analytical chemistry

Article

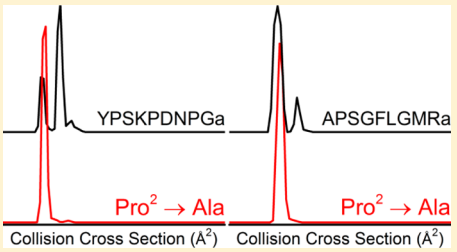
pubs.acs.org/ac

Penultimate Proline in Neuropeptides

Matthew S. Glover,[†] Earl P. Bellinger,[‡] Predrag Radivojac,[‡] and David E. Clemmer*[,†]

[†]Department of Chemistry, [‡]Department of Computer Science and Informatics, Indiana University, Bloomington, Indiana 47405, United States

ABSTRACT: A recent ion mobility spectrometry-mass spectrometry (IMS-MS) study revealed that tryptic peptide ions containing a proline residue at the second position from the N-terminus (i.e., penultimate proline) frequently adopt multiple conformations, owing to the *cis*−*trans* isomerization of Xaa[1]−Pro[2] peptide bonds [*J. Am. Soc. Mass Spectrom.* **2015**, *26*, 444]. Here, we present a statistical analysis of a neuropeptide database that illustrates penultimate proline residues are frequently found in neuropeptides. In order to probe the effect of penultimate proline on neuropeptide conformations, IMS-MS experiments were performed on two model peptides in which penultimate proline residues were known to be important for biological activity: the N-terminal region of human neuropeptide Y (NPY_{1-9}, Tyr[1]-Pro[2]-Ser[3]-Lys[4]-Pro[5]-Asp[6]-Asn[7]-Pro[8]-Gly[9]-NH_2) and a tachykinin-related peptide (CabTRP Ia, Ala[1]-Pro[2]-Ser[3]-Gly[4]-Phe[5]-Leu[6]-Gly[7]-Met[8]-Arg[9]-NH_2). From these studies, it appears that penultimate prolines allow neuropeptides to populate multiple conformations arising from the *cis*−*trans* isomerization of Xaa[1]−Pro[2] peptide bonds. Although it is commonly proposed that the role of penultimate proline residues is to protect peptides from enzymatic degradation, the present results indicate that penultimate proline residues also are an important means of increasing the conformational heterogeneity of neuropeptides.



Although the vast majority of peptide bonds adopt the *trans* configuration in polypeptides, Xaa−Pro bonds adopt both *cis* and *trans* isomers due to the unique structure of the proline side chain.[1−3] Several studies have demonstrated that proline isomerization acts as a conformational switch in which the *cis*−*trans* isomerization of a single Xaa−Pro bond leads to global conformational changes, with the two forms having different functions.[4−8] Proline residues and proline-rich regions of sequences are frequently found in dynamic regions of biomolecules, giving them a high degree of conformational heterogeneity.[5,9] Such systems often do not adopt rigid secondary structures and are commonly described as unstructured or intrinsically disordered.[9,10] It remains challenging to characterize the structure of dynamic proline-containing regions of biomolecules by traditional techniques such as nuclear magnetic resonance spectroscopy and X-ray crystallography, presumably because so many stable conformations are present simultaneously.[11,12]

Several recent studies illustrate that ion mobility spectrometry-mass spectrometry (IMS-MS) provides structural insight into dynamic systems that populate multiple conformations, including proline-containing peptides.[12−17] IMS separates ions based on differences in their shapes and charge states, making it an effective means of delineating differences in conformation.[18,19] When coupled with electrospray ionization (ESI), it is possible to monitor populations of conformations emerging from solution.[12,20,21] The evaporative cooling process of ESI is thought to aid in kinetically trapping solution populations.[22−24] In a recent study, we reported that peptides with a proline residue located at the second position from the N-terminus (i.e., penultimate proline) have a propensity to adopt multiple conformations from the *cis*−*trans* isomerization of Xaa[1]−Pro[2] peptide bonds.[17] We subsequently noticed that penultimate proline residues are frequently located in many bioactive peptides, prompting us to inspect the structural implications of penultimate proline residues.

In this study, the role of penultimate proline residues in neuropeptide conformations is examined, as we found this class of neurochemical signaling molecules to frequently contain a proline residue in the N-terminal penultimate position. A statistical analysis of NeuroPedia,[25] an online database of neuropeptide sequences, is performed to illustrate the frequency of proline position in neuropeptide sequences. We use IMS-MS studies to investigate two model neuropeptide systems, the N-terminal region of neuropeptide Y (NPY_{1-9}, Tyr[1]-Pro[2]-Ser[3]-Lys[4]-Pro[5]-Asp[6]-Asn[7]-Pro[8]-Gly[9]-NH_2) and a tachykinin-related peptide from *Cancer borealis* (CabTRP Ia, Ala[1]-Pro[2]-Ser[3]-Gly[4]-Phe[5]-Leu[6]-Gly[7]-Met[8]-Arg[9]-NH_2), in detail. We selected these two systems because they both have penultimate proline residues that are proposed to be important for their biological activity.[26−28]

