International Journal of Mass Spectrometry 298 (2010) 111–118

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

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

A study of ion-neutral collision cross-section values for low charge states of peptides, proteins, and peptide/protein complexes☆

Francisco A. Fernandez-Lima, Ryan C. Blase, David H. Russell*

Department of Chemistry, Texas A&M University, College Station, TX 77843-3255, USA

ARTICLE INFO

Article history:
Received 14 May 2009
Received in revised form 2 October 2009
Accepted 12 October 2009
Available online 21 October 2009

Keywords:
Peptide and protein complex secondary structure
Protein oligomers
Ion-helium collision cross-section
Ion mobility spectrometry
Protein docking
Cluster analysis

ABSTRACT

Here, we report ion-helium collision cross-sections (CCS) for a number of peptide, small protein, and peptide/protein ionic complexes. The CCS values reported here are compared to previously reported results [1,2]. We also compare values for low charge state species, i.e., $[M+H]^+$ and $[M+2H]^{2+}$, formed by MALDI with values for high charge state species formed by ESI, and the measured CCSs are compared with predicted CCS for solid-state and solution phase structures and calculated structures obtained by using a protein–protein structure algorithm generator, based on a combined Biomolecular complex Generation with Global Evaluation and Ranking [3] and Multi Dimensional Scaling [4].

Published by Elsevier B.V.

1. Introduction

Protein folding and protein–protein interactions in a physiological environment depend on numerous factors (e.g., electrostatic interactions, intermolecular interactions with the solvent, entropic effects, intramolecular hydrogen bonding and Van der Waals interactions). In native environments protein folding and protein–protein interactions occur under a range of conditions including aqueous solutions as well as hydrophobic and low dielectric membranes [5–7]. Thus, solvent-free studies provide a means of separating solvation effects from intrinsic properties of the protein side-chain and protein–protein site interactions. For example, solvent-free studies may provide new approaches to studies of "non-native" structure that can be used to evaluate peptide and protein folding/unfolding models [8] as well as nucleation and growth [9].

Fundamental studies of gas-phase macromolecules, specifically studies of conformation and thermo-chemical properties of anhydrous peptides and proteins have grown rapidly since the introduction of ESI and MALDI [10–13]. For example, hydrogen/deuterium (H/D) exchange reactions are frequently used to probe the number of exchangeable hydrogen atoms as a measure of the folding of a targeted protein [14–19]. IR spectroscopic techniques have been developed to study secondary structural elements, based on comparisons of theoretically estimated vibrational frequencies of candidate structures, but this approach is often limited to the study of small peptides [20,21]. Ion mobility spectrometry (IMS) combined with theoretical simulations has proven to be the most versatile technique for conformational analysis of intermediate and equilibrium structures of biomolecules by measuring the ion-neutral collision cross-section (CCS) of molecular ions [1,2,22]. IMS combined with theoretical calculations is a powerful technique for probing the equilibrium between conformational states of the macromolecular ions [23–26]. A major difficulty in simulating processes occurring in IMS experiments is the long timescale over which structural rearrangements may occur (∼few milliseconds) whereas molecular dynamics simulations are typically limited to nanoseconds [27–29]. IMS measured CCSs correspond to an average of all thermodynamically available conformations accessible during the experimental measuring time or drift time. Although a number of theoretical methods have been developed to reduce the computation time [30–38], further improvements are needed to efficiently correlate the theoretical results with available experimental data for peptide systems [39]. The theoretical limitation becomes even more pronounced as the molecule of interest increases from small model peptide ions to large ionic proteins or protein complexes; the later systems demand the use of non-dynamic, theoretical strategies to effectively sample the conformational space [40–43].

☆ This article is part of a Special Issue on Ion Mobility.

* Corresponding author. Tel.: +1 979 845 3345.
E-mail address: russell@mail.chem.tamu.edu (D.H. Russell).

1387-3806/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.ijms.2009.10.009

The goal of this study is to provide a high confidence level ion-helium CCS database of peptides, proteins and peptide/protein complexes. These new data will expand the region of reference conformational space to protein and protein complexes, and complement previous reports on CCS databases for peptide systems [1,2]. The advantages of a combined theoretical and experimental analysis of the molecular complex conformational space is illustrated for the insulin oligomer (n = 1–7) using a protein–protein structure algorithm generator, based on a combined biomolecular complex generation with Global Evaluation and Ranking [3] and Multi Dimensional Scaling [4] analysis.

2. Experimental method

The experimental details of the MALDI-IM-MS instrumentation and data acquisition used in this study have been described elsewhere [44–47]. Briefly, ions are formed in an IM drift cell by matrix-assisted laser desorption ionization (MALDI) using a microcrystal Nd:YAG laser (355 nm, Powerchip Nanolaser, JDS Uniphase Corp.), operating at a pulse rate of 200–300 Hz. Ions entering the drift cell are separated on the basis of drift time through the cell maintained at room temperature, a helium pressure of ~3.0 Torr and field strength/pressure ratio (E/p) of 10–40 V cm^{-1} Torr^{-1}. Ions exiting the IM drift cell are focused (by a multi-element Einzel lens) into an orthogonal-TOF ion source. The TOF is biased at a potential of −6 kV and the ions are extracted by applying a voltage pulse to the TOF push/pull electrodes (+675 V/−675 V, respectively); the TOF extraction potential is pulsed at a rate ~2 kHz, and typical mass resolution for the reflectron TOF is 1500–3000. In all experiments, the temperature of the bath gas is *ca.* 300 K. The experiments were performed using a laser power near the ion desorption threshold to minimize peak broadening of the arrival time distribution (ATD) due to space charge effects [48].

All chemicals were purchased from Sigma–Aldrich (St. Louis, MO) and used without further purification. The analytes consist of neuropeptide Y 1–24 (Human), melittin (honey bee venom), insulin β-chain (bovine), β-amyloid 1–40, insulin (bovine pancreas), ubiquitin (bovine red blood cells), cytochrome C (horse heart), RNase A (bovine pancreas), lysozyme (chicken egg white), and apomyoglobin (horse heart). All analytes were dissolved in distilled water at a concentration of 1.0 mg/ml and mixed (1:1, v:v) with 5 mg/ml α-cyano-4-hydroxycinnamic acid (CHCA), dissolved in 60% acetonitrile, 38% distilled water, 1% of 10% trifluoroacetic acid solution (TFA), and 1% ammonium phosphate solution. The mixture of matrix and analyte was spotted onto a 96-well MALDI target plate. The sample spots were prepared in quadruplicate to guarantee spot-to-spot reproducibility and triplicate CCS measurements were performed as independent experiments.

