

# Tandem Mass Spectrometry of Very Large Molecules: Serum Albumin Sequence Information from Multiply Charged Ions Formed by Electrospray Ionization

Joseph A. Loo, Charles G. Edmonds, and Richard D. Smith\*

Chemical Sciences Department, Pacific Northwest Laboratory, Richland, Washington 99352

Serum albumin proteins,  $M_r$ , ~66 kDa, from 10 different species (bovine, human, rat, horse, sheep, goat, rabbit, dog, porcine, and guinea pig) have been studied by electrospray ionization mass spectrometry (ESI-MS) and tandem MS using a triple-quadrupole instrument. The effectiveness of collisional activation for the multiply charged albumin ions greatly exceeds that for singly charged ions, allowing an extension by a factor of at least 20 to the molecular mass range for obtaining sequence-specific product ions by tandem MS. Efficient dissociation is largely attributed to "preheating" in the interface Coulombic instability and the large number of collisions. Increasing the electric field in the intermediate pressure region, between the nozzle-skimmer elements of the atmospheric pressure/vacuum interface, allows fragmentation of the multiply protonated (to 96+) molecules produced by ESI. The most abundant dissociation product ions assigned have a low charge state (2+ to 5+) and are attributed to "b<sub>n</sub>" mode species from cleavage of the -CO-N- peptide backbone bonds. Particularly abundant dissociation products originate from regions near residues  $n = 20-25$  from the NH<sub>2</sub> terminus for parent ions of moderate charge (~50+). Collisionally activated dissociation (CAD) mass spectra from porcine serum albumin, in contrast to the other albumins, also gave prominent singly charged "y<sub>n</sub>" fragments formed from cleavages near the COOH terminus. Tandem mass spectrometry (MS/MS) of the multiply charged molecular ions, and of fragment species produced by dissociation in the interface (i.e., effective MS/MS/MS), produced similar "b<sub>n</sub>" species and served to confirm spectral assignments. We also show that ESI mass spectra allow a qualitative assessment of protein microheterogeneity and, in some cases, resolution of major contributions. The physical and analytical implications of the results are discussed, including the identification of possible errors in previously published sequences.

## INTRODUCTION

Mass spectrometry continues to mature as an analytical methodology for biomolecules of increasingly greater molecular mass (1-5). In the past several years, new ionization techniques have been developed that have allowed extension to biomolecules with relative molecular mass ( $M_r$ ) in excess of 100 kDa. Matrix-assisted laser desorption (LD) (6-8) coupled to time-of-flight (TOF) mass spectrometers has produced mass spectra for proteins as large as catalase ( $M_r$ , 236 kDa) (6). Another ionization method utilized for large biomolecules is electrospray ionization (ESI) (9-14), which produces multiply charged molecules from highly charged droplets formed by electrostatic nebulization at near atmospheric pressure. Sufficient charges on analyte molecules are often retained so that conventional quadrupole mass spectrometers of limited mass-to-charge ratio ( $m/z$ ) can be used for the study of proteins, typically as multiply protonated molecules, with  $M_r$  of

at least 133 000 Da (12) and more recently to nearly 200 000 Da (14). No fragmentation of covalent bonds is typically evident in ESI mass spectra since the conditions required to obtain nearly complete desolvation are less severe than necessary for dissociation (15). Fragmentation can be induced in a separate step using collisionally activated dissociation (CAD), generally with tandem mass spectrometry (MS/MS), by applying an additional stage of mass ( $m/z$ ) selection prior to the dissociation step. In contrast to TOF mass analyzers used for laser desorption, quadrupole instruments are well suited to tandem mass spectrometry experiments. Thus, the use of ESI-MS/MS for structural characterization of large peptides and proteins with existing instrumentation appears promising (13, 15-18).

Tandem mass spectrometry has been utilized for many years to obtain structural information from organic molecules (19, 20). MS/MS experiments coupled to fast atom bombardment (FAB) have been fruitful for small peptides to  $M_r$ , 3000 Da (21-23). Complete amino acid sequence information for larger peptides and proteins is obtainable by FAB-MS/MS after selective enzymatic digestion (24-26). However, CAD of singly charged molecules becomes less efficient with increasing molecular mass (27-29). The decreased efficiency of molecular ion formation, substantial "matrix" background, the modest extent of initial molecular ion internal excitation, and practical limitations on the average number of collisions allowed using conventional sector instruments conspire to prevent application of these methods at higher  $m/z$ . We have previously demonstrated (16) that multiply charged molecular ions from ESI of a 26-residue peptide (melittin) subjected to CAD conditions yields singly and multiply charged product ions that can be correlated with the known sequence. Multiply charged ions appear to be more susceptible to CAD than their singly charged counterparts at the same collision energy, allowing larger molecules to be addressable by tandem mass spectrometry (13, 15-18). More recently we have extended these methods to ~14 kDa in studies of ribonuclease A (18). Thus, in addition to molecular mass determination (9-14), the polypeptide sequence or related information can potentially be obtained for molecules much larger than the nominal "mass" limit of the quadrupole instrument due to the multiple charging phenomena.

In this report we expand on our initial communication regarding MS/MS of ESI-generated multiply charged molecular ions from serum albumins with  $M_r$ , ~66 000 Da (13, 30). We show that the efficient dissociation of these large molecules can be induced and their product ions correlated to the available sequence. In addition to CAD processes in the conventional collision chamber of the triple-quadrupole instrument, effective dissociation in the atmospheric pressure/vacuum interface of the mass spectrometer can also be induced (31), leading to sequence-specific fragmentation. Although only limited primary sequence information is obtained, the potential of tandem MS with electrospray ionization for large biomolecules is demonstrated. More generally,

these studies provide new insights regarding the nature of activation and dissociation processes at the large molecule limit.

## EXPERIMENTAL SECTION

The ESI source and mass spectrometer have been previously described (12, 13, 15–18). Briefly, the analyte solution, in a 100- $\mu\text{m}$ -i.d. fused silica capillary, mixes with a flowing liquid methanol sheath electrode (32) at the tip of the ESI source. A potential of +4 kV is applied to the sheath electrode to produce a fine mist of highly charged droplets. The flow rates of the analyte solution and liquid sheath are independently controlled by separate syringe pumps (Harvard Apparatus, South Natick, MA, and Sage Instruments, Cambridge, MA) at flow rates of 0.2–0.5 and 2.5  $\mu\text{L min}^{-1}$ , respectively. Ions are sampled by a differentially pumped nozzle/skimmer interface of our design, similar to that used on our previously described single-quadrupole instrument (31), to the triple-quadrupole MS (TAGA 6000E, Sciex, Thornhill, Ontario, Canada) with  $m/z$  limit of 1400. A lens element, typically at +650 V, mounted in front of the nozzle/skimmer assembly, improves the ion-sampling efficiency. Desolvation of the droplets is accomplished by a countercurrent stream of dry nitrogen gas and through more energetic collisions in the interface region. The collisional heating in the interface region is controlled by varying the voltage difference between the nozzle and skimmer elements ( $\Delta\text{NS}$ ); a typical bias used is +135 to +200 V before dissociation of the analyte molecule occurs. Thus, the activation in this region occurs through a large number of collisions which have increasing energy and decreasing frequency due to the decreasing pressure. The pressure in the atmosphere/vacuum interface (in the stagnant region displaced from the beam axis) and in the mass spectrometer is typically  $\sim 1$  Torr and  $6 \times 10^{-6}$  Torr, respectively.

All serum albumin samples were commercially obtained from Sigma Chemical Co. (St. Louis, MO) and were used without further purification. Their origin, catalog number, and lot number are as follows: bovine (A-7638, 77F9369), human (A-8763, 67F9351), rat (A-6414, 97F93211), porcine (pig) (A-1173, 87F9307), sheep (A-4289, 37F93071), horse (A-5280, 66F9381), rabbit (A-0764, 19F9301), dog (A-9263, 108F9327), goat (A-4164, 37F93081), and guinea pig (A-6539, 38F3001). Crystallized and lyophilized samples from Sigma that are essentially fatty acid free (<0.005%) and/or globulin free ( $\leq 1\%$ ) were analyzed (if available). Only a "Fraction V" (modified Cohn low-temperature ethanol precipitation) sample of dog albumin was available from Sigma. Since our goal was to evaluate the nature and potential of collisional dissociation processes, no attempt was made to minimize sample consumption. Analyte solutions (50–100 pmol  $\mu\text{L}^{-1}$ ) were prepared in distilled water with 5% acetic acid. Multiply charged ions from horse myoglobin were used to calibrate the  $m/z$  scale of the mass spectrometer. The argon collision gas target thickness was approximately  $1 \times 10^{14}$  molecules  $\text{cm}^{-2}$  for tandem MS experiments.

