ELSEVIER

Journal of Chromatography B, 699 (1997) 133–147

JOURNAL OF CHROMATOGRAPHY B

Review

Rules relating electrophoretic mobility, charge and molecular size of peptides and proteins

N.J. Adamson, E.C. Reynolds*

Biochemistry and Molecular Biology Unit, School of Dental Science, The University of Melbourne, 711 Elizabeth Street, Melbourne, Vic. 3000, Australia

Abstract

The absence of supporting media in free solution high-performance capillary electrophoresis (HPCE) makes it an ideal system for the study of the relationship between electrophoretic mobility (μ_{em}) and the molecular size and charge of proteins and peptides. In this review, the theory of electrophoresis, developed for rigid, insulating, spherical particles, is modified to develop models for the electrophoretic behaviour of proteins and peptides. For a given set of experimental conditions, μ_{em} of a protein/peptide is proportional to its charge (q) and is inversely proportional to its Stoke's radius (r). Furthermore, μ_{em} is most sensitive to changes in q and, as a consequence, the reliability of equations relating μ_{em} to protein/peptide q and r is dependent upon the accurate calculation or determination of q. For convenience, q and r of proteins and peptides are generally expressed in terms of calculated valence (Z_c) and molecular mass (M), respectively, both of which can be determined from the amino acid sequence of the protein/peptide. However, the calculation of q using Z_c is made more complex by the effects of electrostatic charge suppression, such that Z_c is an overestimation of actual charge. Charge suppression becomes increasingly significant as the protein/peptide charge increases, such that, for peptides, the relationship between q and Z_c can be approximated by a logarithmic function. The μ_{em} for peptides, therefore, can be approximated by the equation: $\mu_{em}=\ln(Z_c+1)/K M^s$ where s varies between 1/3 and 2/3, and K is a constant that is valid for a particular set of experimental conditions. The rather simplistic compensation for charge suppression in this equation is inadequate for proteins where the magnitude of charge suppression is greater and the mechanisms are more complex. For proteins, the relationship suggested for the prediction of μ_{em} from Z_c and M is: $\mu_{em}=Z_c/KF_z M^s$ where s again varies between 1/3 and 2/3 and F_z is a pH-independent proportionality factor defined as the quotient, Z_c/Z_a, with Z_a being actual protein valence. The factor F_z can be determined empirically, however, it is valid only for the particular set of experimental conditions under which it is determined. For peptides, the mass exponent, s, approaches 1/3 when the peptides have high charge densities and open structures. However, s approaches 1/2 for peptides with lower charge densities that are capable of more randomized motion during electrophoresis. Finally, s approaches 2/3 for proteins, suggesting that the frictional forces acting on a protein undergoing electrophoretic motion are proportional to the surface area of these larger, more rigid, structures. In conclusion, the development of relationships between μ_{em}, M and Z_c for peptides and proteins offers a powerful tool, not only for predicting electrophoretic mobility, but also for optimising HPCE separations, studying structural modifications (e.g. phosphorylation, glycosylation, deamidation, etc.), and for the investigation of surface charge characteristics and conformation. © 1997 Elsevier Science B.V.

Keywords: Reviews; Peptides; Proteins

*Corresponding author.

PII S0378-4347(97)00202-8

Contents

1. Introduction

Electrophoresis has proven to be a very effective technique for the purification of proteins and peptides as well as for the measurement of their physicochemical properties, such as molecular mass, isoelectric point and electrophoretic mobility (μ_{em}) [1]. In recent years, the development of free solution electrophoresis in capillary tubing has facilitated the use of very high voltages for the separation of biomolecules with greater than 10^6 theoretical plates [2]. High performance capillary electrophoresis (HPCE) in free solution has developed into a technique that is simple operationally, yet results in extremely efficient, rapid separations of biomolecules.

The absence of supporting media in free solution HPCE makes it an ideal system for the study of the relationship between μ_{em} and the molecular size and charge of proteins and peptides. Since the advent of HPCE, several articles (eg. [3–13]) have been published reporting the use of HPCE to characterise the physicochemical properties of proteins and peptides.

The relationship between the electrophoretic mobility of a particle and its charge and size has been of considerable interest in colloid science and has been the subject of a number of empirical and theoretical treatments [14]. In this review, the theory of electrophoresis, developed for rigid, insulating, spherical particles, is modified to develop models for the electrophoretic behaviour of proteins and peptides. The practical limitations of the theory are discussed and the validity of the theoretical models for proteins and peptides are assessed in context with existing empirical data generated using a variety of electrophoretic techniques,such as HPCE [3–13], paper electrophoresis [15] and Doppler velocimetry [16].

2. Theory

2.1. The electric double layer

A charged particle in solution attracts ions of opposite charge, which results in the accumulation of counter-ions near the particle surface and the formation of an ionic atmosphere [17]. Within this ionic atmosphere, there are two distinct regions of charge. Directly adjacent to the surface of the particle, there is a relatively immobile layer containing ions, which can include water molecules in aqueous solutions. The boundary encompassing this immobilised layer is referred to as the surface of shear and the electric potential at this surface, relative to its value in the bulk solution, is known as the zeta(ζ)-potential or electrokinetic potential [14]. It is this surface which is the major factor in determining the electrophoretic mobility of the particle. The charge within the surface of shear attracts counter-ions from the bulk solution and results in the formation of an oppositely charged outer ionic atmosphere. The charged surface of shear and the outer atmosphere are called the electric double layer, of which the thickness is $1/\kappa$, where κ is defined by the Debye–Hückel theory of ionic solutions [7] as:

$$\kappa = 2^{1/2} Ne(\varepsilon_o \varepsilon RT)^{-1/2} I^{1/2} \quad (2.1)$$

where e is the electronic charge, N is Avogadro's number, ε_0 is the permittivity of free space, ε is the dielectric constant of the medium, R is the gas constant, T is the absolute temperature and I is the ionic strength. I is further defined as $1/2\Sigma C_i z_i^2$ where C_i is the ion concentration and z_i is the valence of fluid ions. The term $1/\kappa$ is also known as the Debye length or, as the formation of the electric double layer results in shielding of the particle, it can be known as the shielding length.

2.2. Electrophoretic mobility

2.2.1. Definition and empirical determination

For a charged particle migrating in an electrical field of strength E (V/cm) with steady-state velocity ν_{em} (cm/s), the electrophoretic mobility (μ_{em}) of the particle is defined as the velocity per unit electrical field strength [14]:

$$\mu_{em} = \nu_{em}/E \quad (2.2)$$

In most electrophoresis systems, the measurement of μ_{em} is made more complex by electroosmosis [1]. This phenomenon occurs when a stationary, charged surface is surrounded by an electrolyte and is placed in an electric field. Counter-ions in the electrolyte accumulate at the surface, forming an electrical double layer [14]. Under the influence of the electric field, these counter-ions migrate towards an electrode, resulting in bulk flow of the electrolyte. This electroosmotic flow can be characterised in the same way as μ_{em} and is defined [2] as:

$$\mu_{eo} = \nu_{eo}/E \quad (2.3)$$

where μ_{eo} is the coefficient of electroosmotic flow and ν_{eo} is the electroosmotic flow velocity. The apparent electrophoretic mobility of charged particles measured empirically (μ_m) in many electrophoresis systems therefore comprises the effective electrophoretic mobility (μ_{em}) component and the electroosmotic component, such that:

$$\mu_m = \mu_{em} + \mu_{eo} \quad (2.4)$$

The electroosmotic component of μ_m can be determined simply by measuring the migration velocity of an uncharged particle (under identical conditions as those for the particle in question), which is substituted for ν_{eo} in Eq. (2.3), along with the known value for E. Eq. (2.4) can then be used to determine μ_{em}. As HPCE has been used extensively for the measurement of μ_{em} of proteins and peptides, the equations pertinent to this technique will be discussed in more detail.

