International Journal of Mass Spectrometry 330–332 (2012) 35–45

Contents lists available at SciVerse ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

A database of alkali metal-containing peptide cross sections: Influence of metals on size parameters for specific amino acids

Jonathan M. Dilger, Stephen J. Valentine, Matthew S. Glover, Michael A. Ewing, David E. Clemmer*

Department of Chemistry, Indiana University, Bloomington, IN 47405, United States

ARTICLE INFO

Article history:
Received 19 March 2012
Received in revised form 27 April 2012
Accepted 1 May 2012
Available online 29 August 2012

Keywords:
Ion mobility spectrometry
Intrinsic amino acid size parameter
Peptide ion structure

ABSTRACT

Ion mobility/mass spectrometry techniques have been used to generate a cross section database containing 1772 entries (147 singly-, 1325 doubly-, and 300 triply-charged) for protonated and alkalated tryptic peptide ions. Such a large number of values make it possible to assess the influence of alkali metal cations [where the cation (M^+) corresponds to Li^+, Na^+, K^+, or Cs^+] on peptide ion conformation. Peptide ion sizes generally increase with increasing cation size relative to the respective singly- or doubly-protonated species. Intrinsic size parameters for individual amino acid residues for alkali $[Pep+M+H]^{2+}$ and $[Pep+2M]^{2+}$ ions are similar to those obtained from the $[Pep+2H]^{2+}$ ions. However, polar residues (Asp, Glu, Asn, Gln, His, and carboxyamidomethylated Cys), as well as Met, appear to be substantially smaller for metal-containing $[Pep+M+H]^{2+}$ species compared with $[Pep+2H]^{2+}$ species. This suggests these residues form tight binding interactions with these metals. A discussion of these results and the implications regarding structure are provided.

1. Introduction

Metal ions play key roles in enzymatic activity, cellular metabolism, and structural stabilization [1–4]. Because of this, structural studies aim to determine the locations of binding sites [5–7], the nature of metal ion coordination [8–12], and the role that metal–ligand interactions have on structure and function [2–4]. Recently, the structures of metal-containing peptides and proteins in the absence of solvent have been investigated [13–24]. These studies are possible because of the development of soft ionization sources such as electrospray ionization (ESI) [25]. Interestingly, the addition of metals to electrospray solutions often influences the ionization efficiency of macromolecules [26–33], and upon activation, metal ions also influence fragmentation patterns [34–44]. These results suggest the existence of highly-specific metal-binding sites.

In this paper we examine the influence of alkali metal ions (Li^+, Na^+, K^+, and Cs^+) on peptide structure by ion mobility spectrometry (IMS) techniques. We report the first large database of collision cross sections for metal-containing peptides. The cross section of an ion is related to its structure; compact geometries have relatively small cross sections compared with more open geometries [45–47]. In the present work we use the database of cross sections to extract general trends in the sizes of the peptides upon association with different metals. The large scale of the database makes it possible to extract information about the average contributions of specific amino acids to the cross section (when containing these different metal ions). Several databases of cross sections for protonated peptides have been presented [48–51]. These data have been used to derive intrinsic size parameters (ISP)s [48–52] as well as volumes [53] for specific amino acid residues in the absence of solvent. Comparisons of these existing data to the metal ion database provide insight to relatively detailed interactions that arise because of the presence of the metal ion.

We begin our analysis by comparing measurements for protonated ions with values reported as a database of 660 cross sections for protonated peptides [50]. The database presented below contains 1772 cross section entries (818 protonated and 954 alkalated). The current measurements are carried out under different source conditions, using a substantially higher-resolving power drift tube. Thus, one difference between the data presented here and prior work is that we observe more peaks from the higher-resolution measurement for individual peptides. With this said, overall there is good agreement between the values reported previously and those reported here for protonated species giving us high confidence in the accuracy of values for the metal-containing species.

In IMS measurements, it is often assumed that identical cross sections for closely-related species are indicative of structural homology. The ISPs derived from our database suggest when substituting a metal for a proton that this is not always the case. We illustrate this idea by considering the measured cross sections of 208.2 ± 0.6 Å^2 and 207.2 ± 0.2 Å^2 for the doubly-charged forms of the Met-Gln-Ile-Phe-Val-Lys peptide that differ only by

* Corresponding author.
E-mail address: clemmer@indiana.edu (D.E. Clemmer).

1387-3806/$ – see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.ijms.2012.05.001

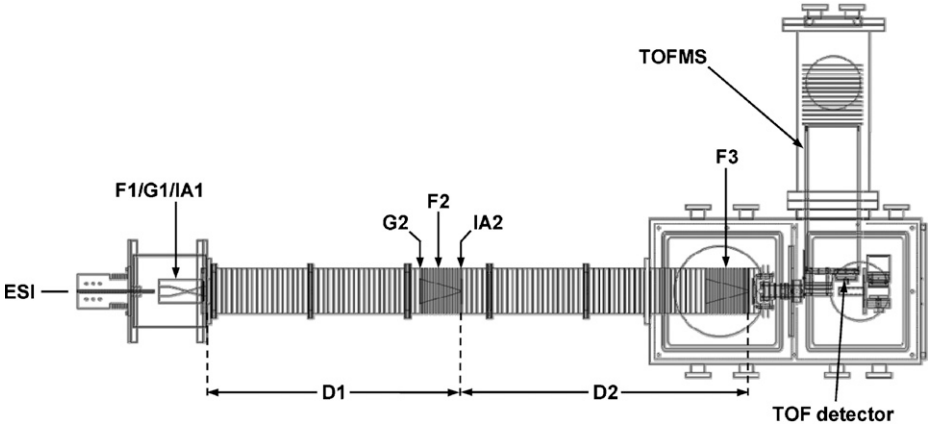

Fig. 1. Schematic diagram of the IMS–IMS–TOF instrument utilized for characterization of all digests. The instrument is comprised of a source region, a 1.8 m drift tube, and a time-of-flight mass spectrometer for subsequent mass analysis. The operational details of the instrument are contained within the text.

inclusion of Na^+. At first glance the nearly identical cross sections suggest that the polypeptide chain exists in a fixed conformation that is not influenced by the charge carrying species. However, the values of ISPs that are derived for all peptides suggest that the cross sections are the same because some residues interact differently with the Na^+ site. Molecular modeling studies of the protonated and sodiated forms are used to illustrate this interpretation.

Finally, the overall approach we have taken here is inspired by studies of the periodic trends in the reactivity of metal ions with small molecules [54,55]. This work shows the key dependence of reaction rates on metal ion kinetic energy and electronic state – benchmark studies that led to a detailed understanding of metal–ligand thermochemistry and reaction mechanisms [56,57]. An understanding of the periodic trends in binding makes it possible to predict values when no data exists [58]. This approach of studying ion-molecule reactions across a region of the periodic table was pioneered by Peter Armentrout's group and is sometimes affectionately referred to as "Armentrouting" a molecule or system by students and colleagues in this field. Here, we extend this experimental philosophy to assess general trends in alkali metal–peptide interactions that influence ion structure.

2. Experimental

General. Nested IMS–MS measurements were performed on a home-built instrument that is described elsewhere [59–61]. A schematic diagram of the instrument is shown in Fig. 1. Mixtures of peptides for tryptic digests of individual proteins are electrosprayed into an hour-glass geometry ion funnel (F1) [62,63], where they accumulate and are periodically pulsed into the drift tube assembly (D1 and D2) (150-μs-wide pulses were used in these studies). Ions are transmitted through the drift tube containing ~3 Torr of He buffer gas (300 K) under the influence of a uniform electric field (~10 V cm^{-1}). To prevent ion losses in the middle of this instrument the ion packets are focused through two ion funnels (F2 and F3). Mobility-separated ions exit the drift tube through a differentially-pumped region and are detected by time-of-flight MS.

Formation of mixtures of peptide ions. Large mixtures of peptides were produced by tryptic digestion of the 24 protein (and protein subunits) listed in Table 1. Peptides were obtained by dissolving 10–100 mg of the whole protein in 2 mL of 200 mM ammonium bicarbonate buffer solution containing 6 M urea. Dithiothreitol (DTT, Sigma–Aldrich D8255-5G) was added in a 40:1 (DTT:protein) molar ratio to each protein solution, followed by a two-hour 37 °C water bath for cleavage of disulfide bonds. Iodoacetamide (IAM, Sigma I6125-5G) was added in an 80:1 (IAM:protein) molar ratio to each protein solution, followed by a two-hour 4 °C ice bath in complete darkness for carboxyamidomethylation of cysteine thiol groups (residue modification denoted as Cys* or C*). Cysteine was added in a 40:1 (Cys:protein) molar ratio and allowed to react at room temperature for 30 min to quench the excess IAM. Finally, TPCK-treated trypsin (Sigma–Aldrich T1426-100MG) was added to provide a 1:50 trypsin:protein ratio (wt:wt). A 200 mM ammonium bicarbonate buffer solution was used for each digestion reaction with a maximum urea molarity of ~2 M. Trypsin reaction mixtures were placed in a 37 °C water bath and allowed to react for 24 h to ensure efficient enzymatic digestion.

Solid-phase extraction was performed using 3 μL Oasis HLB Cartridges (Waters Corporation P/N WAT094226) placed in a vacuum manifold that is maintained at ~5 in Hg. Tryptic digest products were placed into a Centrivap concentrator for evaporation until dryness. The dry peptide powders were resuspended in 1 mL of 50:50 ACN:H_2O solution and diluted to ~0.1 mg mL^{-1} in the same solution; acid was omitted to reduce competition of metal cationic substitution with free protons in solution. The final concentration of peptide content was estimated assuming a 25% recovery of initial protein product. To produce cationized peptides, several different concentrations of metal acetates (obtained from Aldrich or Sigma) were added to a typical ESI solution and examined. A ~0.5 mM metal acetate solution (containing ~0.1 mg mL^{-1} total peptide) was found to produce optimal metal adduction and ESI stability. Separate solutions were prepared to favor protonated peptides.