## RESULTS AND DISCUSSION

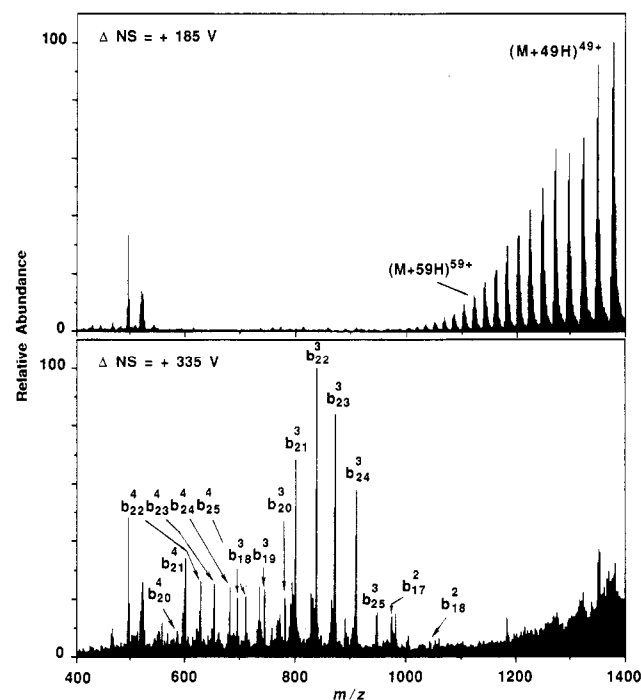
**Electrospray Ionization Mass Spectra of Serum Albumins.** Albumin is an acidic, stable, 66-kDa monomeric protein that is among the most studied in biochemistry (33). Complete sequences are available for human, bovine, and rat albumin (33), sheep (34, 35) and porcine (36) albumin, and partial amino acid sequences (typically the  $\text{NH}_2$  terminal residues) are known for species such as dog (37), rabbit (38), goat (38), and horse (38, 39). Albumin is well conserved among various species; a homology (amino acid sequence alignment) of 80% between human and bovine, 80% between rat and human, and 63% total conservation among the three species is maintained (33). A total of 18 out of the first 31 residues are identical among albumins from human, horse, bovine, rat, sheep, goat, porcine, rabbit, and hamster species. Table I lists the first 31 "known" residues (from the  $\text{NH}_2$  terminus) for nine of the ten albumins studied.

Serum albumins are readily amenable to ESI-MS analysis for  $M_r$  determination, as demonstrated by our laboratory (12, 13, 30, 40) and a number of other researchers (41–47). Figure 1 (top) shows the ESI mass spectrum for rabbit serum albumin

**Table I. Published  $\text{NH}_2$  Terminal Amino Acid Sequences for Serum Albumins<sup>a</sup>**

Species	Sequence																														
	1	5	10	15	20	25	30																								
Bovine	D	T	H	K	S	E	I	A	H	R	F	K	D	L	G	E	E	H	F	K	G	L	V	L	I	A	F	S	Q	Y	L
Human	D	A	H	K	S	E	V	A	H	R	F	K	D	L	G	E	E	N	F	K	A	L	V	L	I	A	F	A	Q	Y	L
Rat	E	A	H	K	S	E	I	A	H	R	F	K	D	L	G	E	Q	H	F	K	G	L	V	L	I	A	F	S	Q	Y	L
Porcine	D	T	Y	K	S	E	I	A	H	R	F	K	D	L	G	E	Q	Y	F	K	G	L	V	L	I	A	F	S	Q	H	L
Sheep	D	T	H	K	S	E	I	A	H	R	F	N	D	L	G	E	E	N	F	K	G	L	V	L	I	A	F	S	Q	Y	L
Goat	D	T	H	K	S	E	I	A	H	R	F	N	D	L	G	E	E	H	F	K	G	L	V	L	I	A	F	S	Q	Y	L
Horse	D	T	H	K	S	E	I	A	H	R	F	N	D	L	G	E	K	H	F	K	G	L	V	L	V	A	F	S	Q	Y	L
Rabbit	E	A	H	K	S	E	I	A	H	R	F	N	D	V	G	E	Z	H	F	K	G	L	V	L	I	A	F	S	Q	Y	L
Dog	E	A	Y	K	S	E	I	A	H	R	Y	N	D	L	G	E	E	H	F	R	G	L	V	L							

<sup>a</sup>Standard one-letter code used for amino acids: A = Ala, R = Arg, N = Asn, D = Asp, C = Cys, E = Glu, Q = Gln, G = Gly, Z = Glx, H = His, I = Ile, L = Leu, K = Lys, M = Met, F = Phe, P = Pro, S = Ser, T = Thr, W = Trp, Y = Tyr, V = Val.



**Figure 1.** Electrospray ionization mass spectra of rabbit serum albumin ( $M_r \sim 66$  100 Da) with nozzle/skimmer potential of (top) +185 V and (bottom) +335 V. See text for sequence ion notation.

with  $\Delta\text{NS} = +185$  V. Experience has shown that this  $\Delta\text{NS}$  voltage difference results in a degree of collisional "heating" in the interface sufficient to result in complete desolvation of electrospray ions (but insufficient for cleavage of covalent bonds) over the  $m/z$  400–1400 range. No ions assigned as fragment species were detected under these conditions. Relative molecular masses with precisions on the order of 0.01% or better can be obtained, as demonstrated by the data in Table II for the 10 different serum albumin species. Although the experimental  $M_r$  precision was similar to that determined for smaller polypeptides and proteins by ESI-MS at our laboratory (13) and others (10, 11, 42, 43), the  $M_r$  measurements for serum albumins were generally high by as much as 0.25% compared to  $M_r$  values based on the complete amino acid sequences (available for five of the species listed in Table II). Molecular weight accuracy by ESI-MS has been found to generally be better than 0.01% for most relatively pure polypeptides up to 30 kDa (10–13, 42, 43). Although the purity of the commercially obtainable serum albumins may

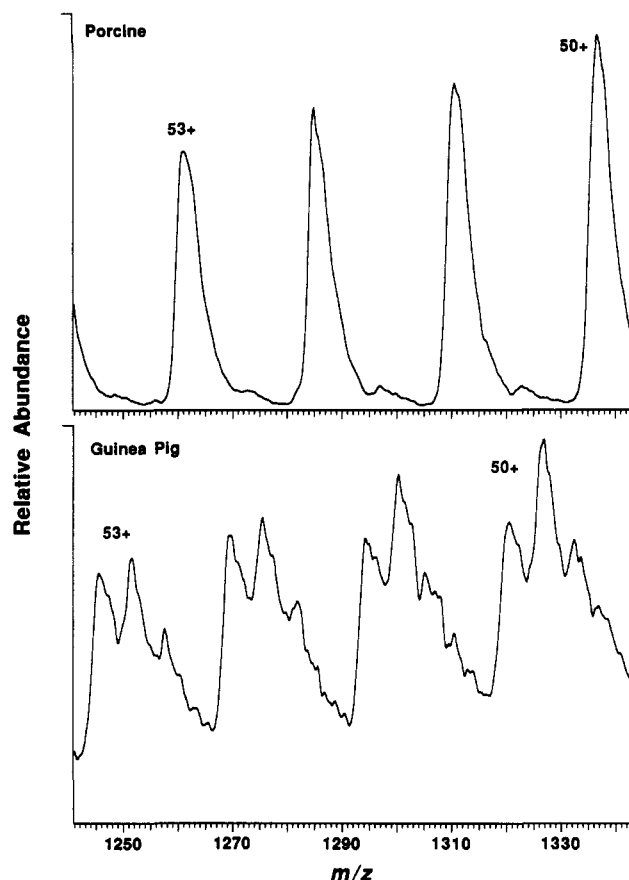
**Table II. Molecular Masses and Peak Widths from Electrospray Ionization Mass Spectra of Serum Albumins**

species	$M_r$ (exptl), Da	$M_r$ (seq), <sup>a</sup> Da	FWHH (50+) <sup>b</sup>
bovine	66 443 ± 5	66 430	4.3
human	66 605 ± 7	66 438	8.2
rat	66 000 ± 6	65 870	5.0
porcine	66 770 ± 5	66 731	4.3
sheep	66 385 ± 6	66 293	5.0
horse	65 667 ± 6		8.7
rabbit	66 148 ± 6		4.0
dog	65 854 ± 7		10.7
goat	66 304 ± 5		(8.9)
	66 578 ± 8		
guinea pig	65 962 ± 8		(15.8)
	66 284 ± 8		
	66 586 ± 7		

<sup>a</sup> Average molecular masses calculated from available complete amino acid (or nucleotide) sequences. <sup>b</sup> Full width in  $m/z$  units at half the peak height for the  $(M + 50H)^{50+}$  multiply charged molecular ion. Values in parentheses indicate that more than one molecular ion component is partially resolved.

be as high as 96% based on the nitrogen content, as Peters (33) points out, microheterogeneity may arise from various contributions, including adducts (fatty acids, bilirubin, glycans, etc.) and structural modifications (cleavages due to traces of proteolytic enzymes in the preparation, amino acid modifications, etc.).

The relative purity of the serum albumins can be qualitatively assessed on the basis of the width of a molecular ion peak and agreement with  $M_r$  based on the amino acid sequence (where available). This requires that spectra be obtained under conditions which result in sufficient activation to remove residual solvent and non-covalently associated molecules. Table II lists the measured widths (in  $m/z$  units) at half the peak maxima for the  $(M + 50H)^{50+}$  species (between  $m/z$  1310 and 1340) for each albumin analyzed. For example, dog serum albumin sample is attributed as having more *major* (i.e., abundant) unresolved components than rat albumin based on the larger peak widths, although all albumin samples showed a significant "tail" to higher  $m/z$ , possibly indicating contamination, or small amounts of residual solution-related adducts, in addition to naturally occurring heterogeneity. Porcine and bovine albumin, which gave the closest agreement with the measured molecular masses, also gave the narrowest peak widths. A few of the albumins, most notably goat and guinea pig albumin, showed additional resolved components, as indicated in Table II. Figure 2 compares a region of the porcine and guinea pig albumin ESI-mass spectra, clearly showing the greater heterogeneity of the latter where at least three major components are evident. Because our  $M_r$  measurements in Table II are based on peak maxima, microheterogeneity could at least partially account for a shift in peak maxima sufficient to explain the difference with  $M_r$  values based on the sequence. Peak widths may not directly correlate to the observed  $M_r$  deviation because of the variable peak shapes between albumin species. Two mass spectra from the same albumin species, but of different origin, may show similar peak widths, yet different experimentally measured  $M_r$  values due to different centroid positions. Using the peak maxima to calculate  $M_r$  emphasizes the major component in the sample, which may not represent the native albumin species. Thus, a properly deconvoluted peak shape (i.e., corrected for instrumental contributions) should provide useful information on the microheterogeneity of such materials. We note, however, that definitive experiments which provide both separation, perhaps based upon capillary electrophoresis (12, 13) of the closely related mixture components, and their separate MS/MS analysis has not yet been demonstrated.