Measurements made by HPCE result in the determination of μ_m, defined in Eq. (2.4). For a peptide migrating with velocity ν_m (cm/s) through a capillary of total length L_t (cm), over which a voltage of magnitude V (volts) is applied, and where the peptide is detected at a distance L_d (cm) from the sample introduction end of the capillary after time t_m (s), μ_m ($cm^2/V.s$) is given by the equation [11]:

$$\mu_m = \nu_m/E = (L_d/t_m)/(V/L_t) \quad (2.5)$$

Similarly, for a neutral molecule migrating with velocity ν_{eo} under identical conditions, μ_{eo} is given by the equation:

$$\mu_{eo} = \nu_{eo}/E = (L_d/t_{eo})/(V/L_t) \quad (2.6)$$

where t_{eo} is the migration time for a neutral molecule whose motion through the capillary is due solely to electroendosmosis. Determination of μ_m and μ_{eo} thus allows Eq. (2.4) to be solved for μ_{em}.

2.2.2. Theory for a rigid, spherical, particle

Much of the theory of electrophoresis was developed for rigid, spherical particles [14,18,19] and has been extrapolated to develop models for the electrophoretic behaviour of proteins and peptides.

In the case of a non-conducting, rigid sphere undergoing steady-state electrophoretic motion, there will be four different forces acting on the sphere at any given time [14]. The first force (F_E) is exerted by the electric field and is defined as:

$$F_E = qE \quad (2.7)$$

where q is the charge on the particle. The second force (F_F) is the retarding force due to friction imparted by the solution on the sphere which is defined as:

$$F_F = -f_c \nu_{em} \quad (2.8)$$

where f_c is the coefficient of friction of the particle. For a rigid particle that is large compared to the molecules surrounding it, f_c is given by Stoke's law:

$$f_c = 6\pi\eta r \tag{2.9}$$

where η is the viscosity of the solution surrounding the sphere and r is the Stoke's radius of the sphere.

The remaining forces acting on the sphere are caused by effects of the electric field on the electric double layer [14]. The ionic atmosphere surrounding the sphere is of the opposite charge and, so, migrates in the opposite direction in the electric field. This electrophoretic motion of the ionic atmosphere induces a retarding force (F_R), which acts upon the sphere and reduces its electrophoretic mobility.

In the absence of an electric field, the ionic atmosphere is a spherical, symmetrical haze around the particle [17]. When an electric field is applied, however, the particle migrates away from the centre of its ionic atmosphere. The counter-ions cannot adjust to the moving sphere, resulting in incomplete formation of the ionic atmosphere in front of the sphere and behind it there is incomplete decay. As a consequence, the centre of the ionic atmosphere constantly lags behind the centre of the sphere, inducing, in most cases, a retarding electrical force (F_{RE}). This phenomenon is called the relaxation effect [14]. For trivalent co-ions, the relaxation effect is positive in the region $0.01 < \kappa r < 1$. At extremely high field strengths (100 000 V/cm), the moving ion can actually outrun its ionic atmosphere, thus reducing shielding, which results in increased electrophoretic mobility of the ion. This is known as the first Wien effect [20].

In general, the forces F_R and F_{RE} are complex functions of the ζ-potential, the dimensions of the sphere, and of the charges, concentrations and the mobilities of the counter-ions in solution [14]. At steady state, the sum of all forces acting on the sphere is zero:

$$F_E + F_F + F_R + F_{RE} = 0 \tag{2.10}$$

The oldest relationship between the electrophoretic mobility and the properties of the double layer is known as the Helmholtz-Smoluchowski equation [21], derived from first principles as:

$$\mu_{em} = \varepsilon\zeta/4\pi\eta \tag{2.11}$$

The equation is reported to be valid for rigid, non-conducting particles of any shape, provided that the value of the viscosity (η) and dielectric constant (ε) in the bulk solution can be applied to the solution in the double layer and the thickness of the double layer is small compared with the radius of curvature of the particle [21]. Hückel [22] then provided a detailed calculation of F_R acting on a spherical particle with the result that the electrophoretic retardation is given by:

$$F_R = (\varepsilon\zeta r - q)E \tag{2.12}$$

Substituting Eqs. (2.7)–(2.9), (2.12) into Eq. (2.11) gives:

$$\mu_{em} = v_{em}/E = (\varepsilon\zeta r + F_{RE})/6\pi\eta r \tag{2.13}$$

Several authors have attempted to address the problem of the relaxation effect for colloidal particles, however, due to the mathematical complexity, the application of the theory is limited [14]. Overbeek [23] and Booth [24,25] demonstrated that the solution of complex mathematical equations could at best be approximated by a power series expansion for which only a few terms could be calculated, however, a few important conclusions could be made. It was concluded that the relaxation effect was negligible in the following circumstances: (i) $\zeta \ll 25$ mV; (ii) $\kappa r \ll 1$ and (iii) $\kappa r \gg 1$. Under these circumstances, Eq. (2.13) reduces to:

$$\mu_{em} = \varepsilon\zeta/6\pi\eta \tag{2.14}$$

The integer difference between the denominators of Eqs. (2.11) and (2.14) was resolved by Henry [26] who analysed the assumptions underlying the Helmholtz-Smoluchowski equation [21] and Hückel's [22] derivation of Eq. (2.10). In the derivation, von Smoluchowski [21] assumed that the direction of the electric field was parallel to the surface of the particle, i.e., when the thickness of the double layer is thin compared with the radius of curvature of the sphere ($\kappa r \gg 1$). Hückel [22] assumed that everywhere in the electric double layer, the lines of the electric field run from the anode to the cathode, i.e., when the thickness of the double layer is large compared with the radius of curvature of the sphere ($\kappa r \ll 1$) or when the conductivity of the particle and solution are equal, which is highly unlikely in practice. Henry [26] recalculated F_R, taking into account the deformation of the electric field and arrived at the equation:

$$\mu_{em} = (\varepsilon\zeta/6\pi\eta) \times X(\kappa r) \quad (2.15)$$

where $X(\kappa r)$ is Henry's function which, mathematically, is a sigmoidal function varying from 1.0 to 1.5 as the quantity κr varies from zero (low ionic strength) to infinity (high ionic strength). Hence, it can be seen that as $\kappa r \rightarrow \infty$, Eq. (2.5) approaches the result of Helmholtz-Smoluchowski [21], defined by Eq. (2.14) and, as $\kappa r \rightarrow 0$, Eq. (2.15) approaches the result of Hückel [22] defined by Eq. (2.13). According to Henry, Eq. (2.14) is valid for particles whose radius is >300 times the thickness of the double layer, whilst Eq. (2.13) is valid for particles whose radius is ≤0.5 times the thickness of the double layer [26]. In the derivation of Eq. (2.15), Henry [26] ignored the contribution of the relaxation effect.

In its present form, Eq. (2.15) is difficult to apply practically, due to the presence of the ζ-potential term. For low values of ζ and κr, the charge, q, of a particle can be approximated [14] by:

$$q = \varepsilon\zeta r(1 + \kappa r) \quad (2.16)$$

therefore, substituting into Eq. (2.15) gives:

$$\mu_{em} = (q/6\pi\eta r) \times X(\kappa r)/(1 + \kappa r) \quad (2.17)$$

For proteins and peptides, the charge is usually low and quite often electrophoresis is carried out at relatively high ionic strength, hence, the ζ-potential can be small and the relaxation effect can be neglected [14]. However, the application of Eq. (2.17) for predicting the electrophoretic mobility of peptides and proteins is limited, as it relies upon the calculation of the actual charge (q) and Stoke's radius of the protein/peptide. The actual charge is most accurately determined from membrane potential measurements [27], whilst the Stoke's radius can be determined from diffusion or sedimentation studies [28]. Tiselius and Svensson [28] demonstrated that the empirically determined electrophoretic mobility of egg albumin at various ionic strengths fitted well with calculated mobility using Eq. (2.17) and measurements of charge and radius based on membrane potentials and diffusion. These methods are, however, tedious, and, as both properties are dependent upon the nature of the solvent, must be determined under identical conditions to those used for electrophoresis. For these reasons, several authors [3–13,15,16] have attempted to relate the electrophoretic mobility of peptides and proteins to their physicochemical properties such as molecular mass (M), number of amino acid residues (n) and calculated charge (q_c) or valence (Z_c). Although, more approximate, these relationships are more practical and have broader applicability, as all of these properties can be determined from the amino acid sequence of the peptide or protein.

