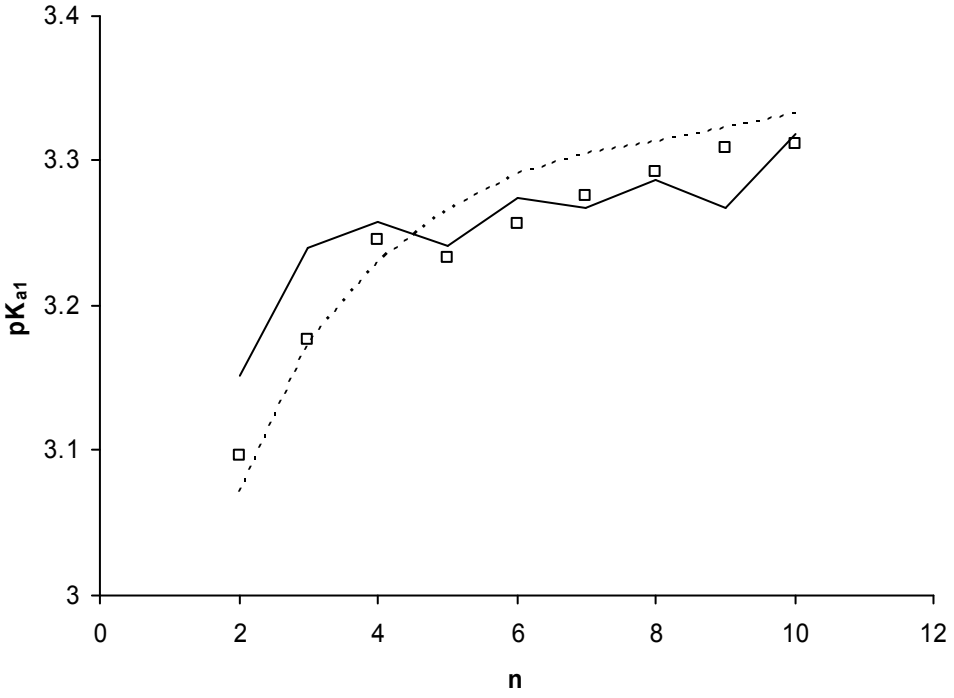

Figure 3.3: pK_{a1} of oligoglycines versus n at 25 °C in aqueous media. Unfilled squares represent experimental pK_{a1} values (C-terminal acid dissociation constants) for oligoglycines from reference 51. The dotted line is from "random" B models with pK_{a1}^0 = 3.39 (the intrinsic pK_{a1} value). The solid line is from the "restricted" B model with pK_{a1}^0 = 3.47.

One objective of the present subsection has been to demonstrate the ability of modeling to accurately predict the pK_{a1} and hence the charge state of oligoglycine peptides. In turn, knowledge of the charge state is a prerequisite to modeling the electrophoretic mobility of a peptide or any particle for that matter. In the next section,

we shall examine the corresponding free solution electrophoretic mobilities in a variety of different buffers at finite ionic strength. Two of these are carried out at low *pH* where the pK_{a1} is important in determining the average charge and hence model mobility of the peptide. However, another buffer system (sodium borate) is at high ionic strength where the important deprotonation reaction involves the N-terminal of the peptide. An analysis similar to that given above can be applied to pK_a(N-term) and has been carried out. For the sake of brevity, however, it shall not be discussed here other than to state pK_a^0(N-term) = 7.81 and 7.71 for "random" and "restricted" models match the experimental pK_a(N-term) values at 25 °C reasonably well [51]. Also, the "high *pH*" studies are carried out at *pH* = 9.3 to 10.7 where the peptide charge is close to -1 in any case.