**Figure 2.** Expanded region of the ESI mass spectra of (top) porcine and (bottom) guinea pig serum albumin with  $\Delta N_S = +185$  V. A three-point smoothing function was applied to the spectra.

Supportive of our suggestion is recent work by Geisow and co-workers (46) who presented ESI-MS data for a commercial sample of human albumin and recombinant human albumin expressed in yeast. Their ESI mass spectrum of the commercially available albumin showed  $M_r$  and peak widths similar to our data. However, the spectrum for the recombinant form showed much narrower peak widths, approximately 2  $m/z$  units at half the peak maximum at  $\sim m/z$  1545 versus  $\sim 6$   $m/z$  units for the commercial sample, indicating a much "cleaner" sample, and yielding a  $M_r$  value only 3 Da from the calculated sequence value. Hirayama et al. (47) recently determined the published amino acid sequence of bovine serum albumin to be incorrect based on FAB LC-MS and MS/MS results from enzymatic digests. The addition of a tyrosine residue to position 156 corresponds to a corrected  $M_r$  value of 66 430 Da, in relatively close agreement to our result. Thus, published sequences may be incorrect, as discussed in the following sections.

Other information is also supportive of the proposed relationship between microheterogeneity and peak shape. Ideally, the peak widths measured at half height for multiply charged species, in which the isotopic contributions are unresolved, are expected to decrease with increasing molecular weight at constant  $m/z$ , and (equivalently) decrease with increasing charge (12, 48). For example, upon the addition of 1,4-dithiothreitol (DTT) to cleave disulfide bonds (49) and increase the extent of ESI charging (13, 40, 50), peak widths (in  $m/z$  units) for sheep albumin decreased from 5.0 for the 50+ ion to 2.7 for the 80+ ion. The peak width at half-height at  $m/z$  1304.9 for the  $(M + 13H)^{13+}$  ion of horse myoglobin ( $M_r$  16950 Da), a relatively pure material, was approximately 3.2  $m/z$  units for the same instrumental resolution. A multiply charged molecular ion in the same  $m/z$  vicinity for a pure 66-kDa protein should have a peak width much less than 3.2. From

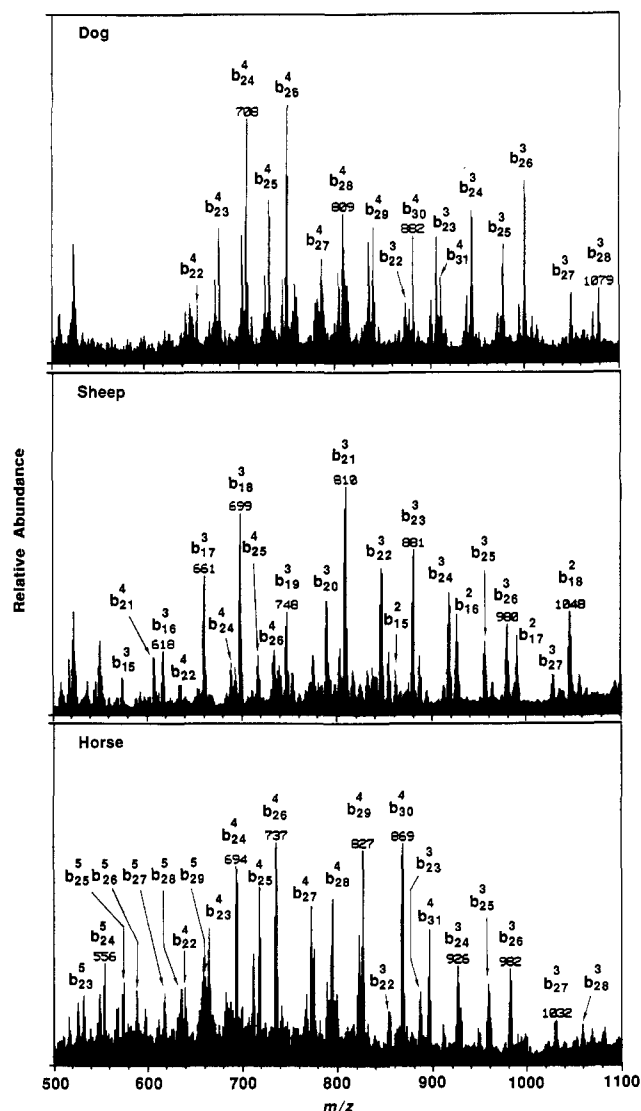
the theoretical isotopic distribution of bovine albumin, a peak width of slightly greater than 0.3 for the  $(M + 50H)^{50+}$  ion would ideally be obtained. Experimentally, a peak width of  $\sim 2 m/z$  units would be expected due to limitations on instrumental performance (i.e., resolution). Most of the albumins studied have peak widths greater than 3.2 (Table II), consistent with a contribution due to sample microheterogeneity. Whether peak width measurements can be used to unambiguously determine ion charge state is uncertain at this time, but such a capability will certainly be limited to lower charge states and be impractical in cases where substantial microheterogeneity exists.

It should be pointed out that the peak widths for albumin molecular ions do not decrease appreciably with increasing  $\Delta NS$ . By increasing the collision energy in the interface region, weakly bound adducts and solvent molecules can be "stripped" from the larger analyte species (15, 31, 51). It has recently been demonstrated that sulfuric acid and phosphoric acid generated in the acidified spray solutions from ubiquitous sulfate and phosphate impurities associate strongly with proteins (51). However, it was also shown that these noncovalently bound adducts can be completely removed by activation in the atmosphere/vacuum interface with modest collision energies. Since no change in peak widths was observed upon further increase in  $\Delta NS$ , even to collision energies sufficient for polypeptide backbone dissociation (see next section), this clearly suggests that the large peak widths in our studies are not due to solvent or other noncovalently attached species.

Although the purity and exact chemical composition of the albumin samples are in question, one or a few components with relatively similar composition are likely to (and apparently) predominate. As will be discussed in the following sections, our results indicate that any chemical modifications that might be present in substantial abundance apparently occur in a region other than the first 30 residues from the  $NH_2$  terminus.

**Dissociation Induced in the Interface Region.** As previously shown (13, 15, 18, 30, 31), dissociation of intact proteins can be induced in the atmospheric pressure/vacuum interface at elevated nozzle/skimmer potentials. We have previously presented preliminary dissociation spectra for bovine, human, and rat albumins (13, 30). Additional mass spectra are shown in Figure 1 (bottom) for rabbit albumin and in Figure 3 for dog, sheep, and horse albumins. All such spectra were obtained using  $\Delta NS = +335$  V and otherwise identical instrumental conditions. Nomenclature used for the designated fragment species is based on the modified Roepstorff (52) notation proposed by Biemann (22), in which lower case letters are utilized and subscripts are used to denote the type of fragmentation and residue position, respectively, counting from the  $NH_2$  terminus for  $a_n$ ,  $b_n$ , and  $c_n$  ions and from the  $COOH$  terminus for  $x_n$ ,  $y_n$ , and  $z_n$  product ions. A superscript is added to indicate the fragment ion charge state. Lack of a superscript denotes a singly charged fragment.

Nearly all the more abundant product ion peaks detected have been assigned as multiply charged  $b_n$  ions arising from fragmentation within the first 30 residues from the  $NH_2$  terminus. The limitation to the terminal regions is consistent with other CAD data of large polypeptides having intermediate charge states and showing fragmentation generally occurring with decreased efficiency at greater distances from the ends of the molecule (13, 15, 16, 18). However, it is important to realize that these spectra are dominated by dissociation of a range of charge states, defined at lower  $m/z$  by the envelope of molecular ion charge states (i.e., those observed at low  $\Delta NS$ ), and at a higher  $m/z$  due to decreased activation efficiency. For native albumins the relevant range of charge states con-



**Figure 3.** ESI mass spectra of (top) dog, (middle) sheep, and (bottom) horse serum albumins with  $\Delta NS = +335$  V.

tributing typically extends from 60+ to perhaps 40+ for the conditions used. Our sequence assignments are facilitated by other factors, such as the expected (18, 30) low probability of observing products due to bond cleavage from regions enclosed by disulfide bonds (62% of bovine albumin is enclosed by its 17 Cys-Cys bonds) because at least two bonds must be broken, increasing the confidence level of our interpretation. Of particular value is the fact that albumins as a class of proteins exhibit a high level of sequence homology. The assumption that dissociation products occur from similar regions of the molecule for all albumin species is a major factor in the tractability of the interpretation procedure. At least two overlapping product ion charge state distributions were observed for all albumin species, providing additional support for our assignments. For example, dissociation of the  $Asn^{18}-Phe^{19}-CO-N-$  bond (giving the  $b_{18}$  ion) of sheep albumin is represented by products at both  $m/z$  1048 (doubly charged) and 699 (triply charged). Cleavage of the  $Ala^{26}-Phe^{27}$  bond ( $b_{26}$  ion) of horse serum albumin is indicated by product ions at  $m/z$  982 (3+), 737 (4+), and 590 (5+). We attribute this distribution of product charge states to the dissociation of parent ions with a variety of charge states. In addition, parent species of the same total charge may have their charges distributed differently throughout the molecule leading to a similar result. It might be possible to aid the interpretation procedure of such spectra by applying deconvolution algo-

**Table III. Electrospray Ionization  $b_n$  Product Ion  $m/z$  for Dog Serum Albumin**

$n^a$	amino acid	3+		4+	
		calcd	exptl	calcd	exptl
22	L	873.3	873.3	655.2	655.3
23	V	906.3	906.5	680.0	680.1
24	L	944.1	944.0	708.3	708.4
25	V	977.1	977.0	733.1	733.1
26	A	1000.8	1000.7	750.8	750.9
27	F	1049.8	1049.9	787.6	787.5
28	S	1078.9	1079.0	809.4	809.6
29	Q/K	1121.6	1121.7	841.4	841.4
30	Y	1176.0	1176.1	882.2	882.4
31	L/I	1213.7	1213.7	910.5	910.5

<sup>a</sup>Residue position from the  $\text{NH}_2$  terminus.

