

Cis–Trans Signatures of Proline-Containing Tryptic Peptides in the Gas Phase

Anne E. Counterman and David E. Clemmer*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

High-resolution ion mobility/time-of-flight techniques were used to measure collision cross sections for 968 tryptic digest peptide ions obtained from digestion of common proteins. Here, we report a mobility signature that aids in identifying proline-containing peptides containing 4–10 residues. Of 129 peptides (≤ 10 residues in length) in the database that contain proline residues, 57% show multiple resolved features in the ion mobility distribution for at least one of the $[M + H]^+$ or $[M + 2H]^{2+}$ ions. These multiple features are attributed to different conformations that arise from populations of cis and trans forms of proline. The number of resolved peaks in the ion mobility distribution appears to be correlated with the peptide ion charge state and the number of proline residues in the peptide.

Of the 20 common residues found in proteins, proline is unique because it contains a secondary amine; its side chain forms a five-membered ring with the backbone nitrogen and α -carbon. The rigidity of the peptide bond leads to cis and trans forms in 94% and 6% abundance, respectively, for peptides in solution.¹ Here, we present evidence that cis and trans forms of peptides also exist in gas-phase ions produced by electrospray ionization.² The structural differences that are induced in the peptides are significant and allow the different forms to often be separated on the basis of differences in their mobilities through a buffer gas. Peptides that do not contain proline do not show this behavior. The ability to rapidly identify those mass-to-charge (m/z) ratios that may contain a proline residue based on an initial mobility signature is a useful complement to assignments based on measurements of m/z ratios.

Currently, little is known about the conformation of proline residues in gas-phase peptides. Comparisons of semiempirical molecular modeling simulations of the singly protonated dipeptides Gly–Pro, Pro–Pro, and Gly–Gly suggest that the most favorable protonation site is the amino terminus.³ In these cases, the C-terminal Pro residue adopts the cis form—favored due to hydrogen bonding between the protonated amino terminus and the carbonyl oxygen of the C-terminal carboxylic acid group. Several studies of fragmentation patterns of proline-containing proteins and peptides have reported the cleavage of the peptide

bond on the N-terminal side of proline residues,⁴ and an extensive database of peptide fragmentation data has been used to develop a relative scale for the fragmentation efficiency of the Xxx–Pro peptide bond.⁵

The work presented in this paper arises from observations from a database of high-resolution ion mobility measurements of tryptic digest peptides.⁶ As the database was examined, we noted that some ions showed multiple resolved features across the ion mobility dimension. Out of distributions measured for 645 small $[M + H]^+$ and $[M + 2H]^{2+}$ tryptic peptide ions (≤ 10 residues), all of the peptides that exhibited peaks at multiple drift times were found to contain proline. No other amino acids were common to all sequences. The appearance of multiple features *only* for sequences containing proline suggests that the separation reflects differences in peptide conformations that arise from the presence of both the cis and trans isomer forms. Molecular modeling simulations have been used to corroborate this idea. The ability to resolve different conformations for a given sequence appears to be related to the number (and presumably positions) of prolines present in the sequence and the charge state of the ion.

EXPERIMENTAL SECTION

Ion Mobility/Time-of-Flight Measurements. Experiments were performed using a hybrid ion mobility/time-of-flight mass spectrometer.⁷ Discussions of tryptic digest preparation⁶ and the general experimental approach are given elsewhere.⁸ The 49 proteins used for digestion (obtained from Sigma Chemical and used without further purification) are as follows: albumin (bovine, chicken, dog, horse, human, pig, sheep), α -amylase (*Bacillus stearothermophilus*), α -casein (bovine), α -chymotrypsin (bovine), α -lactalbumin (bovine), acetyl cholinesterase (bovine), acylase I (pig), aldolase (rabbit), apotransferrin (bovine), aphyrase (potato), alcohol dehydrogenase (horse, yeast), β -casein (bovine), β -lactoglobulin (bovine), carbonic anhydrase (bovine), catalase (bovine),

(1) Stryer, L. *Biochemistry*, 4th ed.; W. H. Freeman and Co.: New York, 1995.
(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64.
(3) Ewing, N. P.; Zhang, X.; Cassady, C. J. *J. Mass Spectrom.* **1996**, *31*, 1345.