Peptide ion assignments. Specific peptides were assigned by comparison of measured m/z values with calculated values obtained from the ExPASy Proteomics Server PeptideMass [64] online tool. For all peptides, assignments up to 2 missed cleavages are considered. An in-house peak-picking software was used with a 0.25 m/z tolerance and an intensity threshold of 5 counts to locate all potential peaks above the noise threshold. Drift time distributions were obtained by integrating all m/z bins across a narrow m/z range corresponding to the monoisotopic mass. The centers of prominent features corresponding to peak maxima are recorded in the database. For metalated samples, these assigned peaks were further cross-checked against peak positions expected for protonated species. Any metalated assignment that might also be assigned as different protonated sequences are eliminated to reduce false positives. The veracity of assignments of peaks as to metal-containing peptides were also bolstered by changes in intensities upon adding metals, as well as shifts in m/z and drift time compared with protonated ions.

Table 1
Protein used for trypsin digestion.[a]

Accession number[b]	Abbreviation	Protein (common name)	Organism	Number of residues	Tissue of origin	Protein family
P49822	ALBU_CANFA	Serum Albumin	*Canis familiaris* (dog)	608	Plasma	ALB/AFP/VDB
P19121	ALBU_CHICK	Serum Albumin	*Gallus gallus* (chicken)	615	Plasma	ALB/AFP/VDB
P35747	ALBU_HORSE	Serum Albumin	*Equus caballus* (horse)	607	Plasma	ALB/AFP/VDB
P08835	ALBU_PIG	Serum Albumin	*Sus scrofa* (pig)	607	Plasma	ALB/AFP/VDB
P00921	CAH2_BOVIN	Carbonic anhydrase 2	*Bos taurus* (bovine)	260	Cytoplasm	alpha-carbonic anhydrase
P02662	CASA1_BOVIN	Alpha-S1-casein	*Bos taurus* (bovine)	214	Mammary gland specific	alpha-casein
P02663	CASA2_BOVIN	Alpha-S2-casein	*Bos taurus* (bovine)	222	Mammary gland specific	alpha-casein
P02666	CASB_BOVIN	Beta-casein	*Bos taurus* (bovine)	224	Mammary gland specific	beta-casein
P00004	CYC_HORSE	Cytochrome c	*Equus caballus* (horse)	105	Mitochondrion matrix	cytochrome c
P00924	ENO1_YEAST	Enolase 1	*Saccharomyces cerevisiae* (Baker's yeast)	437	Cytoplasm	enolase
P01966	HBA_BOVIN	Hemoglobin subunit alpha	*Bos taurus* (bovine)	142	Red blood cells	globin
P02070	HBB_BOVIN	Hemoglobin subunit beta	*Bos taurus* (bovine)	145	Red blood cells	globin
P21871	HBA_COLLI	Hemoglobin subunit alpha	*Columba livia* (Domestic pigeon)	142	Red blood cells	globin
P11342	HBB_COLLI	Hemoglobin subunit beta	*Columba livia* (Domestic pigeon)	146	Red blood cells	globin
O12985	HBAD_COLLI	Hemoglobin subunit alpha-D	*Columba livia* (Domestic pigeon)	140	Red blood cells	globin
P69905	HBA_HUMAN	Hemoglobin subunit alpha	*Homo sapiens* (human)	142	Red blood cells	globin
P68871	HBB_HUMAN	Hemoglobin subunit beta	*Homo sapiens* (human)	147	Red blood cells	globin
P01965	HBA_PIG	Hemoglobin subunit alpha	*Sus scrofa* (pig)	141	Red blood cells	globin
P02067	HBB_PIG	Hemoglobin subunit beta	*Sus scrofa* (pig)	147	Red blood cells	globin
P01948	HBA_RABIT	Hemoglobin subunit alpha-1/2	*Oryctolagus cuniculus* (rabbit)	142	Red blood cells	globin
P02057	HBB_RABIT	Hemoglobin subunit beta-1/2	*Oryctolagus cuniculus* (rabbit)	147	Red blood cells	globin
P68240	HBA_SHEEP	Hemoglobin subunit alpha-1/2	*Ovis aries* (sheep)	142	Red blood cells	globin
P02075	HBB_SHEEP	Hemoglobin subunit beta	*Ovis aries* (sheep)	145	Red blood cells	globin
P02754	LACB_BOVIN	Beta-lactoglobulin	*Bos taurus* (bovine)	178	Mammary gland specific	lipocalin
P00711	LALBA_BOVIN	Alpha-lactalbumin	*Bos taurus* (bovine)	142	Mammary gland specific	glycosyl hydrolase 22
P00698	LYSC_CHICK	Lysozyme C	*Gallus gallus* (chicken)	147	Egg white and polymorphonu-clear leukocytes	glycosyl hydrolase 22
P68082	MYG_HORSE	Myoglobin	*Equus caballus* (horse)	154	Muscle	globin
Q29443	TRFE_BOVIN	Serotransferrin	*Bos taurus* (bovine)	704	Plasma	transferrin
P02789	TRFE_CHICK	Ovotransferrin	*Gallus gallus* (chicken)	705	Plasma	transferrin
P02787	TRFE_HUMAN	Serotransferrin	*Homo sapiens* (human)	698	Plasma	transferrin
P24627	TRFL_BOVIN	Lactotransferrin	*Bos taurus* (bovine)	708	Plasma	transferrin
P62979	UBIQ_HUMAN	Ubiquitin	*Homo sapiens* (human)	76	Cytoplasm	ubiquitin

[a] All proteins were obtained from Sigma and used without further purification. Purities were typically greater than or equal to 70%. See text for description of trypsin digestion protocol.
[b] All information presented was obtained within accession number entries on the Universal Protein Resource (UniProt) at http://www.uniprot.org.

Determination of experimental collision cross sections. The total time required for an ion to traverse through the instrument is equivalent to the sum of the drift time, flight time, and the time required to travel through the instrument's interface regions. For a uniform electric field applied along the drift region, collision cross sections can be calculated directly from the recorded distributions using Eq. (1) [65],

$$\Omega = \frac{(18\pi)^{1/2}}{16}\frac{ze}{(k_bT)^{1/2}}\left[\frac{1}{m_I}+\frac{1}{m_B}\right]^{1/2}\frac{t_dE}{L}\frac{76_0}{P}\frac{T}{273.2}\frac{1}{N} \quad (1)$$

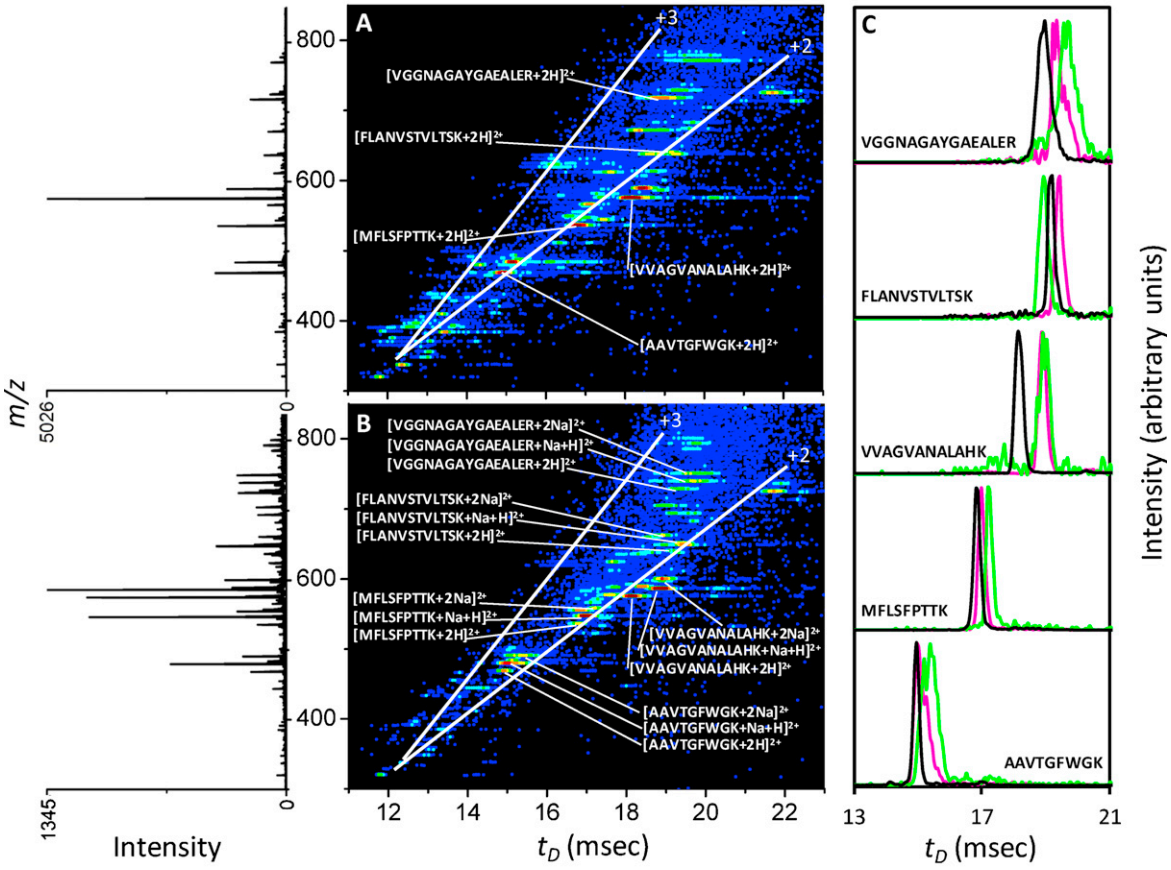

Fig. 2. Nested IMS–MS dot plots for the electrosprayed mixture of peptide ions from the trypsin digest of Hemoglobin from *Ovis aries* (sheep). The datasets have been recorded for samples that contain approximately 0.1 mg mL^{-1} digest in a water:acetonitrile (50:50) solution. The protonated spectrum is shown in panel A and the spectrum obtained for the sample containing sodium acetate is shown in panel B. Mobility families for the doubly- and triply-charged peptide ions are indicated by white lines. Labels have been assigned to five represented peptide ions expected as products of the tryptic digestion. Drift profiles for each of these doubly-charged peptide species are displayed in panel C, with black, pink, and green drift spectra corresponding to $[Pep+2H]^{2+}$, $[Pep+Na+H]^{2+}$, and $[Pep+2Na]^{2+}$ species. Each peak apex in the drift time dimension for the monoisotopic mass is used to derive collision cross section values. A mass spectrum, obtained by integrating all bins across the entire drift time range for each m/z value, is displayed on the left of each plot. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

where ze, k_b, m_I, and m_B refer to the charge of the ion, Boltzmann's constant, the mass of the ion, and the mass of the buffer gas (helium in these experiments). The variables t_d, E, L, P, T, and N correspond to the drift time, electric field, drift tube length, buffer gas pressure and temperature, and the neutral number density of the buffer gas, respectively. The final term in the equation normalizes the mobility to standard temperature and pressure. Due to the presence of non-linear field regions associated with the two ion funnels (F2 and F3), the experimental drift times need to be calibrated. This is done by two methods. First, we calibrate values to those for well-known systems (e.g. bradykinin and polyalanine). Second, we can select ions in the D1 region to vigorously determine a cross section for some ions (as described previously [66]) and calibrate for these values. Both of these calibrations provide highly accurate cross sections. Values determined here for ions that are believed to have the same structures typically agree to within ~1% of values reported previously [50].