3.3 Comparison of Model and Experimental Peptide Mobilities of Oligoglycines.

Experimental studies of the mobility of peptides as a function of ionic strength, *I*, are very limited [19, 22, 50]. Due to the presence of amino and carboxylic acid moieties on the ends of the peptides and possibly as side groups, the charge state of the peptide is strongly *pH* dependent in the ranges from about 2 to 5 and 7 to 9. For most buffer systems in which CE experiments are carried out, it is difficult to vary *I* without varying *pH* unless considerable care is taken in the design of the buffer [51]. If, for example, we wish to maintain *pH* = 3.0 in a NaH_2PO_4/H_3PO_4 buffer, the lower bound on ionic strength is about 0.001 and consists of pure H_3PO_4 in water. Higher ionic strengths at this pH can be achieved by carefully adjusting the absolute and relative amounts of NaH_2PO_4 and H_3PO_4 present. In a series such as the oligoglycines at low *pH*, the effective charge, *Z*, of the peptide is determined by the charge state of the C-terminal carboxylic acid group, or pK_{a1}, and these depend on n. As discussed in the previous sub-section, however, the pK_{a1}

values are known experimentally and can be reproduced reasonably well using the B model. At this point, the only remaining model assumptions concern peptide conformation. As discussed at the end of Section 3.1, two B-model conformational classes are considered, "random" and "restricted". In the "random" class, a broad range of dihedral angles are sampled and many conformations are possible. In the "restricted" class, only those conformations with a D_T^0 exceeding a certain threshold are considered. In the limit of zero ionic strength, the analysis of Section 3.1 indicates that the oligoglycine conformations are compact at low *T*, but expand to adopt a nearly "random" conformation at high *T*. The present section extends this analysis to finite ionic strength and addresses the possibility of specific peptide-buffer interactions beyond simple electrostatic interactions. However, this analysis shall be limited to 25 ºC.

For experimental data, we turn to the study by Survay et al. [19] on the free solution mobility of oligolycines at 25 ºC in a variety of buffers and covering a range of *pH* values. Specifically, low *pH* sodium phosphate, low pH sodium citrate, and high *pH* sodium borate buffers are considered. Table 3.1 summarizes the buffer conditions used in modeling that match experimental conditions [19].

Table 3.1 : Composition of BGE buffers

Designation	Composition	I (M)	pH
Ia	sodium phosphate	0.010	2.03
Ib	"	0.018	3.01
IIa	sodium citrate	0.005	2.57
IIb	"	0.013	3.24
IIIa	sodium borate	0.048	9.3
IIIb	"	0.070	10.1

Care was also taken in the experiments to account for the effects of Joule heating, which is an important consideration since raising the solution temperature by a few degrees significantly reduces the solution viscosity and this, in turn, increases the absolute mobility [19, 22, 51, 62]. For example, Joule heating may raise the average temperature inside a capillary by about 4 ºC [62] and this reduces the solution viscosity by 8 %. Since mobility varies inversely with solution viscosity, Joule heating will cause the absolute mobility to increase by a factor of 1.092 relative to that of a sample at 25 ºC without Joule heating. In quantitative studies such as the present one, where comparisons of absolute model and experimental mobilities are made, it is important to account for sources of systematic error that are as large as this.

Table 3.2 compares experimental mobilities, μ_{exp}, "random", μ_{random}, and "restricted", μ_{rest}, model mobilities in different BGE solutions. The relative discrepancy, E_{random}, is defined

$$E_{random} = \frac{\mu_{exp} - \mu_{random}}{\mu_{exp}} \qquad (12)$$

E_{rest} is defined in a similar manner except μ_{rest} replaces μ_{random} in Eq. (12).

Table 3.2: Experimental and model mobilities in different BGE

n	BGE	$\mu_{exp}^{(1)}$	$\mu_{random}^{(1)}$	$\mu_{rest}^{(1)}$	E_{random}	E_{rest}
2	Ia	25.92	24.44	24.94	+0.057	+0.038
"	Ib	15.57	14.79	16.03	+0.050	-0.030
"	IIa	21.91	20.84	21.86	+0.049	+0.002
"	IIb	11.71	11.60	12.63	+0.009	-0.079