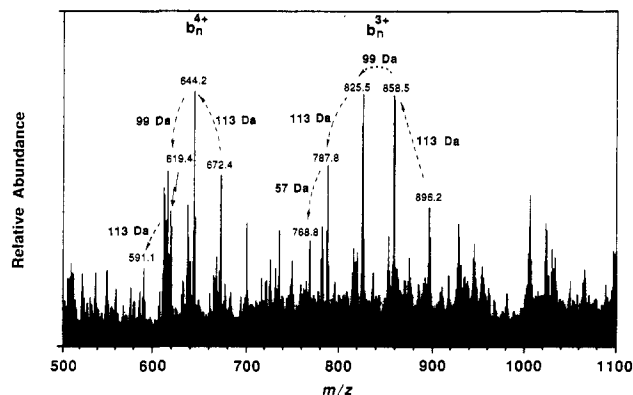
ithms (53), used to "collapse" distributions of multiply charged molecular ions to a single peak, for product ion spectra composed of several fragment ion charge states.

Subtle differences in the degree of charging for the fragment ions for the albumins were also observed and found consistent with the spectral assignments. We have previously demonstrated a good correlation between the number of basic residues and the maximum positive charge for molecular ions observed in ESI mass spectra for a number of polypeptides and proteins (13, 40). In a 5% acetic acid solution, most, if not all, basic groups might be protonated in solution since all  $\text{pK}_a$  values are above 3. Horse, bovine, and rat albumin each have 7 basic amino acid residues (lysine, arginine, and histidine) within the first 28 residues, in addition to the possible  $\text{NH}_2$  terminal charge site, and we observe multiple charging for product ions up to the 5+ state. On the other hand, the most highly charged fragment is 4+ for dog, sheep, rabbit, porcine, goat, and human albumin, which contain either 5 or 6 basic sites within the initial 28 residues.

Distinctive  $m/z$  patterns were obtained for the albumin species analyzed by collisional dissociation as described above. These permitted detailed verification of reported sequences for residues between position ~18 and ~28 from the  $\text{NH}_2$  terminus, depending upon the serum albumin species. For example, a sample of goat serum albumin yielded ESI mass spectra similar to those of sheep albumin at high  $\Delta\text{NS}$  values. Although one difference in the initial 30 residues between goat and sheep albumin has been indicated by conventional Edman degradation sequence analysis (38), as shown in Table I, the ESI mass spectra indicate otherwise. Asparagine resides at position 18 in the case of sheep albumin, whereas histidine is reputed to be present at this site for goat albumin. Thus, a 23-Da shift should be observed and easily identified in the collision spectra for the two species. Instead, identical product ion spectra were obtained with  $\Delta\text{NS} = +335$  V, indicating Asn at position 18 for goat albumin. Peptide-mapping experiments from tryptic digestion followed by separate HPLC analysis and ESI-MS (data not shown) also revealed similarities in their primary structure for the samples examined.

Unfortunately, the additional tyrosine residue at position 156 and the reversal of residues 94–95 from –GluGln– to –GlnGlu– detected for bovine serum albumin by Hirayama et al. (47) were not observed in our work because only the  $\text{NH}_2$  terminus was probed.

In the case of dog serum albumin, only the first 24 residues are available from the literature (37). Along with known sequence homologies, the experimentally observed product ions allowed prediction of the sequence for residues 25–31, as indicated by Table III. Ions  $b_{22}$  to  $b_{24}$  were assigned from calculations for all possible  $m/z$  values based upon the available amino acid sequence. Remaining products were assigned as sequential  $b_n$  dissociation products starting from



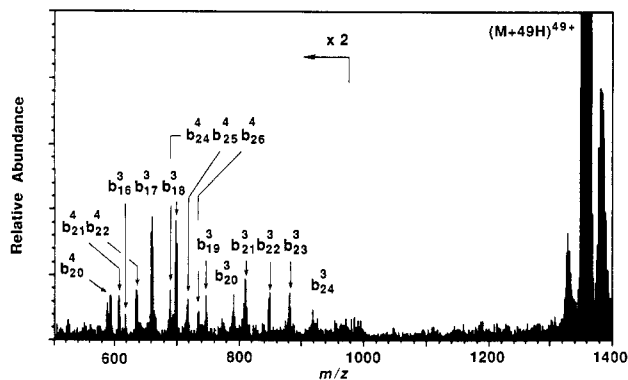
**Figure 4.** Partial ESI mass spectrum of guinea pig serum albumin with  $\Delta\text{NS} = +335$  V.

$b_{24}$ . For example, the ion at  $m/z$  708.3 was assigned as  $b_{24}^{4+}$  with a mass of 2830 Da. The next major ion at higher  $m/z$  is  $m/z$  733.1; its mass for a 4+ ion would be 2929.2 Da, or an increase of 99 Da, corresponding to a valine residue ( $b_{25}^{4+}$ ). An overlapping suite of triply charged ions is also present and was used for confirmation of the assignments. An ion at  $m/z$  977.0 is assigned as the  $b_{25}^{3+}$  product, and so forth to  $b_{31}^{4+}$ . In the case of the sequence for dog albumin, it was not possible with our current instrumentation to distinguish the presence of lysine versus glutamine (0.04-Da difference) at position 29 or leucine versus isoleucine at position 31 from the data.

It should be noted that although major fragmentation generally occurs between residues 18–28 for the serum albumins, allowing for sequence determination in this region, it also allows for a limited verification of the previous 1–17 residues by comparison of the  $M_r$  of the smallest size product ion to the expected value. For example, the smallest  $M_r$  fragment ion observed for dog serum albumin ( $b_{22}$ ) is in good agreement with the  $M_r$  calculated for the first 22 residues. However, this does not provide data for the residue sequence. Nor does it rule out possible substitutions that sum to the same  $M_r$  (e.g., Asp for Leu (+2 Da) and Val for Thr (–2 Da)). MS/MS of the multiply charged product ions could possibly provide this additional information (see following section).

Interpretation of the collision mass spectrum for guinea pig serum albumin was slightly more difficult. To our knowledge, no published sequence information is available. In addition, as indicated in Table II and Figure 2, three resolvable components differing by approximately 300 Da were suggested by the ESI mass spectrum. Also evident was a set of low abundance peaks between  $m/z$  750 and 1100 in the low-voltage  $\Delta\text{NS}$  spectra that correlate to a molecular species of  $M_r$  13.9 kDa (18+ to 13+ charges) and may also contribute to the collisional dissociation spectra. The spacings between the most abundant peaks in the spectrum at high  $\Delta\text{NS}$  (Figure 4), ascribed as series of triply and quadruply charged product ions, indicate a sequence of amino acids separated in mass by 57, 113, 99, and 113 Da. This suggests a portion of the sequence to be Gly-Leu(Ile)-Val-Leu(Ile) (although evidence for Gly is weaker because of the poor signal-to-noise ratio for the 4+ ion). This is similar to residues 21–24 of most of the other albumin species listed in Table I. Many peaks in the same spectral region remain unassigned because of the complexity of this mixture. Simplification of the mixture by preseparation and analysis of the individual constituents may provide for a more complete sequence determination.

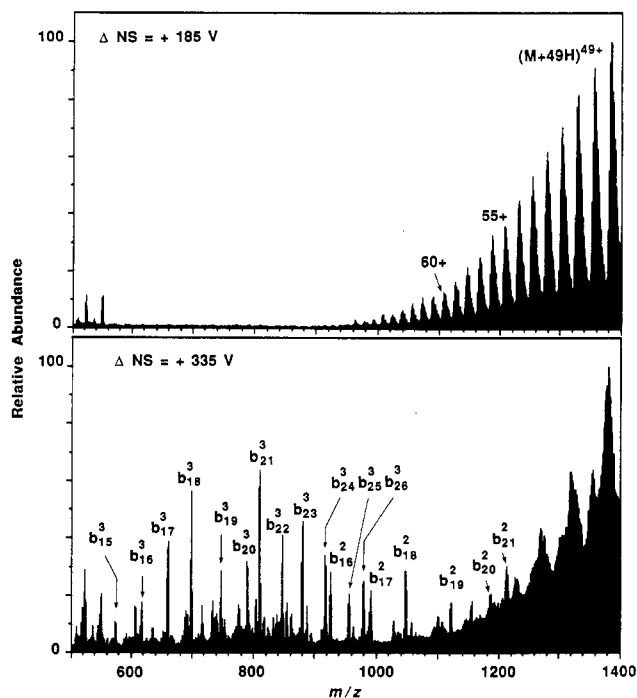
**Tandem Mass Spectrometry.** Dissociation of selected molecular ion charge states for proteins as large as 66 kDa has been achieved in this study, a factor ~20 larger than feasible for singly charged ions using equivalent methods. Results from collisionally activated dissociation tandem MS of the  $(M + 49\text{H})^{49+}$  ion of sheep albumin is shown in Figure 5. The



