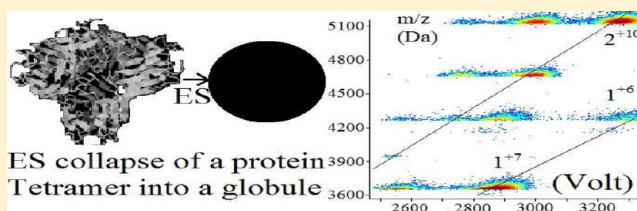


Capillary and Coulombic Effects on the Gas Phase Structure of Electrosprayed Concanavalin A Ions and Its Clusters  $C_n^{+z}$  ( $n = 1-6$ )J. Fernandez de la Mora,<sup>\*,†</sup> R. Borrajo-Pelaez,<sup>†,‡</sup> and M. Zurita-Gotor<sup>§,⊥</sup><sup>†</sup>Mechanical Engineering Department, Yale University, New Haven, Connecticut 06520-8286, United States<sup>‡</sup>SEADM, Parque tecnológico de Boecillo 205, Valladolid, Spain<sup>§</sup>Departamento de Ingeniería Espacial y Mecánica de Fluidos, Universidad de Sevilla, Sevilla 41092, Spain

**ABSTRACT:** Ion mobility spectrometry (IMS) coupled to mass spectrometry (MS) is used to study the gas phase collision cross section  $\Omega(z, n)$  in  $\text{CO}_2$  of multimers  $C_n$  ( $n = 1-4, 6$ ) of concanavalin A, whose tetramer  $C_4$  has a crystal structure resembling four tetrahedrally arranged globules.  $C_n^{+z}$  ions electrosprayed from aqueous solutions of triethylammonium formate ( $\text{Et}_3\text{AF}$ ) are moderately charged (up to  $z = 6$  and 17 for  $n = 1$  and 6) and produce narrow mobility peaks. Charge states down to  $z = 1$  obtained with a charge-reducing radioactive  $^{63}\text{Ni}$  source are studied for the dimer and the tetramer via pure IMS (no MS). The mobilities are independent of pH in the range 6–8, controlled by addition of triethylamine to the  $\text{Et}_3\text{AF}$ . The measured compactness group  $\Omega(z, n)/n^{2/3}$  is practically independent of  $n$  and  $z$ , whereas mobility calculations with clusters of touching spheres show that it should vary with  $n$  by 20–30% for a variety of scattering models. This contrast suggests that, irrespective of ambiguities on the scattering model, all multimers adopt globular shapes, precluding in particular a tetrahedral tetramer. Acetic acid solutions (87 mM aqueous) yield ions with substantially higher  $z$ , mostly with broad mobility distributions. Exceptionally high  $z$  tetramers ( $z = 25-29$ ) and trimers have narrowly defined mobilities with compact but nonspherical shapes. Addition of 2–4 mM  $\text{Et}_3\text{AF}$  to the 87 mM aqueous acetic acid solution yields narrowly defined mobilities almost identical at all  $z$  values to those from the  $\text{Et}_3\text{AF}$  buffer, although with higher charge states showing also a transition to nonspherical shapes. We conclude that all gas phase clusters charged below a Rayleigh-like charge,  $z_R$ , are globular without regard to solution conditions, some undergoing a sharp shape transition at a critical  $z = z_R$ . We confirm that gas phase protein cross sections differ from those expected from the crystal structure, with a trend to compact probably driven by their high surface energy (and opposed by Coulombic stresses). The Rayleigh-like shape transitions are similar to those arising in linear homopolymers, although not as sharply defined. They yield a surface energy for protein matter almost as high as the surface tension of water. This quantitative conclusion is corroborated by prior data on cytochrome *c* and apomyoglobin (also showing a critical shape transition) as well as measurements of the maximum charge versus mass in aggregates of dipeptides.



## 1. INTRODUCTION

Ion mobility spectrometry (IMS) coupled to mass spectrometry (MS) provides a powerful tool to study the gas phase structure of ions, including biological species. When these ions are produced by electrospray<sup>1</sup> (ES) from solutions in which they are in a native state, a number of features of the solution structure appear to be preserved:<sup>2</sup> for example, analogs of acid denaturation,<sup>3,4</sup> noncovalent protein–ligand bonds,<sup>5</sup> the biological activity of electrosprayed viruses<sup>6,7</sup> and enzymes,<sup>8</sup> ring structures,<sup>9</sup> etc. Conversely, a variety of experimental observations dating back to the earliest IMS studies of gas phase protein ions have provided strong indications that collision cross sections are more characteristic of the gas phase ion than of its preexisting solution conformation. First, mobility peaks of mass-selected gas phase ions are most often rather broad, even when examined at high mobility resolution and when electrospraying from nondenaturing solutions.<sup>10</sup> This behavior apparently reflects the coexistence of a multitude of gas phase conformations, whereas only one would be expected for the crystal structure. Second, unfolded gas phase

conformations are produced by electrospraying highly charged ions from denaturing solutions. However, when the charge on these unfolded ions is reduced by seeding the gas with vapors of strong bases, they revert to folded structures as compact as those from low-charge ions electrosprayed from nondenaturing solutions.<sup>11</sup> These observations show that not only the unfolded-refolded gas phase structures but also the almost identical originally folded gas phase structures are characteristic of the gas phase.

In a study of the effect of temperature on the mobility of electrosprayed cytochrome *c*, the +5 ion was found to remain compact up to the highest temperature studied (571 K), but the +6 and +7 ions underwent a structural transformation that was completed in the intervals 423–474 and 373–423 K, respectively.<sup>12</sup> The direct interpretation of these results is that a sharp structural transformation of the gas phase ion arises at a

Received: November 7, 2011

Revised: July 10, 2012

Published: July 10, 2012

critical value,  $z^*$ , which decreases with gas temperature. These critical transformations are associated with the loss of stability of the compact structure due to Coulombic repulsion and are evidently inherent to the gas phase ion. Other gas phase peculiarities of cytochrome *c* have been observed directly by holding the +9 ion in a Paul trap for variable periods and measuring the time evolution of its cross section.<sup>13</sup> These early studies showed that the cross section measured in He at the lowest charge state are typically well below ( $\sim 15$ – $20\%$ ) those computed for the crystal structure.

There is, of course, nothing surprising about the fact that a sufficiently plastic ion would tend to adopt a gas phase structure dependent not only on  $z$  and the bath temperature, but also on the nature of the surrounding gaseous or liquid environment. Many studies have assumed that such characteristic gas phase structures exist and have also computed them, at least for ions of moderate size.<sup>14</sup> Notwithstanding this evidence, the assumption that solution structures should in some way be preserved in the gas phase, at least under certain favorable conditions, remains firmly ingrained in this field. Even the authors having produced some of the best evidence available for the contrary view have remained somewhat attached to the notion that gas phase structures reflect solution structures (i.e., *It seems likely that the sharp peak observed for the +6 to +8 charge states represents the native conformation in solution*<sup>10</sup>). An interesting justification for this view is the possible rigidity or crystallization of the solution structure that would require crossing an activation barrier before it can evolve into the equilibrium gas phase structure. We actually know that a barrier for gas phase protein ion compaction exists and is substantial in some cases (apomyoglobin) but not in others (cytochrome *c*).<sup>15</sup>

This early evidence of structural changes resulting from the transfer of protein ions into the gas phase has been reinforced in more recent studies.<sup>16–18</sup> Our group has provided additional experimental evidence of substantial gas phase ion compaction, including a reduction of  $\sim 40\%$  in the cross section with respect to the crystalline structure of the large protein complex GroEL ( $m \sim 800$  kDa).<sup>19</sup> Our approach, however, has been controversial because of a number of difficult experimental and interpretive problems. First, as discussed later, gas phase protein structures are not firmly known. Second, the relation between structure and mobility is also ambiguous, even for a given ion structure. For a *hard* sphere, the mobility,  $K$ , is given by kinetic theory as

$$\frac{K}{z(1 + m_g/m)^{1/2}} = \frac{3}{4\xi} \frac{e}{p\Omega} \sqrt{\frac{\pi kT}{8m_g}} \quad (1)$$

where  $kT$  is the thermal energy,  $e$  is the ion charge,  $p$  is the gas pressure,  $m$  is the ion mass, and  $m_g$  is the mass of the bath gas molecules. For a sphere of diameter  $d$ , when one accounts for the finite diameter,  $d_g$ , of the gas molecules, the cross section,  $\Omega$ , is

$$\Omega = \pi(d + d_g)^2/4 \quad (2)$$

The coefficient  $\xi$  takes different values, depending on the scattering law used. For the most common model, when the sphere is perfectly smooth and undergoes elastic and specular collisions with the gas background,  $\xi$  is unity.

$$\xi_{\text{HS}} = 1 \quad (3a)$$

More generally, the quantity  $\xi > 1$  is an empirical factor introduced by Millikan and his students<sup>20,21</sup> to account for a substantial enhancement on the drag found in their experiments with micrometer oil drops in the free molecule limit. Its measured value is

$$\xi_M = 1.36 \quad (3b)$$

not only for air, but also for a variety of gases, such as He and CO<sub>2</sub>.<sup>20–25</sup> For unclear reasons, this factor is rarely used in the determination of protein cross sections from measured mobilities. In some such studies,  $\xi$  has simply been taken to be unity.<sup>26–29</sup> In others (starting with the work of Jarrold and colleagues previously discussed), it has been calculated by the so-called exact hard sphere scattering model (EHSS),<sup>30</sup> in which each atom in the ion is supposed to interact as an ideal smooth hard sphere undergoing elastic and specular collisions with the gas. The cross section,  $\Omega_{\text{EHSS}}$ , found in their work based on eq 1 (and the assumption that  $\xi = 1$ ) differs from the geometrical cross section of the ion, calling again for the use of a drag enhancement factor  $\xi > 1$ . The value of this factor is not unambiguous until a geometric cross section is defined. This has often been based in EHSS calculations on the *mean projected area*,  $\Omega_{\text{PA}}$ , of the molecule, where all possible orientations are considered equiprobable. In the case of spheres,  $\Omega_{\text{PA}}$  coincides with the geometrical cross section, eq 2, so the factor  $\xi$  arising in Millikan's formula, eq 1, and the ratio

$$\xi_{\text{EHSS}} = \Omega_{\text{EHSS}}/\Omega_{\text{PA}} \quad (3c)$$

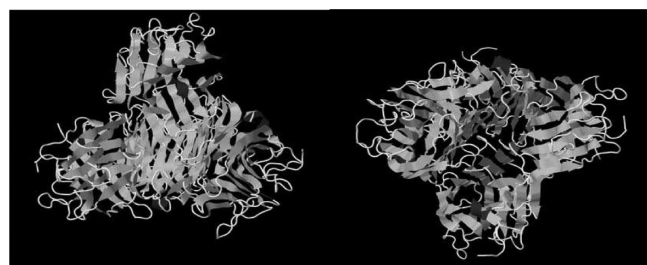
are directly comparable. A successful comparison has been recently provided in two independent calculations for the crystal structure of the large protein complex GroEL, both yielding  $\xi_{\text{EHSS}} = 1.36$ .<sup>26,19</sup> On one hand, this serves to validate experimentally the EHSS model in the limit of large ions. On the other, it confirms the general applicability of Millikan's drag law not just to oil drops, but also to many other materials, including protein ions. Earlier model calculations for large fullerene aggregates extrapolated to the limit of zero curvature had given the slightly smaller value  $\xi_{\text{EHSS}\infty} = 1.3$ .<sup>30</sup>

For nonspherical geometries, there is neither a rigorous connection between mobility and projected area<sup>31</sup> nor experimental measurements comparable to Millikan's. Rigorous comparisons between model and experiments are still possible in the case of sufficiently symmetric objects for which the drag tensor is isotropic. In this case, eq 1 still holds with Millikan's  $\xi$  when substituting  $4\Omega$  with the total surface area of the body (the wetted area).<sup>31</sup> The same result is approximately valid for shapes moderately close to a sphere or to another geometry sufficiently symmetric to have an isotropic drag tensor.<sup>31</sup>

Although previous ambiguities relating the value of the constant  $\xi$  appear now to be relatively settled, the interpretation of mobility data still faces a number of difficulties. First, standard nonspherical gas phase protein shapes are not available to test model calculations, so the only hard experimental data available relating one standard structure to a given mobility corresponds to spheres. The gas phase structure of moderate size ions can be modeled numerically,<sup>14</sup> enabling the computation of the mobility of that structure, and its comparison with a measured mobility. The approach has been reasonably successful with cluster ions and biological ions of moderate sizes. But its reliability is insufficient to provide a gas phase shape standard for proteins with 50 kDa–1 MDa masses, which could repeatably be introduced in the gas phase

with robust and well established shape and cross section. These problems are compounded in the case of large proteins by the difficulty in performing EHSS calculations.<sup>32</sup>

In the present article, we reexamine the possible gas phase compaction of protein ions by a method independent of any assumption on the value of  $\xi$ , via relative rather than absolute mobility measurements. Our approach is to compare the mobility of a globular protein to that of its aggregates (dimer, trimer, tetramer, ...). This task may be carried out in several ways. First, one can select proteins whose native state consists of an aggregate of monomers of known crystal structure. This is done here with concanavalin A, which, like other lectins, forms naturally as a tetramer with a mass of  $\sim 102.4$  kDa. Two views of the crystal structure of this complex are shown in Figure 1,<sup>33</sup>



**Figure 1.** Two views of the crystal structure of concanavalin A showing its tetrahedral structure. From protein data bank.<sup>33</sup>

where one can directly grasp its tetrahedral organization. If the compaction hypothesis were to apply generally and if an ES ion of this tetramer were to be spheroidized, its mobility would be substantially larger than that of the tetrahedral structure. An idea of how much the mobility would increase upon spheroidization of the native tetramer can be obtained by comparing the mobility of a sphere having four times the volume,  $V_1$ , of the monomer with that of a tetrahedrally arranged group of four touching spheres, each of volume  $V_1$ . The calculation and the proposed comparison are made in this article, confirming that the crystalline tetrahedral tetramer is substantially compacted in the gas phase. Other observations on concanavalin ions confirm the notion that their geometry is readily changed by capillary and electrostatic forces naturally present in the gas phase. This point is shown in particular by the abrupt loss of a compact globular shape at a critical charge,  $z^*$ , observed here for the monomer and tetramer of concanavalin A, similarly to earlier findings in cytochrome *c* and apomyoglobin. The measured  $z^*$ , its dependence on protein mass ( $z^* \sim m^{1/2}$ ) and bath gas temperature  $T$ , and the maximal charge versus mass found in aggregates of small peptides may be rationalized by a Rayleigh-like model. This analogy provides a quantitative value of the surface energy,  $\gamma_{\text{prot}}$  of protein matter, which all available data show approaches the surface tension of water ( $\gamma_{\text{prot}} \sim 0.93\gamma_{\text{water}}$ ) and is much larger than the corresponding value on homopolymers.