The ion-helium collision cross-sections (CCSs) were determined by using the method described by Mason and McDaniel [49]. Briefly, CCSs are obtained by recording the ATD at different applied voltages (V) to the IM cell and plotting the mean ATD versus 1/V. The ATD-intercept of this linear plot equals the time the ions spend outside the IM cell (t_o), which permits the determination of the CCS values as a function of the drift time ($t_d = \langle \mathrm{ATD} \rangle - t_o$) from the equation:

$$\mathrm{CCS} = \frac{(18\pi)^{1/2}}{16} \frac{ze}{(k_b T)^{1/2}} \left(\frac{1}{m_i} + \frac{1}{m_g} \right)^{1/2} \frac{t_d E}{L} \frac{760}{P} \frac{T}{273.2} \frac{1}{N_0} \qquad (1)$$

where z and e are number of charges and elementary charge, k_b is Boltzmann's constant, T is drift gas temperature, m_g and m_i are the bath gas and ion masses, E is the electric field, L is drift length, p is the helium bath gas pressure, and N_0 is bath gas number density. The CCSs are reported including the propagated error of ATD, t_o determination, and standard deviation of CCS at different E/p values. A fullerene mixture was used as a CCS reference [50,51].

3. Theoretical and computational details

The structures of protein complexes were predicted using a Biomolecular complex Generation with Global Evaluation and Ranking (BiGGER) program [3]. Briefly, BiGGER uses an efficient algorithm to search and select a set of candidate complexes. Each candidate is then evaluated according to the estimated probability of being an accurate model of the complex. This allows the user to rank all models according to a single criterion and focus on the more likely ones if no additional information is available. In the first stage, a step-by-step search through all the possible docking configurations is performed, followed by a stringent filtering criterion that will retain only a small fraction of the possible structures. The maximum number of docking candidates was set to 5000; the docking search was performed using angular steps of 15°, and docking resolution and added radius of 1.0 Å. In the second stage, each solution is evaluated using a neural network to estimate its likelihood of being an accurate model of the complex. The overall evaluation of docked solutions is given by the Interaction Global Score. This is the estimated probability (in percentage) that a docking solution with that Neural Network response is an accurate model of the complex as a function of the surface docking score, side-chain contact filter, electrostatics interaction score, and hydrophobic score [8,52]. Despite modeling some side-chain flexibility of the individual docking units, BiGGER assumes that there are no significant conformational changes during complex formation. To overcome this problem, the top 100 configurations were further optimized using the Open Force Field program in Accelrys Cerius^2, by Accelrys Inc. [53]. The Consistent Force Field (version 2002) was employed in the geometry optimization [54]. Following geometry optimization, the theoretical ion-neutral, collision cross-sections were calculated using the Trajectory Method implemented in the MOBCAL software, with all the interaction potentials taken as the default values [50,51].

The search for candidate structures of the protein–protein complexes was performed following the scheme contained in Fig. 1. An initial set of candidate structures measured experimentally is pre-defined as docking input elements. After each cycle, the highest score structure is selected as a potential docking input for the next cluster size. All potential combinations of docking pre-

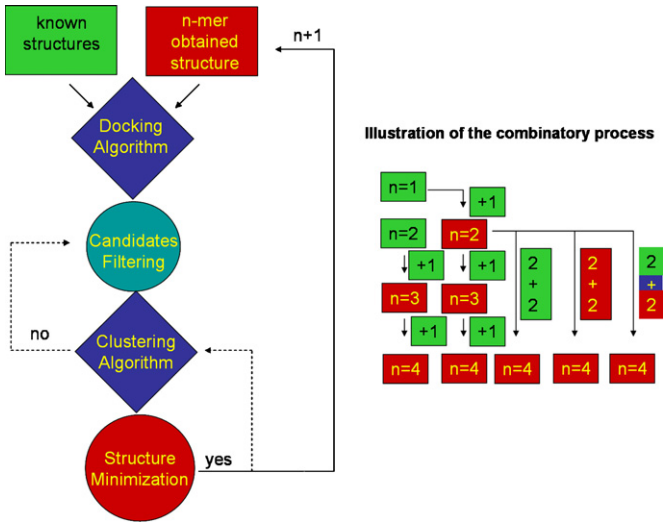

Fig. 1. Scheme of the protein–protein candidate structure generation algorithm. On the right, the combinatory process is illustrated for the case of pre-defined protein monomer and dimer structures as docking elements (green). New docking elements are obtained from each combinatory output (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1
Ion-helium collision cross-section values for single-charged peptides, proteins, and peptide and protein complexes.

Analyte	*m*/*z*	*n*-mer	Average CCS (A^2)	Average K_o ($cm^2\ V^{-1}\ s^{-1}$)
Neuropeptide Y (1–24)				
Monomer	2657.9	1	417 ± 36	1.29 ± 0.096
Dimer	5314.8	2	651 ± 34	0.83 ± 0.041
Trimer	7971.7	3	771 ± 34	0.70 ± 0.028
Tetramer	10628.6	4	932 ± 39	0.58 ± 0.026
Pentamer	13285.5	5	1044 ± 46	0.51 ± 0.026
Hexamer	15942.4	6	1159 ± 35	0.46 ± 0.012
Heptamer	18599.3	7	1192 ± 105	0.48 ± 0.038
Octamer	21256.2	8	1227 ± 103	0.46 ± 0.040
Melittin				
Monomer	2847.5	1	469 ± 30	1.15 ± 0.057
Dimer	5694	2	774 ± 35	0.69 ± 0.015
Trimer	8540.5	3	881 ± 68	0.62 ± 0.038
Tetramer	11,387	4	1009 ± 92	0.56 ± 0.040
Insulin β-chain				
Monomer	3496.9	1	501 ± 27	1.07 ± 0.043
Dimer	6992.8	2	730 ± 30	0.74 ± 0.028
Trimer	10488.7	3	872 ± 41	0.62 ± 0.030
Tetramer	13984.6	4	1016 ± 40	0.53 ± 0.026
Pentamer	17480.5	5	1167 ± 101	0.47 ± 0.039
Hexamer	20976.4	6	1270 ± 103	0.45 ± 0.033
Heptamer	24472.3	7	1370 ± 122	0.41 ± 0.035
Octamer	27968.2	8	1436 ± 134	0.40 ± 0.038
β-Amyloid (1–40)				
Monomer	4330.9	1	560 ± 35	0.96 ± 0.067
Dimer	8660.8	2	773 ± 46	0.70 ± 0.055
Trimer	12990.7	3	950 ± 57	0.57 ± 0.044
Tetramer	17320.6	4	1118 ± 60	0.48 ± 0.031
Pentamer	21650.5	5	1300 ± 114	0.44 ± 0.041
Hexamer	25980.4	6	1377 ± 115	0.41 ± 0.034
Insulin				
Monomer	5734.6	1	654 ± 37	0.82 ± 0.044
Dimer	11468.2	2	930 ± 41	0.58 ± 0.029
Trimer	17201.8	3	1119 ± 52	0.48 ± 0.028
Tetramer	22935.4	4	1356 ± 101	0.43 ± 0.033
Pentamer	28,669	5	1486 ± 110	0.40 ± 0.037
Hexamer	34402.6	6	1592 ± 144	0.38 ± 0.050
Heptamer	40136.2	7	1705 ± 141	0.35 ± 0.037
Ubiquitin				
Monomer	8565.8	1	821 ± 45	0.66 ± 0.040
Dimer	17130.6	2	1147 ± 57	0.47 ± 0.029
Trimer	25695.4	3	1410 ± 129	0.41 ± 0.045
Tetramer	34260.2	4	1580 ± 122	0.37 ± 0.041
Cytochrome C				
Monomer	12,232	1	983 ± 58	0.55 ± 0.030
Dimer	24,463	2	1303 ± 101	0.41 ± 0.031
RNase A				
Monomer	13683.4	1	983 ± 52	0.55 ± 0.039
Dimer	27365.8	2	1407 ± 107	0.41 ± 0.043
Trimer	41048.2	3	1838 ± 132	0.34 ± 0.040
Lysozyme				
Monomer	14306.1	1	1017 ± 56	0.53 ± 0.031
Dimer	28611.2	2	1464 ± 106	0.39 ± 0.035
Trimer	42916.3	3	1895 ± 107	0.33 ± 0.032
Apomyoglobin				
Monomer	16952.5	1	1077 ± 77	0.50 ± 0.052
Dimer	33,904	2	1587 ± 144	0.37 ± 0.040