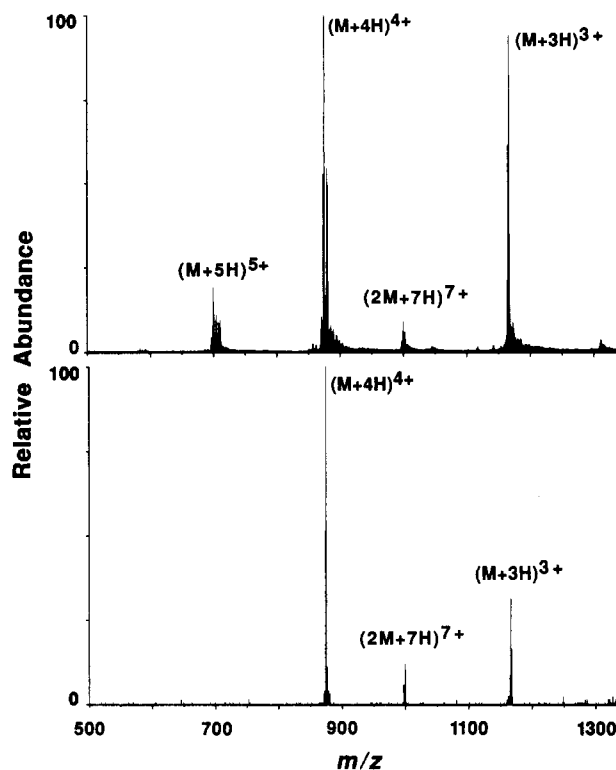
**Figure 5.** Tandem mass spectrum of the  $(M + 49H)^{49+}$  molecular ion of sheep albumin with  $E_{lab} = 6860$  V (argon collision target).

fragment ions identified in the tandem MS experiments were generally of the same  $m/z$  values produced by CAD in the nozzle/skimmer interface (compare with Figure 3). These CAD studies used argon as the neutral target and a laboratory-frame translational energy ( $E_{lab} = q\Delta V$ , where  $q$  is the net ion charge and  $\Delta V$  is the voltage difference between the skimmer and the quadrupole collision cell) of 6.86 keV. The resolution in the first quadrupole was lowered to increase parent ion intensity for subsequent CAD, such that the peak width at half-maximum was 7–8  $m/z$  units compared to 5  $m/z$  units used in Table II.

In addition to the product ions at lower  $m/z$ , all MS/MS studies of albumins gave product ions near the parent species, such as shown in Figure 5, with peaks at approximately  $m/z$  1329 and 1384 for sheep albumin. There are several possible origins for these peaks, three of which are discussed below. Their  $m/z$  values closely correspond to intact molecular ions with 50+ and 48+ charges, respectively. Charge-transfer/charge-stripping reactions, although postulated previously (17, 54), seem unlikely for the experimental conditions. A second possibility is that these peaks might also arise due to fragmentation producing a species of approximately 33.2 kDa, roughly half the intact molecule. Support for this is found in the ESI mass spectrum shown in Figure 6 (bottom) for the collisional dissociation of sheep albumin at  $\Delta NS = +335$  V. Above the broad rising baseline extending to the instrumental  $m/z$  limit are very broadened peaks at  $m/z$  1384, 1329, 1277, and 1230 that correspond to a 33.2-kDa fragment (with 24+ to 27+ charges, respectively). Alternatively, dissociation of a 49+ ion from a 66.4-kDa molecule may also produce a complementary ion pair (i.e., two charged fragments that in sum total the parent ion) of nearly equal size, 33.2 kDa, with 24+ and 25+ charges. One possible dissociation site to explain a 33.2-kDa fragment would be in the vicinity of Glu<sup>291</sup>, in a region near the middle of the primary sequence that is unbound by Cys–Cys bridges (i.e., only a single bond cleavage is necessary to liberate two charged fragments). A third possibility, which we consider most likely, is that these ions arise from dissociation of a noncovalently bound dimer [i.e.,  $(2M + 98H)^{98+}$ , whose  $m/z$  position would be essentially the same as the  $(M + 49H)^{49+}$  ion]. It is well established that multimeric proteins generally separate into their noncovalently bonded species with normal ESI-MS conditions (12, 13), and one could expect a dimer would fragment readily. Such a dimer species may be present in low abundance ( $\leq 5\%$  of the monomer on the basis of Figure 2) due to the relatively high sample concentration. Closer inspection of ESI mass spectra of bovine serum albumin obtained on  $m/z > 1400$  limit quadrupole instruments (43, 47) show peaks at high  $m/z$  in between the major peaks, similar to Figure 2 for porcine albumin, that may be indicative of dimeric species. We have observed dimeric species of smaller peptides, such as bovine



**Figure 6.** ESI mass spectra of sheep serum albumin ( $M_r \sim 66\,400$  Da) with nozzle/skimmer potential of (top) +185 V and (bottom) +335 V.



**Figure 7.** (Top) Electrospray ionization mass spectra of bovine insulin B-chain (oxidized) and (bottom) ESI tandem mass spectrum showing dissociation of the  $(2M + 7H)^{7+}$  ion ( $E_{lab} = 525$  eV).

insulin A-chain (oxidized,  $M_r$  2533 Da) and B-chain (oxidized,  $M_r$  3496 Da). MS/MS of the  $(2M + 7H)^{7+}$  of B-chain insulin shown in Figure 7 yields the  $(M + 3H)^{3+}$  and  $(M + 4H)^{4+}$  monomeric ions. Similarly, dissociation of the  $(2M - 7H)^{7-}$  ion from A-chain insulin gives the  $(M - 3H)^{3-}$  and  $(M - 4H)^{4-}$  products. We have also recently interpreted similar earlier results for cytochrome *c* and myoglobin in terms of such dimer dissociation processes (55). While this explanation seems most likely for the albumins, definitive distinction among these



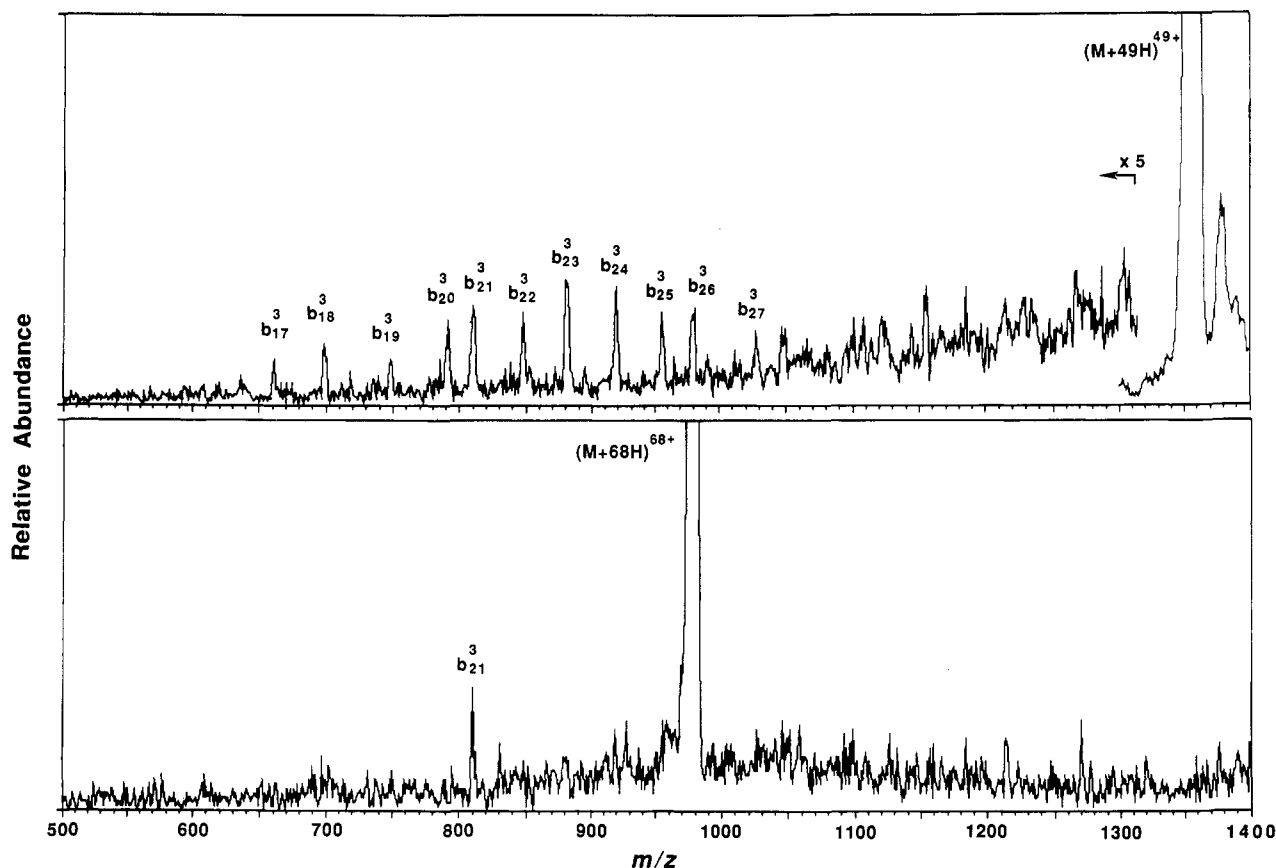


Figure 9. MS/MS spectra of the  $(M + 49H)^{49+}$  (top) and  $(M + 68H)^{68+}$  (bottom) parent ions from disulfide-reduced sheep albumin.