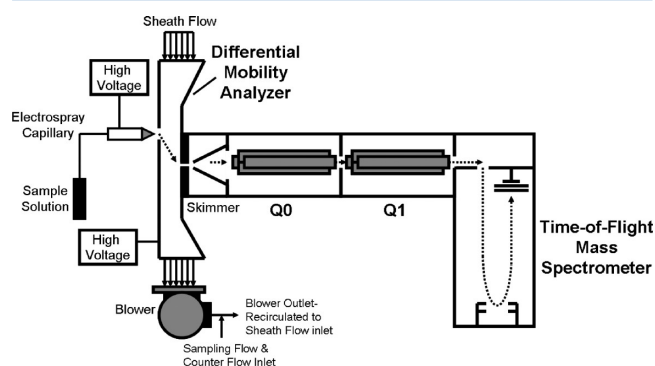
## 2. EXPERIMENTS AND METHODS

**Materials.** Concanavalin A (C-2010) was purchased from Sigma-Aldrich (St. Louis, MO, USA). The protein was first dissolved in water (HPLC grade, J. T. Baker) and then cleaned with Nanosep centrifugal filters (Pall Corporation; Ann Arbor, MI, USA) with a filter cutoff molecular weight of 10 kDa, less than half of the monomeric molecular weight of concanavalin A. After this desalting, the protein was dissolved at  $50 \mu\text{M}$

concentration in different aqueous buffers meant to achieve various charge state ranges of the electrosprayed ions. These buffers included 100 mM triethylammonium formate ( $\text{Et}_3\text{AF}$ ; Fluka 90359), 20–100 mM ammonium acetate (AA; Sigma-Aldrich A-7330), and 87 and 525 mM acetic acid ( $\text{HOAc}$ , J. T. Baker 9508-00). To control the pH in some samples, small amounts of ammonia (J. T. Baker, 29% aqueous solution) or triethylamine (Sigma T-0886) were added at concentrations below  $4 \mu\text{M}$ . The pH was measured with pH-indicator strips (EMD, USA). The tetraheptylammonium bromide used to produce a mobility standard was from Fluka (87296).

**Ionization.** The electrospray source used to ionize proteins has been described previously.<sup>34</sup> It involves a silica capillary (Polymicro Technologies, 1068150383) having an inner diameter of  $40 \mu\text{m}$  and an outer diameter of  $360 \mu\text{m}$ , tapered manually in house to  $50 \mu\text{m}$  at the tip, where a Taylor cone is formed. The electrospray source was operated in the cone-jet mode in a chamber with two glass windows, which allowed visual inspection of the cone when the capillary tip was sufficiently away from the analyzer inlet. Voltages in the range of 2–3 kV were applied to the sample (contained in a polypropylene vial) to electrify the tip by conduction through the liquid in the capillary. The liquid flow rate delivered from the reservoir to the tip was controlled by raising to 0.1 bar the pressure in the vial containing the sample. After monitoring the current of the electrospray operating in stable cone-jet mode (confirmed by watching it through the window with a microscope), the capillary tip was progressively brought closer to the entrance slit of the mobility analyzer to maximize the ion signal transmitted to the mass spectrometer. To maintain a constant electrospray current (to preserve the stability of the electrospray, now optically inaccessible), the needle voltage was reduced as the tip moved forward.

**DMA-MS and DMA with Radioactive Charge Neutralizer.** Most of the measurements in this work were performed with a commercial parallel plate differential mobility analyzer (DMA P4, SEADM S.L., Boecillo, Spain) in series with a quadrupole time-of-flight mass spectrometer (QStar XL, MDX Sciex, Toronto, Canada). A schematic of the tandem instrument is given in Figure 2. Its operation has been



**Figure 2.** Schematic of the tandem DMA-MS instrument (Courtesy of Dr. C.J. Hogan).

previously described.<sup>35,36</sup> The bath (sheath flow) gas in the DMA was dry  $\text{CO}_2$ , recirculated in a closed loop by a blower. The mobilities to be reported are therefore in  $\text{CO}_2$  under near atmospheric pressure. A constant  $\text{CO}_2$  flow of  $0.7 \text{ L min}^{-1}$  was introduced to the closed gas circuit, of which  $0.5 \text{ L min}^{-1}$  was sampled by the mass spectrometer. The remaining  $0.2 \text{ L min}^{-1}$



exited the inlet slit of the DMA toward the ionization chamber, in the form of a counterflow stream. The use of CO<sub>2</sub> reduces the high tendency of purely aqueous electrosprays to produce corona discharges. The counterflow gas opposes entry of neutral vapors into the DMA and contributes to solvent evaporation of the ES drops before they enter the DMA. The temperature of the sheath gas, measured at the DMA inlet, was 34.5 °C. A voltage difference from 1.5 to 4.5 kV was applied between the two DMA electrodes to selectively transmit ions of controlled mobilities into the mass spectrometer. A mass spectrum was acquired at each fixed voltage, then the voltage was stepped up (typically by 5 V) and a new mass spectrum was acquired. The full set of mass spectra corresponding to the various voltages was then stored as a single DMA-MS file. The declustering potential, DP (voltage difference between the inlet pinhole and the skimmer), in the mass spectrometer ion path was between 50 and 350 V.

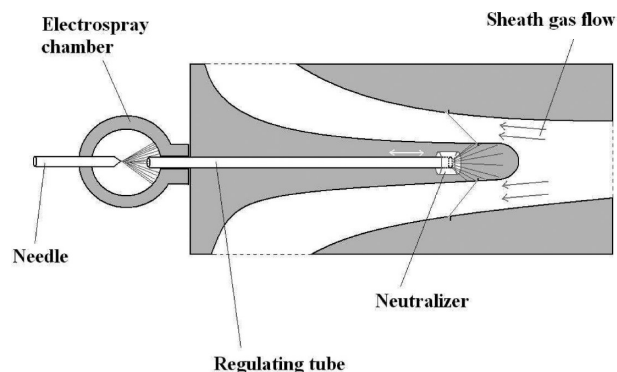
Since the DMA is a linear mobility spectrometer,<sup>35,36</sup> it was possible to determine the absolute mobility,  $K_i$ , of ions appearing in a spectrum at voltage  $V_i$  by a simple calibration consisting of measuring the voltage  $V_s$  between DMA electrodes that was necessary to transmit a standard ion of known mobility  $K_s$ . This calibration was carried out daily following experiments. The strict inverse relation between mobility and voltage then gives

$$K_i = \frac{K_s V_s}{V_i} \quad (4)$$

The ion used as the mobility standard was tetraheptylammonium<sup>+</sup> (THA<sup>+</sup>), whose mobility in ambient temperature CO<sub>2</sub> has been reported by Ude<sup>37</sup> as  $K_s = 0.664 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . No mobility standard was available at the DMA temperature of 34.5 °C, but the mobility shift must be intermediate between the polarization limit ( $K_s \sim T/p$ ) and the hard sphere limit ( $K_s \sim T^{1/2}/p$ ). Therefore,  $K_s \sim T^{(3\pm 1)/4}$ . Using a correction proportional to  $T^{3/4}$  would increase  $K_s$  by 2.5%, with a maximal error below 1%. Accordingly,  $K_s = 0.681 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  ( $\pm 0.83\%$ ). Using this value of  $K_s$  in eq 4 yields protein mobilities at 34.5 °C. To compare them with measurements taken at room temperature, we assume the protein mobility varies approximately as for a hard sphere ( $K_i \sim T^{1/2}/p$ ) and, therefore, needs to be divided by  $T^{-1/2}$ , leading finally to a correction factor of  $[T/T_0]^{(1\pm 1)/4}$  if we use the room temperature standard  $K_s = 0.664 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . Taking  $T/T_0 = 307.5/295$ , the mean correction  $[T/T_0]^{1/4}$  would be 1%. Given the small corrections involved and our present interest in relative mobilities, the reported  $K_i$  will be simply based on eq 4 with Ude's uncorrected standard ( $K_s = 0.664 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ). For future quantitative use of our mobility data, one should preferably determine the correct mobility at 34.5 °C or use the above approximate correction ( $K_s = 0.681 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ ). The DMA voltage assigned to each protein ion was the mean value of the Gaussian fitting of the voltage signal associated with the protonated protein ion mass range. The mobility peak including the full mass peak with its various adducts was considerably broader and will be seen below to give a mean mobility  $\sim 1\%$  smaller than the mass-selected protonated peak.

Additional experiments at low ion charge states were carried out without MS on the basis of a cylindrical DMA operating lamina-ly at a relatively high Reynolds number.<sup>38</sup> The sample ions were introduced into the DMA after passing through a Ni-63 radioactive source, termed the *neutralizer*, used to reduce the

initial charge state. As shown in Figure 3, the electrospray was generated in a closed chamber, and the resulting ions were



**Figure 3.** Sketch of the apparatus for charge-reduced mobility measurement.

extracted toward the DMA through a stainless steel tube (length, 127 mm; o.d., 3.175 mm; i.d., 1.27 mm). The opposite extreme of this tube was located at a variable position inside the charge neutralizer, a coaxial cylindrical Ni piece (length, 6.98 mm; o.d., 7.62 mm; i.d., 7.36 mm) coated on its inner face with Ni-63 (10 mCi). High-energy beta particles emitted by the radioactive neutralizer ring ionize the gas inside the active cylinder such that the resulting anions reduce the charge level of the positively charged protein ions. The stainless steel tube delivering the CO<sub>2</sub> carrying the multiply charged proteins could be moved along the direction of its axis within the Ni-63 source. This controlled the level of charge reduction, from none up to a certain maximum that depends on the sample flow rate.

The DMA used was a gift from its developer (Nano-Engineering Corporation; model MiniDMA). It is peculiar in achieving neutralization inside a bullet-shaped electrode, relatively close to the inlet slit (Figure 3). The DMA circuit was operated in recirculation mode, with a sheath flow of CO<sub>2</sub> sufficiently small to permit analysis of low-mobility ions formed at small  $z$ . The DMA ion outlet was connected to an electrometer, followed by a pump drawing a sample outlet flow rate ranging between 0.5 and 2.5 L min<sup>-1</sup>. Mobility calibration in this case is more complex because of the large disparity in mobilities between the singly charged proteins and the THA<sup>+</sup> standard. Because the blower velocity needed to be small for the DMA to cover the smallest mobilities, the DMA temperature was relatively close to room temperature, and was not monitored.

Direct mobility calibration under these conditions is problematic as a result of the substantial width of the THA<sup>+</sup> peak and the presence of other ionic peaks produced by the radioactive source. Calibration was therefore performed indirectly as follows: First, we measured a mobility spectrum of the protein ions with a weak level of neutralization. Immediately after, under the same flow conditions but without neutralization, we obtained a spectrum of THA<sup>+</sup>, which appeared at 1533 V [i.e.,  $V_s K_s = 1533 \times 0.664 \text{ cm}^2/\text{s}$  in eq 4]. To verify the identity of this standard, we also measured the mobility of the dimer ion (THABr)/THA<sup>+</sup> and checked that the ratio of mobilities between the monomer and dimer took the value reported by Ude.<sup>37</sup> This then provided the mobilities of several high-charge protein ions. The sheath gas velocity in the DMA was then reduced, the level of charging was decreased, and a new mobility spectrum was taken. The mobility scale

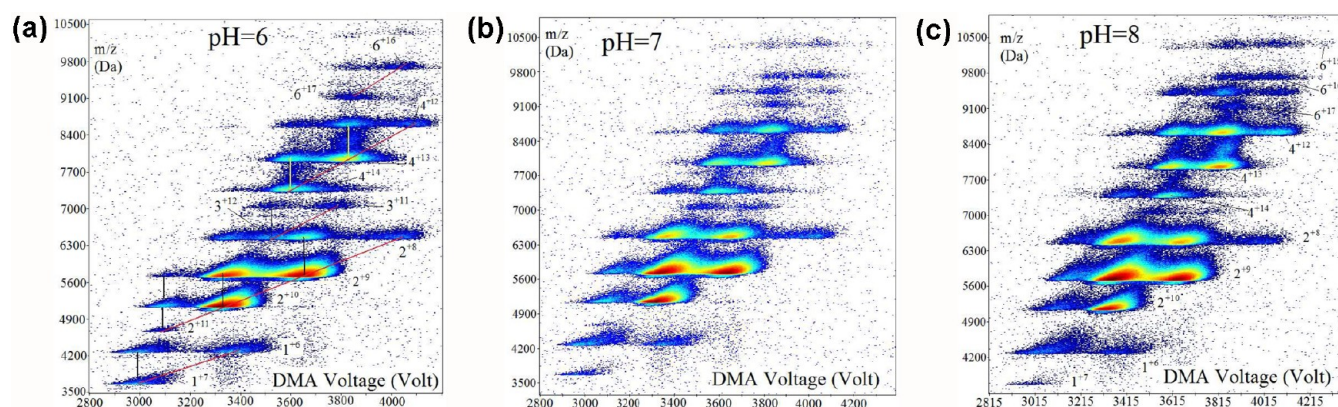
Table 1. Computed Mobilities,  $K_n$ , for Symmetric Aggregates of  $n$  Touching Spheres