defined inputs are considered. For example, the scheme contained in Fig. 1 considers the structure of a protein monomer (n1) and dimer (n2) as pre-defined docking elements. For $n = 2$, the candidature structure pool is composed of the dimer structure and the top 100 (n1 + n1) structures. The highest score structure from the 100 n1 + n1 structures is then defined as the n1n1 docking element and used in the search for larger cluster sizes. That is, the search for trimer (n3) candidates will consider n1n1, n2, and n1 as docking elements, which yield n1n1 + n1 and n2 + n1 as potential combinations. In the insulin example, since the number of potential combinations increases geometrically with the cluster size, the number of new defined docking elements after the optimization cycle for a given combination was simplified to the highest score candidate (i.e., only 1 structure); nevertheless, this algorithm has the potential to set the number of new defined elements by the number of structurally different cluster families. The docking elements for the insulin monomer (n1) and dimer (n2) were obtained from X-ray diffraction data [55].

A user-developed algorithm was used to classify the resulting optimized structures in clusters. The cluster analysis is based on

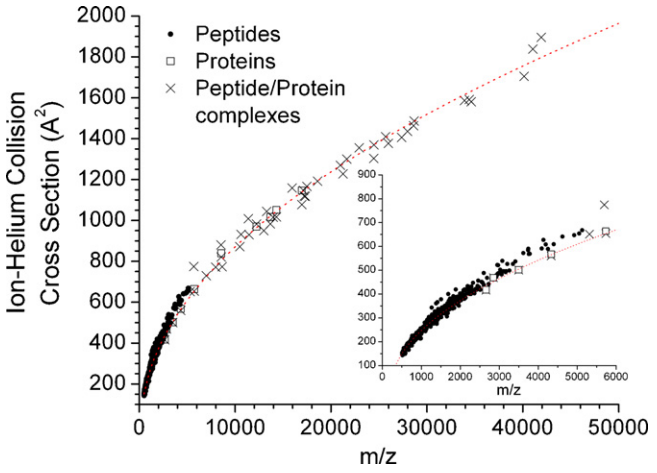

Fig. 2. Plot of ion-helium collision cross-sections vs. m/z for single-charged peptide and peptide and protein ionic complexes. In the inset, CCS for single-charge peptide ions obtained by digesting nine standard proteins with three proteolytic enzymes (trypsin, chymotrypsin, and pepsin) is shown [1,2]. The total fitting function is $CCS = 8.8(m/z - 210)^{1/2}$.

the comparison of all structures and subsequent classifications in clusters, i.e., structures are classified such that a given cluster contains a set of structures that are more similar to one another than to members of other clusters. Comparison between the similarities of two structures is made using the average root mean square deviation (RMSD) using the SUPPOSE program [56]. Average RMSD values are computed from the pair-wise RMSD measured between corresponding backbone atoms of two structures, once the corresponding atoms have been identified and the molecules have been rotated and translated as rigid bodies to the best match. A multidimensional scaling (MDS) method was used to generate an n-dimensional projection of all the structures, where the inter-point distances are "close" to their average RMSD values [4]. Briefly, the dimension, n, of the MDS space was chosen such that the inter-point distances and the RMSD values correspond within a stress value below 5% [4]. The structures were classified in the n-dimensional MDS space using a hierarchical clustering algorithm (HCA) [57]. The HCA classifies the structures without a predetermined number of clusters, which permits users to determine the natural grouping with interactive visual feedback (e.g., dendrograms, 3D projections and/or color mosaic). The optimum number of clusters can be defined by the user or can be determined using an iterative algorithm from a user-defined RMSD tolerance criteria within a cluster.

4. Results and discussion

Ion mobility-mass spectrometry separates gas-phase ions on the basis of ion-neutral collision cross-section (CCS) and m/z ratios. One of the objectives of the present study is to compile a database of CCS values for single-charged peptide and protein ionic complexes, which can be used as calibration standards for CCS values (Table 1). Establishing a CCS database of macromolecules and molecular complexes is especially important for techniques such as differential mobility analyzers (DMA) [58], field asymmetric waveform ion mobility spectrometers (FAIMS) [59,60], and T-wave ion mobility spectrometers [61,62], because the data obtained from such experiments cannot be used for first principle CCS calculations.

Fig. 2 contains a plot of CCS versus m/z for a large number of single-charged, peptide and protein ions, as well as oligomeric peptide and protein ionic complexes. The dataset includes CCS values of previously reported single-charged peptide ions obtained by digesting nine standard proteins with three proteolytic enzymes (trypsin, chymotrypsin, and pepsin) [2]. Inspection of Fig. 2 shows that peptide ions and peptide/protein ionic complexes fall on the same trend line in the CCS versus m/z plot, and this average CCS trend line can be described as a square-root function of m/z: $\langle CCS \rangle = 8.8(m/z - 210)^{1/2}$. We have previously observed that in the case of peptide ions, a small number of ion signals deviate from the average CCS trend line, while other molecular classes fall on very different trend lines [2,63]. For calibration purposes, we recommend the use of the CCS data reported in Table 1. The broad distribution of CCSs contained in Fig. 2 illustrates the structural diversity of the peptide and peptide ion complexes while highlighting the utility of the IM-MS separation to identify a single molecular class. For example, the CCS for melittin $[M+H]^+$ ions (see Table 1) lies above (by approximately 4%) the average peptide ion CCS trend line. The deviation from the expected value could be an indication of helical character for the melittin $[M+H]^+$ ion, which in its native state is ~88% helical [64–66]. In addition, helical character was previously used to explain the results for gas-phase MS fragmentation reactions and MALDI hydrogen/deuterium (H/D) exchange [19,67], and formation of melittin oligomers with helical character has been also observed in basic or high ionic strength solutions using NMR studies [68].