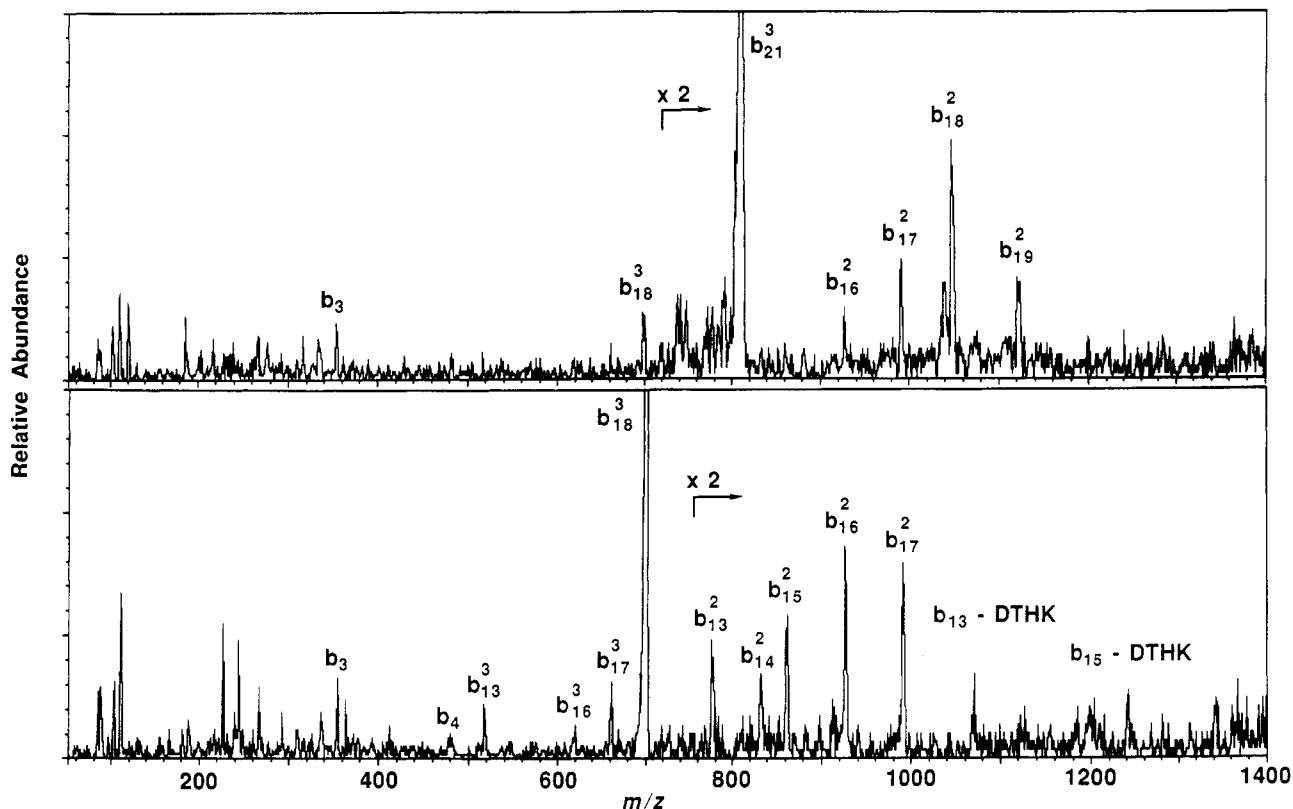
lombic effects depends upon the internal energy of the molecular ions; at high levels of excitation branching ratios are increasingly determined by the preexponential "frequency factor" associated with the dissociation pathway (57). Bunker and Wang (58) have previously shown that for uncharged (or singly charged) molecules, in the absence of Coulombic effects, the frequency factor (which is related to the reduced mass of the products for a simple linear homopolymer) is greatest for cleavage near the chain termini, and at the large molecule limit approaches zero for particular cleavages near the chain center. Our calculations (57) show that consideration of Coulombic forces completely reverses this situation if a sufficiently high charge state is obtained. Of particular relevance, however, is that at levels of charge and excitation qualitatively similar to those of our MS/MS studies, Coulombic effects can range from small to dominant (57). Thus higher charge states may dissociate more rapidly, but more importantly, may dissociate to yield products due to cleavages nearer the molecule's center which are not evident for lower charge state molecular ions. This may explain, at least in part, the fact that products evident for the  $(M + 49H)^{49+}$  MS/MS spectra are not prominent for the  $(M + 68H)^{68+}$  spectra, even though CAD efficiencies are greater.

For the albumin species, most of the fragmentation is centered around residue 20 from the  $NH_2$  terminus, with the various multiple charge distributions providing information on a series of  $\sim 10$  sequential residues. These originate in the molecular region known as subdomain IA (33). The three-dimensional structure for human serum albumin was recently obtained with 4-Å resolution (59, 60). It is tempting to correlate such 3-D structural information with observed gas-phase dissociation processes, but more information and additional experimentation is required before this will be feasible.

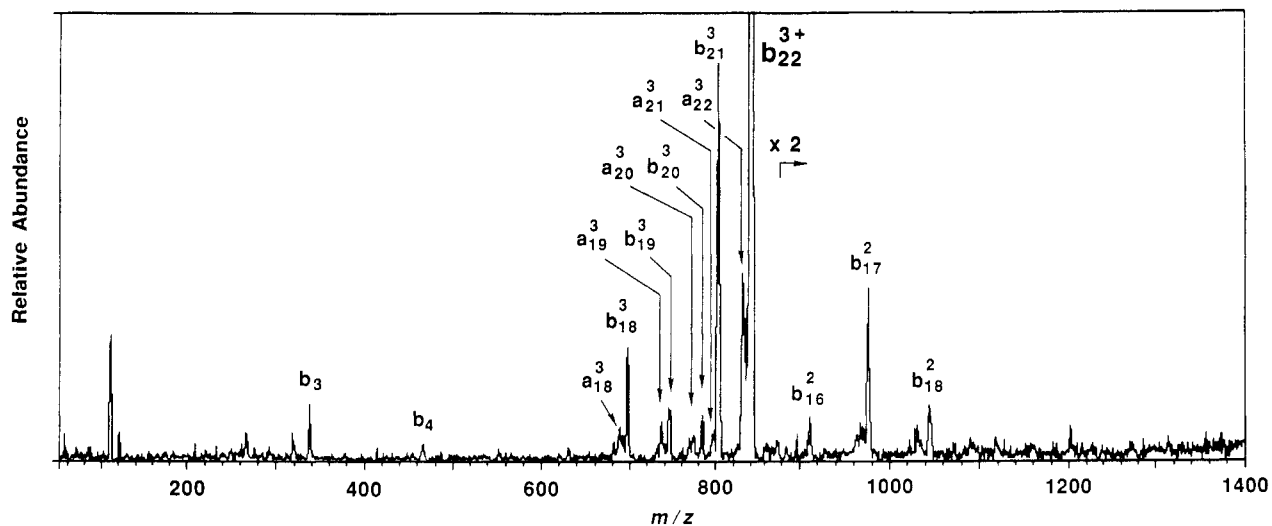
It is unclear how higher order protein structure (in solution, and perhaps even gas phase) may also affect the dissociation process. Chait and co-workers (61, 62) and our laboratory (63)

have demonstrated the effects of protein conformations (three-dimensional structure) on the resulting ESI mass spectra. Factors such as disulfide reduction (13, 18, 40) and solvent conditions such as pH (61, 63) and organic solvent content (62, 63) can have dramatic roles in the higher order structure and the observed charge distribution in ESI mass spectra. We have also previously observed possible higher order structural effects on collisional dissociation spectra for ribonuclease A ( $M_r$  13 682 Da) (18) by comparison of MS/MS spectra of the native versus disulfide-reduced forms for the same parent charge state. It is arguable that removal of structural inhibitions, such as the reduction of disulfide bonds, allows the molecule to relax to a more extended conformation. Whether the addition of more positive charges forces the molecule to an even more "stretched" form due to electrostatic charge repulsion will likely depend on both thermodynamic and kinetic constraints.

**CAD/MS/CAD/MS.** Tandem mass spectrometry can be further extended by CAD of fragment ions produced at an elevated  $\Delta NS$  bias in the atmospheric pressure/vacuum interface. These experiments, effectively MS/MS/MS or  $MS^3$ , can be used to confirm product ion assignments and, in more favorable cases (15, 18), extend the amount of sequence information. This experiment is analogous to MS/MS of small peptides generated by enzymatic digestion of a larger protein for the purpose of complete sequence determination (21-26). CAD/MS/CAD/MS experiments for sheep albumin were performed on the triply charged  $b_{18}$  and  $b_{21}$  fragments with  $\Delta NS = +335$  V (Figure 10). Doubly charged  $b_{16}$ - $b_{19}$ ,  $b_{18}^2$ , and singly charged  $b_2$ - $b_4$  and  $y_2$ - $y_3$  are found for the  $b_{21}^3$  parent ion. CAD of  $b_{18}^3$  yields  $b_{13}^2$ - $b_{17}^2$ ,  $b_{16}^3$ - $b_{17}^3$ , and  $b_3$ - $b_4$ , along with weak signals indicating internal fragmentation designated as ( $b_{13}$  - DTHK) to ( $b_{16}$  - DTHK) (see figure caption for nomenclature). Although only limited additional sequence information was obtained relative to MS/MS of the molecular ions (compare, with Figure 3, middle) due to the currently



**Figure 10.** MS/MS spectra of the  $b_{21}^3$  (top) and  $b_{18}^3$  (bottom) product ions of sheep albumin with  $\Delta NS = +335$  V and  $E_{lab} = 429$  eV for CAD. Internal fragmentation due to additional loss of the initial four residues, Asp Thr His Lys is denoted by " $b_n$ -DTHK", where D = Asp, T = Thr, H = His, and K = Lys.



**Figure 11.** ESI-MS/MS spectrum of the  $b_{22}^{3+}$  sequence ion of rabbit albumin with  $\Delta NS = +335$  V and  $E_{lab} = 420$  eV.

limited sensitivity and resolution, such studies provide information confirming the original product ion assignments and serve to demonstrate the potential of extended MS analysis (i.e.,  $MS^n$ , where  $n \geq 3$ ) for more complete structural analysis.