$n^d$	(a) $d_g = 0$															
	1	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	2 <sup>b</sup>	3 <sup>b</sup>	4 <sup>b</sup>	5 <sup>b</sup>	6 <sup>b</sup>	2 <sup>c</sup>	3 <sup>c</sup>	4 <sup>c</sup>	5 <sup>c</sup>	6 <sup>c</sup>
$nK_n/K_1$	1	1.07	1.133	1.186	1.256	1.312	1.094	1.187	1.279	1.364	1.445	1.091	1.182	1.272	1.355	1.434
$K_1/K_n/n^{2/3}$	1	1.178	1.273	1.338	1.362	1.385	1.152	1.215	1.241	1.254	1.257	1.155	1.22	1.248	1.262	1.267

$n^d$	(b) $d_g/d_{\text{monomer}} = 0.1$															
	1	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	2 <sup>b</sup>	3 <sup>b</sup>	4 <sup>b</sup>	5 <sup>b</sup>	6 <sup>b</sup>	2 <sup>c</sup>	3 <sup>c</sup>	4 <sup>c</sup>	5 <sup>c</sup>	6 <sup>c</sup>
$nK_n/K_1$	1	1.095	1.178	1.249	1.339	1.405	1.121	1.239	1.354	1.457	1.555	1.118	1.234	1.346	1.447	1.543
$K_1/K_n/n^{2/3}$	1	1.151	1.224	1.271	1.277	1.293	1.124	1.164	1.172	1.174	1.169	1.126	1.169	1.18	1.182	1.178

<sup>a</sup>Perfectly elastic and specular reflections. <sup>b</sup>Complete thermal accommodation and random re-emission. <sup>c</sup>Epstein's model taking 9 and 91% of the gas molecules to rebound according to the first and second models, respectively. <sup>d</sup> $n = 2$ , two touching spheres;  $n = 3$ , touching spheres on the vertices of an equilateral triangle;  $n = 4$ , touching spheres on the vertices of a regular tetrahedron;  $n = 5$ , touching spheres on the vertices of two regular tetrahedra joined at the base;  $n = 6$ , touching spheres on the vertices of an octahedron.



**Figure 4.** DMA-MS spectra for concanavalin A at various pH. Ion abundance is shown in logarithmic color scale versus DMA voltage (horizontal axis, inversely proportional to ion mobility,  $K$ ) and  $m/z$  (vertical axis). The protein is electrosprayed at 50  $\mu\text{M}$  concentration from aqueous 100 mM  $\text{Et}_3\text{AF}$ , with triethylamine addition for pH control. The inclined lines drawn on the left figure correspond to monomers, dimers, trimers, tetramers, and hexamers from bottom to top. MS declustering potential: DP = 200 V.

under the new condition was determined by matching the coincident series of protein peaks appearing in the two spectra.

### 3. MOBILITY CALCULATIONS FOR AGGREGATES OF SPHERES

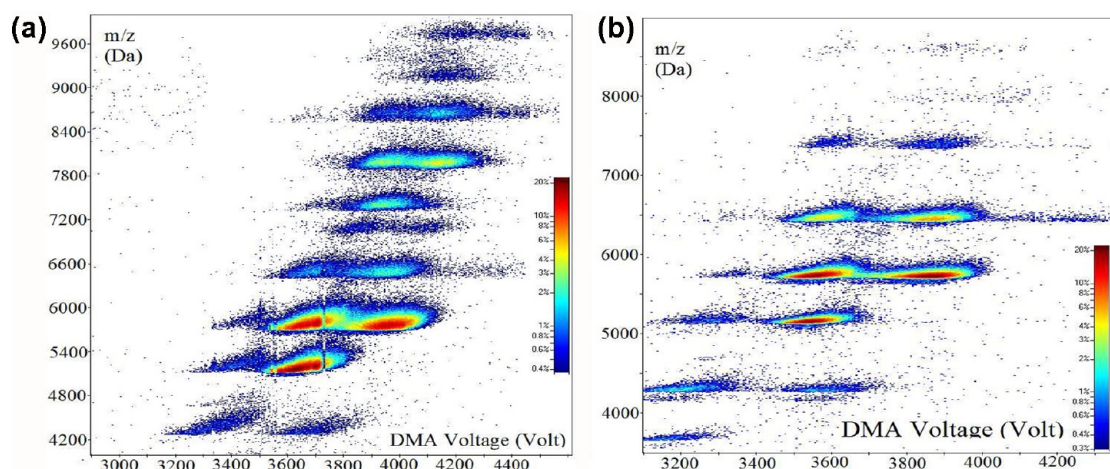
Frictional forces acting on close-packed aggregates of 1–6 identical spherical particles (see Table 1 for the precise configurations) translating at a constant relative speed,  $U$ , with respect to the mean gas velocity were computed as the net flux of linear momentum over an imaginary bounding sphere (or any other convex surface) enclosing the particle and moving with it. Using a test particle Monte Carlo method,<sup>39</sup> the net momentum flux is, in turn, evaluated as the average over the individual contributions of a large number of simulated gas molecule trajectories obtained as follows: First, an initial gas molecule velocity and point of departure from the enclosing sphere are randomly sampled from an equilibrium distribution of gas velocities, which in a frame attached to the particle becomes a shifted Maxwellian centered around  $U$ . The trajectory of this gas molecule is then followed until it leaves the domain of integration, when the momentum it has transmitted to the ion is computed. We note that in the free molecule limit of interest here, gas molecules follow linear trajectories except upon collision with the ion surface. Two surface scattering models were used. In the first, collisions are treated as elastic and specular. In the second, colliding gas molecules are diffusely re-emitted after fully accommodating at the particle surface. Neither of these scattering models is realistic for aggregates of proteins (even of spherical proteins).

However, although they differ drastically, both predict qualitatively similar trends as a function of the number of aggregates, so all provide a semiquantitative measure of the ratios between the mobilities of the cluster of  $n$  spheres and a single sphere.

The model used by Epstein to match Millikan's experimentally measured  $\xi = \xi_M = 1.36$  for spheres is a hybrid of the two models considered, taking 9 and 91% of the gas molecules to rebound according to the first and second models, respectively. Friction forces are linear with particle velocity for low Mach numbers (particle velocity small compared with the gas mean molecular speed) and can be expressed in terms of a drag tensor. Determining the mobility requires consideration of the coupling between rotational and translational motion; however, if all orientations are assumed equiprobable, the mobility can be computed as  $(K_1^{-1} + K_2^{-1} + K_3^{-1})/3$ , where  $K_1$ ,  $K_2$ , and  $K_3$  are the eigenvalues of the drag tensor. This simplifying assumption is relatively unimportant for the most relevant cases of the single sphere and the tetrahedrally organized four-sphere aggregate, in which symmetry considerations<sup>31</sup> as well as numerical results show that the drag tensor is isotropic ( $K_1 = K_2 = K_3$ ). For such configurations, the friction is independent of ion orientation and is in the direction opposing particle motion. In addition, the mobility becomes just the inverse of the single scalar defining the drag tensor.

The leading order contribution to the drag vanishes in the low-drift Mach limit. This results in poor convergence of the Monte Carlo algorithm.<sup>40</sup> Here, we avoid this difficulty by





**Figure 5.** Effect of declustering voltage DP on DMA-MS spectra for concanavalin A in 100 mM Et<sub>3</sub>AF buffer: (left) DP = 150 V and (right) DP = 250 V.

linearizing the shifted Maxwellian distribution function, which gives the low-field mobility limit.<sup>41</sup> Results from these calculations are included in Table 1a under the assumption that the effective diameter of the gas molecule is zero ( $d_g = 0$ ). Similar calculations are included in Table 1b for the case when  $d_g$  is 10% of the diameter of the monomeric sphere, which is quite close to the expected ratio for CO<sub>2</sub> and the monomer of concanavalin A.

#### 4. EXPERIMENTAL RESULTS

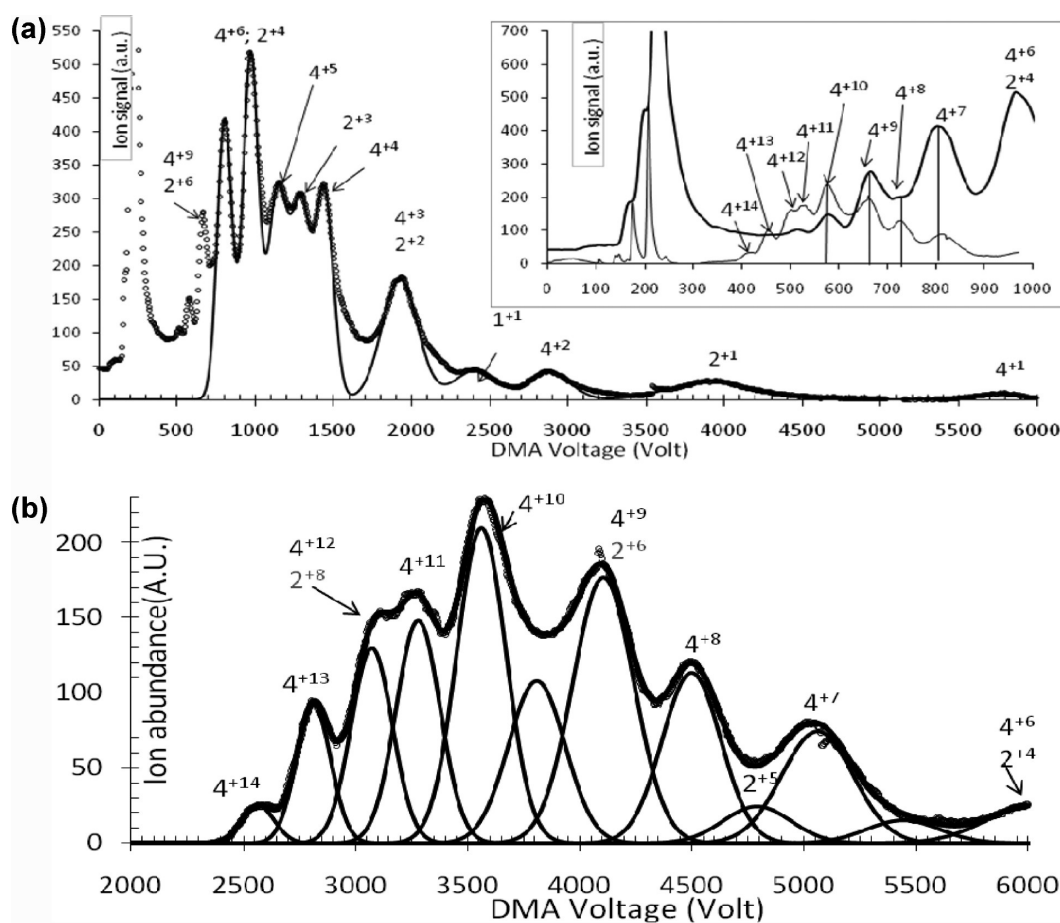
**Effect of pH.** Concanavalin A is known to take its native tetrameric form at pH > 7 and to be predominantly dimeric under acidic conditions.<sup>42</sup> Some of our mobility measurements will be taken for convenience under acid or neutral conditions. For this reason, we have first studied electrospray ions at varying pH and concluded that the gas phase mobility is independent of pH. To reach relatively low charge states, a buffer of 100 mM Et<sub>3</sub>AF was used, adding varying quantities of triethylamine to control the pH from 6 (0% triethylamine) to 8 (0.02% triethylamine). Higher charge states (not shown) were obtained from a 100 mM ammonium acetate buffer, where the pH was controlled by adding NH<sub>3</sub>, from 0% (pH = 6) to 0.3% (pH = 8). Two-dimensional (2D) DMA-MS spectra corresponding to the 100 mM Et<sub>3</sub>AF buffer are shown in Figure 4 for pH 6, 7, and 8. These images give ion abundance (logarithmic color scale) versus  $m/z$  (vertical axis) and DMA voltage (horizontal axis, proportional to inverse ion mobility  $1/K$ ). The various peaks are identified as  $n^{+z}$ , in terms of their state of aggregation,  $n$ , and number of elementary charges,  $z$ . Figure 4 shows that the ion mobilities for all the aggregates are essentially independent of pH. The same is found for the ammonium acetate buffer. For this reason, subsequent measurements will use simple buffers without addition of ammonia or amine. Although the dimer series  $2^{+z}$  is dominant in all cases, a substantial increase in the abundance of the tetramer is clear at increasing pH. Another noteworthy feature of Figure 4 is the considerably narrower peaks of the tetramers with respect to the dimers (in both mass and mobility). Note also two weak trimer peaks at  $\sim 7000$   $m/z$  and three hexamers above 8400  $m/z$ .

**Charge Stripping and Declustering Voltage.** The DMA-MS spectra are complicated by the frequent loss of one or more charges following the mobility measurement but prior

to the mass measurement. We have previously described this effect, which is due to collisional heating of the ions, partly in the radio frequency (RF) electric fields within the MS ion guide,<sup>34–36,43</sup> and partly in the free jet due to the declustering potential difference, DP, between the critical orifice and the skimmer. As a result, at a given mobility in a DMA-MS spectrum, one encounters the parent peak at its original mass,  $M$ , and charge state,  $z_0$  (at  $m/z = M/z_0$ ), as well as one or more ghost peaks directly above the original peak (at  $m/z \sim M/(z_0 - 1)$ ,  $M/(z_0 - 2)$ , etc.). Some of these transitions are indicated with vertical lines in Figure 4a. Diagonal red lines are shown joining the main (nonghost) ions for each state of aggregation. Note that the mobility is measured in CO<sub>2</sub> at atmospheric pressure and is not affected by subsequent activation processes leading to charge stripping. Only the measured  $m/z$  is affected. Since the mobility peaks are relatively narrow, this complication rarely leads to fatal loss of information in our measurements. In some cases, charge stripping provides extra information. For instance, of the two relatively weak ions appearing at  $m/z = 6894$  (most distinctly in Figures 4a and 5a), the least mobile corresponds to the trimer with 11 charges,  $3^{+11}$ , and the most mobile is the ghost of  $3^{+12}$ . The parent ion,  $3^{+12}$ , cannot be distinguished directly because it falls right on top of two peaks at  $\sim 6400$   $m/z$ , corresponding to ghost peaks for  $2^{+10}$  and  $2^{+9}$ . In this case,  $3^{+12}$  is undetectable directly, but its mobility can be determined thanks to its ghost peak.