(4) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *83*, 6233. Martin, S. A.; Biemann, K. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 213. Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Anal. Chem.* **1993**, *65*, 425.
(5) Brei, L.; Wysocki, V. Poster presented at the 2001 ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, IL.
(6) Counterman, A. E.; Zientara, G. A.; Herbst, S. A.; Clemmer, D. E., manuscript in preparation.
(7) Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998**, *70*, 2236. Henderson, S. C.; Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **1999**, *71*, 291.
(8) For recent reviews, see: Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577. Hoaglund Hyzer, C. S.; Counterman, A. E.; Clemmer, D. E. *Chem. Rev.* **1999**, *99*, 3037.

conalbumin (chicken), concanavalin A (jack bean), creatine phosphokinase (rabbit), cytochrome *c* (horse), enolase (yeast), glucose oxidase (*Aspergillus niger*), L-glutamic dehydrogenase (bovine), hemoglobin (bovine, human, pig, rabbit, sheep), insulin (bovine), invertase (yeast), κ -casein (bovine), lactotransferrin (bovine), lysozyme (chicken), myoglobin (horse), papain (papaya), peroxidase (horseradish), L- α -phosphatidylcholine (chicken), rhodanese (bovine), ribonuclease A (bovine), transferrin (human), trypsinogen (bovine), ubiquitin (bovine), and urease (jack bean). Samples of GPGG and human β -casomorphin (YPFVEPI) were obtained from Sigma Chemical (both at 97% purity) and used without further purification. Peptide solutions (prepared as $\sim 5 \times 10^{-4}$ – 5×10^{-3} M total peptide in 49:49:2 water/acetonitrile/acetic acid) were electrosprayed² into a differentially pumped desolvation region, and short pulses (200 μ s) were introduced into a helium-filled drift region (~ 150 Torr, 28 °C) for mobility experiments. As the packet of ions traverses the 0.5849-m drift region under the influence of a weak, uniform electric field (171.0 V·cm⁻¹), different species are separated due to differences in mobilities. Compact ions (having small collision cross sections)⁹ have higher mobilities than ions having more extended conformations.¹⁰ Additionally, high-charge-state ions have higher mobilities than low-charge-state ions.¹¹ As ions exit the drift tube, they enter an orthogonal geometry reflectron time-of-flight mass spectrometer. With this configuration, mobility and flight time distributions can be recorded in a nested fashion, discussed previously.⁷ Sequence assignments are based on comparison of m/z ratios determined from time-of-flight data recorded in the mass spectrometer⁷ with monoisotopic values that are calculated for expected peptides. The m/z ratios are obtained using the PeptideMass tool available through ExPASy.¹² Data are often presented as two-dimensional plots of drift and flight times (or m/z ratios). It is also useful to plot the data on a drift time \times charge state (z) scale. This provides a feeling about relative differences in the overall cross section of different charge states.

Theoretical Peak Shapes. Comparisons of calculated distributions to experimental peaks provide information about the number of different conformations that are present. The calculated distribution for transport of a single geometric form through the drift tube is given by⁹

$$\Phi(t) = \int \frac{C}{(Dt)^{1/2}} (v_D + L/t) \left[1 - \exp\left(\frac{-r_0^2}{4Dt}\right) \right] \times \exp\left(\frac{-(L - v_D t)^2}{4Dt}\right) P(t_p) dt_p \quad (1)$$

where $\Phi(t)$ is the intensity of ions passing through the exit aperture as a function of time, D is the diffusion constant, v_D is the measured drift velocity, r_0 is the radius of the entrance

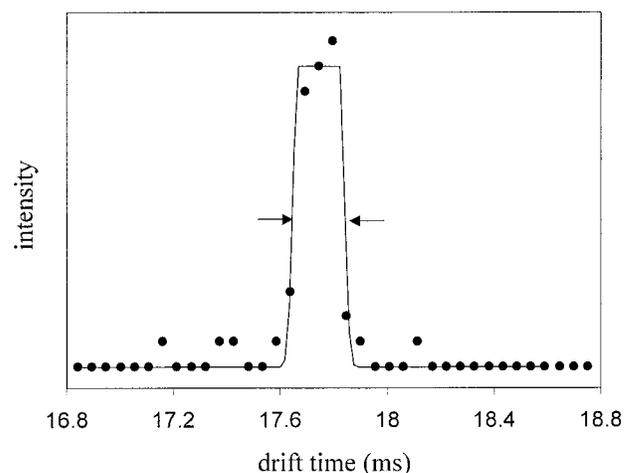


Figure 1. Calculated distribution (dashed line) obtained from eq 1 superimposed on a plot of the experimental drift time distribution (solid line and open circles) for [KIEK + H]⁺ (measured from a tryptic digest of bovine hemoglobin). The arrows indicate the full width at half-maximum intensity of the distribution.

aperture, $P(t_p) dt_p$ is the distribution function for the pulse of ions entering the drift tube, and C is a constant.¹³ Figure 1 shows a comparison of the expected distribution with the experimentally measured distribution for [KIEK + H]⁺, a non-proline-containing peptide from our database. In this case, the width of the calculated distribution is dominated by the duration of the input pulse. There is close agreement between the calculated and experimental peak shapes. This suggests that only a single conformer type is present. However, it is also possible that multiple conformations having identical mobilities are present. Experimental peaks that are significantly broader than the calculated distribution indicate either that multiple conformations exist or that structural changes occur as ions travel through the buffer gas. Calculated peak shapes can be used as an indication of the minimum number of states that could lead to a given experimental distribution.