3. Results and discussion

Nested IMS–MS measurements and observation of charge-state families for protonated and sodiated peptides. Distributions of peptide ions (either protonated, or containing Li^+, Na^+, K^+, or Cs^+) were produced by ESI. These ions were introduced into the IMS–MS instrument and data were recorded for 5 min for each sample. Example IMS–MS distributions for protonated and sodiated forms of tryptic peptides from sheep hemoglobin are shown in Fig. 2. As reported previously [67], peaks for protonated peptides fall into families that arise because of similarities in the size-to-charge ratio of the ions. Peaks found in the datasets for alkali-containing species fall into similar families. Primarily doubly-charged species and to a lesser extent singly- and triply-charged peptides are observed, consistent with the idea that most peptides contain two highly basic sites (the N-terminal amino group and the C-terminal Lys or Arg side chain). We expect protons to be stabilized by interactions at the basic C-terminal-residue side chain (butylamine in the case of Lys, propylguanidine group in the case of Arg) and the N-terminal amine. A third proton may be incorporated along the peptide backbone or associated with other basic residues, which is expected to be especially important for peptide sequences that arise as a product of missed cleavages (and those containing the His residue). It is less clear where charges will reside upon substitution of alkali metals for a proton. Prior work has shown that substitution of Na^+ for H^+ in polyalanine chains reduces the abundance of globular structures [68] and stabilizes helices through interactions with multiple carbonyl groups on the C-terminal end of the polypeptide chain [13]. We anticipate that larger alkali cations will be stabilized through multiple interactions with electronegative groups.

Changes in drift times upon substituting Na^+ for H^+ in peptide sequences. All of the digests and metals studied here lead to noticeable changes in the drift time distributions compared with

protonated systems. An example of typical changes can be observed from the data in Fig. 2. The five most intense peaks correspond to the AAVTGFWGK, MFLSPTTK, VVAGVANALAHK, FLANVSTVLTSK, and VGGNAGAYGAEALER sequences. For these sequences, we observe the $[Pep+2H]^{2+}$, $[Pep+Na+H]^{2+}$, and $[Pep+2Na]^{2+}$ ions. Drift times generally increase upon substituting Na^+ for H^+. For example, substitution of a single Na^+ to the MFLSPTTK, FLANVSTVLTSK, VVAGVANALAHK, and VGGNAGAYGAEALER sequences leads to a relative drift time increase of 0.9%, 1.4%, 4.0%, and 1.6%, respectively. The AAVTGFWGK sequence appears to contract by 0.3% upon incorporating single sodium. Peaks for the doubly-sodiated AAVTGFWGK, MFLSPTTK, VVAGVANALAHK, and VGGNAGAYGAEALER ions are shifted even more (2.5%, 1.9%, 4.1%, and 3.5%, respectively). In this comparison, the FLANVSTVLTSK ion peak shifts to shorter drift times (by a factor of 1.5%) compared with the respective $[Pep+2H]^{2+}$ ion. Sometimes, the drift times for doubly-sodiated species are observed to be smaller than their counterparts, as is the case for FLANVSTVLTSK.

Summary of cross sections obtained from analysis of a single digest. A sense of the distribution of species that are characterized in the cross section database can be obtained by examining Table 2 which lists all cross sections for doubly-charged peptide ions obtained from the single sheep hemoglobin digest. We find 29 peptide ions observed across all doubly-charged configurations. Of these peptide sequences, 22 contain a C-terminal lysine and 7 contain a C-terminal arginine with 6 sequences containing missed cleavages. Here, cross sections for the major features are reported for 26 $[Pep+2H]^{2+}$, 45 $[Pep+M+H]^{2+}$, and 49 $[Pep+2M]^{2+}$ ions. Of these 94 metal-containing peptides, 74 are observed to increase in cross section (35 $[Pep+M+H]^{2+}$ and 39 $[Pep+2M]^{2+}$) as compared to the $[Pep+2H]^{2+}$ ions. No relative decreases in cross section are noted for the $[Pep+2Cs]^{2+}$ ions.

Summary of the complete database. Complete lists of cross sections for all singly-, doubly-, and triply-charged peptide ions identified across all charge configurations are located in Supplemental Tables 1–3, respectively. These supplemental tables provide information on number of residues, peptide sequence, protein source, and molecular mass information for all peaks identified for each observed charge configuration. Cross sections are reported for 78 singly- (containing 1–9 residues), 405 doubly- (containing 2–17 residues), and 140 triply-charged (containing 4–22 residues) peptide sequences. The majority of these tryptic peptides contain a C-terminal lysine residue (59 singly-, 269 doubly, and 94 triply-charged) followed by those peptides containing a C-terminal arginine residue (15 singly-, 125 doubly-, and 43 triply-charged). Of the peptide ions observed, sequences containing internal lysine or arginine residues are fairly common, comprising 34.8% of all assigned sequences (5 singly-, 124 doubly-, and 86 triply-charged). Here, we report a total of 1772 cross section entries (147 singly-, 1325 doubly-, and 300 triply-charged) for protonated and alkalated tryptic peptide ions. Of these cross sections, 818 are charged only by protons (67 $[Pep+H]^+$, 500 $[Pep+2H]^{2+}$, and 251 $[Pep+3H]^{3+}$). The remaining reported measurements incorporate alkali cations, with 262 containing Li^+ (28 $[Pep+Li]^+$, 145 $[Pep+Li+H]^{2+}$, 84 $[Pep+2Li]^{2+}$, and 5 $[Pep+Li+2H]^{3+}$), 316 containing Na^+ (20 $[Pep+Na]^+$, 158 $[Pep+Na+H]^{2+}$, 110 $[Pep+2Na]^{2+}$, and 28 $[Pep+Na+2H]^{3+}$), 222 containing K^+ (8 $[Pep+K]^+$, 126 $[Pep+K+H]^{2+}$, 75 $[Pep+2K]^{2+}$, and 13 $[Pep+K+2H]^{3+}$), and 154 containing Cs^+ (24 $[Pep+Cs]^+$, 80 $[Pep+Cs+H]^{2+}$, 47 $[Pep+2Cs]^{2+}$, and 3 $[Pep+Cs+2H]^{3+}$).

Assessing the accuracy of cross sections reported here. Although our solution conditions were designed to favor alkalated peptide ions (e.g., metal acetate solutions with no added acid), protonated peptide ions are still observed and are reported within our database. Therefore, it is possible to compare many cross section values for specific protonated peptide ions with measurements reported previously [50]. Fig. 3 shows a plot of all singly- and doubly-protonated measurements contained within both datasets. A comparison of cross sections for these peptides is observed to fall along a line for the entire range of measurements. Within these measurements, the relative change in cross section for 53 singly-protonated peptide is calculated to be $0.8 \pm 0.6\%$. Similar agreement is observed for the major features of 89 doubly-protonated peptide ions reported within both datasets.

While most values measured here are in excellent agreement with prior measurements, there are some differences. One difference arises because of variations in relative abundance. Here, some cross sections for minor features agree with prior results. One possible explanation for the differences in major features is that the previous measurements [50] were recorded using an injected-ion drift tube. In that work, the ion injection process causes a rapid heating/cooling cycle which may alter conformer populations relative to the source distributions. Outliers in our reported measurements are generally associated with relatively minor features. It is also clear that an improvement in the experimental resolving power is responsible for some of the observed differences. Several peptide ions appear to have been previously observed as the average of multiple features that are resolved in this work. Overall there is excellent agreement between the cross sections reported here and our previous work [50].

Understanding the range of cross sections for species with similar masses. Comparisons of cross sections for species with similar masses provides some insight about the range of different ion structures that are present. Fig. 4 shows cross sections for the singly-charged species (all alkali cations) and doubly-charged peptides (for only the Cs^+-containing ions) as a function of molecular weight. Note that all doubly-charged values are reported in the database but are not included here because the plots become congested and are less instructive. For all charge configurations, a strong correlation of increasing cross section with mass is observed. At a given mass, comparisons with calculated values for model structures shows that the smallest values have tightly-packed globular geometries while the largest values have relatively open, elongated structure. For small peptides ($m \sim 1000$), cross sections vary by $\sim\pm5\%$. As the peptide size increases so does the range of possible structures. At $m = \sim 1200$, ions display $\sim\pm6\%$ range of cross sections. This increases to $\sim\pm8\%$ for $m = \sim 1400$.

Variations in cross section with cation size. Another trend is that cross sections increase with increasing cation size, as is observed in Fig. 4. Literature values for the effective radii of the alkali cations are $r(Li^+) = 0.76$ Å, $r(Na^+) = 1.02$ Å, $r(K^+) = 1.38$ Å, and $r(Cs^+) = 1.67$ Å [69]. Compared with singly-charged $[Pep+H]^+$ species, singly-charged metal-containing species are larger by an average of 4.1 ± 5.5 Å^2 (Li^+), 4.9 ± 4.4 Å^2 (Na^+), 5.3 ± 4.6 Å^2 (K^+), and 9.6 ± 7.0 Å^2 (Cs^+), corresponding to an $4.3 \pm 6.7\%$, $4.6 \pm 6.3\%$, $4.6 \pm 3.8\%$, and $7.4 \pm 4.7\%$ change, respectively. These changes are much greater than the cation sizes, requiring that addition of a metal lead to a substantial change in structures of many of the amino acid chains.