The effect of the declustering voltage can be seen in Figure 5, taken in a 100 mM Et<sub>3</sub>AF buffer at two different declustering voltages: 250 V on the right and 150 V on the left. In addition to enhancing charge stripping, an increase in the DP considerably sharpens the mass peaks by removing adducts, which focuses them closer to the mass of the purely protonated ion (right panel). There is a slight but clear shift of some of the ghost peaks toward the right. This does not mean that charge stripping makes the structure of these ions less compact, since the mobility is measured before charge reduction. It shows, rather, that some of the ions initially less mobile (perhaps with a higher level of clustering or carrying more protonated amines and fewer protons) are more prone to lose charge.

**Mobility Measurements at Reduced Charge States.** The lowest charge states measured, going down to  $z = 1$ , were obtained with a buffer of 100 mM Et<sub>3</sub>AF by means of charge reduction in the radioactive neutralizer, followed by a DMA (no

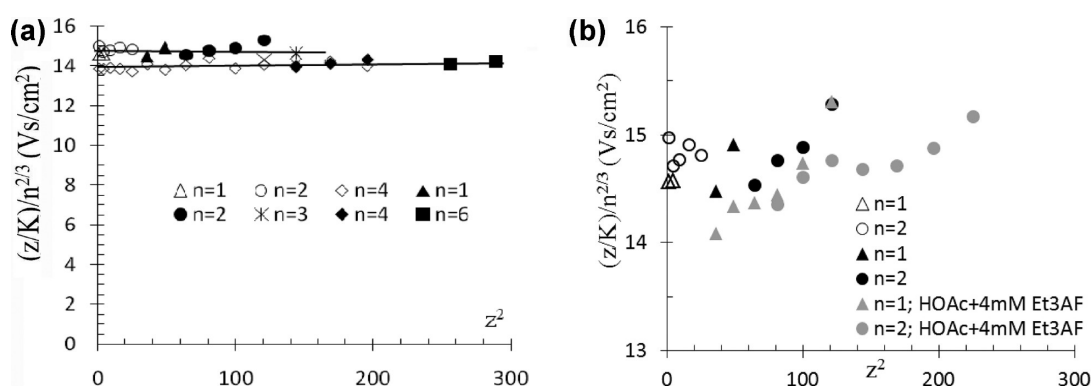


**Figure 6.** Charge-reduced mobility spectra (ion abundance in arbitrary units versus DMA voltage in volts) of concanavalin A electrosprayed from 100 mM Et<sub>3</sub>AF buffer, showing mainly tetramer and dimer ions: (top) wide mobility range with high level of neutralization; (bottom) lower neutralization level at higher resolution.

**Table 2.** Peak Voltages,  $V$ , Identified for Charge-Reduced Ions of Concanavalin A, and Corresponding Approximate Charge States,  $V_n/V$ , Seen To Approximate Integer Numbers<sup>a</sup>

$z$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$V$ (volt)														
$n = 2$	3916	1924	1287.5	975	775									
$n = 4$	5753	2873	1924	1438	1139	974	819	728	664	576	530	497	455	415
$V_n/V$														
$n = 2$	1	2.035	3.042	4.016	5.053									
$n = 4$	1	2.002	2.99	4.002	5.051	5.91	7.02	7.90	8.66	9.99	10.86	11.57	12.64	13.86

<sup>a</sup> $V_4 = 5753$ ,  $V_2 = 3910$ .  $V_s = 1533 \times 0.1619$  V.

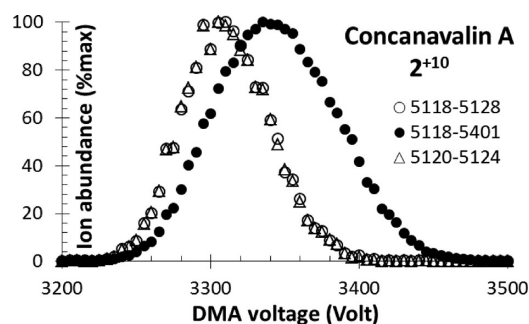


**Figure 7.** Compactness parameter  $z/K/n^{2/3}$  versus square charge state  $z^2$  for clusters  $C_n$  of concanavalin for various states of aggregation  $n$ .  $z/K/n^{2/3}$  is proportional to the collision cross section normalized by the surface area of a sphere with the cluster volume. Left, 100 mM Et<sub>3</sub>AF: open symbols are for charge-reduced data; solid symbols and asterisks are for natural charge states. Right: Close up of data for monomers and dimers compared with those in an 87 mM aqueous acetic acid buffer with 4 mM Et<sub>3</sub>AF (gray symbols).

MS). Figure 6 shows two such mobility spectra, covering respectively the lower (top panel) and higher (lower panel) charge states. The inset in Figure 6 superposes the two spectra after multiplying by 0.162 the horizontal axis of the high resolution spectrum (lower panel). The excellent matching found serves to calibrate the low- $z$  spectrum (top panel) with respect to the other (lower panel), calibrated directly against the  $\text{THA}^+$  standard. The relatively involved assignment of ions is discussed in the Appendix and is summarized in Table 2 for the dimer and tetramer series. Briefly,  $1^{+1}$ , appearing at 2400 V, is the only identifiable member of the  $n = 1$  series. The only directly identifiable members of the  $n = 2$  series are  $2^{+1}$ ,  $2^{+3}$ , and  $2^{+5}$ , although  $2^{+2}$  and  $2^{+4}$  have mobilities almost coincident with  $4^{+3}$  and  $4^{+6}$ , and their indirectly obtained approximate mobilities are also included in Table 2.

Figure 7 compares the natural charge results drawn from Figure 4a (solid symbols) with those obtained from Figure 6 for charge-reduced ions (open symbols), both for aqueous solutions with 100 mM  $\text{Et}_3\text{AF}$  at pH = 6. The ratio  $z/K$ , proportional to the cross section  $\Omega$ , is normalized by  $n^{2/3}$ , proportional to the surface area of a sphere with the cluster volume. The ratio  $z/K/n^{2/3}$  represented is therefore a measure of the compactness of the clusters. Note that this measure of compactness is not exact, not even for the case of spheres, since the correction for  $d_g$  suggested by eq 2 has not been included. However, because the objective of this study is to rely as little as possible on any particular scattering hypothesis,  $z/K/n^{2/3}$  is an ideal model-independent measure of compactness. The computed  $z/K/n^{2/3}$  normalized with the corresponding value for the monomer is included in the third row of Table 1 for  $n = 1$ –6, showing a substantial dependence on  $n$  for aggregates of spheres for all scattering models. This holds both when accounting and not for the finite diameter of the gas molecule, although in the first case, the  $n$  dependence is lost for  $n > 4$ . The fact that the data show only a rather small  $n$  dependence indicates that the real clusters are much more compact than if they were made up of  $n$  undeformed touching spheres. The agreement between the DMA-MS data with natural charge states and the charge-reduced pure DMA data can be tested only for the cases in which the two data types are available ( $n = 1, 2, 4$ ). It is excellent, given the differences in the two measurements. A first variation already quantified at  $\sim 1\%$  is due to the different temperatures. Furthermore, the DMA-MS mobility is based only on the protonated ions, while the pure DMA measurement includes all the impurities clustered on the protein.

The positive slopes in the core of the broad structures of the DMA-MS peaks in Figure 5a show that clustering (increased  $m/z$  for a given peak) shifts the mobility distribution to higher voltages (smaller  $K$ ). The magnitude of this effect is quantified in Figure 8 by comparing mass-resolved mobility distributions for the dimer at  $z = 10$  on the basis of either a narrow mass range (open circles) or the full mass peak including all clusters (solid circles). The narrow mass range corresponds approximately to the full isotopic envelope of the purely protonated protein, although there is no measurable broadening of the mobility peak between  $m/z$  windows of 4 and 10 Da. Both show  $\text{fwhm} = 2.1\%$ , unusually narrow for a protein, and approaching the resolving power of the DMA. The mobility peak, including all impurities clustered around the protein, is considerably wider ( $\text{fwhm} = 3.1\%$ ), but its maximum is displaced by only 1% from that for the mass-selected purely protonated ion.



**Figure 8.** Mobility spectra for the  $2^{+10}$  ions of Figure 4a for the  $m/z$  windows indicated in the caption, showing a  $\sim 1\%$  decrease in mobility for the broad range 5118–5401  $m/z$ , including all clusters.

Some of the data scatter in Figure 7 is readily understood in terms of limitations of the measurement system. For instance, the high dispersion of the  $n = 1$  data for  $z = 6$  or 7 relative to  $z = 1$ –2 is due to imperfect definition of the mobility peak resulting from the wide mobility distributions and the small ion intensity in the original DMA-MS data (Figures 4a and 5a). The open diamonds for the tetramer show anomalously high ordinates at  $z = 9$  and 12 (see also Table 2). The  $z = 9$  datum can be seen in Figure 6 to have an unassigned shoulder to its right, which evidently limits the accuracy with which  $K$  can be inferred. The  $4^{+12}$  peak has a similar problem because it is barely resolved from  $4^{+11}$ .

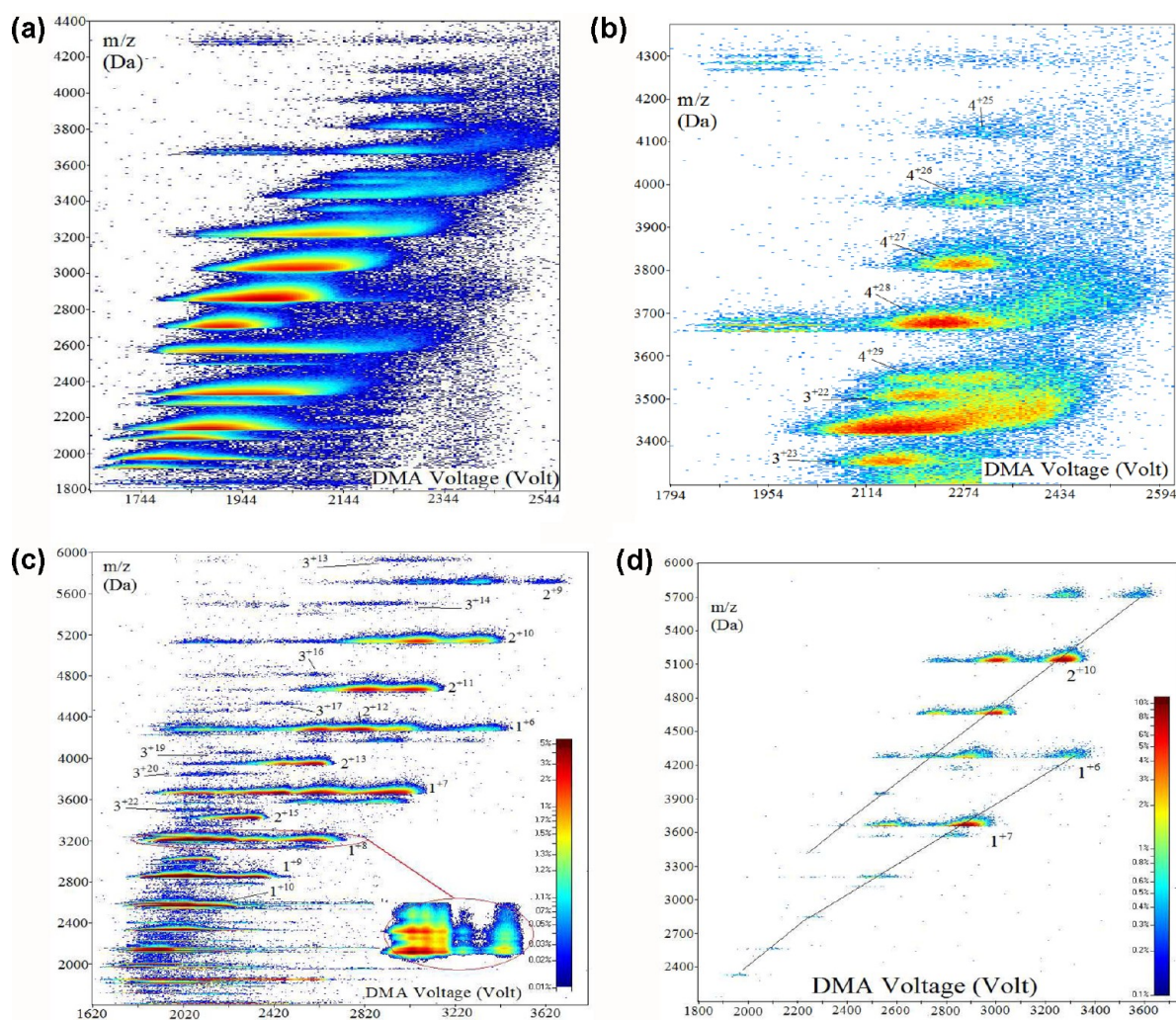
More important than the slight differences between high and small  $z$  data seen in Figure 7 is the fact that both natural and  $^{63}\text{Ni}$ -reduced charge state measurements yield very much the same cross section per unit cluster area ( $\sim z/K/n^{2/3}$ ) for  $n = 1, 2$ , and 4 clusters. The few charge states available for the trimer and hexamer ions also fall within the same narrow range of  $z/K/n^{2/3}$  values. The measured low  $z$  values of  $z/K/n^{2/3}$  normalized with the value for the monomer are 1.021 and 0.95 for  $n = 2$  and 4, respectively. The predicted ratios for symmetric aggregates of spheres are 1.178–1.152 for  $n = 2$  and 1.338–1.241 for  $n = 4$ , with the range of values shown associated with the extreme elastic and inelastic models used (footnotes a and b in Table 1). Particularly noteworthy is the fact that the tetramer has a compactness parameter,  $z/K/n^{2/3}$ , 5% larger than the monomer, whereas its tetrahedral form is expected to have a value between 24 and 34% smaller. In view of the substantial  $n$  dependence expected in  $z/K/n^{2/3}$  for aggregates of spheres and the relatively modest variation encountered, we conclude that, to a first approximation, all the clusters are densified into spheroidal shapes, with little interstitial volume surviving between the various monomers within an aggregate. The measured mobilities of concanavalin A clusters therefore support our prior conclusion<sup>19</sup> that a large compaction takes place when native protein ions are transferred from neutral aqueous solutions to the gas phase. The structures measured are accordingly characteristic of the gas, rather than simply inherited from pre-existing solution structures.