“	IIIa	-23.98	-21.96	-22.34	+0.084	+0.068
“	IIIb	-25.35	-21.72	-21.94	+0.143	+0.134
3	Ia	22.21	20.24	22.59	+0.089	-0.017
“	Ib	14.77	13.11	15.35	+0.112	-0.039
“	IIa	19.39	17.97	20.40	+0.073	-0.052
“	IIb	11.43	10.44	12.46	+0.087	-0.090
“	IIIa	-20.66	-17.75	-19.97	+0.141	+0.033
“	IIIb	-20.92	-17.34	-19.46	+0.171	+0.070
4	Ia	19.07	17.52	19.64	+0.081	-0.030
“	Ib	12.90	11.72	13.43	+0.092	-0.041
“	IIa	16.89	15.89	17.89	+0.059	-0.059
“	IIb	10.13	9.53	10.99	+0.059	-0.085
“	IIIa	-18.18	-15.12	-17.11	+0.168	+0.059
“	IIIb	-18.37	-14.66	-16.58	+0.202	+0.097
5	Ia	17.13	15.53	17.47	+0.093	-0.020
“	Ib	11.65	10.57	11.80	+0.093	-0.013
“	IIa	15.22	14.28	15.91	+0.062	-0.045
“	IIb	9.13	8.72	9.65	+0.045	-0.057
“	IIIa	-16.14	-13.24	-15.08	+0.180	+0.066
“	IIIb	-16.23	-12.78	-14.57	+0.213	+0.102
6	Ia	15.52	14.01	15.90	+0.097	-0.024
“	Ib	10.61	9.64	10.83	+0.092	-0.021
“	IIa	13.87	13.01	14.60	+0.062	-0.053

"	IIb	8.38	8.03	8.92	+0.042	-0.064
"	IIIa	-14.07	-11.81	-13.54	+0.161	+0.038
"	IIIb	-14.54	-11.36	-13.02	+0.219	+0.104

(1) – mobilities are in 10^{-9} m^2 V^{-1} sec^{-1}

As discussed in Section 3.1, the "restricted" model fits the experimental data [22] very well at 25 °C in the limit of zero ionic strength irrespective of the type of buffer present [19, 22]. The "restricted" model is quite compact and yields a substantially smaller average hydrodynamic radius relative to the "random" model. In the phosphate buffer, the "restricted" model clearly fits the experimental data better than the "random" model. What this is telling us is phosphate buffer ions, present at finite concentration, do not appear to significantly alter peptide conformation relative to its "zero salt" value. For the citrate ions, the situation is more complicated. In this case the "restricted" model tends to overestimate the absolute mobility while the "random" model tends to underestimate it. An average solution conformation intermediate between the compact "restricted" and more extended "random" models would be consistent with this result. This, in turn, indicates that citrate ions are interacting with the peptide and possibly inducing it to adopt a more extended, random conformation.

The behavior of the oligoglycines in the high *pH* borate buffer shows the greatest discrepancy with both models and is the hardest to understand. Note that the absolute experimental mobilities are substantially higher than either the "random" or "restricted" model values even though the "restricted" model values are closer. Since the "restricted"

model already corresponds to a very compact structure, it is not physically possible that the borate ions are forcing the oligoglycines into an even more compact state. Under the high pH conditions employed in the experiment, the peptide charge itself is expected to be -1 or very close to that value. What then could explain the large $|\mu_{exp}|$ values? One possibility involves complex formation of the anionic peptide with borate anion. This would produce a complex with an effective charge of -2 (assuming complex formation of a single borate anion). This increased absolute charge would result in an increase in $|\mu_{exp}|$ which would be consistent with experiment. Similar borate-anion complex formation has been implicated in the mobility studies of organic anions [63, 64]. In a companion paper, we explore this possibility in greater depth.