An important consideration in the analysis presented below is the identification of distributions that must contain multiple structures. For the purposes of this analysis, we require that (1) multiple peaks are clearly resolved or (2) that the overall peak width is at least a factor of 1.8 greater than the peak width of the calculated distribution for single conformer transport and that the peak shape is irregular (e.g., there is evidence for a shoulder). The latter requirement—that peak widths be at least a factor of 1.8 greater than the width of a calculated distribution—ensures that the determination of the presence of an additional conformer is confirmed by ion intensity recorded in at least three adjacent acquisition windows in the mobility dimension.

Molecular Modeling Simulations. Simulations of [GPGG + H]⁺ were performed to test whether significant differences in collision cross sections arise between peptide conformations in which proline residues are cis or trans. Calculations were performed using a simulated annealing protocol (similar to that described previously)¹⁴ and the AMBER force field in the InsightII software package.¹⁵ Briefly, the simulated annealing method

(9) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York, 1988.

(10) von Helden, G.; Hsu, M.-T.; Kemper, P. R.; Bowers, M. T. *J. Chem. Phys.* **1991**, *95*, 3835. Jarrold, M. F.; Constant, V. A. *Phys. Rev. Lett.* **1991**, *67*, 2994. Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1995**, *117*, 10141.

(11) Valentine, S. J.; Counterman, A. E.; Hoaglund, C. S.; Reilly, J. P.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1213.

(12) Wilkins, M. R.; Lindskog, I.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Hochstrasser, D. F.; Appel, R. D. *Electrophoresis* **1997**, *18*, 403.

(13) Under low-field conditions, the diffusion constant is related to the measured mobility by the expression $D = Kk_b T/ze$, where K is the mobility, k_b is Boltzmann's constant, and ze is the charge.

(14) Counterman, A. E.; Clemmer, D. E. *J. Am. Chem. Soc.* **1999**, *121*, 4031.

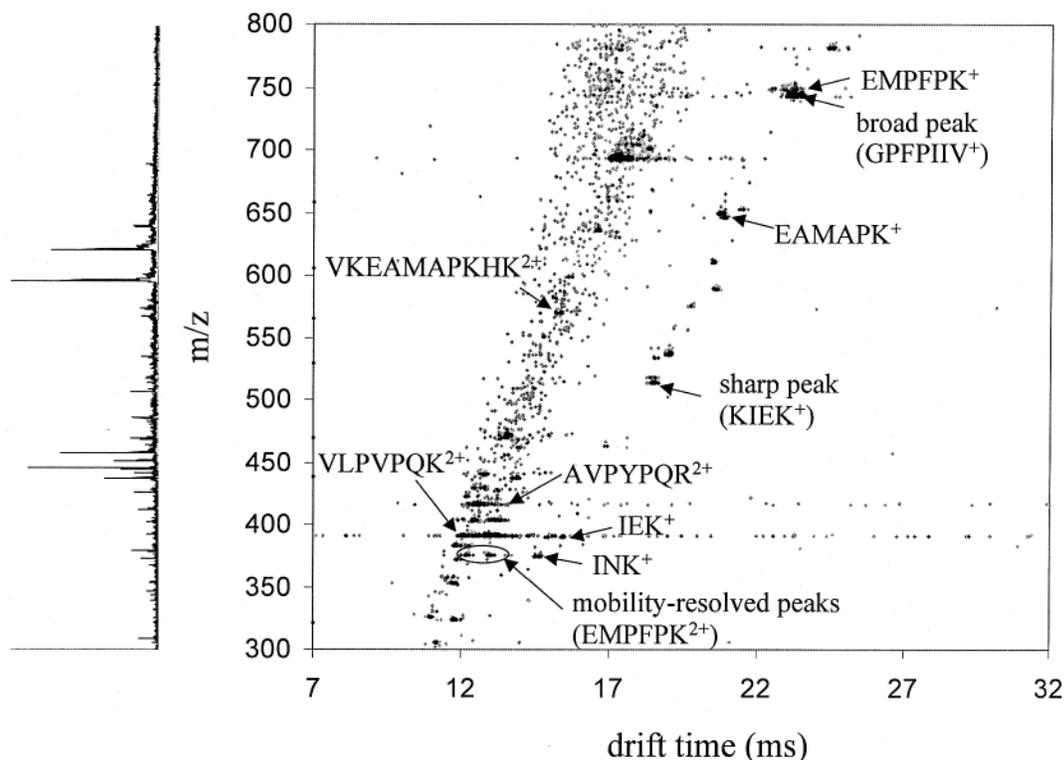


Figure 2. Ion mobility/time-of-flight data set for a mixture of tryptic peptides from digestion of β -casein. Data were recorded using a helium buffer gas pressure of 154.60 Torr and an electric field strength of $171.0 \text{ V}\cdot\text{cm}^{-1}$. Data points are represented by using gray scale to indicate different intensity cutoffs as follows: ≥ 2 counts, light gray; ≥ 4 counts, medium gray; and ≥ 5 counts, black.