## 5. SPRAYING CONCAVALIN A FROM AQUEOUS ACETIC ACID WITH $\text{Et}_3\text{AF}$

One hypothetical explanation for our findings here and prior studies discussed in the introduction can be formulated as follows:

*Gas phase equilibrium conformation hypothesis: The gas phase cross section,  $\Omega$ , of a protein electrosprayed from a*





**Figure 9.** DMA-MS data for concanavalin A in 87 mM aqueous acetic acid buffer with slight additions of  $\text{Et}_3\text{AF}$ . Top, no  $\text{Et}_3\text{AF}$ , including on the right a detail on trimers and tetramers. Bottom left: 2 mM  $\text{Et}_3\text{AF}$ , with a detail of the mass peak at  $\sim 3200$   $m/z$ . Bottom right: 4 mM  $\text{Et}_3\text{AF}$ , with lines distinguishing the monomer and dimer series.

solution is a property of the protein, which may depend on its charge state, but is independent of its prior structure in solution.

$$\Omega = \Omega(z) \quad (5)$$

In the present section, we test whether the *equilibrium hypothesis* just advanced also applies to concanavalin A dissolved in denaturing solutions, as already shown to be the case for cytochrome *c* and, with some activation requirements, also apomyoglobin. An alternative *counter-hypothesis* could be formulated as follows:

*Counter-hypothesis: The solution structure of a dissolved protein is preserved after electrospraying it into the gas phase.*

To further test these two hypotheses, we shall follow the approach of Jarrold and colleagues of acidifying the solution while moderating the final charge by addition of a base, although our base will be introduced as relatively small quantities of  $\text{Et}_3\text{AF}$  in the solution. One motivation for this test is provided by the observation of shape transitions in ions of polyethylene glycol (PEG),<sup>44,45</sup> which are much better defined than those found in protein experiments.<sup>10–12</sup> These critical transitions can be studied with greater flexibility in

polymers than in proteins because (i) they take place at less extreme charge states achievable with many buffers; (ii) the solution structure is not critically dependent on the solvent and the buffer, nor is it biologically relevant or contentious; (iii) the mass variable can be controlled in a quasi-continuum fashion, providing a precise definition of the  $z^*(m)$  relation; (iv) the mobility peaks are very narrow, revealing typically only one dominant structure at each  $z$  and  $m$ ; (v) not just one but many critical shape transitions are clearly defined;<sup>44,45</sup> (vi) the rich structural landscape found in polymers corresponds undoubtedly to equilibrium shapes adopted in the gas phase, as shown by a variety of observations<sup>44</sup> as well as modeling studies.<sup>14,46–48,45</sup>

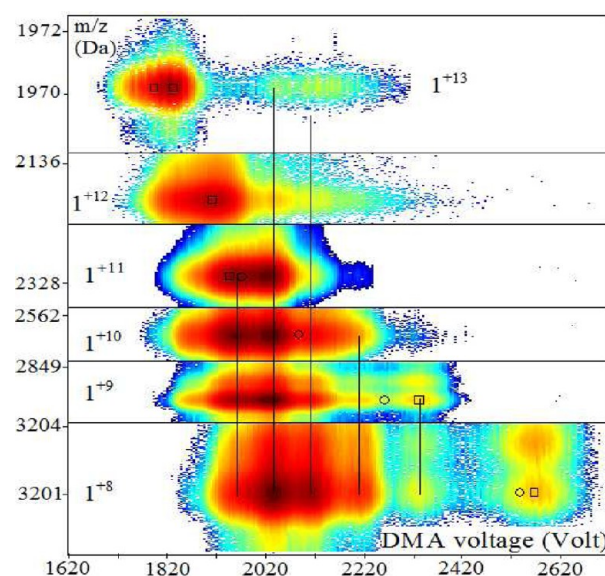
We have accordingly run a series of electrosprays of concanavalin A in 87 mM aqueous acetic acid solutions, including 0, 2, and 4 mM  $\text{Et}_3\text{AF}$ . Corresponding DMA-MS spectra are shown in Figure 9. The purely acidic buffer produces high charge states, mostly with rather broad mobility structures (Figure 9a). Exceptionally, as shown in the close-up of Figure 9b, the tetramers exhibit relatively narrow mobility distributions. Interestingly, although acidification was seen in the experiments at varying pHs to moderate the level of aggregation, our present much lower pH does not suppress the

formation of tetramers. In view of the decreasing complexity of Figure 9a–d as the concentration of Et<sub>3</sub>AF is increased, we shall discuss them in order of decreasing salt concentration.

Addition of 4 mM Et<sub>3</sub>AF (Figure 9d) yields low charge states quite comparable to those seen in the pure Et<sub>3</sub>AF buffer (down to 1<sup>+6</sup> in Figure 5), including a drastic reduction of signal, although providing a wider range of high charge ions (up to 1<sup>+11</sup>, vs 1<sup>+7</sup> in Figure 4). The data for the acid plus 4 mM salt solution are compared in Figure 7b with those in the 100 mM Et<sub>3</sub>AF acid-free buffer. The figure shows almost identical mobilities, with slightly more compact conformations for the new buffer. It is clear from the data in Figure 9, top, that the proteins are denatured in the 87 mM acidic buffer, and it is unlikely that this situation would be entirely reversed in solution by a small addition of Et<sub>3</sub>AF. It is equally clear that the main effect of Et<sub>3</sub>AF in nondenaturing buffers is to reduce the charge state of the protein ions, yet this effect almost surely does not take place in solution. The amine reduces the charge state of the ions either as they are released from the drop or shortly thereafter (presumably as a result of the much weaker binding of a protein to a protonated amine than to a proton<sup>49,50</sup>). This is a surface effect and must therefore be also active in our acidic solutions. We are therefore faced with two situations with radically different solution environments producing comparable charge states and almost identical gas phase structures. The experiment with 87 mM acid and 4 mM Et<sub>3</sub>AF therefore strengthens the *gas phase equilibrium hypothesis* and weakens the *counter-hypothesis*, very much as in prior work with cytochrome *c* and apomyoglobin.

The spectrum of Figure 9c for 2 mM addition of Et<sub>3</sub>AF includes essentially the same low-charge states as the 4 mM salt solution, both situations showing fairly close mobilities for these low charges. The high mobility end from the spectrum of the 2 mM solution, however, shares many of the complexities of the purely acidic solution. These peculiarities include long horizontal features extending broadly toward the left, particularly at *m/z* associated to 1<sup>+z</sup> ions, and going down to *z* values as small as 5. There are also long unresolved features extending toward the right. There is, however, an important difference in that these broad structures are featureless in the purely acidic solution but include numerous discernible peaks at 2 mM salt addition (see inset in Figure 9c, magnifying the region near *m/z* ~ 3200 *m/z*). The various vertically stacked peaks in the inset are associated with different charging agents, starting at the bottom with the purely protonated protein (8H<sup>+</sup>). Up to five such vertically stacked peaks can be seen on the elongated high-mobility feature on the left of the inset. The peak most to the right in that inset corresponds to 1<sup>+8</sup>.

Assignment of peaks in Figure 9c (2 mM Et<sub>3</sub>AF) is relatively straightforward only up to *z* = 9. The presence of additional resolved mobility peaks in this figure, absent from Figure 9d (4 mM Et<sub>3</sub>AF), is evident in the close-up of the DMA-MS spectra assembled in Figure 10, as well as in the mass-selected mobility spectra of Figure 11. Assignment of *n* and *z* for *z* ≥ 10 is nontrivial as a result of this real multiplicity of structures, further confounded by their vertical repetition due to charge stripping. Figure 10 includes square symbols indicating the approximate location in (*m/z*, *V*) space of the protein monomer ions identified in the simpler spectrum of Figure 9d (4 mM Et<sub>3</sub>AF). These more reliably assigned peak positions from the 4 mM Et<sub>3</sub>AF solution are not reproduced exactly in the 2 mM Et<sub>3</sub>AF solution. They are, nonetheless, relatively close, except for a nontrivial difference of ~2.7% for 1<sup>+9</sup>. The

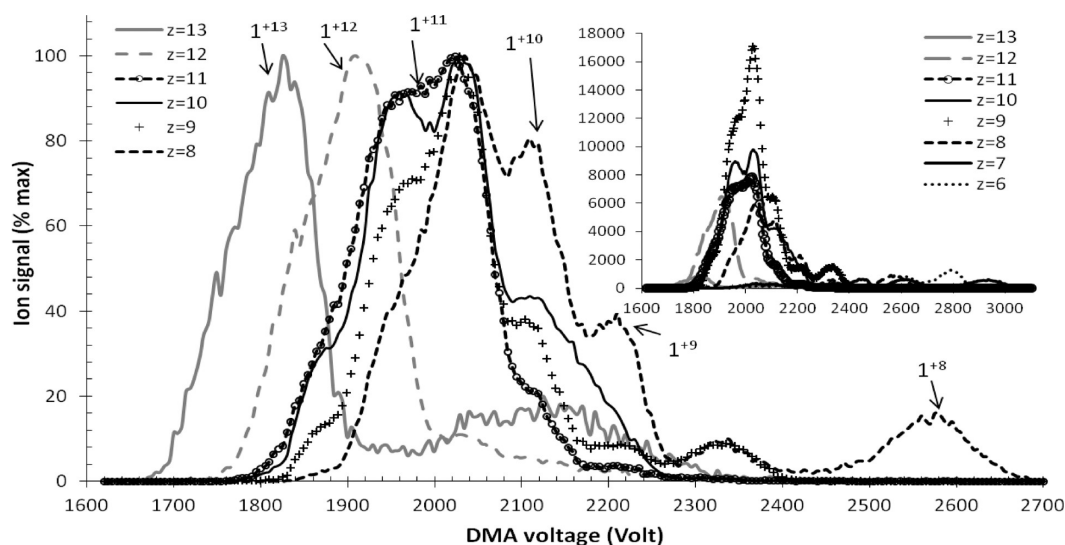


**Figure 10.** DMA-MS mobility spectra from concanavalin A in 87 mM acetic acid solution with 2 mM Et<sub>3</sub>AF. Each of the horizontal strips included is for a narrow *m/z* range with lower bound associated with the purely protonated form of 1<sup>+z</sup> ions. Ion assignments are indicated with squares, and conformations from the 4 mM Et<sub>3</sub>AF solution are marked with circles.

peak positions for *z* ≥ 8, however, are somewhat ambiguous because these charge states are represented by very few ion counts for the 4 mM Et<sub>3</sub>AF solution (Figure 9d). The peaks appearing above and to the right of each of these squares cannot be ghost ions resulting from charge loss. They must rather be associated to low-mobility ions characteristic of the acid solution, which have not been able to achieve the more compact conformations facilitated by the 4 mM Et<sub>3</sub>AF buffer. For instance, the narrow peak seen at 2330 V and at the *m/z* of 1<sup>+10</sup> (very weak), 1<sup>+9</sup> and 1<sup>+8</sup> originates at *z* ≥ 10, and has a considerably smaller mobility than the 1<sup>+10</sup> ion appearing in the buffer with 4 mM Et<sub>3</sub>AF. At the highest charge state at which its ghosts may be identified (1<sup>+10</sup>), this peak is represented almost exclusively by the protonated ion, suggesting that, as it went through the DMA, it contained several Et<sub>3</sub>NH<sup>+</sup> ions that were readily lost via *z* = 10 → 9 → 8 transitions in the ion guides of the mass spectrometer. Another prominent ion appearing in the 2 mM Et<sub>3</sub>AF buffer but not at 4 mM peaks at 2030 V. It originates at a rather small mobility probably at *z* = 13, yet undergoes many charge loss transitions. The inset in Figure 11 shows mobility spectra without normalizing ion abundance, revealing that this unassigned ion is dominant at charge states 8–11.

The observation of more than one structure at given *z* gives the impression that several protein conformations coexist at the largest charge states. Acid buffers have often been seen to yield broad mobility peaks of mass-selected ions, previously interpreted as including several or even many different configurations. This is very much what we observe in the absence of Et<sub>3</sub>AF. The effect of the 2 mM Et<sub>3</sub>AF buffer to separate these numerous peaks is, however, more unusual and sheds some new light. First, only a few within this multitude of peaks truly belong to a given charge state, most resulting from charge loss transitions in the ion guides. Although this phenomenon is highly dependent on the details of the mass spectrometer, the wide use of RF focusing between the IMS





**Figure 11.** Mass-selected mobility spectra in 87 mM acetic acid solution with 2 mM Et<sub>3</sub>AF, primarily for the monomer of concanavalin A. Many more peaks appear than in Figures 8 and 4. The inset gives absolute signals.