Discussion 4

The availability of extensive FSCE [19, 22, 51] and other transport studies [55] of low molecular weight oligoglycines make them ideal systems in which to study peptide solution conformations and also peptide interactions with different BGE solutions. From experimental peptide mobilities in the limit of zero ionic strength and peptide concentration, $\mu^{0\infty}$, coupled with modeling, it is possible to examine how peptide conformation varies with temperature and degree of polymerization [22]. This comparison has shown that although a "random" conformation is adopted by low molecular weight oligoglycines at high temperature, a substantially more compact structure is adopted at low temperature (15 and 25 °C). For $n > 6$, a model α helical conformation is consistent with experiment. For $n = 3$ through 6, compact cyclic structures are proposed. Independent support of this for $n = 3$ to 5 comes from gas phase IR studies [57]. It is also shown that the B model is capable of reproducing reasonably

well the pK_a of the C-terminal of low molecular weight oligoglycines by the charge regulation mechanism [48, 60, 61]. This is necessary in order to predict the charge state of the peptide at low *pH*. At finite concentrations of BGE, it is concluded that different buffer ions interact differently with oligoglycines and may or may not alter the solution conformation of the peptide. In low *pH* phosphate buffers, the "restricted" conformational model that accounts for $\mu^{0\infty}$ accounts for the mobility at finite buffer concentration. In this case, the buffer ions have no apparent influence on peptide conformation. In low *pH* citrate buffers, the oligoglycines appear to adopt an average structure that is intermediate between the "restricted" and "random" models. In high *pH* borate buffers, it is proposed that anion-anion complex formation must be involved in order to account for the large absolute mobilities observed experimentally.

It is worthwhile to discuss other possible explanations of the variation of A (Eqs. (1-3)) with temperature for the oligoglycines. Basically, we have assumed that the variation is due entirely to a conformational change in the peptides. It should be emphasized that A is a hydration radius and Eqs. (1-3) are strictly valid for a continuum solvent that obeys stick hydrodynamic boundary conditions on the particle surface. Even for glycine, measurements of D_T have shown that A increases monotonically from 2.235 to 2.332 Å as T is raised from 1 to 37 °C [65]. If we define,

$$\alpha = \left(\frac{\partial A}{\partial T}\right)_P \qquad (13)$$

then α equals +0.0026 Å/°C for glycine. The temperature dependence of A for glycine can be attributed to some variation in hydration or alternatively a slight expansion in conformation with increasing temperature. In modeling, however, we have not varied the side bead radii of glycine with temperature to account for this although that could be

done. Doing so, however, would not change our conclusions. From the results of reference 15, α = 0.0079, 0.0137, 0.0161, 0.0172, and 0.0193 Å/°C as n increases from 2 to 6. Since these are considerably greater than α for glycine, accounting for the change in hydration or "expansion" of individual glycines with temperature would not alter our conclusions.

A second factor to consider is inter-particle (peptide-peptide) interactions that would be concentration dependent. For peptides of low but finite concentration, c_0, the mobility in the limit of zero ionic strength, μ^{∞}, is related to the translational diffusion constant at finite peptide concentration, D_T, and at zero concentration, D_T^0, by

$$\mu^{\infty} = \frac{eZ}{k_B T} D_T = \frac{eZ}{k_B T} D_T^0 (1 + k_D c_0) = \mu^{0\infty} (1 + k_D c_0) \qquad (14)$$

In Eq. (14), $\mu^{0\infty}$ is the mobility in the limit of zero ionic strength and zero peptide concentration and k_D is a constant related to inter-particle interactions (see, for example, reference 66). Strictly speaking, what is actually reported in references 19 and 22 is μ^{∞} and not $\mu^{0\infty}$. If k_D were both large and strongly temperature dependent, the apparent variation in A with temperature could be due to inter-particle interactions and not a conformational change. For diglycine at 25 °C, $k_D \cong -0.76$ M^{-1} [55]. In reference 22 all oligoglycines are present in the solution, but migrate as discrete bands with each band corresponding to a particular n value. Thus, the local (within a band) concentration should be the concentration of a particular oligoglycine. The solutions in reference 22 contain the equivalent of 20 mM in monomers not considering the peak dilution during the CE experiment. Furthermore, in these solutions the monomers are polymerized to oligomers with n ranging from 2 to over 10. Thus, the (local) concentration of a particular oligoglycine in a particular band will be much smaller than 20 mM. Assuming