2.2.3. Application of theory to the electrophoresis of proteins and peptides

Several assumptions have been made in the development of a rigid, spherical particle model for electrophoresis [14] and these assumptions should be considered when applying this model to the electrophoresis of peptides and proteins. Firstly, the model is based on the particle being a non-conducting, rigid, sphere at infinite dilution. Electrophoresis of proteins and peptides is usually conducted at low analyte concentrations and, so, interactions between the solute molecules can be neglected. Furthermore, the conductivity of peptides and proteins would be low compared to the solution in which electrophoresis is performed and, so, the assumptions concerning the shape of the electric field in the electric double layer made for $\kappa r \ll 1$ are reasonable. The spherical model has been adhered to in the investigation of the mobility of peptides [9]. This can be attributed to the assumption that the random motion of these molecules could result in them behaving like spheres, provided that there was no net overall orientation in the electric field, and that the electric double layer may smooth out irregular shapes. Random motion is reasonable at the moderate field strengths (10 V/cm) used in conventional electrophoresis, however, field strengths of up to 400 V/cm are achieved in HPCE, which can impart orientation restraints on rod-like molecules [29].

For proteins, Abramson et al. [30] report a further improvement on Eq. (2.17), which accounts for protein asymmetry by a tabulated function f/f_o:

$$\mu_{em} = (q/6\pi\eta r) \times X(\kappa r)(1 + \kappa r_i)/(1 + \kappa r + \kappa r_i)(f/f_o) \quad (2.18)$$

where r_i is the average radius of ions in the ionic atmosphere. Grossman and Soane [29], with their

work on the rod-shaped tobacco mosaic virus (TMV), have demonstrated the use of shape factors to account for the orientation of non-spherical particles in an electric field and the effect on the frictional force, however, such details on the structure of peptides and proteins, particularly under the influence of an electric field, are often unknown. Several workers have suggested that when determining the electrophoretic mobility of particles with rough or irregular surfaces, the radius of the perturbances on the surface should be considered rather than the overall radius of the particle [14]. Abramson [31] demonstrated that the mobility of certain proteins does not change when they are adsorbed onto a carrier particle, even though the radius of the kinetic unit increases appreciably. This suggests that the proteins do not change shape considerably on adsorption, and that the radius of the adsorbed protein is the dominant shape factor. However, Overbeek and Wiersema [14] suggest that there are effects of the large particles on the overall friction and that overlap of the electric double layers on the individual perturbances would lead to an increase in the effective electrophoretic radius with decreasing ionic strength.

The assumption of rigidity is most valid for proteins, due to the presence of secondary and tertiary structure. Peptides are more flexible structures and, as previously discussed, this random motion might smooth out any irregular shapes. However, whilst the average shape over time may quite closely approximate a sphere, the study on TMV [29] demonstrates the relationship between the friction factor and the orientation of the rod-shaped viral particle (i.e. μ_{em} vs. E) to be non-linear. This would result in an underestimation of μ_{em} by the model for peptides and proteins with rod-shaped structures, unless a preferred orientation were used, as described by Abramson [30] who assumed that one-third of the particles oriented along the direction of the field and two-thirds oriented perpendicular to the field.

In general, the rigid, spherical model discussed previously has been shown to have broader applicability than would have been first thought, considering the assumptions made in its development [14]. Many of the problems related to the assumptions have been shown to be negligible (Brownian motion, non-linear terms), avoidable experimentally (non-univalent electrolytes) or could be incorporated into the model (non-spherical particles). The ability to neglect the relaxation effect under certain conditions reduces the complexity of the equation with respect to non-spherical particles, Brownian motion, variations in the dielectric constant of the particle and mixtures of electrolytes [14].

3. Models for proteins and peptides

Offord [15] reported an early attempt to simplify the fundamental model for the electrophoresis of a spherical, rigid, insulating particle, discussed in Section 2.2.2, for application to a range of small peptides. In this study, Offord [15] related electrophoretic mobility to peptide molecular mass by assuming that during electrophoresis the peptide approximates a sphere of constant density, such that its radius would be proportional to $V_s^{1/3}$ and, therefore, would be proportional to $M^{1/3}$, where the molecular mass (M) of the peptide is proportional to the volume of the sphere (V_s) the peptide approximates. According to Stoke's law, the frictional force experienced by the peptide would be proportional to r and, therefore, should be proportional to $M^{1/3}$. Offord [15] proposed that if the frictional force experienced by a peptide undergoing electrophoretic motion was due to its oppositely charged ionic atmosphere causing a backflow of solvent in the immediate vicinity of the migrating molecule, then the retarding force could be considered as a result of shear across a small element of liquid close to the molecule. This retarding force would then be some function of the surface area of the molecule (i.e. r^2) and, therefore, would be proportional to $M^{2/3}$. Alternatively, Grossman et al. [6], in the development of their empirical relationship between electrophoretic mobility and peptide charge and size, chose to treat the peptide molecule as a classical polymer in solution. From studies of synthetic polymers [32], it has been demonstrated that the radius of gyration is proportional to the square root of the number of polymer units multiplied by the length of a single unit. In this instance, the frictional force exerted on the peptide would be proportional to $M^{1/2}$. There-

fore, in summary, we have the following models for the dependence of electrophoretic mobility of peptides and proteins in solution on their charge and molecular mass:

Stoke's law: $\mu_{em} \propto q/M^{1/3}$ (3.1a)

Classical polymer: $\mu_{em} \propto q/M^{1/2}$ (3.1b)

Offord's surface area: $\mu_{em} \propto q/M^{2/3}$ (3.1c)

Several workers have applied these models to the electrophoresis of peptides and proteins with somewhat mixed results. As mentioned previously, Offord [15] performed a plot of log μ_{em} versus log M for over 100 peptides (150–3500 rel. mol. mass units) at pH 1.9 and 6.5 (excluding Cys-containing peptides at pH 1.9 and His- and Cys-containing peptides at pH 6.5), which resulted in a series of lines of slope $-2/3$, corresponding to groups of peptides with similar calculated valences ($Z_c = 1, 2, 3, 4$). Nyberg et al. [3] extended Offord's [15] relationship to include enzymic degradation products of Substance P (SP). The chain-length of the peptides studied ranged from two to eleven amino acids and their migration times relative to that of SP exhibited a linear relationship with $M^{2/3}/Z_c$. The work of Deyl and Rohlieck [4] further supported Eq. (3.1c). In this study, the relative migration times of peptides from cyanogen bromide cleavage of collagen types I and III were plotted against $M^{2/3}/Z_c$, yielding a linear relationship. The peptides, analysed in 2.5 mM sodium tetraborate, pH 10.5, varied in molecular mass from 13 000–25 100 and contained 18 to 30 acidic residues. Frenz et al. [5] attempted to correlate migration times of tryptic peptides of recombinant human growth hormone (rhGH) with $M^{2/3}/Z_c$. However, in this study, the authors gave no indication of the goodness of fit of the data and did not attempt to examine the alternative relationships defined in Eqs. (3.1a) and (3.1b). Grossman et al. [6] have reported a semi-empirical relationship of the form:

$$\mu_{em} \propto \ln(Z_c + 1)/n^{0.43} \quad (3.2)$$

where, as previously stated, n is the number of amino acid residues. Grossman et al. [6] firstly used a series of peptides of identical length but with varying charge to examine the relationship between μ_{em} and Z_c. The authors observed a linear relationship between μ_{em} and Z_c for $Z_c < 0.5$, however, at higher values of Z_c, the relationship became non-linear with the effect of any additional valence on electrophoretic mobility decreasing. This relationship between μ_{em} and Z_c was best described by a logarithmic function. This result contradicts fundamental theory which states that a direct proportionality exists between electrophoretic mobility and charge (see Eq. (2.17)). This direct proportionality has been confirmed for a number of proteins using membrane potential measurements to determine charge [27,28,30]. The lack of direct proportionality between μ_{em} and Z_c for the Grossman et al. [6] data is likely to be attributable to their calculation of charge using peptide primary sequence and the Henderson-Hasselbach equation, which allows the construction of the theoretical titration curve for each peptide from which Z_c can be determined for a given pH. Based on the Henderson-Hasselbach equation, Sillero and Ribeiro [33] derived the following expression for the calculation of Z_c:

$$Z_c = \sum_{n=1-4} P_n / 1 + 10^{pH - pK(P_n)} - \sum_{n=1-5} N_n / 1 + 10^{-pK(N_n) - pH} \quad (3.3)$$

where P_n and N_n are the integral number of cationic (i.e. $P_1 = tNH_2$, $P_2 = His$, $P_3 = Arg$, $P_4 = Lys$) and anionic (i.e. $N_1 = tCOOH$, $N_2 = Asp$, $N_3 = Glu$, $N_4 = Cys$, $N_5 = Tyr$) amino acid residues, respectively, and $pK(P_n)$ and $pK(N_n)$ are the ionization potentials of these amino acids. The calculation of Z_c for proteins and peptides that exhibit post-translational modifications such as phosphorylation and glycosylation can also be achieved using Eq. (3.3) by treating these ionisable groups in an analogous fashion to the amino acids. The accuracy of this method is largely dependent upon the estimation of appropriate ionization constants, $pK(P_n)$ and $pK(N_n)$. It is not valid to use the ionization potentials for free amino acids for the calculation of Z_c for proteins and peptides, as these can differ significantly to those for amino acyl residues in the protein or peptide [9]. Firstly, peptide bond formation induces an electrostatic change in the charge on the amino and carboxyl groups at the chain termini, resulting in a shift in their pK_a values from approximately 9.5 to 8.1 and 2.2 to 3.2, respectively. Secondly, the environment in which an amino acid resides can have a dramatic effect on its

ionization potential. In proteins, amino acids may reside in a multitude of microenvironments, such as at the surface of the protein in an aqueous environment where the dielectric constant is high or within the interior of the protein in a relatively hydrophobic environment where the dielectric constant is low. Also, induced effects from proximate amino acyl residues can produce a shift in pK_a values. Finally, ionization constants are effected by the ionic nature of the solution. Table 3 lists ionization constants developed for amino acyl residues using biosynthetic human insulin (BHI) and human growth hormone (hGH) [9]. Although relatively few reports on the average values of ionization constants for amino acyl residues in proteins and peptides have been published [9,34,35], some values vary by up to 1 pH unit (Table 1). Therefore, whilst the Henderson-Hasselbalch equation has been demonstrated to be reasonably accurate for prediction of protein and peptide pI values, the equation is not always accurate for calculation of valence at pH values that are far removed from the pI [7,8]. The non-linear relationship between μ_{em} and Z_c observed by Grossman et al. [6] also suggests limitations in the Henderson-Hasselbalch equation for prediction of actual valence (Z_a). At high values of Z_c, the increase in mobility observed by Grossman et al. [6] did not reflect the increase in peptide charge. Tanford and Kirkwood [36] have shown that overestimation of actual valence by titration measurements is due to electrostatic charge suppression. Therefore, even when using appropriate ionization constants, the calculation of valence using the Henderson-Hasselbalch equation is likely to result in overestimation of Z_c under conditions where the protein has appreciable charge. Compton [7] introduced a proportionality constant (F_z), to account for differences between Z_c and Z_a through the expression $Z_a = Z_c/F_z$. Examination of the nature of F_z using classic expressions for charge suppression [8] revealed a logarithmic dependency of F_z on Z_a (cf. Grossman et al. [6]), M and I (through (κ) with pH independence, such that:

$$\log(F_z - 1/2) = Z_a K_5 K_6 (M^{1/3} + \kappa K_5 M^{2/3}) + \log 1/2 \quad (3.4)$$

where $K_5 = (3v/4\pi N)^{1/3}(f/f_o)$ and $K_6 = (0.868e^2 N)/(2\varepsilon\varepsilon_o RT)$. Substituting the expression for the Debye–Hückel parameter κ (defined in Eq. (2.1)) into Eq. (2.17) and incorporating F_z, Compton [7] showed that μ_{em} could be defined by the equation:

$$\mu_{em} = \frac{K_1 Z_c / F_z}{K_2 M^{1/3} + K_3 M^{2/3} I^{1/2}} \quad (3.5)$$

where $K_1 = eX(\kappa r)$, $K_2 = 6\pi\eta(f/f_o)(4\pi N/3v)^{-1/3}$ and $K_3 = 6\pi\eta(2^{1/2}Ne)(\varepsilon_o\varepsilon RT)^{-1/2}(f/f_o)^2(4\pi N/3v)^{-2/3}$ and v is the protein partial specific volume. In this derivation, Compton [7] expresses the Stoke's radius of the protein in terms of the more useful, albeit more approximate, protein molecular mass through the equation:

$$r = (3Mv/4\pi N)^{1/3}(f/f_o) \quad (3.6)$$

The frictional ratio, f/f_o, usually varies from 1.0 (ideal behaviour) to 1.7 for globular proteins having non-spherical dimensions, to greater than three for cylindrical proteins, such as myosin. The partial specific volume observed for natural proteins is in the range 0.70–0.75 cm^3/g and is often taken to be 0.75 cm^3/g without any measurements or calcula-

Table 1
Ionization constants of amino acyl residues

Amino acyl residue	C-terminal	N-terminal	Side-chain
Ala (A)	3.20[a] (3.20)[b]	8.20 (8.20)	
Arg (R)	3.20 (3.20)	8.20 (8.20)	12.50 (12.00)
Asn (N)	2.75 (3.20)	7.30 (8.20)	
Asp (D)	2.75 (3.20)	8.60 (8.20)	3.50 (4.00)
Cys (C)	2.75 (3.20)	7.30 (8.20)	10.30 (9.00)
Gln (Q)	3.20 (3.20)	7.70 (8.20)	
Glu (E)	3.20 (3.20)	8.20 (8.20)	4.50[c] (4.50)
Gly (G)	3.20 (3.20)	8.20 (8.20)	
His (H)	3.20 (3.20)	8.20 (8.20)	6.20 (6.40)
Ile (I)	3.20 (3.20)	8.20 (8.20)	
Leu (L)	3.20 (3.20)	8.20 (8.20)	
Lys (K)	3.20 (3.20)	7.70 (8.20)	10.30 (10.40)
Met (M)	3.20 (3.20)	9.20 (8.20)	
Phe (F)	3.20 (3.20)	7.70 (8.20)	
Pro (P)	3.20 (3.20)	9.00 (8.20)	
Ser (S)	3.20 (3.20)	7.30 (8.20)	
Thr (T)	3.20 (3.20)	8.20 (8.20)	
Trp (W)	3.20 (3.20)	8.20 (8.20)	
Tyr (Y)	3.20 (3.20)	7.70 (8.20)	10.30 (10.0)
Val (V)	3.20 (3.20)	8.20 (8.20)	

[a] Data from Rickard et al. [9].
[b] Values in parentheses from Matthew [34].
[c] In casein, pK_a (γ-carboxyl of glutamate) = 4.85 [38].

tions [37]. For peptides, the data is limited, however, there appears to be a similar range (0.65–0.75 cm^3/g) for the partial specific volume of these molecules. Compton [7] determined the value of the proportionality constant, F_z, to be 5.66 for the chimeric L6 IgG monoclonal antibody (cL6) and further demonstrated the usefulness of this relationship by optimising the HPCE resolution of cL6 from its deamidated analogue. In this study, Compton et al. [7] constructed a calculated mobility and valence titration curve for cL6 and deamidated cL6 and then selected a pH at which the difference in their mobilities was maximum. Compton and O'Grady [8] have suggested that for eight proteins ranging in molecular mass from 14 186–147 760, at 0.07 ionic strength, $\upsilon = 7.3 \cdot 10^{-4}$ m^3/kg, $X(\kappa r) = 1.05$, and $f/f_o = 1.0$, electrophoretic mobility should have a molecular mass dependency of approximately $M^{1/2}$. In fact, for these data, we plotted μ_{em} against $F_z q/M^{1/2}$ and obtained a significant correlation ($p < 0.01$), with an R^2 value of 0.713. However, a significant correlation ($p < 0.01$) was also obtained for a plot of μ_{em} against $F_z q/M^{2/3}$ ($R^2 = 0.703$), but a non-significant correlation ($p > 0.10$) was obtained for a plot of μ_{em} against $F_z q/M^{1/3}$ ($R^2 = 0.374$). The validity of the proportionality constant, F_z, is demonstrated by the correlation coefficients obtained when F_z was excluded from the relationship. For plots of μ_{em} against $q/M^{1/2}$, $q/M^{2/3}$ and $q/M^{1/3}$, R^2 values of 0.125, 0.222 and 0.072, respectively, were obtained.