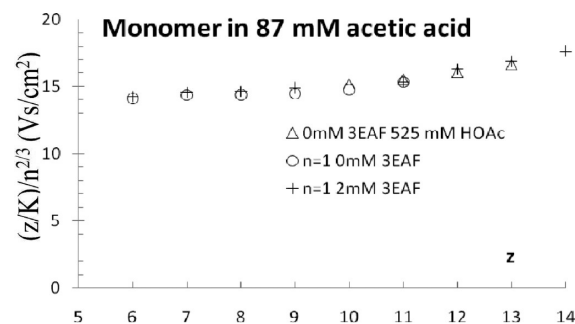
and the TOF-MS stage of other modern instruments suggests that this complication may similarly apply to many earlier IMS-MS investigations on protein structure (though probably not to the early studies at Jarrold's and Clemmer's groups).

Further insights on the origin of the extra 2 mM solution peaks missing at 4 mM salt follow from the following observations. As already discussed, the extra conformations have lower mobilities than the compact ions of the same  $z$  seen at 4 mM salt. They are also far more prone to lose charge, as evident from their many vertical repetitions. For instance, the monomer peak at 2210 V, originating probably at  $z \geq 13$ , reaches maximal abundance at  $z = 9$  and is distinctly present even at  $z = 6$ . In contrast to these 7 charge loss events, the compact ions from the 4 mM salt solution rarely lose more than two charges. It is therefore apparent that these extra ions having high charge levels take elongated shapes of low mobility. These extended configurations expose a number of weakly bound charges (probably Et<sub>3</sub>NH<sup>+</sup>), which are more easily removed by RF heating in the ion guides than from more compact configurations. Interestingly, a relatively simple peak pattern is recovered at the  $m/z$  of  $1^{+13}$ , the highest charge at which an unfragmented protein monomer can be recognized for the 2 mM Et<sub>3</sub>AF buffer.

The unusual stability toward charge loss of the most mobile peak at this mass is manifested in the absence of corresponding ghost peaks below it in Figure 10. This is probably due to its existence only as purely protonated. This most mobile peak at 1970  $m/z$  is clearly associated with the most compact form of  $1^{+13}$ . The high mobility tail to its left (in contrast to its sharp right boundary) is probably the ghost of  $1^{+14}$ , which is not seen at its original mass because it is completely converted into  $1^{+13}$  within the MS. One can with modest error assign the peak voltage of  $1^{+14}$  at  $\sim 1770$  V. The dominant feature at the  $m/z$  corresponding to  $1^{+12}$  presents many of the trends just noted for  $1^{+13}$ , with a high mobility tail from the ghost of  $1^{+13}$  and a sharper right tail only slightly contaminated by weak ghosts from lower mobility structures. This major peak can be safely assigned to the most compact structures of  $1^{+12}$ . It differs from  $1^{+13}$  in including not just the protonated ion but also numerous other charging agents and also (consequently) in yielding numerous ghost peaks at higher masses. The situation at the

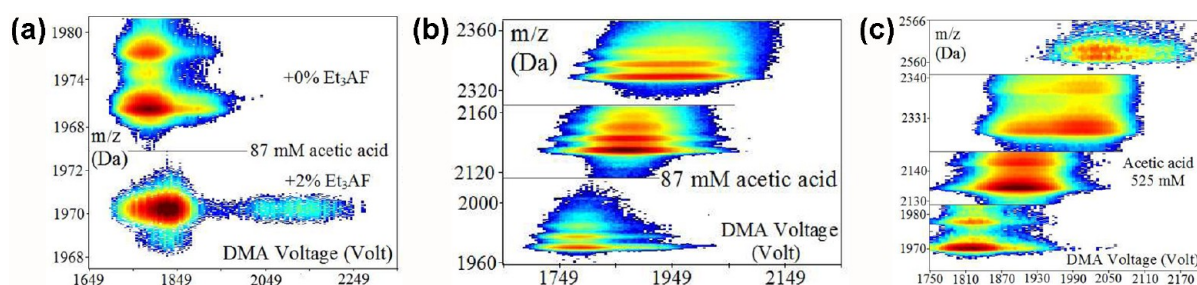
mass of  $1^{+11}$  is now substantially different due to contamination at 2030 V from the ghost of the low-mobility ions seen in Figure 10 at the  $m/z$  of  $1^{+13}$ . This ion originates, in fact, at  $z \geq 14$  because it is clearly present at 1830  $m/z$ , the expected  $m/z$  for  $1^{+14}$ . Unlike the compact  $1^{+14}$  ion, this low-mobility ion is not completely converted. The same feature seen at the  $m/z$  of  $1^{+13}$  and  $1^{+14}$  in the voltage range from 2010 to 2200 V is also present at 1708  $m/z$ , the expected mass of  $1^{+15}$ . These high charge states therefore do survive only in less compact configurations.

We now return to the assignment of a mobility to the compact form of the  $1^{+11}$  ion. This ion cannot be the dominant peak appearing at 2030 V, already seen to be a low-mobility ghost from higher charge states. It must therefore be the next-most-abundant peak appearing immediately to its left (1960 V), since all features at still higher mobilities are ghosts from the compact forms of  $1^{+13}$  and  $1^{+14}$ . In conclusion, the foregoing discussion gives us considerable confidence in our assignment of mobilities to the compact ions  $1^{+11}$ – $1^{+14}$  formed from the 2 mM Et<sub>3</sub>AF solution, which are not directly polluted by low-mobility ions originating at higher  $z$ . The assignment of voltages to the ions  $1^{+6}$ – $1^{+9}$  is straightforward. An ambiguity remains for the  $1^{+10}$  ion, which falls in the middle of the *polluted* region. Figure 12 compares the cross section (normalized with  $n^{2/3}$ ) found in the 2 and 4 mM strongly acidic solutions for the



**Figure 12.** Comparison of relative cross sections for the monomer of concanavalin A for 0 and 2 mM Et<sub>3</sub>AF in 87 mM acetic acid (circles and crosses) and in 525 mM acetic acid with no salt.





**Figure 13.** Close up of DMA-MS spectra of concanavalin A in acetic acid buffer at the  $m/z$  for  $1^{+13}$ ,  $1^{+12}$ , and  $1^{+11}$ . Left,  $10^{+13}$ , 87 mM acetic acid, with and without 2 mM  $\text{Et}_3\text{AF}$ . Center and right: charge states 10–13 in pure acetic acid solutions 87 mM (center) and 525 mM (right).

monomer of concanavalin A. The two agree rather well. Both show also a clear departure from the relatively constant low- $z$  cross section limit starting somewhere between  $z = 8$  and  $z = 11$ .

We now reconsider the purely acidic solutions based on the insights just gained from  $\text{Et}_3\text{AF}$  solutions. It is difficult to make precise statements on ion mobilities based on the data of Figure 9a because of the lack of well-defined mobility peaks. The only exception (in addition to the tetramer results already discussed) is at the highest charge state at which a compact monomer is observed, corresponding to  $1^{+13}$ . As shown in Figure 13a, it gives two moderately broad peaks at 1784 and 1889 V. The comparison also presented in the figure with the 2 mM salt solution peak previously assigned to  $1^{+14}$  and  $1^{+13}$  shows that the same two peaks are combined here. Their relative abundances are inverted, but this is to be expected, given the trend of the salt to reduce the charge state of the protein. We may consequently assert that, at  $z = 13$  and 14, the pure acid buffer yields very much the same compact ions obtained when adding 2 mM  $\text{Et}_3\text{AF}$ . Figure 13b assembles three magnified pieces of the DMA-MS spectrum for the three peaks corresponding to  $m/z$  for  $1^{+13}$ ,  $1^{+12}$ , and  $1^{+11}$ . The structures are broader than in Figure 10, but if one discounts the contamination at the high mobility end of each peak associated with the ghost from the peak below it having a higher  $z$ , the resulting mobilities are approximately the same for the acid solutions with and without 2 mM salt addition.

In view of the ambiguities in the mobility associated with the wide mobility peaks in acid solutions with no  $\text{Et}_3\text{AF}$ , we have studied other acetic acid solutions and taken spectra at smaller declustering voltages. Figure 13c includes DMA-MS data for a  $10\ \mu\text{M}$  solution of concanavalin A in water with 3% acetic acid (524 mM), using  $\text{DP} = 50\ \text{V}$ . The mobility peaks are in this case considerably better defined than in the 87 mM acid solution, providing more precise mobility measurements. The data for this purely acidic solution are compared in Figure 12 with those from acid +  $\text{Et}_3\text{AF}$ , showing a minimal shift.

In conclusion, we confirm the well-known fact that the solution has a strong effect on the charge state,  $z$ , and on the aggregation level,  $n$ , of electrosprayed protein ions. We further find that some solutions produce much better defined mobility peaks than others, although the broad structures seen in acidic solutions may be partly due to the prevalence at high  $z$  of the confounding effect of charge stripping. Acid solutions do produce highly unstable low-mobility peaks of high  $z$  not formed from other solutions. This has traditionally been viewed as evidence of gas phase survival of solution structures. In reality, the experiment shows just that acid solutions are more effective in populating high  $z$  states in the gas, but says nothing about whether the resulting low gas phase mobility is

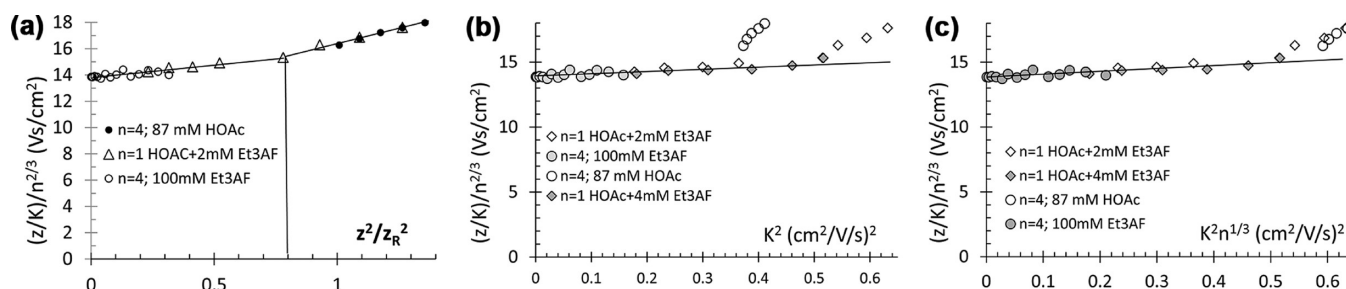
characteristic of the solution or is, rather, the equilibrium gas phase configuration associated with a high  $z$ . Furthermore, whenever several  $C_n^{+z}$  peaks originating from different solutions are well-defined, the associated mobilities have minimal dependence on solution conditions. The *gas phase equilibrium hypothesis* is therefore compatible with all the measurements made on well-resolved peaks, both in nondenaturing and in denaturing solutions. This holds also for all the well resolved elongated gas phase structures seen, whose mobilities depend only on the charge state and not on solution characteristics. These structures must therefore be due to Coulombic stretching arising in the gas phase at high  $z$ , rather than to preservation of preexisting solution structure across the two phases.

## 6. CRITICAL LOSS OF STABILITY OF GLOBULAR STRUCTURES AND THE SURFACE ENERGY OF PROTEIN MATTER

The simplest hypothesis that can be put forward to explain the strong tendency found for compaction of protein molecules and their clusters into spheroidal shapes (and their sudden extension above a critical  $z$ ) is that the phenomenon is driven by capillary forces, as if gas phase proteins were akin to either liquid drops or plastic materials subject to surface tension. The application of this macroscopic notion to individual macromolecules such as proteins may at first sight seem objectionable, so its conceptual defense and experimental validation will be attempted in this section. In the case of polymer chains, we have previously argued that their surface energy,  $\gamma$ , is a quantity that may be measured and computed.<sup>44</sup> This concept is also very useful to understand why gas phase polymer ions undergo abrupt loss of globular shape at an approximately constant  $z^2/m$  ratio. This is exactly analogous to the Rayleigh limit of a liquid drop, which may remain spherical only as long as its charge is below a critical value,

$$z^2 < z_R^2 = 48\pi\gamma\epsilon_0/(\rho e^2) \quad (6)$$

where  $m$  is the ion mass (in kg rather than Da),  $m/\rho$  is the drop volume,  $e$  is the elementary charge,  $\epsilon_0$  is the electrical permittivity of vacuum, and  $\gamma$  is the surface tension of the liquid (or the effective surface energy of the cluster). The notion of the Rayleigh limit is not restricted to liquid drops. It is generally applicable to any deformable object whose potential energy is dominated by a surface and a Coulombic contribution. This is generally the case at macroscopic sizes and remains approximately applicable, even at diameters of only a few nanometers. In this case, the Rayleigh limit (6) marks the critical charge above which the spherical shape no longer



**Figure 14.** Several representations of the dependence of the compactness parameter  $(z/K)/n^{2/3}$  versus the charge state for the monomer and tetramer of concanavalin A.

minimizes the potential energy, whence the equilibrium shape ceases to be compact.