c_0 equals 1 mM and we choose k_D appropriate for diglycine, then $k_D c_0$ = -0.00076. Inserting this in Eq. (14) above, it is seen that the contribution of the concentration dependent term to μ^∞ and hence A is below 0.1 %. On this basis, it seems unlikely that inter-particle interactions are responsible for the observed variation in A with temperature. As an extension to modeling previously conducted for translational diffusion constants of macromolecules [39], the rotational diffusion constant was now modeled. Understanding the rotational diffusion constant allows for a complete understanding of how molecules move thru gels and solutions. As explained above a molecule is rotated about an x, y, and z-axis. Once an electric force is exerted upon a molecule it will rotate in a medium as well as interact with its surroundings. Using an effective medium model, the rotational motion can be modeled. Solving for the rotational force as a tensor and a 3 by 3 matrix the Eigen values for the force can be solved and related to the rotational relaxation time. Using experimental data [36], the rotational relaxation time for DNA of 622 base pairs was identified and theoretically modeled. The model has the ability to within a good approximation model DNA in a gel of concentrations less than 0.010 g/ml. For a strand of DNA with 622 base pairs, the model was successful and identified a persistent length of 65nm, and a gel fibre radius for agarose gel to be 2.5 nm. At higher concentrations of gel however, the model overestimates the relaxation time, which a proposed explanation is the direct interaction between the gel and the DNA. This factor has not yet been explored.

Conclusion 5

In conclusion, we affirm that fundamental electro hydrodynamic modeling of electrophoresis is capable of accounting for the FSCE behavior of many systems. The

present study of oligoglycines and previous studies of peptides and related systems bear this out [9, 10, 38, 41-50]. The motivation for this effort is both to aid in designing conditions optimal for a particular separation and also using FSCE as a method for structure elucidation. In the subsequent paper, this is extended to peptides in certain BGE solutions where complex formation between peptide and components in the BGE may occur.

References

[1] Allison, S.A., Pei, H., Twahir, U., Baek, S., Lee, M., Nguyen, V., Garcia, J., J. Phys. Chem. B. 2009, 113, 13576-13584

[2] Allison, S.A., Pei, H., Twahir, U., Baek, S., Lee, M., Nguyen, V., Brown, J., Wu, H., Electrophoresis 2010, 31, 920-932

[3] Allison, S.A., Twahir, U., Wu, H., Pei, H., Journal of Colloid and Interface Science, 352, (2010) 1-10

[4] Stellwagen C. N., Colloid and Surfaces, Physiochemical and Engineering Aspects 209, 2002, 107-122

[5] Brinkman, H.C.; Appl. Sci. Res. A, 1947, 1, 27-34

[6] Debye, P.; Bueche, A.M., J. Chem. Phys., 1948, 16, 573-579

[7] Allison, S., Pei, H., Twahir, U., Wu, H., Cottet, H., J. Sep. Sci. 2010, 33, 2430- 2438

[8] Adamson, N.J., Reynolds, E.C., *J. Chromatogr. B*, 1997, *699*, 133-147.

[9] Messana, I., Rossetti, D.V., Cassiano, L., Misiti, F., Giardina, B., Castagnola, M., *J. Chromatogr. B*, 1997, *699*, 149-171.

[10] Cifuentes, A., Poppe, H., *Electrophoresis*, 1997, *18*, 2362-2376.

[11] Kasicka, V., *Electrophoresis*, 2006, *27*, 142-175.

[12] Kasicka, V., *Electrophoresis*, 2008, *29*, 179-206.

[13] Foret, F., *Electrophoresis*, 2009, *30*, 534-539.

[14] Kasicka, V., *Electrophoresis*, 2010, *31*, 1-25.

[15] Grossman, P.D., Colburn, H.H., Lauer, H.H., *Anal. Biochem.*, 1989, *179*, 28-33.

[16] Rickard, E.C., Strohl, M.M., Nielsen, R.G., *Anal. Biochem.*, 1991, *197*, 197-207.

[17] Cifuentes, A., Poppe, H., *J. Chromatogr. A*, 1994, *680*, 321-341.

[18] Castagnola, M., Cassiano, L., Messana, I., Nocca, G., Rabino, R., Rossetti, D.V., Giardina, B., *J. Chromatogr. B*, 1994, *656*, 87-97.