On the other hand, the $[M+H]^+$ ion of neuropeptide Y 1–24 presents a large negative deviation in CCS (−4.3%), that could be related to the lower helical content (<40%) observed in native state studies [69]. As the molecular size increases, the influence of the helical/sheet content on the CCS may be overshadowed by other structural features, i.e., side-chain and intramolecular interactions. The CCS for insulin $[M+H]^+$ ions agrees very well with the expected value, whereas the CCS for apomyoglobin $[M+H]^+$ ions has a large negative deviation (−5.4%) from the average CCS. In their native states, these proteins contain ~46% and ~74% helical character, respectively [55,70,71]. Quite possibly the intramolecular disulfide bonding of insulin makes the total secondary structure spherical/globular, where the helical contribution to a more elongated form (and larger CCS) has been reduced by these side-chain interactions [55]. Similarly, for apomyoglobin the side-chain interaction with the heme group in the native state forces the $[M+H]^+$ ion to adopt a globular/spherical compact configuration surrounding the heme group [70,71]. Overall, side-chain and intramolecular interactions may also stabilize the secondary structure of the ionic complex in "non-native" states, while not conserving the original secondary structure and/or helical/sheet content. Only large differences in the secondary structure of native versus "non-native" states will significantly contribute to the packing density or CCS, making higher resolution IMS experiments necessary to successfully observe these effects for large peptides and proteins. The CCS values contained in Table 1 are in good agreement with reported data obtained using DMA analyzers [72–74], where the electromobility particle diameter correspond to the diameter of a sphere if the molecular complex is assumed to have a globular shape.

The CCS values reported in Table 1 are for single-charged, protonated molecules, $[M+H]^+$, which are always the most abundant ions formed by MALDI. In some cases, double-charged ions, $[M+2H]^{2+}$ ions can be observed in MALDI-IM-MS experiments (e.g., results of McLean et al. [75] and Blase et al. [76]), and measured CCSs for some of the $[M+2H]^{2+}$ ions are compared with previously reported CCS values obtained by using ESI (Fig. 3) [77,78]. Note that the collision cross section increases as the charge state increases. Thus it appears that as the charge state increases Coulombic interactions also increase, forcing the molecular structure to separate the charge sites, which in turn increases the total conformational space and leads to larger CCS values. Comparison of the CCS values with reported Protein Data Bank structures obtained by using nuclear magnetic resonance (solution phase structure) and X-ray diffraction (crystal structure) results shows that the best agreement is

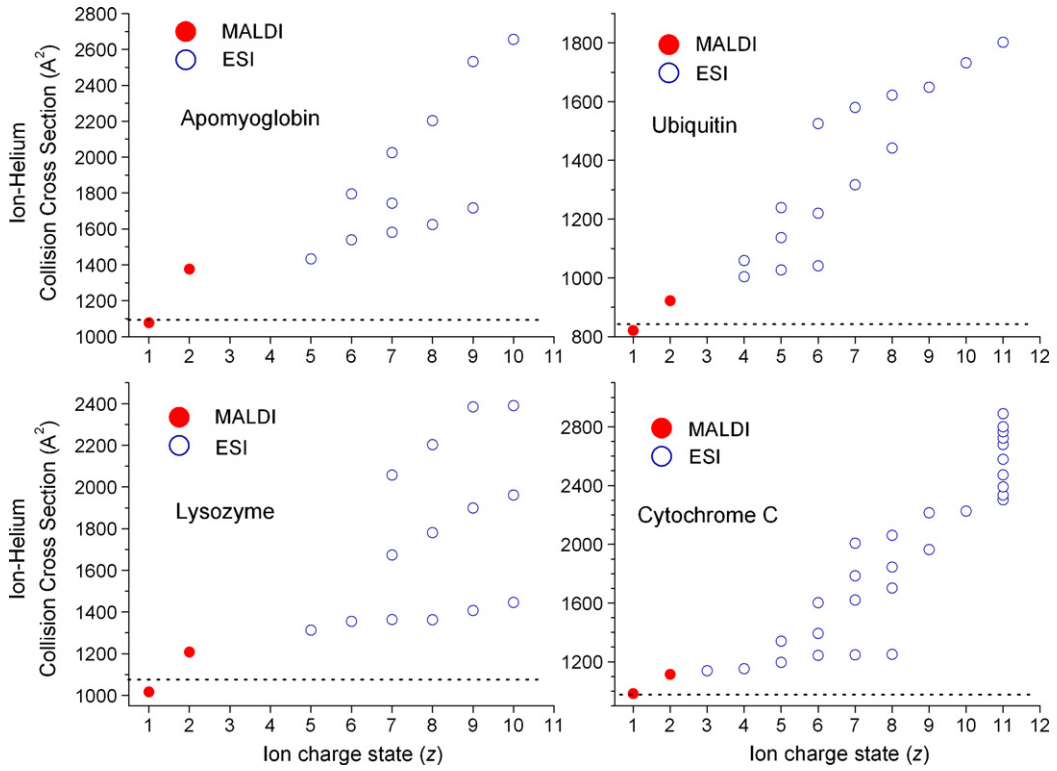

Fig. 3. Ion-helium CCS variation as a function of the ion charge state, *z*. Single- and double-charged ions were produced using a MALDI source, and the higher order charge states were previously reported using an ESI source [77,78]. The dashed lines represent the corresponding average CCS values of nuclear magnetic resonance (solution phase structure) and X-ray diffraction (crystal structure) reported structures [70,71,79–85].

observed for the lower charge states CCS values obtained using a MALDI source (Fig. 3) [70,71,79–85]. Larger deviations for CCS for the higher charge states may also arise if higher energy configurations or "non-native" states are sampled. For example, denaturing of proteins in the gas-phase at high charge states has been previously observed in ESI-IM-MS experiments [86,87].

Many of the peptides studied in the present work are also observed as oligomeric structures (see Table 1). The CCS of the protein complexes do not deviate significantly from the predicted trend line, and as discussed earlier can be used to infer structural information of the ionic protein complexes. That is, the CCS provides an experimental verification for the packing density of the oligomeric structures, which may ultimately be correlated with the peptide/protein aggregation mechanism, or as shown recently by Ruotolo et al. could be used to infer denaturation of peptide/protein ionic complexes [86,87].

In previous studies we have shown that IMS data are greatly enhanced when complemented with theoretical studies of candidate structures by direct comparison of the theoretical and experimental CCS values [27,88–90]. In this study we use a protein–protein structure algorithm generator to compare experimental and theoretical CCS values for insulin oligomers (n = 1–7). The spherical/globular shape of the insulin monomer makes it a good candidate to illustrate the utility of this method.