Figure 12 shows that the cross section of the monomer ion varies relatively weakly with  $z$  for  $z < 10$  and far more strongly for  $z \geq 11$ . The same monomer data as well as those for the tetramer are represented in Figure 14a in terms of the ratio  $(z/z_R)^2$ , where  $\gamma/\rho$  is based on the bulk values for room temperature water, for which<sup>51</sup>

$$z_R = 0.0779m^{1/2} \quad (m \text{ in Da}) \quad (7)$$

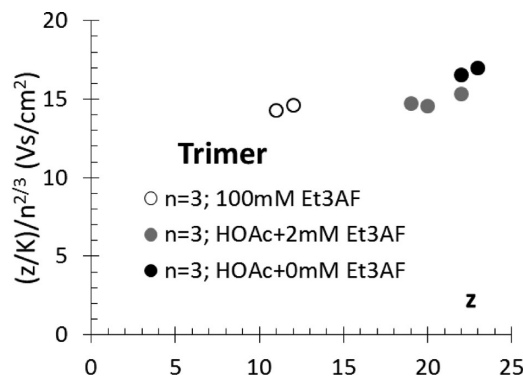
The Figure shows that the change in slope takes place for  $(z/z_R)^2$  somewhere between 0.8 and 1, with almost coincident behavior for the monomer and the tetramer. Note that charge states well above the Rayleigh charge are achieved (up to  $z^2/z_R^2 = 1.3$ ). Despite this apparent universal scaling in terms of the Rayleigh ratio, the change in cross section is considerably less abrupt than for polyethylene glycol ions. A more drastic transition is observed when the square of the ion mobility (Figure 14b) substitutes for the horizontal variable  $z^2$ . The reason is that  $K$  is strictly proportional to  $z$  as long as the cross section is constant. But past the critical conditions at which the cross section begins to increase substantially,  $K$  increases more slowly than  $z$ , so the slow transition of Figure 14a is concentrated in Figure 14b over a narrower range of the variable  $K^2$  than that of the variable  $z^2$  and is consequently sharper. The new variable  $K^2$  has another interesting advantage in that  $K^2 \sim z^2/d^4$  where  $d$  is the diameter of the ion (assumed to be spherical). But  $z^2/d^4$  is proportional to the characteristic energy associated with the ion–dipole interaction. Consequently, the small variation of the cross section with  $K^2$  seen at small  $K^2$  could be interpreted as due to ion dipole interaction (polarization). This would then explain why the dimer and tetramer data collapse approximately on the left region of the plot.

Returning now to the critical transition, Figure 14b shows convincingly that a shape change takes place quite abruptly at a certain critical condition, as expected for a Rayleigh-like transition. Although the variable  $K^2$  is ideal to demonstrate the onset of a critical shape transition, the corresponding critical values of  $K^2$  are different for the monomer and the tetramer. To demonstrate that this transition is not only abrupt but is, in fact, a Rayleigh-like transition, we slightly modify the horizontal variable to make it coincide with a Rayleigh ratio in the limit when the cross section is independent of  $z$ . As already noted,  $K^2 \sim z^2/d^4$  so that  $K^2 n^{1/3} \sim K^2 d \sim z^2/d^3 \sim z^2/m \sim (z/z_R)^2$ . These considerations therefore suggest that the transitions for the monomer and the tetramer seen in Figure 14b would collapse into a single curve in terms of the variable  $K^2 n^{1/3}$ . This is the horizontal variable used in Figure 14c, where the

monomer and tetramer data do, indeed, fall on top of each other. We therefore confirm that there is a sharply defined critical shape transition for the monomer and tetramer of concanavalin A at a fixed ratio  $(z/z_R)^2 \sim 0.92$ . If we took the Rayleigh analogy in earnest, we would conclude that the ratio  $\gamma/\rho$  for these protein clusters is 92% of the value for water. Assuming that the value found for concanavalin A is characteristic of protein matter, we would conclude that

$$(\gamma/\rho)_{\text{protein}} \sim 64.4 \text{ cm}^3/\text{s}^2 \quad (8)$$

The two most extensive data sets showing a critical shape transition for the tetramer and the monomer may be complemented by more scanty data sets for the dimer and trimer. The dimer data include a (very weak but unambiguous) datum at  $z = 19$  ( $m/z^2 = 149$ ) and show also a departure from the globular trend line at about the same  $z/z_R$  as the monomer and tetramer. These data are not included in Figure 13 because the departure from critical conditions is slight and the presence of a critical transition is not as compelling as for  $n = 1$  and 4. The trimer data for a given solution include, at most, two or three charge states. They are collected in Figure 15, showing a



**Figure 15.** Representation of the few data available for the trimer of concanavalin A showing a shape transition at  $z \sim 21$ .

shape transition at  $z$  between 21 and 22. In this case  $z_R = 21.75$ , so taking  $z^* = 21$ , the critical condition corresponds to  $(z^*/z_R)^2 = 0.93$ , very close to the value found for the monomer and tetramer series.

The absence of dimer ions at  $z$  values slightly above  $z^*$  deserves some discussion, because it differs from the behavior for  $n = 1, 3$ , and 4. The dimer shape in the drop is considerably farther from spherical than that of any of the other clusters, so it would naturally tend to achieve a greater level of charging above the Rayleigh limit. The absence of these highly charged dimers must hence be attributed not to the inability of the spray

to produce them, but to their inability to stick together once formed.

Taken globally, the critical transition information obtained from our concanavalin data seems to be rather solid. However, each of the data sets for the various aggregates has undeniable limitations. For instance, the tetramer data present a large gap between the high charges from the purely acidic solution ( $z = 25$ – $29$ ) and the low charges from the  $\text{Et}_3\text{AF}$  buffer ( $z = 1$ – $14$ ). Although the supercritical data seem at decreasing  $z$  to converge toward the low- $z$  cross section, there is no direct proof that they actually do so; nor are there data at the critical point. The high- $z$  monomeric data involve the probably correct but not entirely error-proof convoluted assignment process already discussed. The dimer data surpass only slightly the Rayleigh limit, providing weaker evidence of a critical behavior than the monomer trimer and tetramer data. Finally, the trimer series shows a clear critical transition in the two acidic buffers but includes very few data points for each solution. Given these limitations, to further support the validity of the notion of the surface energy of protein matter and its quantitative value (eq 8), we have analyzed other related literature data.

The singular protein surface energy found is consistent with earlier observations on the stability of highly charged clusters of aminoacids, first noted for quadruply charged clusters of 24 arginine molecules.<sup>52</sup> Stable multiply charged clusters of dipeptide mixtures with  $n$  and  $z$  values as large as 75 and 7, respectively, have also been reported.<sup>53</sup> In these situations, the aggregates are not held together by covalent bonds, so these high charge levels have been rationalized as due to relatively strong hydrogen bonding between the  $n$  monomer molecules composing the cluster.<sup>54,55</sup>

Here, we reinterpret these dipeptide cluster data<sup>53</sup> in light of the Rayleigh analogy. The smallest  $m/z^2$  observed are collected in Table 3, together with several related values for our protein

$$(\gamma/\rho)_{\text{dipeptide}} > 0.86(\gamma/\rho)_{\text{water}} \quad (9)$$

We note that the Rayleigh limit applies only to objects free to reduce their potential energy by changing shape. For deformable bodies, the spherical geometry minimizes energy (surface + Coulombic) only for  $z < z_R$ . The Rayleigh limit then marks the limiting condition for a spherical shape to be stable. If a peptide cluster were a rigid solid, it could, in principle, retain an arbitrary charge without deforming; however, we know from the critical shape transitions observed for concanavalin A monomers and tetramers that protein matter is, indeed, sufficiently deformable to cease to be globular above the Rayleigh limit. Therefore, we expect that dipeptide clusters will be similarly deformable. The result (eq 9) is therefore firm, giving a close quantitative confirmation of our more precise conclusion (eq 8).

For aggregates of concanavalin A, on the basis of the data of Figure 13, we can locate the instability of the globular shapes at about  $(n, z) = (1, 12)$  and  $(4, 24)$ , both corresponding to  $m/z^2 \sim 178$ . As noted before, this implies  $(\gamma/\rho)_{\text{protein}} \sim 0.93(\gamma/\rho)_{\text{water}}$ , quite comparable to the lower limit of 86% inferred from the dipeptide data. A noteworthy difference between concanavalin A and the dipeptide clusters, however, is the ability of the former to accept charge levels well above the Rayleigh limit not only of the protein, but even of water. This is evident in the most highly charged protein clusters seen, with  $(n, z) = (1, 14)$ ,  $(3, 23)$ , and  $(4, 29)$ , whose  $m/z^2$  is well below 165 (Table 3). To a lesser extent, the same is seen for the dimer data of Figure 9c. As argued elsewhere,<sup>55</sup> charge states above the Rayleigh limit of water are impossible for spherical water drops. However, an evaporating water drop encapsulating a nonspherical protein will necessarily adopt a nonspherical envelope shape near the end of the evaporation process, enabling higher charge levels than possible in a spherical drop. Therefore, the fact that some of the  $C_n^{+z}$  ions are charged above the Rayleigh limit shows that they are nonspherical in solution.

As already noted, data demonstrating critical transitions for the conformation of smaller proteins as a function of  $z$  have been reported before by Jarrold and colleagues.<sup>10–12,56</sup> They showed a compact branch of the  $\Omega(z)$  curve with a modest  $z$  dependence, followed by a sudden rise at a critical  $z$ . Because their ions received considerable activation on being injected into the mobility cell, this critical  $z$  was dependent on experimental conditions. However, at low injection energies, ions produced from unacidified solutions showed the compact mobility up to  $z = 8$  (though with less compact conformations being already dominant at  $z = 8$ ). No compact structure was seen at  $z = 9$ . In this case ( $m \sim 12\,000$ ),  $z_R \sim 8.5$ , providing a lower bound for the ratio  $(\gamma/\rho)_{\text{cyc}}/(\gamma/\rho)_{\text{water}} > 8/8.5 = 0.94$ . This ratio is very close to our ratio 0.93 for monomers and tetramers of concanavalin A. Similar work with apomyoglobin<sup>15</sup> shows that the compact  $\Omega(z)$  branch extends up to  $z = 10$ , which in this case coincides with  $z_R$ . The situation, however, differs from that of cytochrome  $c$  in that some activation is required for the elongated structure to become compact following charge reduction (perhaps because this protein is not as plastic as the others). This circumstance suggests that room temperature apomyoglobin behaves more like a solid than a deformable material, whereby quantitative conclusions regarding surface tension are somewhat risky. In the case of cytochrome  $c$ , a moderate level of activation very slightly decreased the cross section. The thermal denaturation experiments of cytochrome  $c$  previously discussed,<sup>12</sup> showing

**Table 3. Mass versus Charges Observed in Clusters of Peptides and Concanavalin A**

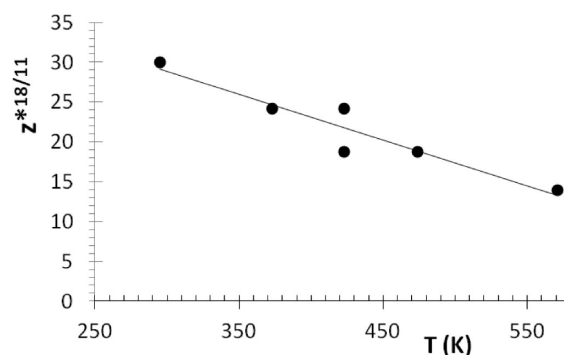
dipeptides <sup>54</sup>			$C_n^{+z}$		
$z$	$(m/z)_{\text{min}}$	$(m/z^2)_{\text{min}}$	$n$	$z$	$m/z^2$
2	455.3	228	1	12	178 <sup>b</sup>
3	577.6	193	1	14	131 <sup>c</sup>
4	852	213	4	24	178 <sup>b</sup>
7	1700	243 <sup>a</sup>	4	29	122 <sup>c</sup>
			3	21	174 <sup>b</sup>
			3	23	145 <sup>c</sup>
			2	19	142 <sup>c</sup>

<sup>a</sup>This high  $m/z^2$  is indicative of a broad trend, as individual ions were not distinguishable. <sup>b</sup> $z$  at critical shape transition. <sup>c</sup>Maximal  $z$  observed.

clusters. Note first that the Rayleigh limit (7) for a nanodrop with the bulk  $\gamma/\rho$  of water corresponds to  $m/z^2 = 165$  Da. For reference, the corresponding ratio for PEG is  $\sim 500$  Da. All the dipeptide clusters have  $m/z^2$  over 193 Da, somewhat below the Rayleigh limit for water. This result is to be expected, since these clusters are formed as charged residues from water drops, which cannot be above the Rayleigh limit. Therefore, even if the charge-carrying ability of the clusters themselves were greater than that of water, there is no mechanism to deposit this charge. If we now assume that the minimum observed  $m/z^2$  can be interpreted as a Rayleigh-like limit, the intrinsic  $(\gamma/\rho)_{\text{peptide}}$  would have to be at least  $165/192$  ( $\sim 0.86$ ) times that of water:



that  $z_R$  decreases with temperature, can be similarly interpreted as indicating that the surface energy of the protein decreases, as expected, with increasing temperature. As shown in Figure 16,



**Figure 16.** Representation of the critical charge  $z^*$  for destabilization of the compact configuration of cytochrome *c* versus bath gas temperature,  $T$ , for the data of reference 12.

the approximate data available for  $z^*(T)$  yield a reasonably smooth dependence of  $\gamma_{\text{prot}}(T)$  compatible with commonly accepted scaling laws for the surface tension of organic materials, including a linear dependence of  $\gamma^{9/11}$  on temperature [i.e., equation (16.8) of ref 57]. Note that the graph implies a substantial 2-fold variation of  $\gamma^{9/11}$ .

## 7. CONCLUSIONS

We have measured the electrical mobility  $K(z, n)$  in  $\text{CO}_2$  for various cluster ions,  $C_n^{+z}$ , of concanavalin A. We cover an unprecedented range of charge states,  $z$ , produced by electrospraying from aqueous solutions over a wide range of pH, with and without the charge-reducing buffer  $\text{Et}_3\text{AF}$ , including additional charge reduction down to  $z = 1$  with a radioactive source.