[19] Survay, M.A., Goodall, D.M., Wren, S.A.C., Rowe, R.C., *J. Chromatogr. A*, 1996, *741*, 99-113.

[20] Janini, G.M., Metral, C.J., Issaq, H.J., Muschik, G.M., *J. Chromatogr. A*, 1999, *848*, 417-433.

[21] Solinova, V., Kasicka, V., Koval, D., Hlavacek, J., *Electrophoresis* 2004, 25, 2299-2308.

[22] R. Plasson, H. Cottet, *Anal. Chem.*, 2005, *77*, 6047-6054.

[23] Offord, R.E, *Nature (London)*, 1966, *211*, 591-593.

[24] Sitaram, B.R., Keah, H.H., Hearn, M.T.W., *J. Chromatogr. A.*, 1999, *857*, 263-273.

[25] Janini, G.M., Metral, C.J., Issaq, H.J., *J. Chromatogr. A.*, 2001, *924*, 291-306.

[26] Jalali-Heravi, M., Shen, Y., Hassanisadi, M., Khaledi, M.G., *J. Chromatogr. A.*, 2005, *1096*, 58-68.

[27] Mittermayr, S., Chovan, T., Guttman, A., *Electrophoresis*, 2009, *30*, 890-896.

[28] Debye, P., Huckel, E., *Physik. Zeitschr.*, 1924, *25*, 49-52.

[29] Huckel, E., *Physik. Zeitschr.*, 1924, *25*, 204-210.

[30] Henry, D.C., *Proc. R. Soc. London*, 1931, *133A*, 106-140.

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

[32] Overbeek, J. Th. G., *Kolloid Beih.*, 1943, *54*, 287-364.

[33] Booth, F., *Proc. R. Soc. London*, 1950, *203A*, 514-551.

[34] Pitts, E., *Proc. R. Soc. London*, 1953, *217A*, 43-70.

[35] Wiersema, P.H., Loeb, A.L., Overbeek, J. Th. G., *J. Colloid Interface Sci.*, 1966, *22*, 78-99.

[36] O'Brien, R.W., White, L.R., *J. Chem Soc. Faraday Trans. 2*, 1978, *74*, 1607-1626.

[37] Stigter, D., *J. Phys. Chem.*, 1978, *82*, 1417-1423; 1424-1429.

[38] Allison, S.A., Chen, C., Stigter, D., *Biophys. J.*, 2001, *81*, 2558-2568.

[39] Yoon, B.J., Kim. S., *J. Colloid Interface Sci.*, 1989, *128*, 275-288.

[40] Allison, S.A., *J. Coll. Interface Sci.*, 2005, *282*, 231-237.

[41] Kim, J.Y., Ahn, S.H., Kang, S.T., Yoon, B.J., *J. Colloid Interface Sci.*, 2006, *299*, 486-492.

[42] Allison, S.A., *Macromolecules*, 1996, *29*, 7391-7401.

[43] Mazur, S., Chen, C., Allison, S.A., *J. Phys. Chem. B*, 2001, *105*, 1100-1108.

[44] Allison, S.A., Carbeck, J.D., Chen, C., Burkes, F., *J. Phys. Chem. B*, 2004, *108*, 4516-4524.

[45] Allison, S.A., Xin, Y., Mitchell, H., in: Pandali, S.G. (Ed.), *Recent Developments in Macromolecules*, Research Signpost, Kerali, India 2005, *8*, 25-46.

[46] Xin, Y., Mitchell, H., Cameron, H., Allison, S.A., *J. Phys. Chem. B*, 2006, *110*, 1038-1045.

[47] Xin, Y., Hess, R., Ho, N., Allison, S.A., *J. Phys. Chem. B*, 2006, *110*, 25033-25044.

[48] Pei, H., Xin, Y., Allison, S.A., *J. Sep. Sci.*, 2008, *31*, 555-564.

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

[50] Grass, K., Bohme, U., Scheler, U., Cottet, H., Holm, C., *Phys. Rev. Lett.*, 2008, *100*, 96104-1 – 96104-4.