involves repeatedly heating initial extended conformations to 1000 K over 1 ps, equilibrating at 1000 K for 2 ps, and then cooling to 300 K over 1 ps; 100 trial geometries are generated from each run. All trial conformers were protonated at the N-terminus.³ Collision cross sections were calculated from trial geometries using the exact hard spheres scattering (EHSS) method,¹⁶ which has been calibrated to values from the trajectory approach¹⁷ as described in an earlier work.¹⁴ Agreement of calculated cross sections to within 2% of experimental values is required for conformers to be considered as viable candidate structures.

RESULTS AND DISCUSSION

Typical Ion Mobility/Time-of-Flight Data for Proline-Containing Tryptic Peptides. Figure 2 shows an ion mobility/time-of-flight data set for a tryptic digest of β -casein, a milk protein that contains 35 proline residues in the 224-residue sequence. These data show similarities to distributions that were recorded for 48 other tryptic digest mixtures.⁶ Of the 10 ions assigned in the region shown, 7 sequences contain at least one proline residue. Inspection of the data shows three types of peaks in the mobility distribution associated with individual flight times: (1) sharp peaks, such as the one found at $m/z = 517.35$ for $[\text{KIEK} + \text{H}]^+$ (calculated $m/z = 517.33$); (2) broad features [e.g., the distribution at $m/z = 742.34$, assigned to $[\text{GPFPIIV} + \text{H}]^+$ (calculated $m/z = 742.45$); and (3) multiple resolved peaks such as those found at $m/z = 374.66$ (assigned to $[\text{EMPFPK} + 2\text{H}]^{2+}$, calculated $m/z =$

374.69). In regions containing broad features or multiple resolved peaks, the m/z values correspond to peaks that can be assigned to proline-containing sequences—for instance: $[\text{GPFPIIV} + \text{H}]^+$ [$m/z(\text{expt}) = 742.34$, $m/z(\text{calc}) = 742.34$]; $[\text{EMPFPK} + 2\text{H}]^{2+}$ [$m/z(\text{expt}) = 374.66$, $m/z(\text{calc}) = 374.69$]; $[\text{VLPVPQK} + 2\text{H}]^{2+}$ [$m/z(\text{expt}) = 390.74$, $m/z(\text{calc}) = 390.75$]; and $[\text{AVPYPQR} + 2\text{H}]^{2+}$ [$m/z(\text{expt}) = 415.79$, $m/z(\text{calc}) = 415.73$].

Additional insight can be gained from examination of drift time distributions (obtained by integrating the two-dimensional data set over a flight time range that corresponds to a single m/z ion). The distributions shown in Figure 3 show typical distributions for non-proline-containing $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ peptides that were identified in the β -casein tryptic digest and other data sets. The low-intensity, high-mobility features (present at short drift times) in the $[\text{KIEK} + \text{H}]^+$ slice, and other $[\text{M} + \text{H}]^+$ distributions shown in Figure 3, are associated with chemical noise from the $[\text{M} + 2\text{H}]^{2+}$ family of peaks (and higher charge states) in the two-dimensional data set¹⁸ and are not relevant to the discussion of monomer $[\text{M} + \text{H}]^+$ ions presented here. In general, distributions for both $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ non-proline-containing ions show single, sharp peaks having peak shapes that can be represented by a calculated distribution for the transport of a single conformer (eq 1).⁹ Examination of distributions for 20 additional randomly selected non-proline-containing $[\text{M} + \text{H}]^+$ peptides shows that peak widths are essentially the same as those for distributions calculated for transport of a single conformer (eq 1)—an average factor of 1.01 ± 0.01 greater. Previously reported

(15) InsightII, Biosym/MSI, San Diego, CA, 1995.

(16) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86.

(17) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082.

(18) Counterman, A. E.; Hilderbrand, A. E.; Srebalus Barnes, C. A.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1020.