Fig. 4 contains a comparison between theoretical and experimental CCS for the insulin oligomers. The protein–protein structure algorithm generator assumes that there are no significant conformational changes during complex formation. That is, the secondary structure of the insulin monomer does not significantly change as the individual units assemble. The binding orientations will ultimately define the total conformational state of the protein complex, which can be related to the gas-phase packing density.

Theoretical CCSs are obtained after averaging the highest score structures obtained in the protein–protein structure algorithm generator. The docking parameters (e.g., angular step, docking resolution and added radius) of the protein–protein structure algorithm generator were verified by the observation of the n2 structure of the insulin dimer as one of the 100 highest scores for insulin monomer (n1 + n1) combinations. Good agreement is observed between the predicted and experimentally observed CCS values (<2%). Close inspection reveals smaller CCSs for the n = 3 and n = 6 oligomers (see Fig. 4 inset), which suggests a more compact structure for this species. Fig. 5 contains the most abundant backbone orientation of the insulin trimer (n = 3) obtained from the protein–protein structure algorithm generator. The insulin trimer is obtained from a combination of the insulin dimer (n2) and docked structures of

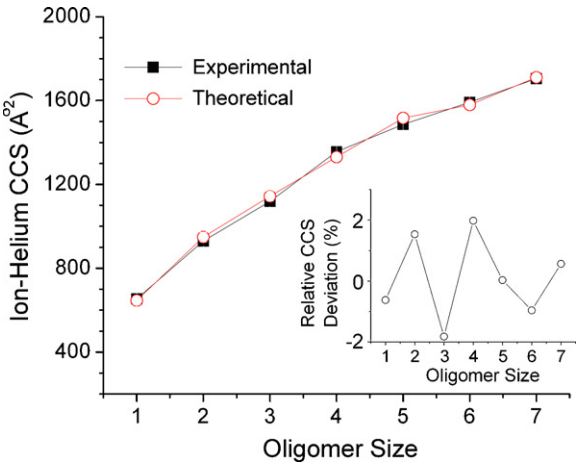

Fig. 4. Theoretical and experimental ion-helium CCS for the insulin oligomers. In the inset, relative CCS deviations to the average experimental CCS values are shown. Notice the negative deviations for the insulin trimer and hexamer. The average theoretical and experimental CCS values agree within 2%.

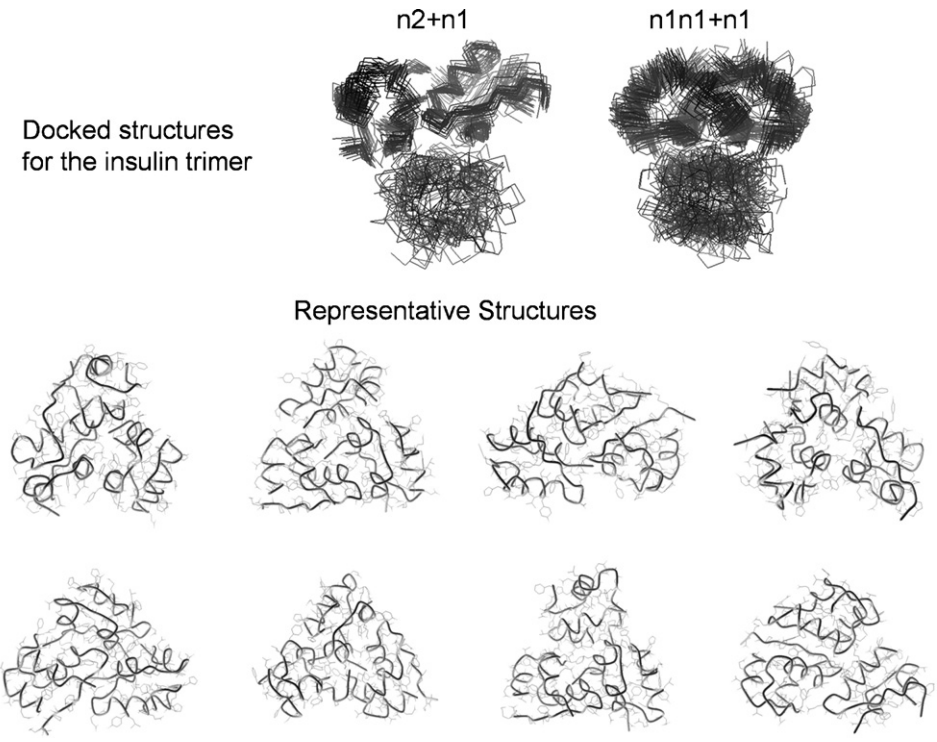

Fig. 5. Illustration of the docking procedure for the insulin trimer using the protein–protein structure algorithm generator. Top: Overlay of the backbone structures obtained from the combination of n2 + n1 and n1n1 + n1. Bottom: Backbone and side-chain representation of the representative structures obtained from the cluster analysis of the highest score structures.

insulin monomers (n1 + n1) with the insulin monomer (n1). Eight representative structures were obtained for the insulin trimer, and the main differences between these structures are related to the orientations of the docking sites. It appears that the spherical/globular shape of the insulin monomers makes the trimer structure more stable owing to a larger number of interactions. Each monomer unit in the trimer structure is attached to two other monomer units. For larger oligomers, multiple docking regions are observed, which in some cases promotes formation of a triangular pyramid.

The hexameric insulin ion also appears to favor a single structural motif, where each unit is coordinated with three insulin monomer units forming a compact structure that resembles two parallel, triangular trimer structures coupled together. This compact hexamer configuration shares some similarities with solution based studies, where the insulin hexamer counterparts are coordinated to Zn metal ions and each insulin monomer unit is coordinated to three other insulin monomer units [91–93]. A high degree of stabilization of the oligomeric complexes with the larger number of docking-site interactions can be directly correlated with negative CCS deviations if the monomer unit has a spherical/globular shape. When the monomer unit deviates from the spherical/globular shape, docking-site interactions between the oligomer counterparts may be sufficiently strong to stabilize the oligomer complex in site-oriented configurations, where their correlation with the CCS may not be straightforward.

5. Conclusions

A high confidence level peptide and peptide/protein ionic complex ion-helium CCS reference database has been compiled, which significantly extends the conformational space of single-charged ions up to ~50 kDa. This database provides essential data for future theoretical studies and contains calibration standards for other gas-phase separation techniques where direct CCS measurements cannot be obtained.