One of the most interesting aspects of Compton's model [7] is that it predicts a complex dependency of electrophoretic mobility on molecular mass and ionic strength. For any given experimental condition, Eq. (3.5) reduces to:

$$\mu_{em} = K_4(Z_c/F_z)M^{-s} \tag{3.7}$$

where s varies from 1/3 to 2/3 and K_4 is an aggregate of K_{1-3}. Compton [7] proposed that the discrepancy observed between the work of Offord [15] ($M^{-2/3}$ model) and Grossman et al. [6] ($M^{-1/2}$ model) is due to this complex dependency of μ_{em} on M and I (see Eq. (3.5)) whereby at high values of I, the dependency approaches $M^{-2/3}$, and at low values of I, the dependency approaches $M^{-1/3}$. Offord [15] studied peptides with molecular masses in the range of 150–3500 at an ionic strength of approximately 0.07, whilst Grossman et al. [6] studied peptides of 330–4527 molecular mass at an ionic strength of approximately 0.01. Other studies on peptides and proteins are summarised in Table 2, in which the mass range of the molecules investigated, the ionic strength, the best relationship for M^{-s}, the correlation coefficient (R^2) and the significance (p value)

Table 2
Summary of studies on the relationship between μ_{em} and the molecular mass of proteins and peptides at various ionic strengths using paper electrophoresis, HPCE and Doppler velocimetry

Reference	M	I	Model	R^2	n_s[a]
Offord [15]	150–3500	0.070	$M^{-2/3}$		
Nyberg et al. [3]	271–1346	0.030	$M^{-2/3}$		
Deyl et al. [4]	13 000–25 100	0.100	$M^{-2/3}$		
Frenz et al. [5]	400–1700	0.100	$M^{-2/3}$	-[c]	27
Grossman et al. [6][b]	330–4527	0.012	$M^{-1/2}$	0.978[d]	40
Rickard et al. [9]	315–22 818	0.100	$M^{-2/3}$	0.899[d]	43
Compton and O'Grady [8]	14 186–147 760	0.070	$M^{-1/2}$	0.713[e]	8
Issaq et al. [10]	160–445	0.050	$M^{-2/3}$	0.999[d]	6
Adamson et al. [11]	1648–3125	0.035	$M^{-2/3}$	0.993[d]	13
Chen et al. [12]	130–460	0.047	$M^{-1/2}$	0.972[d]	37
Adamson and Reynolds [13]	1210–3125	0.035	$M^{-2/3}$	0.909[d]	25
Basak and Ladisch [16]	178–140 000	0.005	$M^{-2/3}$	0.884[d]	51

[a] n_s = sample size (number of peptides/proteins studied).
[b] Actual correlation with $n^{0.43}$, where n = number of residues, however, approximated to $M^{1/2}$ for comparison purposes.
[c] not determined.
[d] $p < 0.001$.
[e] $p < 0.01$.

are given. There appear to be inconsistencies between the results in the literature, in particular the data of Issaq et al. [10] and Chen et al. [12]. In these studies, two groups of peptides of almost identical masses, examined under conditions of identical ionic strength, resulted in different relationships for the dependence of electrophoretic mobility on molecular mass. The study by Issaq et al. [10] was conducted on a series of polyalanines, from dimer to hexamer, which were analysed by HPCE in 50 m*M* phosphate buffer (pH 2.5, $I=0.05$) at 25°C with an applied voltage of 15 kV. Issaq et al. [10] demonstrated a linear relationship between μ_{em} and $Z_c/M^{2/3}$, however, they did not investigate the other models defined in Eqs. (3.1a) and (3.1b). Chen et al. [12] plotted migration time against $M^{1/2}/Z_c$ for a similar series of polyglycines analysed by HPCE under almost identical conditions to those used in the study by Issaq et al. [10], with the exception that the temperature was maintained at 60°C. Chen et al. [12] showed a linear fit for their data using the $M^{1/2}/Z_c$ relationship. We have further analysed these data by applying all three models for molecular mass ($M^{-1/3}$, $M^{-1/2}$, $M^{-2/3}$) as well as a simple inverse proportionality model (M^{-1}) to the polyalanine and polyglycine peptide series of Issaq et al. [10] and Chen et al. [12], respectively. The data of Issaq et al. [10] were found to correlate most closely ($R^2=0.999$) with the $M^{-1/3}$ model compared with the $M^{-1/2}$ model ($R^2=0.998$), $M^{-2/3}$ model ($R^2=0.995$) and the M^{-1} model ($R^2=0.986$). The data of Chen et al. [12] were found to correlate most closely ($R^2=0.9999$) with the $M^{-2/3}$ model followed by the $M^{-1/2}$ model ($R^2=0.9997$), the $M^{-1/3}$ model ($R^2=0.998$) and the M^{-1} model ($R^2=0.997$). Whilst the dependency of electrophoretic mobility on the size of the peptide, as approximated by *M*, is clear, there appears to be little difference between the various models for these series of small homologous peptides. The study by Chen et al. [12] is more comprehensive than that of Issaq et al. [10] in that the authors extended the range of peptides used to include a further fourteen dipeptides (of which ten contained N-terminal glycine), thirteen tripeptides, four tetrapeptides and one pentapeptide (comprising alanine only), demonstrating a closer fit to the data with the $M^{-1/2}$ model ($R^2=0.972$) than the $M^{-2/3}$ ($R^2=0.734$) and $M^{-1/3}$ ($R^2=0.476$) models. One aspect of this study by Chen et al. [12] is that the majority of peptides studied contained no ionizable residues, such that, under the conditions of low pH used, the only appreciable charge was conferred by the N-terminal α-amino group. It is likely that charge suppression effects would be minimal for these peptides and that the calculated valence would be a reasonable approximation of the actual valence. In effect, the value of F_z for these peptides would approach unity. The dependence of F_z on *M* makes the interpretation of those models developed for peptides and proteins more complex (Table 2), in which calculated valence was used and charge suppression was likely to be significant. Interestingly, for the data of Grossman et al. [6], Chen et al. [12] and Compton and O'Grady [8], for which charge suppression was either taken into account or negligible, the $M^{-1/2}$ model was most appropriate. In the remaining studies, listed in Table 2, for which charge suppression effects were likely to be significant, the $M^{-2/3}$ model gave the best fit for the data. It is likely that the higher value of the molecular mass exponent, *s*, for this model more adequately compensates for the increase in F_z as *M* increases. Using the simplified expression for Eq. (3.4):

$$\log F_z = aZ_a M^p + b \tag{3.8}$$

and substituting Eq. (3.8) into Eq. (3.5). Compton and O'Grady [8] derived the following expression for electrophoretic mobility:

$$\mu_{em} = \frac{K_1 Z_c 10^{-aMb+c}}{K_2 M^{1/3} + K_3 M^{2/3} I^{1/2}} \tag{3.9}$$

which, for a particular protein, functionally reduces to:

$$\mu_{em} = Z_c / K_5 M^s \tag{3.10}$$

where *s* varies between 1/3 and 2/3 and K_5 is an aggregate of K_{1-3}. The form of Eq. (3.10) is identical to that of the models developed for peptides and proteins in which charge suppression was not taken into account [3–13,15,16]. The general form of Eq. (3.10) can be tested by performing a plot of log (Z_c/μ_{em}) against log *M*, which should yield a straight line, the slope of which would be equal to the value of *s*. Using data generated previously in this laboratory for tri- and diphosphorylated peptides from

enzymic digests of bovine casein [13] (Table 3), we observed quite a poor linear relationship ($R^2=0.619$) for a plot of log (Z_c/μ_{em}) against log M. When the data was replotted taking into account charge suppression using the Grossman et al. [6] non-linear function, $\ln(Z_c+1)$, in place of Z_c, a much improved correlation was obtained ($R^2=0.986$) with a slope of $s=0.345$. We have previously demonstrated [13] that the electrophoretic mobility of casein phosphopeptides (CPP), containing the cluster sequence -Ser(P)-Ser(P)-Ser(P)-Glu-Glu-, forms a separate linear relationship to that of the tri- and diphosphorylated casein peptides when plotted against $Z_c/M^{2/3}$. Using a value of $s=0.345$, we have plotted μ_{em} against $\ln(Z_c+1)/M^{0.345}$ for cluster and tri/diphosphorylated peptides, independently, as shown in Fig. 1. Table 4 presents data comparing the $\ln(Z_c+1)/M^{0.345}$ model with the $Z_c/M^{2/3}$ model, $\ln(q+1)/n^{0.43}$, $Z_c/M^{1/3}$ and $Z_c/M^{1/2}$ models [13]. The $\ln(Z_c+1)/M^{0.345}$ model gave the best correlation for both sets of peptides and most adequately satisfies the boundary condition that as Z_c approaches zero, μ_{em} must also approach zero. This boundary condition was inadequately satisfied in the Z_c/M^s models. The value of 0.345 for the mass exponent suggests that the Stoke's radius model ($M^{-1/3}$) is best for these highly negatively charged peptides. It is likely that the high negative charge density of these peptides imparts a

Table 3
Physicochemical properties and effective electrophoretic mobilities of various phosphopeptides from enzymic digests of casein

Peptide[a]	M[b] (Da)	Z_c[c]	$\mu_{em}\cdot 10^4$ (cm^2/V.s)
Cluster CPP			
α_{S2}-CN-4P(f53-70)	2162	−13.06	3.615[d]
α_{S1}−CN-5P(f59-77)	2464	−15.11	3.682
α_{S1}−CN-5P(f59-79)	2721	−15.04	3.535
[Glp^{59}]α_{S1}-CN-5P(f59-79)	2704	−15.04	3.587
[Met(O)^{60}]α_{S1}-CN-5P(f59-79)	2738	−16.04	3.566
α_{S2}-CN-4P(f46-70)	3009	−16.06	3.484
α_{S2}-CN-4P(f2-20)	2490	−12.99	3.351
β-CN-4P(f7-25)	2357	−11.99	3.299
β-CN-4P(f2-25)	2969	−14.91	3.484
β-CN-4P(f1-25)	3125	−13.91	3.185
α_{S1}-CN-5P(f61-79)	2462	−15.04	3.838
α_{S1}-CN-5P(f61-78)	2334	−16.04	3.991
α_{S2}-CN-4P(f53-69)	2033	−14.06	3.986
α_{S2}-CN-4P(f2-19)	2327	−12.99	3.710
α_{S1}-CN-4P(f61-73)	1697	−13.04	3.995
β-CN-4P(f7-21)	1900	−12.99	3.748
α_{S1}-CN-4P(f61-69)	1088	−10.99	4.443
[Ser(P)^{53}]α_{S2}-CN-4P(f52-63)	1640	−12.99	4.061
β-CN-4P(f12-21)	1396	−10.99	4.130
Tri- and diphosphorylated CPP			
α_{S2}-CN-2P(f126-135)	1281	−8.04	4.113
α_{S1}-CN-2P(f43-52)	1243	−8.80	4.346
α_{S1}-CN-2P(f37-52)	2010	−8.80	3.690
α_{S1}-CN-2P(f41-50)	1210	−6.00	3.759
α_{S1}-CN-2P(f41-52)	1452	−7.00	3.785
[Ser(P)^{41}]α_{S1}-CN-3P(f40-52)	1565	−9.00	4.004

[a] Primary structure of peptides; β-CN-4P(f1-25), RELEELNVPGEIVESLSSSEESITR; α_{S1}-CN-5P(f59-79), QMEAESISSSEEIVPNSVEQK;α_{S2}-CN-4P(f46-70), NANEEEYSIGSSSEESAEVATEEVK; α_{S2}-CN-4P(f2-20), NTMEHVSSSEESIISQETY; α_{S2}-CN-2P(f126-135), EQLSTSEENS; α_{S1}-CN-2P(f37-52), VNELSKDIGSESTEDQ; where S = Ser(P).

[b] Calculated using free acid residue weights.

[c] Calculated using Ser(P) $pK_{a2}=6.48$ [38]; Lys $pK_{a\varepsilon}=10.30$; Arg pK_a (Guanidino)=12.50 [1]; Glu $pK_{a\gamma}=4.85$; Asp $pK_{a\beta}=4.85$ [38]; N-terminus and C-terminus pK_a values are from Rickard et al. [9].

[d] Maximum coefficient of variation, 3.2%; mean coefficient of variation, 0.93%; cm=capillary length (72 cm), V=volts applied (30 kV).

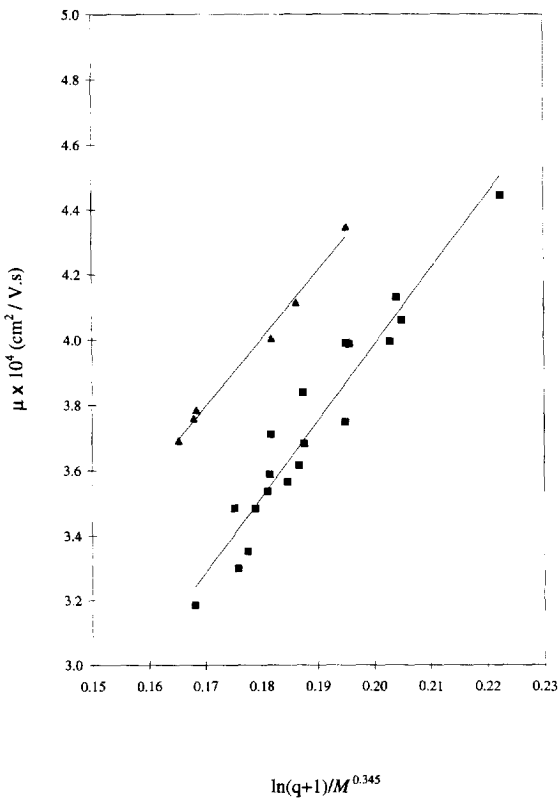

Fig. 1. Effective electrophoretic mobility μ vs. $\ln(q+1)/M^{0.345}$ for casein phosphopeptides containing the cluster sequence -Ser-(P)-Ser(P)-Ser(P)-Glu-Glu- (■, $R^2=0.924$, y-intercept $=6.68\cdot10^{-5}$ cm^2/V.s) and di- and triphosphorylated casein peptides (▲, $R^2=0.990$, y-intercept $=2.61\cdot10^{-5}$ cm^2/V.s).