Fig. 4 also shows significant increases in cross section for doubly-charged $[Pep+Cs+H]^{2+}$ and $[Pep+2Cs]^{2+}$ peptide ion species. Compared to the doubly-protonated $[Pep+2H]^{2+}$ species, the $[Pep+M+H]^{2+}$ ions are larger by an average of 1.5 ± 7.8 Å^2 (Li^+), 3.1 ± 8.8 Å^2 (Na^+), 5.1 ± 8.2 Å^2 (K^+), and 7.6 ± 8.7 Å^2 (Cs^+), corresponding to a change of $0.6 \pm 2.9\%$, $1.3 \pm 3.3\%$, $1.9 \pm 3.1\%$, and $3.0 \pm 3.8\%$, respectively. For the $[Pep+2M]^{2+}$ species, the observed increases are 2.0 ± 8.1 Å^2 (Li^+), 2.7 ± 7.3 Å^2 (Na^+), 5.8 ± 8.2 Å^2 (K^+), and 15.0 ± 9.1 Å^2 (Cs^+), corresponding to a change of $0.9 \pm 3.5\%$, $1.2 \pm 3.0\%$, $2.4 \pm 3.4\%$, and $6.5 \pm 5.4\%$, respectively. Compared to the doubly-protonated $[Pep+2H]^{2+}$ species, larger cross sections are observed upon incorporation of alkali cations for 65.6% (Li^+), 65.0% (Na^+), 72.5% (K^+), and 83.8% (Cs^+) for the $[Pep+M+H]^{2+}$ series and 61.9% (Li^+), 65.7% (Na^+), 81.1% (K^+), and 100% (Cs^+) for the

Table 2
Cross sections for doubly-charged peptides from the tryptic digest of *Ovis aries* (sheep).[a]

Number of residues	Assigned sequence[b]	Peptide mass[c]	Cross Section (Å^2)[d]								
			$[Pep+2H]^{2+}$	$[Pep+Li+H]^{2+}$	$[Pep+Na+H]^{2+}$	$[Pep+K+H]^{2+}$	$[Pep+Cs+H]^{2+}$	$[Pep+2Li]^{2+}$	$[Pep+2Na]^{2+}$	$[Pep+2K]^{2+}$	$[Pep+2Cs]^{2+}$
2	LR	287.20	118.2 (0.6, 4)								
2	VK	245.17	112.3								
2	YR	337.18	126.6 (0.3, 2)								
4	AHGK	411.22	138.8 (1.0, 2)		138.1						
4	SNVK	446.25									
5	AAWGK	531.28	156.6 (0.4, 3)								
5	AHGKK	539.32	162.8 (0.9, 4)								
5	GHGEK	526.25	160.2					169.4			
6	HLDDLK	739.39									
6	VKAHGK	638.39	177.1 (0.8, 4)								
7	MLTAEEK	820.40	221.6								
7	VAAALTK	672.42	186.2 (0.5, 2)	192.8				192.8			
7	VDPVNFK	817.43	201.0 (0.7, 7)		208.9			219.0	207.4 (0.5, 3)		
7	VLSAADK	702.39	186.2 (0.5, 2)					201.2	203.5		
9	AAVTGFWGK	935.49	225.2	239.6	224.5	235.5	238.8	237.7	230.9	240.0	249.1
9	LHVDPENFR	1125.56	257.5 (1.1, 3)		256.6 (0.6, 2)				261.6	261.4 (0.6, 2)	266.8
9	LRVDPVNFK	1086.62	263.9 (0.6, 7)	254.7 (2.3, 3)	255.3 (1.8, 6)	261.8 (0.7, 4)	263.5 (1.0, 3)	256.5	261.9	258.7 (0.6, 2)	270.5
9	MFLSFPTTK	1070.55	254.3 (0.6, 3)	257.9 (0.6, 2)	256.5 (1.0, 3)	258.6 (1.5, 3)	259.8 (0.6, 2)	248.0	259.2 (1.2, 2)	263.3 (0.6, 2)	266.8
10	LLVVYPWTQR	1273.72	291.7 (0.8, 5)	293.9	294.6 (1.1, 4)	297.0 (1.5, 3)	302.2		280.2 (0.0, 2)	285.2 (0.7, 2)	309.7
10	VLDSFSNGMK	1096.52	250.0	259.3					261.0	268.3	264.9
11	KVLDSFSNGMK	1224.62	283.8	264.9	265.6	273.8	280.8		270.1	271.1	
11	VDEVGAEALGR	1114.56			259.2						
11	VLSAADKSNVK	1130.63	258.2	261.1	264.7	270.1	278.9	272.4	266.5	269.2	278.9
12	FLANVSTVLTSK	1278.72	289.7 (1.0, 5)	297.1 (0.7, 2)	293.9 (0.6, 4)	299.9 (2.1, 2)	303.6 (0.7, 2)		285.4 (1.0, 2)	300.8 (0.8, 2)	318.1
12	VVAGVANALAHK	1148.67	274.6 (1.0, 5)	281.3 (0.7, 2)	285.5 (1.2, 5)	277.5 (1.0, 3)	281.7 (0.0, 2)	271.4	285.9 (1.6, 2)	288.0 (0.7, 2)	296.6
13	VKVDEVGAEALGR	1341.73	277.4	294.8		300.3	303.2	289.2	288.4	303.0	307.8
15	VGGNAGAYGAEALER	1433.69	286.6	294.8	291.1	294.8		289.2	296.6	301.2	303.2
16	HHGNEFTPVLQADFQK	1866.90	353.4							358.7	
16	TYFPHFDLSHGSAQVK	1832.88	336.6 (1.2, 3)		352.3	350.1 (0.5, 2)	354.5				

[a] All values were obtained using a coupled ion-mobility time-of-flight home-built instrument. See text and references for full description of the instrument.
[b] Peptide sequence correspond to fragments expected from tryptic digests as obtained from the Peptide Mass tool at http://expasy.org/tools/peptide-mass.html.
[c] Molecular weight is reported as the monoisotopic mass of the peptide without the ionizing species.
[d] Cross sections correspond to the average values of all datasets. Uncertainties are given parenthetically along with the total number of measurements made per charge state and correspond to one standard deviation about the mean when three or more measurements are reported, or as a range when only two measurements are reported.

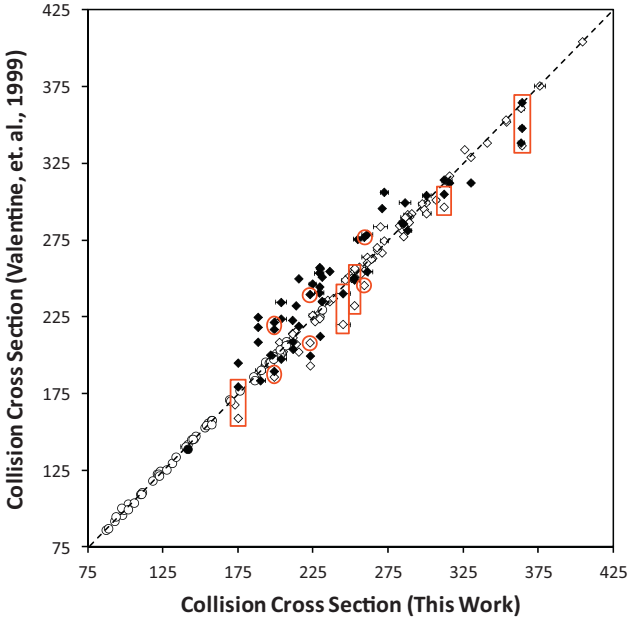

Fig. 3. Comparison of intersecting cross section measurements between the works presented here (*x*-axis) and a previously reported database [50] by Valentine et al. Singly-protonated peptide ions are represented as circles and doubly-protonated peptide ions are represented as diamonds. Closed datapoints are representative of our reported minor features. Vertical and horizontal errors correspond to standard deviations reported within the respective databases. Discrepancies in correlation resulting from improvements in resolving power (red circles) or source conditions (red rectangles) are highlighted. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

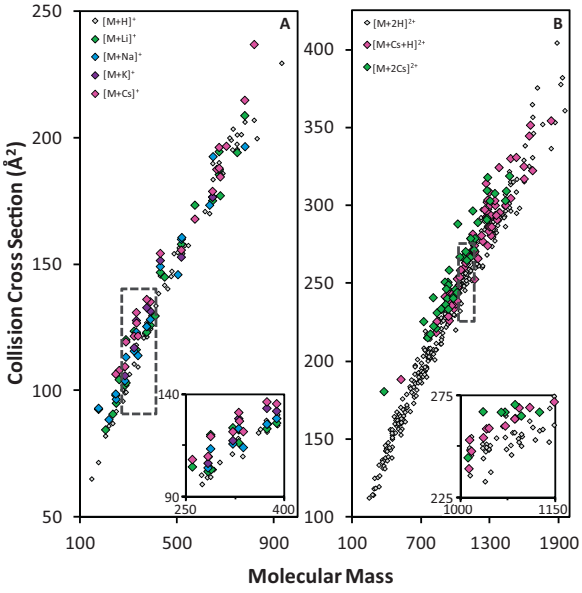

Fig. 4. Cross sections of all recorded tryptic peptides as a function of the parent ion molecular weight for the observed alkali metal cation species. Panel A displays measurements for $[Pep+H]^+$ ions (gray diamonds) compared with $[Pep+M]^+$ ions (green, blue, purple, and pink diamonds for $M^+ = Li^+$, Na^+, K^+, or Cs^+, respectively). Panel B displays measurements for $[Pep+2H]^{2+}$ ions (gray diamonds) compared with $[Pep+Cs+H]^{2+}$ (pink diamonds) and $[Pep+2Cs]^{2+}$ (green diamonds) ions. The insets (defined as the dashed line box) display the variations in cross section within the cationic charged species for a given molecular mass. Observed minor features are not included here. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

$[Pep+2M]^{2+}$ series. Here we note that we've presented results for peptides of sequence length 2–16. For longer peptides, it is possible that the metal interaction will have a lesser effect on the overall peptide structure. For example, elements of secondary structure may be preserved or stabilized upon metal binding [13,68]. Finally, we observe more lower molecular weight species with decreasing cationic size. Presumably, longer sequences are required to solvate alkali cation(s) having larger radii.