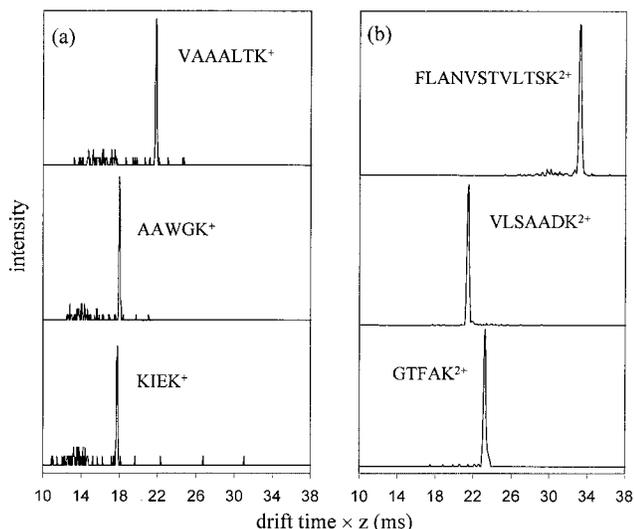


Figure 3. Drift time distributions for non-proline-containing (a) $[M + H]^+$ ions of KIEK (from a tryptic digest of β -casein), AAWGK (bovine hemoglobin), and VAAALTK (bovine hemoglobin); and (b) $[M + 2H]^{2+}$ ions of GTFK (rabbit hemoglobin), VLSAADK (bovine hemoglobin), and FLANVSTVLTSK (bovine hemoglobin). Data are normalized to a helium buffer gas pressure of 150.00 Torr and an electric field strength of $171.0 \text{ V}\cdot\text{cm}^{-1}$.

molecular modeling simulations indicate that $[M + H]^+$ tryptic peptides (5–10 residues in length) generally favor compact globular conformations in which the protonated side chain of the C-terminal Lys residue is solvated by hydrogen-bonding interactions with several backbone carbonyl groups and polar side chains.¹⁴ Similar results have been found for peptides (<10 residues in length) in which an Arg residue at the C-terminus is protonated.^{19,20}

Figure 4 shows distributions for 10 proline-containing $[M + H]^+$ and $[M + 2H]^{2+}$ sequences. Each of these distributions contains multiple features. For example, two peaks are clearly resolved for $[\text{NTPEK} + \text{H}]^+$, at 18.48 and 18.90 ms; $[\text{EAMAPK} + \text{H}]^+$, at 20.08 and 20.40 ms; $[\text{YIPGTK} + \text{H}]^+$, at 20.57 and 21.10 ms; and $[\text{VLPVPQK} + \text{H}]^+$, at 23.53 and 24.06 ms (although we note that the peaks at 18.90, 20.40, 20.57, and 24.06 ms, respectively, are small). There is no evidence for interconversion between these states on the time scale of the experiment. The distribution observed for $[\text{GPFPIIV} + \text{H}]^+$ (which does not contain a C-terminal Lys or Arg residue because it occurs at the C-terminal end of the β -casein sequence) shows a broad peak centered at 22.26 ms, with an unresolved shoulder at 22.52 ms, indicating that at least two types of conformers are present; these structures are not resolved under the present experimental conditions. Two peaks are resolved in four of the $[M + 2H]^{2+}$ distributions: for $[\text{EAMAPK} + 2\text{H}]^{2+}$ at 21.48 and 22.33 ms; $[\text{EMPFK} + 2\text{H}]^{2+}$, at 23.18 and 24.77 ms; $[\text{LRVDPVNFK} + 2\text{H}]^{2+}$, at 29.22 and 30.37 ms; and $[\text{LLVVPWTQR} + 2\text{H}]^{2+}$ at 32.68 and 33.41 ms. The $[\text{VLPVPQK} + 2\text{H}]^{2+}$ distribution shows two broad features centered at 22.86 and 24.77 ms. The peak shapes in this distribution suggest that at least four conformer types may be present.

(19) Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355.

(20) Counterman, A. E.; Clemmer, D. E., unpublished results.

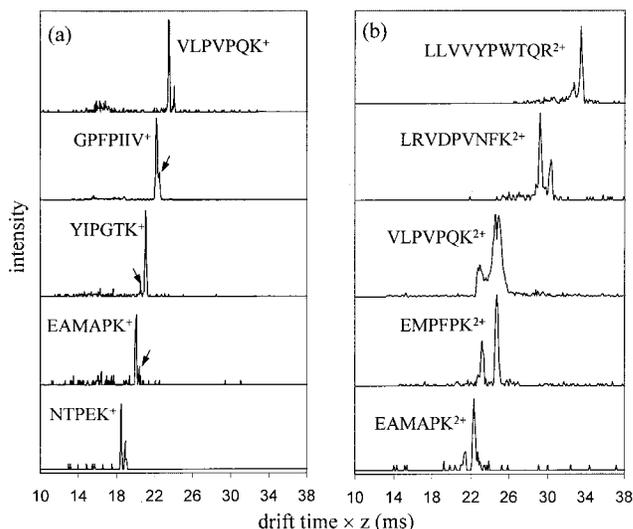


Figure 4. Drift time distributions for proline-containing (a) $[M + H]^+$ ions of LTPEK (from a tryptic digest of bovine apotransferrin), EAMAPK (bovine β -casein), YIPGTK (equine cytochrome *c*), GPFPIIV (bovine β -casein), and VLPVPQK (bovine β -casein) sequences; and (b) $[M + 2H]^{2+}$ ions of EAMAPK (bovine β -casein), EMPFK (bovine β -casein), VLPVPQK (bovine β -casein), LRVDPVNFK (bovine hemoglobin), and LLVVPWTQR (bovine hemoglobin). Arrows are used to show the location of secondary peaks in the $[M + H]^+$ distribution. Data are normalized to a helium buffer gas pressure of 150.00 Torr and an electric field strength of $171.0 \text{ V}\cdot\text{cm}^{-1}$.