- (1) In aqueous  $\text{Et}_3\text{AF}$ , with and without charge reduction, the ratio  $z/K$  (proportional to collision cross section  $\Omega$ ) is well-defined, with a rather small dependence on  $z$ . For all the clusters studied ( $n = 1, 2, 3, 4, 6$ ), the associated compactness group  $\Omega(n, z)/n^{2/3}$  shows much less dependence on  $n$  than expected for aggregates of several protein ions, each preserving its monomeric shape. The aggregates must therefore be compacted into single globules, excluding in particular a gas phase tetramer with the tetrameric crystal structure of concanavalin A.
- (2) Higher charge states are achieved by electrospraying from 87 mM aqueous solutions of acetic acid. Many ion mobility peaks are broad and unsuitable to extract precise mobilities, with the exception of tetramer ions appearing with  $z = 25$ –29 and trimer ions with  $z = 22$ –23. Upon adding 2 mM and 4 mM  $\text{Et}_3\text{AF}$  to the 87 mM acetic acid solution, narrower peaks result, covering continuously from the highest charge states typical of the pure acid buffer down to the lowest  $z$  obtained in the pure  $\text{Et}_3\text{AF}$  buffer.
- (3) The various buffers used span different  $z$  ranges, but within the region where several buffers yield overlapping  $z$  values, the  $\Omega(z, n)$  curves obtained are approximately independent of the buffer. The structures found are accordingly characteristic of the gas phase.
- (4) Above a critical  $z^*(n)$ ,  $\Omega(z)$  increases in a relatively abrupt form for all the aggregates of concanavalin A for

which data are available at large enough  $z$  ( $n = 1$ –4). This  $z^*$  scales approximately as  $n^{1/2}$ , very much like in the Rayleigh limit of a charged drop. This Rayleigh-like behavior yields a surface energy over density ratio  $\gamma/\rho$  that exceeds 90% of the value for water. The same quantitative conclusion is reached on the basis of published data on the critical  $z$  for loss of compact structure in cytochrome *c* and apomyoglobin. Data on the maximal  $z$  versus  $m$  for electrosprayed dipeptide aggregate clusters provide the comparable lower bound  $(\gamma/\rho)_{\text{peptide}} > 0.86(\gamma/\rho)_{\text{water}}$ .

- (5) The tendency of protein clusters to compact into spheroidal shapes at low  $z$  and to lose that compact shape at a critical charge scaling as the  $1/2$  power of its mass provides strong evidence pointing to the approximate validity of the notion of the surface energy of protein or peptide matter.
- (6) The high effective surface energy of proteins and peptides, approaching that of water, explains the experimental difficulty in producing relatively compact yet nonglobular protein ions. In contrast, such ions are readily produced with industrial polymers.

## 8. FINAL BIBLIOGRAPHIC REMARKS

An article independently supporting one of our main conclusions has recently appeared.<sup>58</sup> They use a DMA to measure the mobilities of clusters ( $n = 1$ –4) of cytochrome *c*, lysozyme, myoglobin, ovalbumin, and bovine serum albumin electrosprayed from nondenaturing and denaturing solutions. They report broad mobility peaks ( $8.5\% < \text{fwhm} < 30.4\%$ ) at  $z = 1$ –3, probably due to the special charge-reduction process used (precluding Coulombic explosions) leading to unusually large ES drops and substantial adduction. They conclude that both solutions produce comparable mobilities, although the nondenaturing buffer yields slightly more compact ions, even slightly more compact than expected from the crystal structure.

## ■ APPENDIX ON THE ASSIGNMENT OF CHARGE REDUCED IONS

The peak at 2400 V labeled  $1^{+1}$  in Figure 6a is not part of either the dimer or the tetramer series. Its appearance halfway between  $4^{+2}$  and  $2^{+2}$  suggests that it could also be  $3^{+2}$ . We reject this hypothesis because of the complete absence of the corresponding  $3^{+1}$  ion at 4800 V, while the  $4^{+1}$  and  $2^{+1}$  are present. Another possible assignment of the 2400 V peak is to  $1^{+1}$ . In this case, the ratio of tetramer to presumed monomer mobility would be  $2400/5753 = 0.417$ , slightly larger than the expected area ratio for spheres having volumes in a 4/1 ratio:  $(1/4)^{2/3} = 0.397$ . Assignment of this peak to  $1^{+1}$  would then imply a tetramer slightly more compact than the monomer. This is not unexpected for a compacted tetramer because the crystal structure of the monomer ion has a protruding end that may have survived the compaction process suffered by the gas phase monomer. The tentative assignment of the peak at 2400 V to the monomer is reinforced by the lack of agreement between the data and the Gaussian fit seen in Figure 6a at 1200 V, which signals the probable presence there of a weak contribution of the  $1^{+2}$  ion. Other unassigned peaks are similarly recognizable in Figure 6 from the difference between this Gaussian fit and the data.

The assignment of high charge state peaks is based on the higher resolution data of Figure 6b. Notice first the weakness of

the  $2^{+5}$  ion as compared with  $4^{+7}$  and  $4^{+8}$ , suggesting that the tetramer series is dominant in the pure DMA measurements. This observation is puzzling given the opposite trend seen in the DMA-MS measurements of Figure 5a. But the effect is real, as confirmed by the fact that the coincident peak pairs of structure ( $4^{+3i}$ ,  $2^{+2i}$ ) with  $i$  an integer number are not anomalously high. Furthermore, none of the peaks clearly defined in Figure 6b are in the expected dimer sequence  $V_2/z$ . A likely explanation for these anomalous abundances is in the very small transmission of ions (<1%) in the convoluted path from the ES source to the cylindrical DMA used for the stand-alone DMA measurements. These losses are dominated by the space charge field, which removes predominantly the most mobile ions.

Several other unassigned ions not forming clear peaks can be recognized by gaps in the Gaussian fitting appearing at 3806 and 5440 V. All the major peaks are therefore assigned to the tetramer, and no mobility information beyond  $z = 5$  is available for the charge-reduced dimers. However, because the tetramer and dimer cross sections vary little with  $z$  for all the available charge-reduced data (up to  $z = 14$  and  $5$ , respectively), it is fair to presume that the ratio 1.469 between tetramer and dimer cross sections measured at  $z = 1$  will apply similarly up to at least the pair ( $4^{+12}$ ,  $2^{+8}$ ).

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [juan.delamora@yale.edu](mailto:juan.delamora@yale.edu).

### Present Address

<sup>†</sup>Abengoa Research, Campus Palmas Altas, Sevilla 41012, Spain

### Notes

The authors declare the following competing financial interest(s): Following Yale rules, Juan Fernandez de la Mora declares that he has a personal interest in the company SEADM manufacturing the mobility analyzer used in this research.

## ACKNOWLEDGMENTS

We thank Applied Biosystems and SEADM for their loan of our DMA-MS system (<http://www.eng.yale.edu/DMAMSFacility/>), Yale's W. M. Keck Center for hosting it, and Bruce Thomson for his guidance on mass spectrometry and Q-Star MS issues. We acknowledge innumerable contributions of Chris Hogan, Carlos Larriba, Juan Fernández García at Yale, and Alejandro Casado (SEADM) to the continuous operation of our experimental facility.

## REFERENCES

- (1) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.
- (2) Ruotolo, B. T.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2006**, *10*, 402–408.
- (3) Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 214–217.
- (4) Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1993**, *115*, 6317–6321.
- (5) Xie, Y. M.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432–14433.
- (6) Fuerstenau, S. D.; Benner, W. H.; Thomas, J. J.; Brugidou, C.; Bothner, B.; Siuzdak, G. *Angew. Chem., Int. Ed.* **2001**, *40*, 542–544.
- (7) Hogan, C. J.; Kettleson, E. M.; Ramaswami, B.; Chen, D. R.; Biswas, P. *Anal. Chem.* **2006**, *78*, 844–852.
- (8) Dunn, R. V.; Daniel, R. M. *Philos. Trans. R. Soc. B* **2004**, *359*, 1309–1320.
- (9) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, *310*, 1658–1661.
- (10) Hudgins, R. R.; Woenckhaus, J.; Jarrold, M. F. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165/166*, 497–507.
- (11) Jarrold, M. F. *Acc. Chem. Res.* **1999**, *32*, 360–367.
- (12) Mao, Y.; Woenckhaus, J.; Kolafa, J.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999**, *121*, 2712–2721.
- (13) Badman, E. R.; Myung, S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1493–1497.
- (14) Gidden, J.; Bowers, M. T.; Jackson, A. T.; Scrivens, J. H. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 499–505.
- (15) Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2987–2994.
- (16) Breuker, K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18145–18152.
- (17) Breuker, K.; Jin, M.; Han, X. M.; Jiang, H. H.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1045–1053.
- (18) Steinberg, M. Z.; Elber, R.; McLafferty, F. W.; Gerber, R. B.; Breuker, K. *ChemBioChem* **2008**, *9*, 2417–2423.
- (19) Hogan, C.; Ruotolo, B.; Robinson, C.; Fernandez de la Mora, J. *J. Phys. Chem. B* **2011**, *115*, 3614–3621.
- (20) Millikan, R. A. *Phys. Rev.* **1923**, *22*, 1–23.
- (21) Eglin, J. M. *Phys. Rev.* **1923**, *22*, 161–170.
- (22) Kim, J. H.; Mulholland, G. W.; Kukuck, S. R.; Pui, D. Y. H. *J. Res. Natl. Inst. Stand. Technol.* **2005**, *110*, 31–54.
- (23) Allen, M. D.; Raabe, O. G. *J. Aerosol Sci.* **1982**, *13*, 537–543.
- (24) Allen, M. D.; Raabe, O. G. *Aerosol Sci. Technol.* **1985**, *4*, 269–286.
- (25) Rader, D. G. *J. Aerosol Sci.* **1990**, *21*, 161–168.
- (26) van Duijn, E.; Barendregt, A.; Synowsky, S.; Versluis, C.; Heck, A. J. R. *J. Am. Chem. Soc.* **2009**, *131*, 1452–1459.
- (27) Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39*, 1633–1655.
- (28) Scarff, C. A.; Thalassinos, K.; Hilton, G. R.; Scrivens, J. H. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3297–3304.
- (29) Counterman, A. E.; Valentine, S. J.; Srebalus, C. A.; Henderson, S. C.; Hoaglund, C. S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 743–759.
- (30) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86–91.
- (31) Fernandez de la Mora, J. *J. Aerosol Sci.* **2002**, *33*, 477–489.
- (32) Shvartsburg, A. A.; Mashkevich, S. V.; Baker, E. S.; Smith, R. D. *J. Phys. Chem. A* **2007**, *111*, 2002–2010.
- (33) Reeke, G. N., Jr.; Becker, J. W.; Edelman, G. M. *J. Biol. Chem.* **1975**, *250*, 1525–1547.
- (34) Hogan, C. J.; Fernandez de la Mora, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8079–8090.
- (35) Rus, J.; Moro, D.; Sillero, J. A.; Royuela, J.; Casado, A.; Estevez-Molinero, F.; Fernandez de la Mora, J. *Int. J. Mass Spectrom.* **2010**, *298*, 30–40.
- (36) Fernandez de la Mora, J. The differential mobility analyzer (DMA): Adding a true mobility dimension to a preexisting API-MS. In *Ion Mobility Spectroscopy – Mass Spectrometry: Theory and Applications*; Wilkins, C., Trimpin, S. Eds.; Taylor and Francis: New York, 2011; Chapter 5 (pp.105–136).
- (37) Ude, S. Measurement and properties of nanometer particles in the gas phase. *Ph.D. Dissertation*, Yale University, 2004.
- (38) Fernandez de la Mora, J. Sub-3 nm aerosol measurement with DMAs and CNCs. In *Aerosol Measurements*; Baron, P., Pramod, K., Willeke, K., Eds.; John Wiley & Sons, Inc.: New York, 2011; Chapter 32.
- (39) Bird, G. A. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*; Oxford University Press: New York, 1994.
- (40) Chan, P.; Dahneke, B. E. *J. Appl. Phys.* **1981**, *52*, 3106–3110.
- (41) Zurita-Gotor, M. Rational Numerical Simulations in Sol Reaction Engineering. *PhD Thesis*, Yale University, 2004.
- (42) Zand, R.; Agrawal, B. B. L.; Goldstein, I. J. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 2173–2176.
- (43) Hogan, C. J., Jr.; Fernández de la Mora, J. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 158–172.

- (44) Ude, S.; Fernandez de la Mora, J.; Thomson, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 12184–12190.
- (45) Larriba-Andaluz, C.; Fernandez de la Mora, J. *J. Phys. Chem. B* **2012**, *116*, 593–598. See also Larriba, C. Production of ions & particles via simple and compound electrosprays in vacuum, gases or liquids (polar and non-polar). *Ph.D. Thesis*, Yale University, Mechanical Engineering Department, September 2010
- (46) Trimpin, S.; Plasencia, M.; Isailovic, D.; Clemmer, D. E. *Anal. Chem.* **2007**, *79*, 7965–7974.
- (47) Trimpin, S.; Clemmer, D. E. *Anal. Chem.* **2008**, *80*, 9073–9083.
- (48) De Winter, J.; Lemaire, V.; Ballivian, R.; Chirot, F.; Coulembier, O.; Antoine, R.; Lemoine, J.; Cornil, J.; Dubois, P.; Dugourd, P.; et al. *Chem.—Eur. J.* **2011**, *17*, 9738–9745.
- (49) Kebarle, P.; Verkerk, U. H. *Mass Spectrom. Rev.* **2009**, *28*, 898–917.
- (50) Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *Anal. Chem.* **2009**, *81*, 369–377.
- (51) Fernandez de la Mora, J. *Anal. Chim. Acta* **2000**, *406*, 93–104.
- (52) Meng, C. K.; Fenn, J. B. *Org. Mass Spectrom.* **1991**, *26*, 542–549.
- (53) Counterman, A. E.; Hildebrand, A. E.; Srebalus Barnes, C. A.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1020–1035.
- (54) Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, *406*, 11–35.
- (55) Fernandez de la Mora, J. *J. Am. Soc. Mass Spectrom.* **2011**, submitted
- (56) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997**, *119*, 2240–2248.
- (57) Guggenheim, E. A. *J. Chem. Phys.* **1945**, *13*, 253–261.
- (58) Maïßer, A.; Premnath, V.; Ghosh, A.; Nguyen, T. A.; Attoui, M.; Hogan, C. J., Jr. *Phys. Chem. Chem. Phys.* **2011**, *13*, 21630–21641.