We also describe the use of a protein–protein structure algorithm generator to estimate CCSs for protein ions as well as protein ionic complexes. On the basis of these studies it appears that intramolecular interactions stabilize the secondary structure of the ionic complexes in native and/or "non-native" states. The end result is that only large differences in secondary structure of native versus "non-native" states will significantly contribute to the measured CCS, thus higher resolution IMS experiments will be needed to observe these effects for large molecular systems. We anticipate that a 3–4-fold increase in the IMS resolution will permit the study of docking-site orientations on the peptide and protein complex assembly, disassembly and exchange of subunits between protein complexes.

The advantages of a protein–protein structure algorithm generator were illustrated by the comparison of theoretical and experimental CCS of insulin oligomers, where deviations in the CCS values were associated with structural motifs of the protein complex. This work provides a step forward for the development of theoretical algorithms to efficiently sample the protein complex conformational space. Future efforts will be focused on the implementation of higher-level docking strategies and the study of protein complexes where side-chain groups and/or "structural pockets" have been modified to induce specific molecular attachment as a method to study drug delivery mechanisms and their effect on the biological function of protein complexes.

Acknowledgments

This work was supported by the Robert A. Welch Foundation (A-1176), the U.S. Department of Energy, Division of Chemical Sciences, BES (DE-FG-04ER-15520) and the NSF-MRI (DSI0821700) grants. We would like to thank Dr. Lisa M. Perez from the Laboratory for Molecular Simulations at Texas A&M University for the helpful discussions and support during the development of the computational methods.