Normalization to generate reduced cross sections that are independent of mass. To account for differences in cross section that arise from differences in mass, it is useful to normalize the system such that differences in cross section are associated with amino acid composition. This is accomplished by dividing each cross section by the average cross section of a peptide of the same mass (determined from a second-order polynomial fit to all cross sections as a function of mass). The result is a reduced cross section. As described previously [49], this removes variations in cross section that occurs with size. Reduced cross sections here typically range from ~0.90 to ~1.10. Note that this normalization to reduced cross section differs from our original definition, which normalizes to a fit of polyalanine cross sections [49].

Derivation of intrinsic size parameters for individual amino acids. One of the advantages of assembling such a large dataset is that it is possible to extract average properties associated with peptide structure. It is interesting to consider the contributions of individual amino acids to the total cross section. If we assume that structures of tryptic peptides are similar we can derive the intrinsic contributions to size of individual residues. The ISPs are especially interesting because they appear to depend on which metal ion is attached. Thus, they provide some insight into which amino acids interact strongly with each metal.

Values of ISPs are derived by solving a set of equations relating the frequency of occurrence of each residue to the reduced cross section of each peptide represented in the dataset, as defined by Eq. (2):

$$\sum_{j=1}^{n} X_{ij} p_j = y_i \tag{2}$$

here, i corresponds to a unique peptide sequence in the parameterization set and is equal to 1 through m, where m is the total number of peptides ions within the parameterization set. Likewise, j corresponds to each unknown size parameter and is equal to 1 through n, where n is the total number of amino acid residues contained within the parameterization subsets (i.e., all 20 natural amino acids, where Cys is modified with a carboxyamidomethyl protecting group). The variable X corresponds to the frequency of occurrence of each amino acid (j) in each sequence (i) and p_j corresponds to the unknown intrinsic amino acid size parameter of the jth residue. Finally, all of this is equal to y_i, which is the reduced cross section of each sequence (i). In the present analysis, each equation provides the relationship of the product of residue frequencies and size parameter with each reduced cross section. This system of equations is solved for the average, best-fit ISP values (p_j) using a least-squares regression model [70]. Uncertainties for each of the derived parameters are determined from the square root of the variance for each parameter and represent one standard deviation.

Subsets of cross sections used for ISP calculations were limited to measurements of pentapeptides and longer sequences containing a C-terminal lysine or arginine. Only the most intense feature in the drift time dimension is used for these derivations. Table 3 shows a summary for all ISPs calculated from all doubly-charged peptide ions. These values are similar to values derived for singly-protonated peptide ions [49] and for doubly-protonated peptide ions [48,50]. In general, all of these datasets show an increase to cross section for nonpolar aliphatic amino acids

Table 3
Intrinsic size parameters for individual amino acid residues.[a]

Residue	Parameterization set								
	$[Pep+2H]^{2+}$ [b]	$[Pep+Li+H]^{2+}$ [c]	$[Pep+Na+H]^{2+}$ [d]	$[Pep+K+H]^{2+}$ [e]	$[Pep+Cs+H]^{2+}$ [f]	$[Pep+2Li]^{2+}$ [g]	$[Pep+2Na]^{2+}$ [h]	$[Pep+2K]^{2+}$ [i]	$[Pep+2Cs]^{2+}$ [j]
Ala	0.98 (0.01)	1.00 (0.03)	1.03 (0.03)	1.00 (0.03)	1.07 (0.04)	0.97 (0.04)	1.02 (0.03)	1.07 (0.03)	1.01 (0.08)
Val	1.05 (0.02)	1.07 (0.03)	1.10 (0.04)	1.18 (0.03)	1.08 (0.04)	1.07 (0.05)	1.06 (0.04)	1.06 (0.04)	1.10 (0.07)
Ile	1.08 (0.02)	1.11 (0.05)	1.10 (0.05)	1.22 (0.04)	1.11 (0.06)	1.16 (0.07)	1.05 (0.06)	1.14 (0.06)	1.24 (0.13)
Leu	1.13 (0.02)	1.17 (0.03)	1.20 (0.03)	1.18 (0.03)	1.14 (0.03)	1.13 (0.05)	1.16 (0.03)	1.16 (0.03)	1.02 (0.10)
Met	1.17 (0.04)	0.95 (0.09)	0.96 (0.08)	1.12 (0.08)	1.02 (0.10)	0.84 (0.12)	1.07 (0.10)	1.06 (0.08)	0.81 (0.17)
Phe	1.01 (0.02)	1.09 (0.05)	1.11 (0.04)	1.07 (0.04)	1.09 (0.05)	1.06 (0.07)	1.08 (0.05)	1.10 (0.05)	1.16 (0.13)
Tyr	0.98 (0.03)	0.94 (0.05)	1.10 (0.07)	1.12 (0.07)	0.99 (0.07)	0.94 (0.06)	0.97 (0.06)	0.98 (0.05)	0.85 (0.11)
Trp	0.95 (0.04)	1.08 (0.09)	1.00 (0.07)	1.03 (0.07)	0.93 (0.09)	0.93 (0.10)	0.68 (0.14)	0.85 (0.10)	1.10 (0.23)
Asp	0.91 (0.03)	1.02 (0.06)	0.91 (0.06)	1.01 (0.06)	0.87 (0.08)	0.77 (0.09)	0.86 (0.06)	0.91 (0.07)	0.95 (0.11)
Glu	0.94 (0.02)	0.93 (0.04)	0.88 (0.04)	0.93 (0.03)	0.91 (0.05)	0.84 (0.07)	0.89 (0.04)	0.90 (0.04)	0.88 (0.10)
Asn	0.92 (0.03)	0.91 (0.06)	0.90 (0.05)	0.87 (0.05)	0.91 (0.07)	0.90 (0.07)	0.90 (0.06)	0.89 (0.06)	1.04 (0.10)
Gln	0.96 (0.03)	0.90 (0.06)	0.91 (0.05)	0.88 (0.05)	0.89 (0.06)	0.93 (0.11)	0.90 (0.06)	0.91 (0.06)	0.87 (0.17)
Ser	0.96 (0.02)	0.91 (0.05)	0.90 (0.05)	0.98 (0.04)	0.99 (0.06)	0.97 (0.07)	0.95 (0.05)	1.02 (0.05)	0.84 (0.12)
Thr	0.93 (0.02)	1.02 (0.04)	0.93 (0.04)	0.94 (0.04)	1.02 (0.05)	0.96 (0.06)	0.99 (0.04)	1.01 (0.04)	1.03 (0.09)
Pro	0.99 (0.02)	0.91 (0.05)	0.87 (0.05)	0.92 (0.05)	0.85 (0.06)	1.00 (0.05)	1.00 (0.05)	0.92 (0.05)	0.91 (0.12)
Gly	0.93 (0.02)	0.91 (0.03)	0.97 (0.03)	0.93 (0.03)	0.99 (0.04)	1.02 (0.05)	1.00 (0.04)	0.94 (0.04)	0.93 (0.07)
His	1.07 (0.03)	0.99 (0.06)	1.03 (0.06)	0.96 (0.05)	0.79 (0.07)	0.96 (0.11)	0.97 (0.09)	0.98 (0.07)	0.86 (0.17)
Cys*	0.95 (0.03)	0.73 (0.07)	0.94 (0.08)	0.87 (0.07)	0.85 (0.08)	0.83 (0.10)	0.92 (0.08)	0.70 (0.14)	1.14 (0.22)
Lys	0.98 (0.02)	0.95 (0.04)	0.97 (0.05)	0.84 (0.05)	0.97 (0.06)	1.08 (0.07)	0.93 (0.06)	0.89 (0.06)	1.04 (0.13)
Arg	1.06 (0.03)	0.99 (0.06)	0.82 (0.08)	0.81 (0.07)	0.96 (0.09)	1.13 (0.09)	1.03 (0.07)	0.88 (0.09)	1.01 (0.19)

[a] The amino acid size parameter were derived by solving a system of equations that relate the occurrence frequency of each amino acid and the unknown size parameters to a reduced cross section of each peptide (see text). Uncertainties, given in parenthesis, correspond to one standard deviation about the mean and generally increase for lower frequency residues (such as methionine, tryptophan, and histidine). Additional sources of uncertainty could also be the result of structural variation within the datasets, adding variance in the average contribution of each residue to the cross sections (supported by the observation of multiple conformer populations within individual peptide ion assignments reported within Supplemental Tables 1–3). One final source of uncertainty could be attributed to a very limited number of false assignments of the peptides during the data analysis. Elevated uncertainties reported for the $[Pep+2M]^{2+}$ species presumably result from decreased representation of these peptide ion measurements within the database. ISP values are not reported for the $[Pep+M]^{+}$ and $[Pep+M+2H]^{3+}$ species due to lacking numbers of measurements.

[b] Amino acid size parameters derived from 284 $[Xxxn\ Lys/Arg +2H]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–16$.

[c] Amino acid size parameters derived from 95 $[Xxxn\ Lys/Arg +Li+H]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–14$.

[d] Amino acid size parameters derived from 105 $[Xxx_nLys/Arg +Na+H]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–15$.

[e] Amino acid size parameters derived from 92 $[Xxxn\ Lys/Arg +K+H]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–15$.

[f] Amino acid size parameters derived from 70 $[Xxx_nLys/Arg +Cs+H]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–15$.

[g] Amino acid size parameters derived from 55 $[Xxx_nLys/Arg +2Li]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4$ to 14.

[h] Amino acid size parameters derived from 82 $[Xxx_nLys/Arg +2Na]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–16$.

[i] Amino acid size parameters derived from 62 $[Xxx_nLys/Arg +2K]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–15$.

[j] Amino acid size parameters derived from 42 $[Xxx_nLys/Arg +2Cs]^{2+}$ peptides. Here, Xxx is any naturally occuring amino acid (exception of carboxyamidomethylated Cys residues) with $n = 4–14$.

and a contraction of cross section for the polar aliphatic amino acids. Intermediate effects are observed for aromatic amino acid residues.