rather open structure during electrophoretic migration and that the electrophoretic motion of such a molecule is best satisfied by Stoke's model. One of the most interesting aspects of these data are the separate linear relationships observed for the cluster and tri/diphosphorylated CPP. This implies that the constant K_5 differs for these two sets of peptides and is a higher value for the cluster CPP. Examination of the definitions of K_{1-3} above, of which K_5 is an aggregate, suggests that the difference in the value of K_5 between the two sets of peptides is attributable to differences in the frictional ratio f/f_o. A higher value of K_5 suggests a higher f/f_o ratio, which is characteristic of a more asymmetric molecule [7]. This implies that the cluster CPP is more extended and asymmetric compared with the tri/diphosphorylated CPPs, which may be expected due to the presence of the cluster sequence -Ser(P)-Ser(P)-Ser(P)-Glu-Glu-. The form of Eq. (3.10) appears to be inadequate for these highly negatively charged CPPs, for which charge suppression effects would be significant. This is supported by the data of Basak and Ladisch [16] who plotted (f/f_o) against $Z_c/M^{2/3}$ for a range of proteins and peptides for which $Z_c/M^{2/3}$ varied from -0.0265 to $+0.0260$. The authors reported that the relationship improved when highly charged proteins with $Z_c/M^{2/3}$ less than -0.012 were excluded. The peptides for which $Z_c/M^{2/3}$ was greater than $+0.012$ were short (two–six residues) and of low net charge and unlikely to be greatly affected by charge suppression. It is, therefore, suggested that for peptides, in the absence of F_z data, the general form of Eq. (3.10) be modified to:

$$\mu_{em} = \ln(Z_c + 1)/K_5 M^s \qquad (3.11)$$

For the data of Grossman et al. [6], a plot of log M against $\log[\ln(Z_c+1)/\mu_{em}]$ gave an s value of 0.435. Fig. 2 shows a plot of μ_{em} against $\ln(Z_c+1)/M^{0.435}$ for the Grossman et al. [6] data, which resulted in an R^2 value of 0.985 and a y-intercept of $5.27\cdot10^{-6}$ (cm^2/V.s). Grossman et al. [6] obtained an R^2 value of 0.978 and a y-intercept of $2.47\cdot10^{-5}$ (cm^2/V.s) for a plot of μ_{em} against $\ln(Z_c+1)/n^{0.43}$, indicating that M is slightly better than n as a measure of peptide size. A plot of log M against log $[\ln(Z_c+1)/\mu_{em})$ for the data of Compton and O'Grady [8] on proteins in the mass range 14 186–147 760 gave an s value of 0.243. However, a non-significant correlation ($p>0.20$) was obtained for a plot of μ_{em} against $\ln(Z_c+1)/M^{0.243}$ ($R_2=0.481$). Whilst the general form of Eq. (3.11) appears to work well for peptides, its rather simplistic compensation for charge suppression is inadequate for proteins where the magnitude

Table 4
y-Intercept and correlation coefficient for relationships involving the electrophoretic mobility of phosphopeptides from casein and various models incorporating peptide charge and size

Relationship	Cluster CPP		Tri/diphosphorylated CPP	
	Intercept	R^2	Intercept	R^2
$\mu_{em}\propto\ln(Z_c+1)/M^{0.345}$	$6.68\cdot10^{-5}$	0.924	$2.61\cdot10^{-5}$	0.990
$\mu_{em}\propto Z_c/M^{2/3}$	$1.19\cdot10^{-4}$	0.909	$2.33\cdot10^{-4}$	0.951
$\mu_{em}\propto\ln(Z_c+1)/n^{0.43}$	$7.03\cdot10^{-5}$	0.872	$9.02\cdot10^{-5}$	0.967
$\mu_{em}\propto Z_c/M^{1/3}$	$1.76\cdot10^{-4}$	0.200	$2.48\cdot10^{-4}$	0.645
$\mu_{em}\propto Z_c/M^{1/2}$	$6.09\cdot10^{-5}$	0.708	$2.32\cdot10^{-4}$	0.836

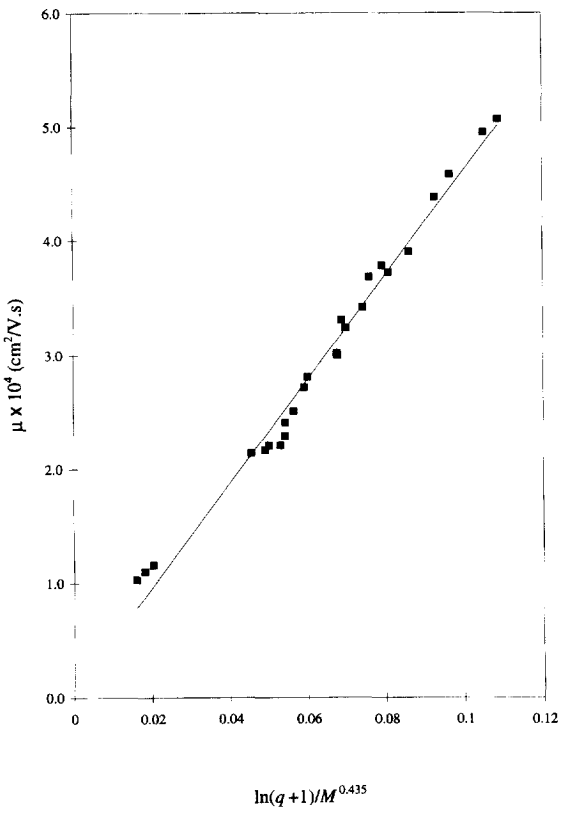

Fig. 2. Effective electrophoretic mobility μ vs. $\ln(q+1)/M^{0.435}$ for the data of Grossman et al. [6] ($R^2=0.985$, y-intercept $=5.27\cdot 10^{-6}$ cm^2/V.s).

of charge suppression is greater and the mechanisms are more complex. Taking charge suppression into account by incorporating the term F_z into the relationship, a plot of log M against log $(Z_c/F_z\ \mu_{em})$ for the data of Compton and O'Grady [8] gave an s value of 0.604 and a subsequent plot of μ_{em} against $Z_c/F_zM^{0.604}$ gave a significant R^2 value of 0.825 and a y-intercept of $3.46\cdot 10^{-5}$ (cm^2/V.s). Further improvement in the correlation might be achieved using protein-specific values of f/f_o and v, rather than approximated values for the calculation of μ_{em}, based on calculated valence, from Eq. (3.5). However, for the data of Compton and O'Grady [8], this value of s of 0.604 gives a better fit ($R^2=0.825$) than the models using a mass exponent of 1/3 ($R^2=0.374$), 1/2 ($R^2=0.713$) or 2/3 ($R^2=0.703$).

It is interesting to compare the molecular mass dependency of electrophoretic mobility for the peptides and proteins studied by Grossman et al. [6], Compton and O'Grady [8] and Adamson and Reynolds [13], all of which have been re-modelled in this review. On re-modelling, the molecular mass dependency for the highly charged CPP [13] approached $M^{-1/3}$, the Stoke's radius model, possibly due to a more asymmetric, open structure imparted by the high negative charge density. For the peptides of Grossman et al. [6], the dependency approached $M^{-1/2}$. The lower charge densities of the Grossman et al. [6] peptides possibly allow a less asymmetric structure and greater random motion during electrophoresis, which is more characteristic of a random polymer. Finally, for the proteins studied by Compton and O'Grady [8], the molecular mass dependency of electrophoretic mobility approached $M^{-2/3}$ for the re-modelled data, suggesting that the frictional forces acting on a protein undergoing electrophoretic motion are proportional to the surface area of these larger, more rigid structures.

4. Conclusions and perspectives

The interpretation of the electrophoretic mobility of peptides and proteins in terms of molecular size and charge is a complex problem which is difficult to unify. From the theory and empirical data available, we can, however, draw a number of important conclusions. The electrophoretic mobility of a protein or peptide is proportional to its actual charge and is inversely proportional to its size. As is the general consensus, the factor that most effects electrophoretic mobility is charge, hence pH is the most significant variable for altering HPCE resolution of peptides and proteins. Accordingly, the most critical parameter when predicting electrophoretic mobility is charge or valence. Due to the effects of electrostatic charge suppression, the calculated valence, determined from amino acid sequence data and appropriate pK_a values for amino acid residues, often overestimates the actual valence of the peptide/protein and can be a source of considerable error in predicted mobility. For peptides, we suggest that the general expression defined in Eq. (3.11) be used for the determination of μ_{em} from Z_c and M. The molecular mass dependency of peptide electrophoretic mobility appears to vary between $M^{-1/3}$ and

$M^{-1/2}$, when charge suppression is taken into account. The lower exponent ($M^{-1/3}$) tends to fit peptides with very high charge densities. Considering the range of assumptions and approximations made in the development of Eq. (3.11), it provides quite a good fit to a wide range of data for peptides, when the correct molecular mass dependency is determined. The form of this equation is such that the molecular mass exponent can be determined from the slope of a plot of log M against log[ln(Z_c+1)/μ_{em}]. The data for proteins is a little less conclusive and it would appear that charge suppression is best treated by using the pH-independent charge suppression factor, F_z. At present, however, there are very few published values of F_z for different proteins and, although the value of F_z for a protein can be determined relatively easily, it is valid only for the specific set of conditions under which it was determined and can be subject to error unless accurate values of the protein friction factor and partial specific volume are used for its calculation [8].