General Aspects of Complete Database. Inspection of the distributions in the current database (containing entries for 968 ions) shows that multiple peaks appear to be more common for small peptides (those containing 10 or fewer residues). There are 645 $[M + H]^+$ and $[M + 2H]^{2+}$ ions that contain 10 or fewer residues. Of these, multiple peaks are only observed for m/z ratios that correspond to peptides containing proline residues. No other amino acid residue is common to all sequences for which multiple peaks are observed. A tabulation of data for 129 proline-containing peptides (containing 10 or fewer residues) yields 73 (~57%) that exhibit clear evidence of multiple peaks. This includes 14 of 33 $[M + H]^+$ (42%) and 59 of 96 $[M + 2H]^{2+}$ (61%) sequences.

Investigation of the Source of Conformational Differences That Arise for Proline-Containing Peptides. The apparent uniqueness of the multiple-peak signature to proline-containing peptides suggests that the separation results from different conformations that arise from the presence of populations of *cis* and *trans* forms of proline. To examine this possibility in more detail, we have recorded data for simple model proline-containing peptides. Figure 5 shows a drift time distribution recorded for $[\text{GPGG} + \text{H}]^+$, a tetrapeptide that has been examined previously by capillary electrophoresis.²¹ Peaks at 10.50 and 10.86 ms are observed. Molecular modeling simulations, initiated from extended structures in which the proline residue was fixed in the *cis* or *trans* configuration, show that the proline orientation influences the resulting peptide conformation and collision cross section. Figure 6 shows cross sections and energies that were calculated from molecular modeling simulations for trial conformers of $[\text{GPGG} + \text{H}]^+$ containing *cis*- and *trans*-Pro configurations. The simulations show that *trans*-Pro favors ringlike geometries that

(21) Moore, A. W., Jr.; Jorgenson, J. W. *Anal. Chem.* **1995**, *67*, 3464.

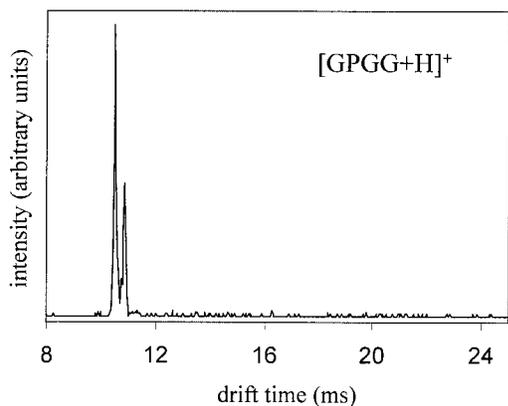


Figure 5. Drift time distribution for $[\text{GPGG} + \text{H}]^+$. Data are normalized to a helium buffer gas pressure of 150.00 Torr and an electric field strength of $171.0 \text{ V}\cdot\text{cm}^{-1}$.

are stabilized by formation of a hydrogen bond between N- and C-terminal groups. The *cis*-Pro $[\text{GPGG} + \text{H}]^+$ ions favor relatively planar conformations, in which the protonated N-terminal amino group forms hydrogen bonds with the C-terminus and two backbone carbonyl groups. The calculated collision cross sections for the conformers that are shown are 94.5 \AA^2 for *trans*-Pro and 98.2 \AA^2 for *cis*-Pro forms, near the experimental values of 93.2 and 96.4 \AA^2 , respectively. From this, we assign the peak at 10.50 ms to the *trans*-Pro geometry and the peak at 10.86 ms to *cis*-Pro geometry.

Other model peptides yield similar results. Overall, the molecular modeling results for model systems support the general

observations made for tryptic peptides. That is, a peptide structure that arises from the presence of *cis*-proline can have a collision cross section very different from that of a peptide structure that contains proline in the *trans* conformation. We note that while most proline-containing peptides appear to show evidence for two (or more) mobility-resolved peaks, many sequences show only a single peak. In these sequences, it is possible that only one proline conformation is favored or that the isomerization has a minimal effect on the collision cross section (e.g., the different structures are not resolved).