Application of ISP values for predicting cross sections. Once ISP values for all amino acid residues have been determined, we can use these values to estimate cross sections for sequences. Fig. 5 shows the ratio of calculated to experimental cross sections for all $[Pep+2H]^{2+}$ and $[Pep+Na+H]^{2+}$ sequences (with five or more residues). We note that because the ISP values are derived from the experimental cross sections the calculated values are strictly a retrodiction. A bonafide prediction of cross sections for unmeasured sequences could also be made, and is one of the benefits of the database.

One way to assess these predictions is to compare these fits to a simple molecular weight fit, as has been described previously [48–50]. The average $[Pep+Na+H]^{2+}$ cross sections from the molecular weight fit show that 41 of 105 calculated values (39%) fall within ±2% of the measured cross sections. On the other hand, 71 of 105 (68%) are within ±2% of the experimental values for the ISP retrodictions yielding a predictive improvement factor of ∼1.7. Predictive improvement factors of ∼1.4, ∼1.2, ∼1.5, and ∼1.4 are found for the doubly-protonated, Li^+-, K^+-, and Cs^+-containing datasets.

Insight into the interactions of alkali cations with specific amino acids. Fig. 6 shows the influence of alkali metal cations on specific amino acid ISPs as a relative difference plot (i.e., $ISP_{[Pep+M+H]}^{2+} - ISP_{[Pep+2H]}^{2+}/ISP_{[Pep+2H]}^{2+}$). It is interesting that families of amino acids appear to interact differently with the cations compared with the protonated species. Here, polar aliphatic residues, Met, Cys* and His are observed to contribute much less to cross sections when alkali cations are present (by values up to ∼9%, ∼20%, ∼24%, and ∼27%, respectively). We interpret this as evidence for tight metal–ligand binding interactions. On the other hand, nonpolar residues (Ala, Val, Ile, Leu) and those containing aromatic side chains (Phe, Tyr, Trp) have larger ISP contributions (up to ∼10% and ∼15%) suggesting that these residues do not interact intimately with the cations. It is also possible that the sequestering of the metal cation by specific peptide regions may force more distal nonpolar and aromatic side chains outward to the helium environment. The decrease in contribution to the cross section observed for proline could be explained by the constriction on the peptide backbone by the residue side chain, allowing other residue side chains and amide backbone carbonyls to coordinate as ligands to the metal cation.

It is interesting that some amino acids appear to interact favorably with different cations of different size. For example, there

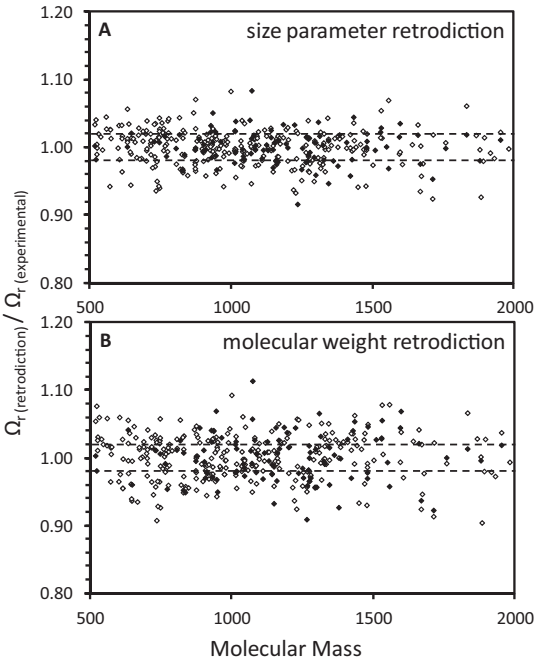

Fig. 5. Prediction accuracy scatter plots for reduced cross sections of 284 $[(Xxx)_nLys/Arg + 2H]^{2+}$ (open diamonds) and 105 $[(Xxx)_nLys/Arg + Na+H]^{2+}$ (solid diamonds) peptide ions. Here, Xxx is any naturally occurring amino acid (except for cysteine residues, which are carboxyaminomethylated) and $n = 4–16$. ISP retrodictions are displayed in panel A and those obtained from a polynomial molecular weight fit to cross section data are displayed in panel B. Dotted lines indicate where retrodicted values agree to within 2% of the experimental measurements.

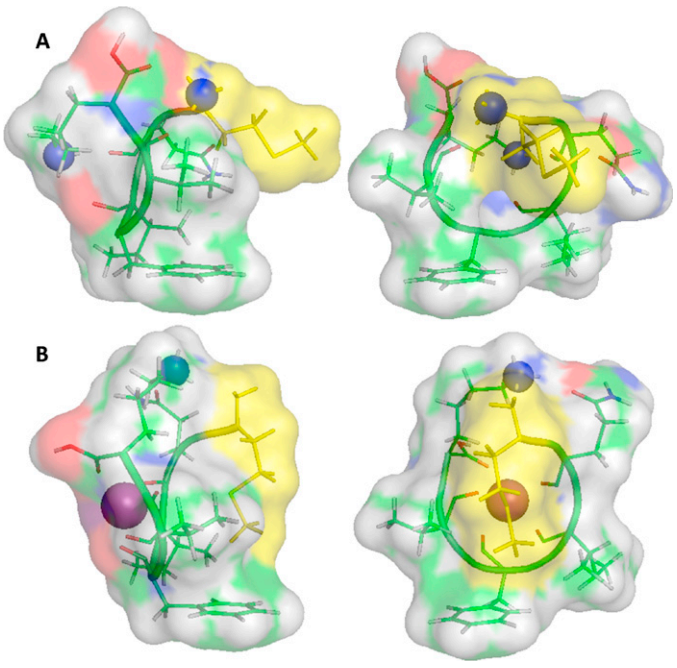

Fig. 7. Representative structures for low-energy conformations of $[MQIFVK+2H]^{2+}$ (panel A) and $[MQIFVK+Na+H]^{2+}$ (panel B). Two separate orientations (rotated 90° along the *y*-axis) are depicted for comparison of surface exposure of the amino acid side chains (particularly methionine, colored yellow) resulting from charge solvation effects. The peptide backbone is rendered as a cartoon with charge-solvating amide carbonyls exhibited. The accessible surface area is transparently overlaid on each structure. Green, red, white, blue, and purple coloration corresponds to surface exposure of carbon, oxygen, hydrogen, nitrogen, and sodium atoms, respectively. Charge sites have been pictured as spheres. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

exists a smaller ISP contribution for interactions of Cs^+ with His and Pro residues than for the much smaller Li^+ cation. At first glance we were surprised that such detailed interactions between individual metals and residues could be extracted from this database. However, we note that because independent datasets are used for each parameterization, there is high confidence in these observed contributions to cross section by the individual amino acid residues and changes imparted by each metal ion.

Implications for use of cross sections in determining structural homology between protonated and alkalated peptides. The differences in ISP values for cationized sequences, compared with protonated peptides, raise an interesting twist in using cross sections to assess structural similarities. One powerful use of cross section values is to suggest structural homology. For example if protonated and sodiated ions of the same sequence have similar cross sections, one might suggest that these ions have similar structures. The ISP values for cationized and protonated ions indicate that this may not be the case. As an example, consider the MQIFVK sequence having $\Omega([Pep+2H]^{2+}) = 208.2$ Å^2 and $\Omega([Pep+Na+H]^{2+}) = 207.2$ Å^2. One might assume that the peptide favors the same structure upon substitution of Na^+. However, our impression from the significant size parameter differences for amino acids that arise upon substitution of Na^+ suggests a different interpretation. The ISPs suggest that the polar aliphatic residue (Gln) as well as Met should decrease in contribution to the cross section (relative to the doubly-protonated species), whereas the nonpolar aliphatic residues (Ile and Val) and the aromatic residue (Phe) will increase in contribution to the cross section. The net effect is that the peptide cross section does not significantly vary because of a cancellation of these opposing effects.

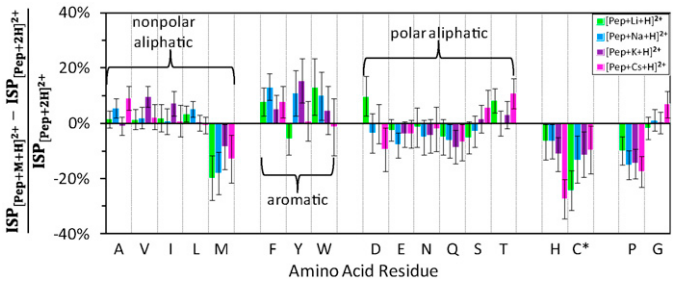

Fig. 6. Relative difference plot in amino acid residue ISP for the alkali $[Pep+M+H]^{2+}$ ion series compared to the ISPs obtained from $[Pep+2H]^{2+}$ ions. Green, blue, purple, and pink bars correspond to M^+=Li^+, Na^+, K^+, or Cs^+, respectively. Uncertainties in the calculation (displayed as bars) represent one standard deviation about the mean and are propagated as the square root of the sum of the products of the squared errors and their partial derivatives. Relative differences for the $[Pep+M+H]^{2+}$ series are calculated by maintaining ISP values for lysine and arginine equivalent to those values derived within the $[Pep+2H]^{2+}$ parameterization set (0.98 and 1.06, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

To illustrate this idea further we have carried out some preliminary molecular modeling studies of these ions. These simulations were performed on the MQIFVK peptide (observed in the tryptic digest of ubiquitin) using the Insight II software package (Accelrys Inc., San Diego, CA) with the consistent valence forcefield (CVFF) using the Discover 3 program.