The development of relationships between μ_{em}, M and Z_c for peptides and proteins offers a powerful tool, not only for predicting electrophoretic mobility, but also for optimising HPCE separations, studying structural modifications (e.g. phosphorylation, glycosylation, deamidation, etc.) and for the investigation of surface charge characteristics and conformation.

5. Abbreviations

BHI	Biosynthetic human insulin
C_i	Concentration of type i ions in electrolyte
cL6	Chimeric L6 IgG monoclonal antibody
CPP	Casein phosphopeptides
e	Electron charge (1.602×10^{-19} C)
E	Electrical field strength (V/cm)
F_E	Force on a charged particle in an electric field
F_F	Electrophoretic drag force
F_R	Electrophoretic retarding force
F_{RE}	Force due to relaxation effect
F_z	Valency proportionality factor (Z_c/Z_a)
f/f_o	Frictional ratio
f_c	Coefficient of friction
HPCE	High-performance capillary electrophoresis
hGH	Human growth hormone
I	Ionic strength
L_d	Length of capillary to detector (cm)
L_t	Total capillary length (cm)
M	Molecular mass
N	Avogadro's number (6.023×10^{23})
N_n	Integral number of anionic residues contained inprotein/peptide
n	Number of amino acid residues
n_s	Sample size
P_n	Integral number of cationic residues contained in protein/peptide
pI	Isoelectric point
pK_a	Ionization constant
p$K(P_n)$	Ionization constant of cationic residue
p$K(N_n)$	Ionization constant of anionic residue
q	Electrostatic charge
q_c	Calculated electrostatic charge
R	Gas constant (8.314 J/K)
R^2	Linear correlation coefficient
r	Stoke's radius
r_i	Average radius of ions in ionic atmosphere
rhGH	Recombinant human growth hormone
SP	Substance P
T	Absolute temperature (K)
t_m	Apparent electrophoretic migration time (s)
t_{eo}	Electrophoretic migration time of uncharged solute whose motion is due solely to electroosmosis
V	Applied voltage (V)
V_s	Volume of sphere
$X(\kappa r)$	Henry's function
Z_a	Actual valence
Z_c	Calculated valence
z_i	Valence of electrolyte ions of type i
ε	Dielectric constant (78.54 for water at 25°C)
ε_o	Permittivity of free space (8.854×10^{-12} C^2/Jm)
κ	Debye–Hückel parameter
μ_{em}	Effective electrophoretic mobility (cm^2/V.s)

μ_{eo}	Coefficient of electroosmotic flow (cm^2/V.s)
μ_m	Apparent electrophoretic mobility (cm^2/V.s)
η	Viscosity (water at 25°C, 8.95×10^{-4} kg/ms)
ν_{em}	Effective steady-state electrophoretic velocity (cm/s)
ν_{eo}	Steady-state electroosmotic flow velocity (cm/s)
ν_m	Steady-state apparent electrophoretic velocity
υ	Partial specific volume
ζ	Zeta-potential

Acknowledgments

This work was supported by an Australian Research Council project, grant No. C495301043.

References

[1] O. Vesterberg, J. Chromatogr. 480 (1989) 3.
[2] J.W. Jorgenson, K. DeArman Lukacs, J. Chromatogr. 218 (1981) 209.
[3] F. Nyberg, M. Zhu, J. Liao and S. Hjertén, in C. Shafer-Nielsen (Editor), Electrophoresis '88, VCH, Weinheim, 1988, p. 141.
[4] Z. Deyl, V. Rohlieck, M. Adam, J. Chromatogr. 480 (1989) 371.
[5] J. Frenz, S.-L. Wu, W.S. Hancock, J. Chromatogr. 480 (1989) 379.
[6] P.D. Grossman, J.C. Colburn, H.H. Lauer, Anal. Biochem. 179 (1989) 28.
[7] B.J. Compton, J. Chromatogr. 559 (1991) 357.
[8] B.J. Compton, E.A. O'Grady, Anal. Chem. 63 (1991) 2597.
[9] E.C. Rickard, M.M. Strohl, R.G. Nielsen, Anal. Biochem. 197 (1991) 197.
[10] H.J. Issaq, G.M. Janini, I.Z. Atamna, G.M. Muschik, J. Lukszo, J. Liq. Chromatogr. 15 (1992) 1129.
[11] N. Adamson, P.F. Riley, E.C. Reynolds, J. Chromatogr. 646 (1993) 391.
[12] N. Chen, L. Wang, Y.K. Zhang, Chromatographia 37 (1993) 429.
[13] N.J. Adamson, E.C. Reynolds, Electrophoresis 16 (1995) 525.
[14] J.Th.G. Overbeek and P.H. Wiersema, in M. Bier (Editor), Electrophoresis; Theory, Methods and Applications, Academic Press, New York, 1967, Ch. 1, p. 1.
[15] R.E. Offord, Nature 211 (1966) 591.
[16] S.K. Basak, M.R. Ladisch, Anal. Biochem. 226 (1995) 51.
[17] P.W. Atkins, Physical Chemistry, Oxford University Press, 1986.
[18] J.T. Edward, in J.C. Giddings and R.A. Keller (Editors), Advances in Chromatography, Marcel-Dekker, New York, Ch. 2, p. 63.
[19] P.C. Hiemenz, Principles of Colloid and Surface Chemistry, Marcel-Dekker, New York, 1977.
[20] H.S. Harned, Physical Chemistry of Electrolyte Solutions, Reinholdt, New York, 1950.
[21] M. von Smoluchowski, Bull. Acad. Sci. Cracovie, (1903) 182.
[22] E. Hückel, Physik. Z. 25 (1924) 204.
[23] J.Th.G. Overbeek, Kolloidchem. Beih. 54 (1943) 287.
[24] F. Booth, Nature 161 (1948) 83.
[25] F. Booth, Proc. Roy. Soc. A203 (1950) 514.
[26] D.C. Henry, Proc. Roy. Soc. A133 (1931) 106.
[27] G.S. Adair, M.E. Adair, Trans. Farad. Soc. 36 (1940) 23.
[28] A. Tiselius, H. Svensson, Trans. Farad. Soc. 36 (1940) 16.
[29] P.D. Grossman, D.S. Soane, Anal. Chem. 62 (1990) 1592.
[30] H.A. Abramson, L.S. Moyer and M.H. Gorin, Electrophoresis of Proteins, Reinholdt, New York, 1942.
[31] H.A. Abramson, J. Gen. Physiol. 15 (1932) 575.
[32] C.H. Tanford, Physical Chemistry of Macromolecules, Wiley, New York, 1961.
[33] A. Sillero, J.M. Ribeiro, Anal. Biochem. 179 (1989) 319.
[34] J.B. Matthew, Annu. Rev. Biophys. Biophys. Chem. 14 (1985) 387.
[35] T. Hirokawa, T. Gojo, Y. Kiso, J. Chromatogr. 390 (1987) 210.
[36] C. Tanford, J.G. Kirkwood, J. Am Chem. Soc. 79 (1957) 5333.
[37] A.A. Zamyatnin, Ann. Rev. Biophys. Bioeng. 13 (1984) 145.
[38] H.E. Swaisgood, in P.F. Fox (Editor), Developments in Dairy Chemistry – 1. Proteins, Applied Science Publishers, New York, London, 1982, Ch. 1.