[51] Plasson, R., Cottet, H., *Anal. Chem.*, 2006, *78*, 5394-5402.

[52] Plasson, R., Cottet, H., *Anal. Chem.*, 2007, *79*, 3020.

[53] Sasisekharan, V., In *Collagen*, Ramanathan, N. (Ed.), Interscience, New York, 1962, p. 39.

[54] Germann, M.W., Turner, T., Allison, S.A., *J. Phys. Chem. A*, 2007, *111*, 1452-1455.

[55] Pei, H., Germann, M.W., Allison, S.A., *J. Phys. Chem. B*, 2009, *113*, 9326-9329.

[56] Flory, P.J., *Statistical Mechanics of Chain Molecules*, J. Wiley, New York, 1969, Chapter 7.

[57] Wu, R., McMahon, T.B., *J. Phys. Chem. B*, 2009, *113*, 8767-8775.

[58] Butler, J.N., *Ionic Equilibrium,* J. Wiley & Sons, New York, 1998, p. 46.

[59] Gluck, S.J., Cleveland, J.A., *J. Chromatogr. A*, 1994, *680*, 43-48.

[60] Lee, L.K., Fitch, C.A., Garcia-Moreno, B.E., *Protein Sci.*, 2002, *11*, 1004-1016.

[61] Sharma, U., Negin, R.S., Carbeck, J.D., *J. Phys. Chem. B*, 2003, *107*, 4653-4666.

[62] Solinova, V., Kasicka, V., Sazelova, P., Barth, T., Miksik, I., *J. Chromatogr. A*, 2007, *1155*, 146-153.

[63] Li, D., Fu, S., Lucy, C.A., *Anal. Chem.*, 1999, *71*, 687-699.

[64] Allison, S.A., Pei, H., Baek, S., Brown, J., Lee, M.Y., Twahir, U.T., Wu, H., *Electrophoresis*, 2010 (in press).

[65] Longsworth, L.G., *J. Phys. Chem.*, 1954, *58*, 770-773.

[66] Tsunashima, Y., Nemoto, N., *Macromolecules*, 1984, *17*, 2931-2933.

Appendix

The basic methodology for the bead model (B model) has been described in detail previously [45-48] and shall not be reproduced here. However, the methodology continues to be improved upon and the purpose of this appendix is to outline an improvement that simplifies the overall procedure. Consider a model bead array in a well defined configuration and let $\underline{\underline{\mu}}$ denote the electrophoretic mobility *tensor* (a 3 by 3 matrix) which we wish to obtain. We begin with Eq. (A34) of reference 41

$$(\underline{\underline{\chi}}_K - \frac{1}{2}\sum_{J=1}^{N} \underline{\underline{W}}^{(2)}_{KJ} \cdot \underline{\underline{d}}_J) \cdot \underline{\underline{\mu}} = \frac{e\kappa}{6\pi\eta_0} \sum_{J=1}^{N} [\underline{\underline{W}}^{(1)}_{JK} \cdot \underline{\underline{g}}_J - \frac{1}{2} h_J \underline{\underline{W}}^{(2)}_{JK}] \qquad (A1)$$

In Eq. (A1), e is the protonic charge, κ is the Debye-Huckel screening parameter defined by Eq. (3), η_0 is the solvent viscosity, $\underline{\underline{g}}_J$ is an *unknown* reduced force tensor for bead J, and N is the number of beads. All other terms in Eq. (A1) are *known* configuration dependent quantities defined previously [46, 48]. We also have the force balance condition,

$$Z \underline{\underline{I}} = \sum_{J=1}^{N} \underline{\underline{g}}_J \qquad (A2)$$

where Z is the effective charge of the bead array and $\underline{\underline{I}}$ is the 3 by 3 identity matrix. Define the 3 by 3 matrix, $\underline{\underline{C}}_K$, and the $3N$ by $3N$ supermatrix, $\underline{\underline{Y}}$,

$$\underline{\underline{C}}_K = \underline{\underline{\chi}}_K - \frac{1}{2}\sum_{J=1}^{N} \underline{\underline{W}}^{(2)}_{KJ} \cdot \underline{\underline{d}}_J \qquad (A3)$$