References

[1] 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, J. Am. Soc. Mass Spectrom. 10 (1999) 1188–1211.
[2] L. Tao, J.R. McLean, J.A. McLean, D.H. Russell, A collision cross-section database of singly-charged peptide ions, J. Am. Soc. Mass Spectrom. 18 (2007) 1232–1238.
[3] P.N. Palma, L. Krippahl, J.E. Wampler, J.J.G. Moura, BiGGER: a new (soft) docking algorithm for predicting protein interactions, Proteins: Struct. Funct. Genet. 39 (2000) 372–384.
[4] T.F. Cox, M.A.A. Cox, Multidimensional Scaling, CRC Press, London, 1994.
[5] D.A. Karp, A.G. Gittis, M.R. Stahley, C.A. Fitch, W.E. Stites, B. Garcia-Moreno, High apparent dielectric constant inside a protein reflects structural reorganization coupled to the ionization of an internal asp, Biophys. J. 92 (2007) 2041–2053.
[6] C.A. Fitch, S.T. Whitten, V.J. Hilser, B. García-Moreno, Molecular mechanisms of pH-driven conformational transitions of proteins: insights from continuum electrostatics calculations of acid unfolding, Proteins: Struct. Funct. Bioinform. 63 (2006) 113–126.
[7] S. Bone, Dielectric properties of biomacromolecules: some aspects of relevance to biological systems, J. Bioelectricity 4 (1985) 389–418.
[8] Y. Wang, H. Zhang, W. Li, R.A. Scott, Discriminating compact nonnative structures from the native structure of globular proteins, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 709–713.
[9] W.F.I.V. Weiss, T.K. Hodgdon, E.W. Kaler, A.M. Lenhoff, C.J. Roberts, Nonnative protein polymers: structure, morphology, and relation to nucleation and growth, Biophys. J. 93 (2007) 4392–4403.
[10] M. Karas, D. Bachmann, F. Hillenkamp, Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules, Anal. Chem. 57 (1985) 2935–2939.
[11] F. Hillenkamp, M. Karas, Matrix-assisted laser desorption/ionisation, an experience, Int. J. Mass Spectrom. 200 (2000) 71–77.
[12] C.M. Whitehouse, R.N. Dreyer, M. Yamashita, J.B. Fenn, Electrospray interface for liquid chromatographs and mass spectrometers, Anal. Chem. 57 (1985) 675–679.
[13] R.B. Cole, Electrospray Ionization Mass Spectrometry;, John Wiley & Sons, New York, 1997.
[14] A. Miranker, C.V. Robinson, S.E. Radford, R.T. Aplin, C.M. Dobson, Detection of transient protein folding populations by mass spectrometry, Science 262 (1993) 896–900.
[15] S.J. Valentine, D.E. Clemmer, Temperature-dependent H/D exchange of compact and elongated cytochrome c ions in the gas phase, J. Am. Soc. Mass Spectrom. 13 (2002) 506–517.
[16] K.A. Resing, A.N. Hoofnagle, N.G. Ahn, Modeling deuterium exchange behavior of ERK2 using pepsin mapping to probe secondary structure, J. Am. Soc. Mass Spectrom. 10 (1999) 685–702.
[17] Y. Deng, Z. Zhang, D.L. Smith, Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics, J. Am. Soc. Mass Spectrom. 10 (1999) 675–684.
[18] O. Nemirovskiy, D.E. Giblin, M.L. Gross, Electrospray ionization mass spectrometry and hydrogen/deuterium exchange for probing the interaction of calmodulin with calcium, J. Am. Soc. Mass Spectrom. 10 (1999) 711–718.
[19] I.D. Figueroa, D.H. Russell, Matrix-assisted laser desorption ionization hydrogen/deuterium exchange studies to probe peptide conformational changes, J. Am. Soc. Mass Spectrom. 10 (1999) 719–731.
[20] A.S. Jaime, G. Monia, V.B. Oleg, R.R. Thomas, Conformation-specific infrared and ultraviolet spectroscopy of tyrosine-based protonated dipeptides, J. Chem. Phys. 127 (2007) 154322.
[21] I. Hunig, K. Kleinermanns, Conformers of the peptides glycine–tryptophan, tryptophan–lycine and tryptophan–glycine–glycine as revealed by double resonance laser spectroscopy, Phys. Chem. Chem. Phys. 6 (2004) 2650–2658.
[22] M.F. Jarrold, Peptides and proteins in the vapor phase, Annu. Rev. Phys. Chem. 51 (2000) 179–207.
[23] B.T. Ruotolo, J.A. McLean, K.J. Gillig, D.H. Russell, The influence and utility of varying field strength for the separation of tryptic peptides by ion mobility-mass spectrometry, J. Am. Soc. Mass Spectrom. 16 (2005) 158–165.
[24] A.B. Kanu, H.H. Hill Jr., Identity confirmation of drugs and explosives in ion mobility spectrometry using a secondary drift gas, Talanta 73 (2007) 692–699.
[25] L.W. Zilch, D.T. Kaleta, M. Kohtani, R. Krishnan, M.F. Jarrold, Folding and unfolding of helix-turn-helix motifs in the gas phase, J. Am. Soc. Mass Spectrom. 18 (2007) 1239–1248.
[26] F.A. Fernandez-Lima, H. Wei, Y.Q. Gao, D.H. Russell, On the structure elucidation using IMS and molecular dynamics, J. Phys. Chem. A 113 (2009) 8221–8234.
[27] Y.Q. Gao, L.J. Yang, On the enhanced sampling over energy barriers in molecular dynamics simulations, J. Chem. Phys. 125 (2006) 114103.
[28] Y.Q. Gao, An integrate-over-temperature approach for enhanced sampling, J. Chem. Phys. 128 (2008) 064105.
[29] Y.Q. Gao, Self-adaptive enhanced sampling in the energy and trajectory spaces: accelerated thermodynamics and kinetic calculations, J. Chem. Phys. 128 (2008) 134111.
[30] C. Bartels, M. Karplus, Probability distributions for complex systems: adaptive umbrella sampling of the potential energy, J. Phys. Chem. B 102 (1998) 865–880.
[31] E. Darve, A. Pohorille, Calculating free energies using average force, J. Chem. Phys. 115 (2001) 9169.
[32] G. Bussi, A. Laio, M. Parrinello, Equilibrium free energies from nonequilibrium metadynamics, Phys. Rev. Lett. 96 (2006) 090601.
[33] D. Hamelberg, J.A. McCammon, Fast peptidyl cis–trans isomerization within the flexible gly-rich flaps of HIV-1 protease, J. Am. Chem. Soc. 127 (2005) 13778–13779.
[34] D. Hamelberg, J. Mongan, J.A. McCammon, Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules, J. Chem. Phys. 120 (2004) 11919.
[35] B.A. Berg, T. Neuhaus, Multicanonical algorithms for first order phase transitions, Phys. Lett. B 267 (1991) 249–253.
[36] A. Mitsutake, Y. Okamoto, Replica-exchange simulated tempering method for simulations of frustrated systems, Chem. Phys. Lett. 332 (2000) 131–138.
[37] Y. Sugita, A. Kitao, Y. Okamoto, Multidimensional replica-exchange method for free-energy calculations, J. Chem. Phys. 113 (2000) 6042.
[38] U.H.E. Hansmann, Parallel tempering algorithm for conformational studies of biological molecules, Chem. Phys. Lett. 281 (1997) 140–150.
[39] M. Damsbo, B.S. Kinnear, M.R. Hartings, P.T. Ruhoff, M.F. Jarrold, M.A. Ratner, Application of evolutionary algorithm methods to polypeptide folding: comparison with experimental results for unsolvated Ac-(Ala-Gly-Gly)5-LysH+, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 7215–7222.
[40] L.P. Ehrlich, R.C. Wade, Protein–protein docking, Rev. Comput. Chem. 17 (2001) 61–97.
[41] A.G. Marshall, I.A. Vakser, Protein–protein docking methods, Protein Rev. 3 (2005) 115–146.
[42] K. Lee, J.-W. Lee, Computational approaches to protein–protein docking, Curr. Proteomics 5 (2008) 10–19.
[43] D.W. Ritchie, Recent progress and future directions in protein–protein docking, Curr. Protein Peptide Sci. 9 (2008) 1–15.
[44] K.J. Gillig, B. Ruotolo, E.G. Stone, D.H. Russell, K. Fuhrer, M. Gonin, A.J. Schultz, Coupling high-pressure MALDI with ion mobility/orthogonal time-of-flight mass spectrometry, Anal. Chem. 72 (2000) 3965–3971.
[45] E.G. Stone, K.J. Gillig, B.T. Ruotolo, D.H. Russell, Optimization of a matrix-assisted laser desorption ionization-ion mobility-surface-induced dissociation-orthogonal-time-of-flight mass spectrometer: simultaneous acquisition of multiple correlated MS1 and MS2 spectra, Int. J. Mass Spectrom. 212 (2001) 519–533.
[46] B.T. Ruotolo, K.J. Gillig, E.G. Stone, D.H. Russell, K. Fuhrer, M. Gonin, J.A. Schultz, Analysis of protein mixtures by matrix-assisted laser desorption ionization-ion mobility-orthogonal-time-of-flight mass spectrometry, Int. J. Mass Spectrom. 219 (2002) 253–267.
[47] K.J. Gillig, B.T. Ruotolo, E.G. Stone, D.H. Russell, An electrostatic focusing ion guide for ion mobility-mass spectrometry, Int. J. Mass Spectrom. 239 (2004) 43–49.
[48] P. Watts, A. Wilders, On the resolution obtainable in practical ion mobility systems, Int. J. Mass Spectrom. Ion Processes 112 (1992) 179–190.
[49] E.W. McDaniel, E.A. Mason, Mobility and Diffusion of Ions in Gases, John Wiley and Sons, Inc., New York, 1973.
[50] 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, J. Phys. Chem. 100 (1996) 16082–16086.
[51] A.A. Shvartsburg, M.F. Jarrold, An exact hard-spheres scattering model for the mobilities of polyatomic ions, Chem. Phys. Lett. 261 (1996) 86–91.
[52] Y. Wang, H. Zhang, R.A. Scott, A new computational model for protein folding based on atomic solvation, Protein Sci. 4 (1995) 1402–1411.
[53] Accelrys Cerius^2 Manual, 4.9 ed., Accelrys Inc., 2003.
[54] U. Dinur, A.T. Hagler, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 2, VCH, New York, 1991, pp. 99–164.
[55] M.H. Nanao, G.M. Sheldrick, R.B. Ravelli, Improving radiation-damage substructures for RIP, Acta Crystallogr. D 61 (2005) 1227–1237.
[56] J. Smith, http://structbio.vanderbilt.edu/~jsmith/, 1997.
[57] J. Seo, B. Shneiderman, Interactively exploring hierarchical clustering results, IEEE Comput. 35 (2002) 80–86.
[58] P.D. Schnier, W.D. Price, R.A. Jockusch, E.R. Williams, Blackbody infrared radiative dissociation of bradykinin and its analogues: energetics, dynamics, and evidence for salt-bridge structures in the gas phase, J. Am. Chem. Soc. 118 (1996) 7178–7189.
[59] A.A. Shvartsburg, K. Tang, R.D. Smith, Understanding and designing field asymmetric waveform ion mobility spectrometry separations in gas mixtures, Anal. Chem. 76 (2004) 7366–7374.
[60] A.A. Shvartsburg, F. Li, K. Tang, R.D. Smith, Distortion of ion structures by field asymmetric waveform ion mobility spectrometry, Anal. Chem. 79 (2007) 1523–1528.
[61] A.A. Shvartsburg, R.D. Smith, Fundamentals of traveling wave ion mobility spectrometry, Anal. Chem. 80 (2008) 9689–9699.
[62] K. Thalassinos, M. Grabenauer, S.E. Slade, G.R. Hilton, M.T. Bowers, J.H. Scrivens, Characterization of phosphorylated peptides using traveling wave-based and drift cell ion mobility mass spectrometry, Anal. Chem. 81 (2009) 248–254.
[63] J.A. McLean, B.T. Ruotolo, K.J. Gillig, D.H. Russell, Ion mobility-mass spectrometry: a new paradigm for proteomics, Int. J. Mass Spectrom. 240 (2005) 301–315.
[64] T.C. Terwilliger, D. Eisenberg, The structure of melittin. I. Structure determination and partial refinement, J. Biol. Chem. 257 (1982) 6010.
[65] T.C. Terwilliger, D. Eisenberg, The structure of melittin. II. Interpretation of the structure, J. Biol. Chem. 257 (1982) 6016.
[66] D. Anderson, T.C. Terwilliger, W. Wickner, D. Eisenberg, Melittin forms crystals which are suitable for high resolution X-ray structural analysis and which reveal a molecular 2-fold axis of symmetry, J. Biol. Chem. 255 (1980) 2578.