Trajectory model cross sections calculated [45] for the low-energy structures generated by simulated annealing are shown in Fig. 7. The $[MQIFVK+2H]^{2+}$ ion conformer is relatively compact with the calculated cross section for the displayed structure (208.2 Å^2) in excellent agreement with the experimental value (208.2 Å^2). In this case, the protonated Lys is solvated by the amide backbone carbonyls of Ile and Gln and the protonated N-terminus interacts with the C-terminus carbonyl and backbone carbonyl associated with Val. The resulting structure is a compact globule that is capped on one face by the protonated Lys side chain with the opposing

face containing the protruding side chains of Gln, Ile, Phe, Val, and Met. The $[MQIFVK+Na+H]^{2+}$ ion conformer (the lowest energy conformation that we found in these studies) has a calculated cross section of 208.4 Å^2, which is also in agreement with experiment (207.2 Å^2). However, in contrast to the protonated peptide, the sodium-containing species has a significantly different structure. Here, Na^+ caps one face of the peptide structure such that it is solvated by four amide backbone carbonyls as well as the C-terminal carbonyl. The protonated Lys side chain is primarily solvated by the Gln side-chain carbonyl as well as the N-terminus. These solvation effects result in a globule with extension of the non-polar (Ile and Val) and the aromatic (Phe) residue side chains toward the molecular surface. Additionally, the Met side chain is observed to collapse on the molecular face opposing Na^+, allowing the two lone electron pairs of the sulfur atom to participate in long-range, ion–dipole interactions with Na^+.

4. Summary and conclusions

This work presents a database of 1772 collision cross sections of protonated and alkalated peptide ions, including: 818 protonated (67 $[Pep+H]^+$, 500 $[Pep+2H]^{2+}$, and 251 $[Pep+3H]^{3+}$), 262 containing Li^+ (28 $[Pep+Li]^+$, 145 $[Pep+Li+H]^{2+}$, 84 $[Pep+2Li]^{2+}$, and 5 $[Pep+Li+2H]^{3+}$), 316 containing Na^+ (20 $[Pep+Na]^+$, 158 $[Pep+Na+H]^{2+}$, 110 $[Pep+2Na]^{2+}$, and 28 $[Pep+Na+2H]^{3+}$), 222 containing K^+ (8 $[Pep+K]^+$, 126 $[Pep+K+H]^{2+}$, 75 $[Pep+2K]^{2+}$, and 13 $[Pep+K+2H]^{3+}$), and 154 containing Cs^+ (24 $[Pep+Cs]^+$, 80 $[Pep+Cs+H]^{2+}$, 47 $[Pep+2Cs]^{2+}$, and 3 $[Pep+Cs+2H]^{3+}$). These cross sections were obtained from the tryptic digestion of 24 proteins with the accuracy assessed by comparisons to previous measurements [50]. Of the sequences reported, 68% terminate with a lysine residue and 35% contain missed cleavages. Such a large number of values enable generalizations on the assessment of peptide ion conformation resulting from solvation effects of the alkali cation(s). For singly-charged (i.e. $[Pep+M]^+$) and doubly-charged (i.e. $[Pep+M+H]^{2+}$ or $[Pep+2M]^{2+}$) peptide ions, the cross section is generally observed to increase with increasing cationic size. Multiple conformers (i.e. minor features) are often observed among the doubly- and triply-charged species, particularly with protonated, Li^+- or Na^+-containing species.

We find that amino acid composition plays a significant role in the observed peptide ion cross sections. Contributions to cross section by individual amino acid residues were assessed for each alkali doubly-charged configuration with the derivation of ISPs. The $[Pep+2H]^{2+}$ ISPs display similar charge solvation effects to previously derived protonated ISPs, [48,49] as individual contributions to cross section are observed to be decreased for polar aliphatic residues and increased for nonpolar aliphatic and aromatic residues. An amplification of these solvation effects are displayed with comparisons of the $[Pep+M+H]^{2+}$ ISPs with the $[Pep+2H]^{2+}$ ISPs with specific amino acid residues interacting favorably with different alkali cations of differing ionic size.

The ISP-derived solvation effects are supported with low-energy structures obtained from molecular dynamics simulations of the MQIFVK peptide in the $[Pep+2H]^{2+}$ and $[Pep+Na+H]^{2+}$ charge configurations. While the empirical measurements for these ions display little change in cross section, structural differences upon substitution of a Na^+ with a H^+ are illustrated with increased interactions of Gln and Met (as well as the peptide backbone) with the charges. Structural dissimilarity is prominently displayed with the collapsing of Met on the molecular face opposing Na^+ within the depicted low-energy structure of $[MQIFVK+Na+H]^{2+}$, in contrast with the significant surface protrusion of Met within the depicted low-energy structure of $[MQIFVK+2H]^{2+}$.

Acknowledgements

JMD gratefully acknowledges the financial support received from the PhD Fellowship provided by Naval Surface Warfare Center, Crane Division. This work is supported in part by funding from the "Next Generation Threat" fund provided by Naval Surface Warfare Center, Crane Division (contract number N00164-08-C-JQ11) as well as funding from the METACyt Grant from the Lilly Endowment.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijms.2012.05.001.

References

[1] E.A. Permyakov, Metalloproteomics, Wiley, New Jersey, 2009.
[2] R.H. Holm, P. Kennepohl, E.I. Solomon, Structural and functional aspects of metal sites in biology, Chemical Reviews 96 (1996) 2239–2314.
[3] J.M. Page, E. Di Cera, Role of Na^+ and K^+ in enzyme function, Physiological Reviews 86 (2006) 1049–1092.
[4] H. Reyes-Caballero, G.C. Campanello, D.P. Giedroc, Metalloregulatory proteins: metal selectivity and allosteric switching, Biophysical Chemistry 156 (2011) 103–114.
[5] J.A. Loo, P. Hu, Interaction of angiotensin peptides and zinc metal ions probed by electrospray ionization mass spectrometry, Journal of the American Society for Mass Spectrometry 5 (1994) 959–965.
[6] O.V. Nemirovskiy, M.L. Gross, Determination of calcium binding sites in gas-phase small peptides by tandem mass spectrometry, Journal of the American Society for Mass Spectrometry 9 (1998) 1020–1028.
[7] L.M. Teesch, R.C. Orlando, J. Adams, Location of the alkali metal ion in gas-phase peptide complexes, Journal of the American Chemical Society 113 (1991) 3668–3675.
[8] P. Hu, J.A. Loo, Gas-phase coordination properties of Zn^{2+}, Cu^{2+}, Ni^{2+}, and Co^{2+} with histidine-containing peptides, Journal of the American Chemical Society 117 (1995) 11314–11319.
[9] P.B. Armentrout, M.T. Rodgers, J. Oomens, J.D. Steill, Infrared multiphoton dissociation spectroscopy of cationized serine: effects of alkali-metal cation size on gas-phase conformation, Journal of Physical Chemistry A 112 (2008) 2248–2257.
[10] M.T. Rodgers, P.B. Armentrout, J. Oomens, J.D. Steill, Infrared multiphoton dissociation spectroscopy of cationized threonine: effects of alkali-metal cation size on gas-phase conformation, Journal of Physical Chemistry A 112 (2008) 2258–2267.
[11] T. Wyttenbach, D. Liu, M.T. Bowers, Interactions of the hormone oxytocin with divalent metal ions, Journal of the American Chemical Society 130 (2008) 5993–6000.
[12] B.A. Cerda, C. Wesdemiotis, Zwitterionic *vs.* charge-solvated structures in the binding of arginine to alkali metal ions in the gas phase, Analyst 125 (2000) 657–660.
[13] C. Wu, J. Klasmeier, H.H. Hill Jr., Atmospheric pressure ion mobility spectrometry of protonated and sodiated peptides, Rapid Communications in Mass Spectrometry 13 (1999) 1138–1142.
[14] M. Kohtani, B.S. Kinnear, M.F. Jarrold, Metal-ion enhanced helicity in the gas phase, Journal of the American Chemical Society 122 (2000) 12377–12378.
[15] B.T. Ruotolo, D.H. Russell, Gas-phase conformations of proteolytically derived protein fragments: influence of solvent on peptide conformation, Journal of Physical Chemistry B 108 (2004) 15321–15331.
[16] B.T. Ruotolo, C.C. Tate, D.H. Russell, Ion mobility-mass spectrometry applied to cyclic peptide analysis: conformational preferences of gramicidin S and linear analogs in the gas phase, Journal of the American Society for Mass Spectrometry 15 (2004) 870–878.
[17] M. Kohtani, M.F. Jarrold, S. Wee, R.A.J. O'Hair, Metal ion interactions with polyalanine peptides, Journal of Physical Chemistry B 108 (2004) 6093–6097.
[18] M.K. Drayß, P.B. Armentrout, J. Oomens, M. Schaefer, IR spectroscopy of cationized aliphatic amino acids: stability of charge-solvated structure increases with metal cation size, International Journal of Mass Spectrometry 297 (2010) 18–27.
[19] Y. Berezovskaya, C.T. Armstrong, A.L. Boyle, M. Porrini, D.N. Woolfson, P.E. Barran, Metal binding to a zinc-finger peptide: a comparison between solution and the gas phase, Chemical Communications 47 (2011) 412–414.
[20] J.T.S. Hopper, N.J. Oldham, Alkali metal cation-induced destabilization of gas-phase protein–ligand complexes: consequences and prevention, Analytical Chemistry 83 (2011) 7472–7479.
[21] J. Pan, L. Konermann, Calcium-induced structural transitions of the calmodulin–melittin system studied by electrospray mass spectrometry: conformational subpopulations and metal-unsaturated intermediates, Biochemistry 49 (2010) 3477–3486.