$$\underline{\underline{Y}} = \begin{pmatrix} \underline{\underline{W}}^{(1)}_{11} & \underline{\underline{W}}^{(1)}_{21} & \cdots & \underline{\underline{W}}^{(1)}_{N1} \\ \underline{\underline{W}}^{(1)}_{12} & \underline{\underline{W}}^{(1)}_{22} & \cdots & \underline{\underline{W}}^{(1)}_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{W}}^{(1)}_{1N} & \underline{\underline{W}}^{(1)}_{2N} & \cdots & \underline{\underline{W}}^{(1)}_{NN} \end{pmatrix} \qquad (A4)$$

In Eq. (A4), $\underline{\underline{W}}^{(1)}_{KJ}$ is a known configuration dependent 3 by 3 matrix defined by Eq. (A.19) of reference [48]. Eq. (A4) is invertible and let $\underline{\underline{Y}}^{-1}$ denote the 3*N* by 3*N* inverse matrix. Define the N^2 3 by 3 matrices, $\underline{\underline{H}}_{JK}$, in terms of $\underline{\underline{Y}}^{-1}$ by

$$\underline{\underline{Y}}^{-1} = \begin{pmatrix} \underline{\underline{H}}_{11} & \underline{\underline{H}}_{21} & \cdots & \underline{\underline{H}}_{1N} \\ \underline{\underline{H}}_{21} & \underline{\underline{H}}_{22} & \cdots & \underline{\underline{H}}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{H}}_{N1} & \underline{\underline{H}}_{N2} & \cdots & \underline{\underline{H}}_{NN} \end{pmatrix} \qquad (A5)$$

Once $\underline{\underline{Y}}^{-1}$ is determined, the $\underline{\underline{H}}_{JK}$ matrices can be constructed. Also define the *N* 3 by 3 matrices, $\underline{\underline{D}}_{L}$,

$$\underline{\underline{D}}_{L} = \sum_{J,K=1}^{N} h_J \underline{\underline{H}}_{LK} \cdot \underline{\underline{W}}^{(2)}_{JK} \qquad (A6)$$

Multiply both sides of Eq. (A1) by $\underline{\underline{H}}_{LK}$ and sum over *K*. Using the definitions of the previous paragraph, it is straightforward to show,

$$\sum_{K=1}^{N} \underline{\underline{H}}_{LK} \cdot \underline{\underline{C}}_{K} \cdot \underline{\underline{\mu}} = \frac{e\kappa}{6\pi\eta_0} \left(\underline{\underline{g}}_{L} - \frac{1}{2} \underline{\underline{D}}_{L} \right) \qquad (A7)$$

Next, sum both sides of Eq. (A7) over *L* and use Eq. (A2) to eliminate the $\underline{\underline{g}}_{L}$ terms. With the additional definition of the 3 by 3 matrix

$$\underline{\underline{P}} = \sum_{L,K=1}^{N} \underline{\underline{H}}_{LK} \cdot \underline{\underline{C}}_{K} \qquad (A8)$$

which is invertible, we finally obtain

$$\underline{\underline{\mu}} = \frac{e\kappa}{6\pi\eta_0} \left[Z \underline{\underline{P}}^{-1} - \frac{1}{2} \sum_{L=1}^{N} \underline{\underline{P}}^{-1} \cdot \underline{\underline{D}}_{L} \right] \qquad (A9)$$

Assuming all orientations of the bead array are sampled with equal probability, the orientationally averaged mobility of a particular configuration can be written

$$\mu = \frac{1}{3}\left(\mu_{xx} + \mu_{yy} + \mu_{zz}\right) \qquad (A10)$$

In the present procedure, the unnecessary step of determining the $\underline{\underline{g}}_K$ terms is avoided.

There is also a correction for the mobility due to "ion relaxation" [31-40, 42-44, 47-49, 64]. For weakly charged peptides, this correction is small. Details on how it is handled in the present work can be found in the companion paper by Allison et al., in this issue.