[67] R.J. Anderegg, D.S. Wagner, C.L. Stevenson, R.T. Borchardt, The mass spectrometry of helical unfolding in peptides, J. Am. Soc. Mass Spectrom. 5 (1994) 425–433.
[68] J.S. Buckley, A.S. Edison, M.D. Kemple, F.G. Prendergast, 13 Calpha-NMR assignments of melittin in methanol and chemical shift correlations with secondary structure, J. Biomol. NMR 3 (1993) 639–652.
[69] S.A. Monks, G. Karagianis, G.J. Howlett, R.S. Norton, Solution structure of human neuropeptide Y, J. Biomol. NMR 8 (1996) 379–390.
[70] D.M. Copeland, A.H. West, G.B. Richter-Addo, Crystal structures of ferrous horse heart myoglobin complexed with nitric oxide and nitrosoethane, Proteins: Struct. Funct. Genet. 53 (2003) 182–192.
[71] R. Maurus, C.M. Overall, R. Bogumil, Y. Luo, A.G. Mauk, M. Smith, G.D. Brayer, A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket, Biochim. Biophys. Acta 1341 (1997) 1–13.
[72] S.L. Kaufman, J.W. Skogen, F.D. Dorman, F. Zarrin, K.C. Lewis, Macromolecule analysis based on electrophoretic mobility in air: globular proteins, Anal. Chem. 68 (1996) 3703–13703.
[73] S.L. Kaufman, Analysis of biomolecules using electrospray and nanoparticle methods: the gas-phase electrophoretic mobility molecular analyzer (gemma), J. Aerosol Sci. 29 (1998) 537–552.
[74] G. Bacher, W.W. Szymanski, S.L. Kaufman, P. Zöllner, D. Blaas, G. Allmaier, Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses, J. Mass Spectrom. 36 (2001) 1038–1052.
[75] J.A. McLean, J.R. McLean, D.H. Russell, Investigation of anhydrous protein structures: the influence of post-translational modifications on tertiary and quaternary structure, in: 19th Annual Symposium of the Protein Society, Boston, August, 2005.
[76] R.C. Blase, F.A. Fernandez-Lima, C. Becker, L.M. Perez, D.H. Russell, Structure of protein aggregates in the gas-phase using MALDI-IM-MS, in: 56th ASMS Conference on Mass Spectrometry and Allied Topics, Denver, 2008.
[77] S.J. Valentine, A.E. Counterman, D.E. Clemmer, Conformer-dependent proton-transfer reactions of ubiquitin ions, J. Am. Soc. Mass Spectrom. 8 (1997) 954–961.
[78] S.J. Valentine, J.G. Anderson, A.D. Ellington, D.E. Clemmer, J. Phys. Chem. B 101 (1997) 3891–3900.
[79] C. Evrard, J. Fastrez, J.P. Declercq, Crystal structure of the lysozyme from bacteriophage lambda and its relationship with V and C-type lysozymes, J. Mol. Biol. 276 (1998) 151–164.
[80] M. Refaee, T. Tezuka, K. Akasaka, M. Williamson, Pressure-dependent changes in the solution structure of hen egg-white lysozyme, J. Mol. Biol. 327 (2003) 857.
[81] S. Vijay-Kumar, C.E. Bugg, W.J. Cook, Structure of ubiquitin refined at 1.8 Å resolution, J. Mol. Biol. 194 (1987) 531–544.
[82] G. Cornilescu, J.L. Marquardt, M. Ottiger, A. Bax, Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase, J. Am. Chem. Soc. 120 (1998) 6836–6837.
[83] S. Benini, A. Gonzalez, W.R. Rypniewski, K.S. Wilson, J.J. Van Beeumen, S. Ciurli, Crystal structure of oxidized *Bacillus pasteurii* cytochrome c553 at 0.97 Å resolution, Biochemistry 39 (2000) 13115–13126.
[84] J. Hasegawa, T. Yoshida, T. Yamazaki, Y. Sambongi, Y. Yu, Y. Igarashi, T. Kodama, K. Yamazaki, Y. Kyogoku, Y. Kobayashi, Solution structure of thermostable cytochrome c-552 from Hydrogenobacter thermophilus determined by 1H-NMR spectroscopy, Biochemistry 37 (1998) 9641–9649.
[85] K. Osapay, Y. Theriault, P.E. Wright, D.A. Case, Solution structure of carbonmonoxy myoglobin determined from nuclear magnetic resonance distance and chemical shift constraints, J. Mol. Biol. 244 (1994) 183–197.
[86] S.-J. Hyung, C.V. Robinson, B.T. Ruotolo, Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes, Chem. Biol. 16 (2009) 382–390.
[87] E.R. Badman, S. Myung, D.E. Clemmer, Evidence for unfolding and refolding of gas-phase cytochrome c ions in a Paul trap, J. Am. Soc. Mass Spectrom. 16 (2005) 1493–1497.
[88] 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, J. Am. Soc. Mass Spectrom. 15 (2004) 870–878.
[89] H.A. Sawyer, J.T. Marini, E.G. Stone, B.T. Ruotolo, K.J. Gillig, D.H. Russell, The structure of gas-phase bradykinin fragment 1-5 (RPPGF) ions: an ion mobility spectrometry and H/D exchange ion-molecule reaction chemistry study, J. Am. Soc. Mass Spectrom. 16 (2005) 893–905.
[90] F.A. Fernandez-Lima, C. Becker, K. Gillig, W.K. Russell, M.A.C. Nascimento, D.H. Russell, Experimental and theoretical studies of (CsI)nCs+ cluster ions produced by 355 nm laser desorption ionization, J. Phys. Chem. A 112 (2008) 11061–11066.
[91] S.I. O'Donoghue, X. Chang, R. Abseher, M. Nilges, J.J. Led, Unraveling the symmetry ambiguity in a hexamer: calculation of the R6 human insulin structure, J. Biomol. NMR 16 (2000) 93–108.
[92] X. Chang, A.M. Jorgensen, P. Bardrum, J.J. Led, Solution structures of the R6 human insulin hexamer, Biochemistry 36 (1997) 9409–9422.
[93] G.D. Smith, W.A. Pangborn, R.H. Blessing, The structure of T6 bovine insulin, Acta Crystallogr. D 61 (2005) 1476–1482.