CAD/MS/CAD/MS experiments have also provided support for peptide sequence assignments for rabbit serum albumin. The ESI mass spectrum of rabbit serum albumin at high  $\Delta NS$  (Figure 1, bottom) shows most of the intense peaks between  $m/z$  500 and 1000 are observed at  $m/z$  values lower than predicted based on its amino acid sequence (38). Triply charged ions corresponding to  $b_{18}^3$  and  $b_{19}^3$  of the published sequence were found. However,  $b_{20}^3$ - $b_{26}^3$  and  $b_{20}^4$ - $b_{25}^4$  ions are  $\sim 15$  Da too low, suggesting the possibility that residue 20 is not Lys (incremental mass = 128 Da) but an alternative residue of mass  $\sim 113$  Da such as Ile or Leu (as indicated in

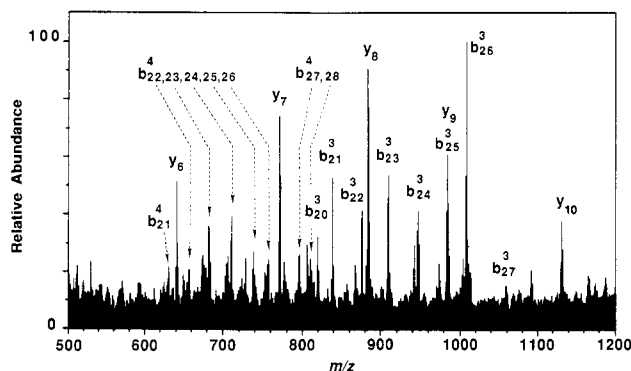
Table IV). Interestingly, position 20 (Lys) is homologous for 7 species reported in this study and for hamster albumin, as reported by Vötsch et al. (38). Our suggestion of an error in the previously suggested sequence is supported by an additional study involving further dissociation of the  $b_{22}^{3+}$  product ion (Figure 11), resulting in doubly charged  $b_{16}$ - $b_{18}$  ions and triply charged  $a_{18}$ - $a_{22}$  and  $b_{18}$ - $b_{21}$  products at  $m/z$  values consistent with Ile/Leu at position 20.

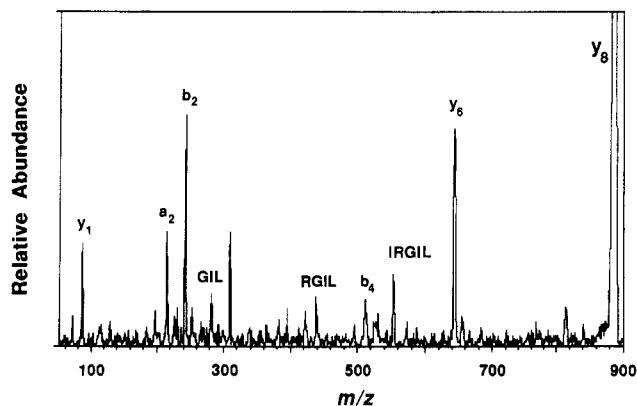
The nucleotide sequence of porcine albumin was recently elucidated to reveal a protein of 583 amino acids (36). An  $M_r$  value from the measured ESI mass spectrum (Table II) is in relatively close agreement with the predicted value (see previous discussion). The interface CAD spectrum ( $\Delta NS = +335$  V) shows the expected 3+ and 4+  $NH_2$  terminal fragment ions ( $b_{20}$ - $b_{28}$ ) (Figure 12). Further dissociation of the  $b_{26}^3$  product ion ( $m/z$  1009) yielded a weak spectrum showing ions at  $m/z$

**Table IV. Electrospray Ionization  $b_n$  Product Ion  $m/z$  for Rabbit Serum Albumin**

$n^a$	amino acid	3+			4+		
		calcd <sup>b</sup>	calcd <sup>c</sup>	exptl	calcd <sup>b</sup>	calcd <sup>c</sup>	exptl
18	H	696.4	696.4	696.5	522.6	522.6	
19	F	745.5	745.5	745.5	559.4	559.4	
20	K <sup>b</sup> (I/L) <sup>c</sup>	788.2	783.2	783.3	591.4	587.6	587.6
21	G	807.2	802.2	802.3	605.7	601.9	602.0
22	L	844.9	839.9	840.0	634.0	630.2	630.2
23	V	878.0	873.0	873.1	658.7	655.0	654.9
24	L	915.7	910.7	910.8	687.0	683.3	683.3
25	I	953.4	948.4	948.5	715.3	711.6	711.5

<sup>a</sup> Residue position from the  $\text{NH}_2$  terminus. <sup>b</sup> According to published sequence (38) with E at position 17. <sup>c</sup> According to best sequence fit of ESI-MS data.





**Figure 14.** CAD/MS/CAD/MS of the  $y_8$  product ion (Ile Glu Ile Arg Gly Ile Leu Ala) of porcine albumin with  $\Delta NS = +360$  and  $E_{lab} = 140$  eV. Internal fragmentation due to loss of residues from both termini of the  $y_8$  ion are denoted by the single-letter codes for the amino acids (e.g., GIL is Gly Ile Leu).

molecules arise due to parent ion charge state (13, 15–18). Ions of low  $m/z$  (higher charge) can be more efficiently activated and dissociated than ions of lower charge state. The dissociation processes must be sufficiently rapid for dissociation during the  $10^{-4}$ -s residence time in the collision quadrupole region.

In this work we have shown that useful structural information is obtainable for multiply charged molecules of at least 66 kDa, a factor of about 20 larger than that reported for singly charged ions. Although the sequence information obtained is presently limited for intact molecules of such size ( $\sim 10$  residues out of  $\sim 580$  amino acids), the potential of the technique is clearly demonstrated. Currently, for complete primary sequence information of proteins by mass spectrometry, the use of proteolytic enzymes followed by FAB-MS/MS (21–26) or even ESI-MS/MS (64) of the resulting peptides is the method of choice. However, the development of CAD-MS/MS for multiply charged ions is still very much in its infancy.

ESI mass spectra and CAD/MS have been used to differentiate among 66-kDa serum albumin proteins from 10 species. In two cases, assignment of interface CAD and conventional MS/MS spectra have unexpectedly indicated differences near the  $NH_2$  terminus with reported amino acid sequences. However, we note that all albumins analyzed were commercially obtained (Sigma) and may show differences from samples previously sequenced by conventional Edman methods (38). Strong support for our product ion assignments was provided by further dissociation of fragment species produced in the atmospheric pressure/vacuum interface. This capability is a uniquely valuable feature of ESI with tandem mass spectrometry and can potentially be used to probe molecular regions unaccessed by the initial CAD process. However, for the albumins, due to the limited  $m/z$  and sensitivity of our quadrupole instrument, only the initial  $NH_2$  terminal 30 residues and, in the case for porcine albumin, 10 residues from the  $COOH$  terminus, could be probed. It is conceivable that ESI-MS/MS may have utility for rapid detection of genetic variants. Over 30 different human serum albumin variants are known, mostly due to single-point mutations in the albumin gene (65). Some have altered residues near the  $NH_2$  terminus or the  $COOH$  terminus that may be probed by ESI-MS/MS.

For molecules of such large  $M_r$ , we find that the dominant fragmentation products for moderate charge states (i.e.,  $\sim +50$ ) are limited to regions near the ends of the molecule, consistent with the RRKM model used by Bunker and Wang (58) for the large-molecule limit. A recent theoretical study (57) suggests that dissociation from the center region will become

more important as charge state increases. We also note that both the MS/CAD/MS and CAD/MS/CAD/MS spectra also suggest an abundance of lower intensity dissociation processes that cannot be suitably examined at present due to limited sensitivity and spectral "congestion" arising from the range of potential products and charge states. In addition, it is possible that larger fragmentation products from regions more interior of the molecule and possible complementary ions may be produced, yet may fall outside our current  $m/z$  limit of detection.

An exciting possibility of ESI is the future use of more sensitive and higher resolution MS methods which might allow much more extensive sequence information to be obtained. It may be possible that ion-trapping methods, such as ion trap MS (41) and Fourier transform ion cyclotron resonance MS (44, 50) allowing sequential  $MS^n$  analyses may provide much more extensive primary structure determination in only a small fraction of time required by conventional methodologies.  $MS^4$  spectra for multiply charged ions from peptides have recently been accomplished with ESI ion trap MS (41, 66). Interpretation of such collisional dissociation spectra for "unknowns" is currently extremely difficult due to the complexity of the spectra and the uncertainty of product ion charge states. With higher resolution methods (e.g., FTMS), charge-state assignments from either resolved isotopic peaks or unresolved isotopic envelopes is possible (44). Indeed, charge-state assignments from resolved isotopic peaks of up to the 18+ molecular ion charge state for equine myoglobin by ESI-FTMS have been published by McLafferty and co-workers (67). Such methods would largely circumvent the sensitivity and resolution limitations of the present triple-quadrupole methods, suggesting the potential for extension to even greater molecular weights.

#### ACKNOWLEDGMENT

We thank J. A. Burghard, D. G. Camp, and R. R. Ogorzalek Loo for experimental assistance and encouragement.