[22] L. Deng, N. Sun, E.N. Kitova, J.S. Klassen, Direct quantification of protein-metal ion affinities by electrospray ionization mass spectrometry, Analytical Chemistry 82 (2010) 2170–2174.
[23] J.R. McLean, J.A. McLean, Z. Wu, C. Becker, L.M. Pérez, C.N. Pace, J.M. Scholtz, D.H. Russell, Factors that influence helical preferences for singly charged gas-phase peptide ions: the effects of multiple potential charge-carrying sites, Journal of Physical Chemistry B 114 (2010) 809–816.
[24] L. Chen, Y.Q. Gao, D.H. Russell, How alkali metal ion binding alters the conformation preferences of gramicidin A: a molecular dynamics and ion mobility study, Journal of Physical Chemistry A 116 (2012) 689–696.
[25] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science 246 (1989) 64–71.
[26] Z.L. Cheng, K.W.M. Siu, R. Guevremont, S.S. Berman, Electrospray mass spectrometry: a study on some aqueous solutions of metal salts, Journal of the American Society for Mass Spectrometry 3 (1992) 281–288.
[27] M. Satterfield, J.S. Brodbelt, Enhanced detection of flavonoids by metal complexation and electrospray ionization-mass spectrometry, Analytical Chemistry 72 (2000) 5898–5906.
[28] K.J. Kock, T. Aggerholm, S.C. Nanita, R.G. Cooks, Clustering of nucleobases with alkali metals studied by electrospray ionization tandem mass spectrometry: implication for mechanisms of multistrand DNA stabilization, Journal of Mass Spectrometry 37 (2002) 676–686.
[29] M. Vairamani, M.L. Gross, G-Quadruplex formation of thrombin-binding aptamer detected by electrospray ionization mass spectrometry, Journal of the American Chemical Society 125 (2003) 42–43.
[30] K. Fukushima, Iwahashi H, 1:1 complex of guanine quartet with alkali metal cations detected by electrospray ionization mass spectrometry, Chemical Communications 11 (2000) 895–896.
[31] C. Wang, S.H. Yang, J. Wang, P. Kroll, K.A. Schug, D.W. Armstrong, Study of complexation between cyclofructans and alkali metal cations by electrospray ionization mass spectrometry and density functional theory calculations, International Journal of Mass Spectrometry 291 (2010) 118–124.
[32] H.J. Sterling, J.D. Batchelor, D.E. Wemmer, E.R. Williams, Effects of buffer loading for electrospray ionization mass spectrometry of a noncovalent protein complex that requires high concentrations of essential salts, Journal of the American Society for Mass Spectrometry 21 (2010) 1045–1049.
[33] T.G. Flick, S.I. Merenbloom, E.R. Williams, Anion effects on sodium ion and acid molecule adduction to protein ions in electrospray ionization mass spectrometry, Journal of the American Society for Mass Spectrometry 22 (2011) 1968–1977.
[34] M. Rožman, S.J. Gaskell, Non-covalent interactions of alkali metal cations with singly charged tryptic peptides, Journal of Mass Spectrometry 45 (2011) 1409–1415.
[35] L.M. Teesch, J. Adams, Fragmentations of gas-phase complexes between alkali metal ions and peptides: metal ion binding to carbonyl oxygens and other neutral functional groups, Journal of the American Chemical Society 113 (1991) 812–820.
[36] W.Y. Feng, S. Gronert, K.A. Fletcher, A. Warres, C.B. Lebrilla, The mechanism of C-terminal fragmentations in alkali metal ion complexes of peptides, International Journal of Mass Spectrometry 222 (2003) 117–134.
[37] D. Bensadek, F. Monigatti, J.A.J. Steen, H. Steen, Why b, y's? Sodiation-induced tryptic peptide-like fragmentation of non-tryptic peptides, International Journal of Mass Spectrometry 268 (2007) 181–189.
[38] Y.M.E. Fung, H. Liu, T.-W.D. Chan, Electron capture dissociation of peptides metalated with alkaline-earth metal ions, Journal of the American Society for Mass Spectrometry 17 (2006) 757–771.
[39] A.J. Kleinnijenhuis, R. Mihalca, R.M.A. Heeren, A.J.R. Heck, Atypical behavior in the electron capture induced dissociation of biologically relevant transition metal ion complexes of the peptide hormone oxytocin, International Journal of Mass Spectrometry 253 (2006) 217–224.
[40] H. Liu, K. Håkansson, Electron capture dissociation of tyrosine O-sulfated peptides complexed with divalent metal cations, Analytical Chemistry 78 (2006) 7570–7576.
[41] H. Liu, K. Håkansson, Divalent metal ion–peptide interactions probed by electron capture dissociation of trications, Journal of the American Society for Mass Spectrometry 17 (2006) 1731–1741.
[42] G.E. Hofmeister, Z. Zhou, J.A. Leary, Linkage position determination in lithium-cationized disaccharides: tandem mass spectrometry and semiempirical calculations, Journal of the American Chemical Society 113 (1991) 5964–5970.
[43] M. Kohler, J.A. Leary, Gas phase reactions of doubly charged alkaline earth and transition metal(II)–ligand complexes generated by electrospray ionization, International Journal of Mass Spectrometry and Ion Processes 162 (1997) 17–34.
[44] H.M. Watson, J.B. Vincent, C.J. Cassady, Effects of transition metal ion coordination on the collision-induced dissociation of polyalanines, Journal of Mass Spectrometry 46 (2011) 1099–1107.
[45] M.F. Mesleh, J.M. Hunter, A.A. Shvartsburg, G.C. Schatz, M.F. Jarrold, Structural information from ion mobility measurements: effects of the long-range potential, Journal of Physical Chemistry 100 (1996) 16082–16086.
[46] D.E. Clemmer, R.R. Hudgins, M.F. Jarrold, Naked protein conformations: cytochrome *c* in the gas phase, Journal of the American Chemical Society 117 (1995) 10141–10142.
[47] G. von Helden, T. Wyttenbach, M.T. Bowers, Conformation of macromolecules in the gas phase: use of matrix-assisted laser desorption methods in ion chromatography, Science 267 (1995) 1483–1485.
[48] S.J. Valentine, M.A. Ewing, J.M. Dilger, M.S. Glover, S. Geromanos, C. Hughes, D.E. Clemmer, Using ion mobility data to improve peptide identification: intrinsic amino acid size parameters, Journal of Proteome Research 10 (2011) 2318–2329.
[49] S.J. Valentine, A.E. Counterman, C.S. Hoaglund-Hyzer, D.E. Clemmer, Intrinsic amino acid size parameters from a series of 113 lysine-terminated tryptic digest peptide ions, Journal of Physical Chemistry B 103 (1999) 1203–1207.
[50] S.J. Valentine, A.E. Counterman, D.E. Clemmer, A database of 660 peptide ion cross sections: use of intrinsic size parameters for bona fide predictions of cross sections, Journal of the American Society for Mass Spectrometry 10 (1999) 1188–1211.
[51] L. Tao, J.R. McLean, J.A. McLean, D.H. Russell, A collision cross-section database of singly-charged peptide ions, Journal of the American Society for Mass Spectrometry 18 (2007) 1232–1238.
[52] S.C. Henderson, J. Li, A.E. Counterman, D.E. Clemmer, Intrinsic size parameters for Val, Ile, Leu, Gln, Thr, Phe, and Trp residues from ion mobility measurements of polyamino acid ions, Journal of Physical Chemistry B 103 (1999) 8780–8785.
[53] A.E. Counterman, D.E. Clemmer, Volumes of individual amino acid residues in gas-phase peptide ions, Journal of the American Chemical Society 121 (1999) 4031–4039.
[54] M.T. Rodgers, P.B. Armentrout, A thermodynamic "vocabulary" for metal ion interactions in biological systems, Accounts of Chemical Research 37 (2004) 989–998.
[55] S.J. Ye, P.B. Armentrout, Absolute thermodynamic measurements of alkali metal cation interactions with a simple dipeptide and tripeptide, Journal of Physical Chemistry A 112 (2008) 3587–3596.
[56] P.B. Armentrout, Electronic state-specific transition metal ion chemistry, Annual Review of Physical Chemistry 41 (1990) 313–344.
[57] E.R. Fisher, J.L. Elkind, D.E. Clemmer, R. Georgiadis, S.K. Loh, N. Aristov, L.S. Sunderlin, P.B. Armentrout, Reactions of fourth-period metal ions (Ca^+–Zn^+) with O_2: metal-oxide ion bond energies, Journal of Chemical Physics 93 (1990) 2676–2691.
[58] P.B. Armentrout, L.F. Halle, J.L. Beauchamp, Periodic trends in transition metal–hydrogen, metal–carbon, and metal–oxygen bond dissociation energies, correlation with reactivity and electronic structure, Journal of the American Chemical Society 103 (1981) 6501–6502.
[59] S.L. Koeniger, S.I. Merenbloom, D.E. Clemmer, Evidence for many resolvable structures within conformation types of electrosprayed ubiquitin ions, Journal of Physical Chemistry B 110 (2006) 7017–7021.
[60] S.L. Koeniger, S.I. Merenbloom, S.J. Valentine, M.F. Jarrold, H.R. Udseth, R.D. Smith, D.E. Clemmer, An IMS–IMS analogue of MS–MS, Analytical Chemistry 78 (2006) 4161–4174.
[61] S.I. Merenbloom, S.L. Koeniger, S.J. Valentine, M.D. Plasencia, D.E. Clemmer, IMS–IMS and IMS–IMS–IMS/MS for separating peptide and protein fragment ions, Analytical Chemistry 78 (2006) 2802–2809.
[62] S.A. Shaffer, K.Q. Tang, G.A. Anderson, D.C. Prior, H.R. Udseth, R.D. Smith, A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry, Rapid Communications in Mass Spectrometry 11 (1997) 1813–1817.
[63] S.A. Shaffer, D.C. Prior, G.A. Anderson, H.R. Udseth, R.D. Smith, An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry, Analytical Chemistry 70 (1998) 4111–4119.
[64] ExPASy Proteomics Server Peptidemass, http://expasy.org/tools/peptide-mass.html.
[65] E.A. Mason, E.W. McDaniel, Transport Properties of Ions in Gases, Wiley, New York, 1988.
[66] S.L. Koeniger, S.I. Merenbloom, S. Sevugarajan, D.E. Clemmer, Transfer of structural elements from compact to extended states in unsolvated ubiquitin, Journal of the American Chemical Society 128 (2006) 11713–11719.
[67] S.J. Valentine, A.E. Counterman, C.S. Hoaglund, J.P. Reilly, D.E. Clemmer, Gas-phase separations of protease digests, Journal of the American Society for Mass Spectrometry 9 (1998) 1213–1216.
[68] J.A. Taraska, A.E. Counterman, D.E. Clemmer, Large anhydrous polyalanine ions: substitution of Na+ for H+ destabilizes folded states, International Journal of Mass Spectrometry 204 (2001) 87–100.
[69] R.D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Crystallographica A32 (1976) 751–767.
[70] S.J. Leon, Linear Algebra with Applications, 3rd ed., Macmillan, New York, 1990.