Influence of Charge State and Number of Prolines on Features Observed. Comparison of distributions for $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ ions formed from the same sequences provides insight about how the protonation state influences these systems. Figure 4 displays two peptides, EAMAPK and VLPVPQK, that show both $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ ions. We also show distributions recorded for $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ ions of YPVEPI as an additional example in Figure 7. The $[\text{YPVEPI} + \text{H}]^+$ distribution shows two peaks, centered at 24.61 and 25.09 ms. However, four peaks are clearly resolved for $[\text{YPVEPI} + 2\text{H}]^{2+}$ distribution (at 24.79 , 25.38 , 25.86 , and 27.32 ms). Drift times (normalized for charge state) obtained for peak centers in the $[\text{YPVEPI} + 2\text{H}]^{2+}$ distribution are also longer than drift times observed for $[\text{YPVEPI} + \text{H}]^+$, indicating that conformations of the $[\text{YPVEPI} + 2\text{H}]^{2+}$ ions are more extended. Examination of other sequences for which both $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ ions are observed shows that, in many cases, more peaks are resolved in the $[\text{M} + 2\text{H}]^{2+}$ distribution (compared with the $[\text{M} + \text{H}]^+$

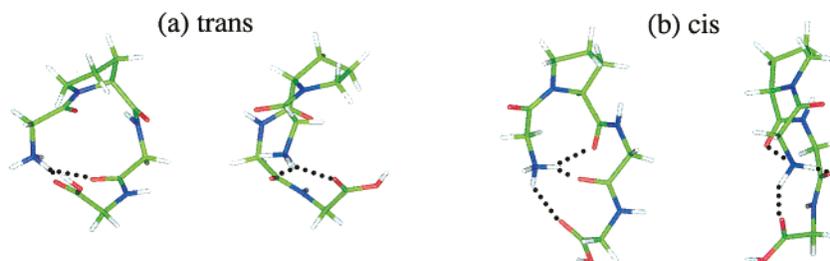
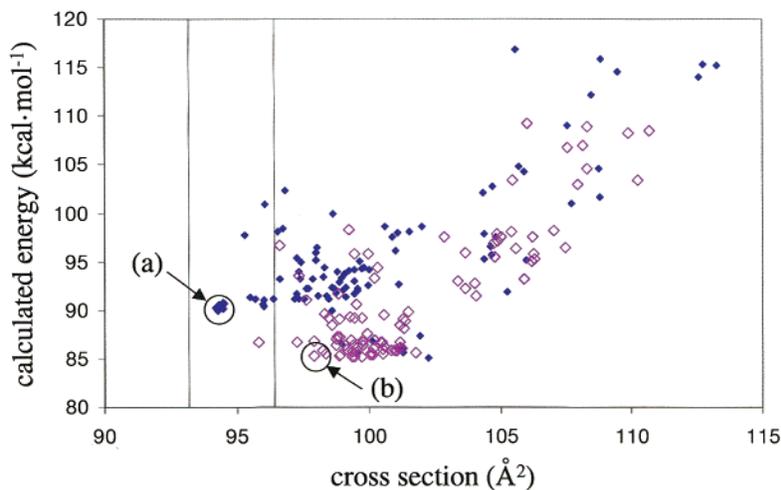


Figure 6. Calculated energies and cross sections from molecular modeling simulations for *cis* (open diamonds) and *trans* (solid diamonds) conformers of $[\text{GPGG} + \text{H}]^+$ (100 conformers from a typical simulated annealing run are shown for each). Conformers a and b are representative geometries obtained from *trans*-Pro and *cis*-Pro $[\text{GPGG} + \text{H}]^+$ simulations, respectively. Energies and calculated cross sections for these ions are indicated by circles and arrows. Solid lines indicate the experimental cross sections (ref 8) determined from positions of the two resolved peaks shown in Figure 5.

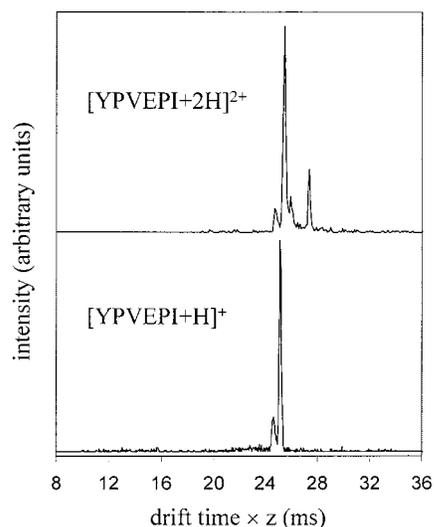


Figure 7. Drift time distributions for $[\text{YPVEPI} + \text{H}]^+$ and $[\text{YPVEPI} + 2\text{H}]^{2+}$. Data are normalized to a helium buffer gas pressure of 150.00 Torr and an electric field strength of $171.0 \text{ V}\cdot\text{cm}^{-1}$.

distribution). It is possible that, in doubly charged ions, protonation or proton bridging involving the proline amide weakens the double bond character of the peptide bond, effectively lowering the rotation barrier associated with the cis \leftrightarrow trans isomerization. Weakening of the peptide bond upon protonation is consistent with quantum mechanical calculations presented by Wysocki and co-workers in support of the "mobile proton" model for peptide dissociation.²² It is also likely in the current study that the propensity for observation of more peaks in $[\text{M} + 2\text{H}]^{2+}$ data is influenced by the repulsive interactions between the protonated sites.