#### LITERATURE CITED

- (1) *The Analysis of Peptides and Proteins by Mass Spectrometry*; McNeal, C. J., Ed.; John Wiley & Sons: Chichester, U.K., 1988.
- (2) *Mass Spectrometry of Biological Materials*; McEwen, C. N., Larsen, B. S., Eds.; Marcel Dekker: New York, 1990.
- (3) Barber, M.; Bordoli, R. S.; Elliott, G. J.; Sedgwick, R. D.; Tyler, A. N. *Anal. Chem.* **1982**, *54*, 645A–657A.
- (4) Barber, M.; Green, B. N. *Rapid Commun. Mass Spectrom.* **1987**, *1*, 80–83.
- (5) Sundqvist, B.; Macfarlane, R. D. *Mass Spectrom. Rev.* **1985**, *4*, 421–460.
- (6) Karas, M.; Bahr, U.; Ingendoh, A.; Hillenkamp, F. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 760–761.
- (7) Karas, M.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Processes* **1989**, *92*, 231–242.
- (8) Beavis, R. C.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 233–237.
- (9) Meng, C. K.; Mann, M.; Fenn, J. B. *Z. Phys. D: At., Mol. Clusters* **1988**, *10*, 361–368.
- (10) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.
- (11) Covey, T. R.; Bonner, R. F.; Shushan, B. I.; Henion, J. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 249–256.
- (12) Loo, J. A.; Udseth, H. R.; Smith, R. D. *Anal. Biochem.* **1989**, *179*, 404–412.
- (13) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882–899.
- (14) Feng, R.; Bouthillier, F.; Konishi, Y.; Cygler, M. Presented at the 39th ASMS Conference on Mass Spectrometry and Allied Topics, May 1991, Nashville, TN.
- (15) Smith, R. D.; Loo, J. A.; Barinaga, C. J.; Edmonds, C. G.; Udseth, H. R. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 53–65.
- (16) Barinaga, C. J.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 160–164.
- (17) Smith, R. D.; Barinaga, C. J. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 54–57.
- (18) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Science* **1990**, *248*, 201–204.
- (19) *Tandem Mass Spectrometry*; McLafferty, F. W., Ed.; John Wiley & Sons: New York, 1983.
- (20) Busch, K. L.; Gilsh, G. L.; McLuckey, S. A. *Mass Spectrometry/Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry*; VCH: New York, 1988.

- (21) Blemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992-998.
- (22) Blemann, K. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99-111.
- (23) Carr, S. A.; Green, B. N.; Hemling, M. E.; Roberts, G. D.; Anderegg, R. J.; Vickers, R. Presented at the 35th ASMS Conference on Mass Spectrometry and Allied Topics, May 1987, Denver, CO (conference proceedings pp 830-831).
- (24) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233-6237.
- (25) Johnson, R. S.; Blemann, K. *Biochemistry* **1987**, *26*, 1209-1214.
- (26) Johnson, R. S.; Mathews, W. R.; Blemann, K.; Hopper, S. J. *Biol. Chem.* **1988**, *263*, 9589-9597.
- (27) Gross, M. L.; Tomer, K. B.; Cerny, R. L.; Giblin, D. E. In *Mass Spectrometry in the Analysis of Large Molecules*; McNeal, C. J., Ed.; Wiley: Chichester, U.K., 1986; pp 171-190.
- (28) Neumann, G. M.; Derrick, P. J. *Org. Mass Spectrom.* **1984**, *19*, 165-170.
- (29) Neumann, G. M.; Sheil, M. M.; Derrick, P. J. *Z. Naturforsch.* **1984**, *39A*, 584-592.
- (30) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chim. Acta* **1990**, *241*, 167-173.
- (31) Loo, J. A.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 207-210.
- (32) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. *Anal. Chem.* **1988**, *60*, 1948-1952.
- (33) Peters, T., Jr. In *Advances in Protein Chemistry*; Anfinsen, C. B., Edsall, J. T., Richards, F. M., Eds.; Academic Press: Orlando, FL, 1985; Vol. 37, pp 161-245.
- (34) Sepulveda, P. M. J. Ph.D. Thesis, University of Texas at Austin, 1972.
- (35) Brown, W. M.; Dzielgielewska, K. M.; Foreman, R. C.; Sanders, N. R. *Nucl. Acids Res.* **1989**, *17*, 10495.
- (36) Weinstock, J.; Baldwin, G. S. *Nucl. Acids Res.* **1988**, *16*, 9045.
- (37) Dixon, J. W.; Sarkar, B. *J. Biol. Chem.* **1974**, *249*, 5872-5877.
- (38) Vötsch, W.; Wagner, H. A.; Anderer, F. A. *Comp. Biochem. Physiol.* **1980**, *66B*, 87-91.
- (39) Chincarini, C. C.; Brown, J. R. *Fed. Proc.* **1976**, *35*, 1334.
- (40) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1990**, *62*, 693-698.
- (41) Van Berkel, G. J.; Glish, G. L.; McLuckey, S. A. *Anal. Chem.* **1990**, *62*, 1284-1295.
- (42) Chowdhury, S. K.; Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 81-87.
- (43) Feng, R.; Konishi, Y.; Bell, A. W. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 387-401.
- (44) Henry, K. D.; McLafferty, F. W. *Org. Mass Spectrom.* **1990**, *25*, 490-492.
- (45) Gallagher, R. T.; Chapman, J. R.; Mann, M. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 369-372.
- (46) Gelsow, M. J.; Harris, R.; Dodsworth, N.; Green, B. N.; Hutton, T. In *Techniques in Protein Chemistry II*; Villafranca, J. J., Ed.; Academic Press: San Diego, CA, 1991; pp 567-572.
- (47) Hirayama, K.; Akashi, S.; Furiya, M.; Fukuhara, K. *Biochem. Biophys. Res. Commun.* **1990**, *173*, 639-646.
- (48) Chan, K.; Wintergrass, D.; Straub, K. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 139-143.
- (49) Cleland, W. W. *Biochemistry* **1984**, *3*, 480-482.
- (50) Henry, K. D.; Williams, E. R.; Wang, B. H.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9075-9078.
- (51) Chowdhury, S. K.; Katta, V.; Beavis, R. C.; Chait, B. T. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 382-388.
- (52) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
- (53) Mann, M.; Meng, C. K.; Fenn, J. B. *Anal. Chem.* **1989**, *61*, 1702-1708.
- (54) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. *J. Phys. Chem.* **1989**, *93*, 5019-5022.
- (55) Smith, R. D.; Loo, J. A.; Ogorzalek Loo, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.*, in press.
- (56) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Ogorzalek Loo, R. R.; Smith, R. D. In *Experimental Mass Spectrometry*; Russell, D. H., Ed.; Plenum: New York, in press.
- (57) Rockwood, A. L.; Busman, M.; Smith, R. D. *Int. J. Mass Spectrom. Ion Proc.*, in press.
- (58) Bunker, D. L.; Wang, F.-M. *J. Am. Chem. Soc.* **1977**, *99*, 7457-7459.
- (59) Carter, D. C.; He, X.-M.; Munson, S. H.; Twigg, P. D.; Gernert, K. M.; Broom, M. B.; Miller, T. Y. *Science* **1989**, *244*, 1195-1198.
- (60) Carter, D. C.; He, X.-M. *Science* **1990**, *249*, 302-303.
- (61) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, *112*, 9012-9013.
- (62) Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 214-217.
- (63) Loo, J. A.; Ogorzalek Loo, R. R.; Udseth, H. R.; Edmonds, C. G.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 101-105.
- (64) Covey, T. R.; Huang, E. C.; Henion, J. D. *Anal. Chem.* **1991**, *63*, 1193-1200.
- (65) Watkins, S.; Madison, J.; Davis, E.; Sakamoto, Y.; Gilliano, M.; Minchiotti, L.; Putnam, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5959-5963.
- (66) Jardine, I.; Hail, M.; Lewis, S.; Zhou, J.; Schwartz, J.; Whitehouse, C. Presented at the 38th ASMS Conference on Mass Spectrometry and Allied Topics, June 1990, Tucson, AZ (conference proceedings pp 16-17).
- (67) Henry, K. D.; Quinn, J. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1991**, *113*, 5447-5449.

RECEIVED for review May 13, 1991. Accepted August 9, 1991. We are grateful for the support of the U.S. Department of Energy, Office of Health and Environmental Research (Grant DE-AC06-76RLO 1830), the National Science Foundation, Instrumentation and Instrument Development Program (Grant DIR 8908096), and the National Institutes of Health, National Center for Human Genome Research (Grant HG 00327). Pacific Northwest Laboratory is operated by Battelle Memorial Institute.

## Nondependence of Diffusion-Controlled Peak Dispersion on Diffusion Coefficient and Ionic Mobility in Capillary Zone Electrophoresis without Electroosmotic Flow

Ernst Kenndler\* and Christine Schwer

Institute for Analytical Chemistry, University of Vienna, Währingerstrasse 38, A 1090 Vienna, Austria

The ratio of the ionic mobility,  $u$ , and the diffusion coefficient,  $D$ , in the expression for the plate number,  $N$ , given by  $N = uU/2D$ , ( $U$  being the applied voltage) for capillary zone electrophoresis without electroosmotic flow was substituted by the Einstein relation between these two parameters. As a consequence, the plate number is dependent exclusively on the charge number of the individual analyte, but not on  $u$  and  $D$ . The validity of this conclusion was demonstrated experimentally with anions of different charge number. Further, the plate number was found to depend on the charge of weak acids according to the degree of dissociation, as expected from theory.

### INTRODUCTION

Peak dispersion in capillary zone electrophoresis can be described by a surprisingly simple theory, when only diffusion in the longitudinal direction, the direction of the electrophoretic migration of the ions, is considered. In this case (when dispersion due to the radial temperature profile within the capillary, and peak broadening caused by electroosmotic or hydrodynamic flow, by extracolumn effects and by the difference in electric conductivity between the analyte zone and the buffering electrolyte are neglected), the variance,  $\sigma_z^2$ , of the solute peak, based on length, is dependent on the diffusion coefficient,  $D$ , and the residence time,  $t$ , given by