An additional correlation is related to the number of prolines present in the sequence. If a single proline is present, no more than two peaks are observed (for example, $[\text{EAMAPK} + 2\text{H}]^{2+}$ in Figure 4); if two prolines are present, up to four features are observed (for instance, $[\text{VLPVPQK} + 2\text{H}]^{2+}$ in Figure 4 and $[\text{YPVEPI} + 2\text{H}]^{2+}$ in Figure 7). This is consistent with the number of possible combinations of cis and trans isomers: one proline can exist either as cis or trans; two prolines may be present as cis–cis, cis–trans, trans–cis, or trans–trans.

Influence of Xxx Identity on Cis/Trans Orientation of the Xxx–Pro Bond. Isomerization of proline has been implicated as a rate-limiting step in protein folding in solution,²³ and studies have demonstrated that the identity of the residue that is adjacent to proline (on the N-terminal side) influences the kinetics of cis \leftrightarrow trans isomerization.^{24,25,26} It is interesting to speculate whether the identity of the Xxx residue (where Xxx corresponds to the

adjacent amino acid on the N-terminal side of the Pro residue) is correlated with the observation of multiple isomeric peptide conformations. To examine this, we considered several different side-chain properties that might influence the cis \leftrightarrow trans isomerization, including the side-chain size (i.e., steric effects) as well as the polarity (a factor that we have found to be important in other studies,²⁷ using 42 $[\text{M} + 2\text{H}]^{2+}$ sequences that contain a single proline residue. There is no significant correlation with size. However, of the 18 peptide sequences examined in which Xxx is nonpolar (Gly, Ala, Val, Ile, Leu, or Phe), only four do not exhibit two peaks. Conversely, 18 of 24 peptides containing a polar residue (Asp, Glu, Ser, Thr, Asn, Gln, Tyr, Trp, or His) at the Xxx position show a single peak. While these results rely on only a few sequences, it appears that there is a preference for polar Xxx residues to exhibit a single peak (consistent with a single conformer type). Finally, we note that for proline-containing peptides that exhibit only a single peak, it is currently unclear whether cis or trans forms are more common.

SUMMARY AND CONCLUSIONS

A database of high-resolution ion mobility distributions for $[\text{M} + \text{H}]^+$ and $[\text{M} + 2\text{H}]^{2+}$ tryptic peptides has been measured and examined. A trend that is apparent involves proline-containing peptides. In many cases, these peptides exhibit multiple peaks in the ion mobility distribution. Examination of these data (as well as data and molecular modeling results for model peptides) indicates that the multiple conformations that are resolved in proline-containing peptides arise from the unique ability of proline to sample cis and trans isomers. For proline-containing tryptic peptides that contain 10 or fewer residues, $\sim 42\%$ of $[\text{M} + \text{H}]^+$ and 61% of $[\text{M} + 2\text{H}]^{2+}$ show evidence for multiple conformers. Here, we have proposed that the observation of multiple peaks in the ion mobility distribution could be used as a signature for those peptides containing prolines. Clearly, the observation of multiple peaks does not require that a proline residue is present, as other isomer forms of peptides can be separated by ion mobility methods.²⁸ However, the use of this approach as a guide in combination with other information about a tryptic mixture (such as m/z values from other peaks), or information from collision-induced dissociation, would seem valuable.

ACKNOWLEDGMENT

The authors gratefully acknowledge the assistance of Gina Zientara and Samantha Herbst in recording the database used for these studies. Financial support for this work comes from the National Institutes of Health (Grant 1R01GM59145) and National Science Foundation (CHE-0078737).

Received for review October 10, 2001. Accepted February 5, 2002.

AC011083K

(22) McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H. *Anal. Chem.* **1993**, *65*, 2859. Somogyi, A.; Wysocki, V. H.; Maier, I. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704.

(23) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953.

(24) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623.

(25) Thuncke, F.; Kalman, A.; Kalman, F.; Ma, S.; Rathore, A. S.; Horvath, C. *J. Chromatogr., A* **1996**, *744*, 259.

(26) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol. Biol.* **1998**, *279*, 449.

(27) Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. *J. Phys. Chem. B* **1999**, *103*, 1203.

(28) Srebalus, C. A.; Li, J.; Marshall, W. S.; Clemmer, D. E. *Anal. Chem.* **1999**, *71*, 3918. Srebalus, C. A.; Li, J.; Marshall, W. S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 352.