Neuropeptide Y is one of the most abundant peptides in both the central and peripheral nervous systems and is involved in regulating a variety of biological processes such as food intake and blood pressure.[29−34] Although the structure is debated, NMR and circular dichroism studies propose that the proline-rich N-terminal region lacks rigid secondary structure and has been described as unstructured.[30−34] Several studies provide conflicting reports as to which specific Xaa−Pro peptide bonds in the N-terminal region adopt both *cis* and *trans*

Received: May 20, 2015
Accepted: July 20, 2015
Published: July 20, 2015

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472

isomers.[30−34] A study of Ala substituted analogues of NPY suggested that the Pro^2 residue is important for binding the Y_1 receptor.[26] Furthermore, neuropeptide Y 3−36 is an endogenous peptide that is formed by cleaving Tyr^1−Pro^2 residues from the N-terminus and has a decreased affinity for an Y_1-like receptor, supporting the idea that Pro^2 is important for receptor-binding selectivity.[27] In addition to serving as a model for penultimate proline isomerization, this system is chosen because it will provide insight into the type of structural information that can be gleaned from ion mobility studies of peptides that have been described as unstructured by traditional techniques.[12,35]

Tachykinin-related peptides (TRPs) are a large family of neuropeptides with five conserved C-terminal residues: Phe-Xaa_1-Gly-Xaa_2-Arg-NH_2.[36] TRPs are involved in a range of physiological processes in the nervous systems of invertebrates, such as excitation of the pyloric rhythm.[28,36] We are interested in this family of peptides because the majority of known invertebrate TRPs have a penultimate proline.[36,37] CabTRP Ia is of particular interest given the known biological significance of the Ala−Pro region of this molecule. CabTRP Ib is a truncated form of CabTRP Ia with the Ala^1−Pro^2 dipeptide cleaved from the N-terminus that was shown to have decreased bioactivity, suggesting that the Ala^1−Pro^2 motif is important for activity.[28]

Below, we focus on understanding the structural implications of the frequent occurrence of penultimate proline residues in neuropeptide sequences through a series of IMS-MS experiments on two model neuropeptide systems. By analyzing N-terminally truncated and Pro → Ala substituted analogues, we demonstrate that penultimate proline residues are important for the conformational heterogeneity of NPY_{1-9} and CabTRP Ia. Molecular modeling simulations are performed to provide insight into the different intramolecular interactions that arise from the *cis*−*trans* isomerization of the Ala^1−Pro^2 peptide bond of CabTRP Ia. These simulations show that *cis* and *trans* configurations result in changes in the H-bonding pattern of the protonated N-terminus, leading to global conformational changes that may be important for biological activity.

EXPERIMENTAL SECTION

Ion Mobility Spectrometry-Mass Spectrometry. IMS-MS and IMS-IMS-MS experiments were conducted on a home-built instrument with a ∼2 m drift tube coupled to a time-of-flight (TOF) mass spectrometer, previously described in detail.[38,39] Ions are produced by electrospray ionization with a Triversa Nanomate (Advion Bioscience, Inc., Ithaca, NY, USA), stored in a Smith-geometry[40] ion funnel, and periodically pulsed into the drift tube, which is filled with 3.0 ± 0.2 Torr He buffer gas and operated with an electric field of ∼10 V·cm^{-1}. Mobility separated ions are transferred through a differentially pumped region and analyzed with an orthogonal geometry TOF-MS in a nested fashion.[41]

IMS-IMS-MS experiments are performed by operating the drift tube as two independent drift regions separated by an ion funnel.[38,39] Ions are mobility selected at the front of the mid funnel and collisionally activated by increasing the electric field across a 0.3 cm region at the end of the funnel. Conformational changes are monitored as a function of activation voltage by allowing annealed ions to mobility separate in the second drift region prior to mass analysis.

Collision Cross-Section Distributions. Drift time (t_D) distributions are measured and converted to collision cross-section (Ω) distributions according to[42]

$$\Omega = \frac{(18\pi)^{1/2}}{16}\frac{ze}{(k_B T)^{1/2}}\left[\frac{1}{M_I}+\frac{1}{M_B}\right]^{1/2}\frac{t_D E}{L}\frac{760}{P}\frac{T}{273.2}\frac{1}{N} \quad (1)$$

where ze is the charge of the ion, k_B is the Boltzmann constant, T is the temperature, E is the electric field, L is the length of the drift tube, P is the pressure, and N is the neutral number density at STP. M_I and M_B are the masses of the ion and buffer gas, respectively. Collision cross section distributions of Ala substituted peptide have been shifted by +2.5 $Å^2$ per Pro → Ala substitution to account for the intrinsic size difference between proline and alanine residues, as previously determined by intrinsic size parameters.[43,44] The 2.5 $Å^2$ correction factor was effective for *cis* and *trans* isomer assignments of triply protonated bradykinin.[13]

Peptide Synthesis. Peptides were synthesized by solid-phase Fmoc synthesis on an Apex 396 peptide synthesizer (AAPPTec, Louisville, KY, USA) similar to procedures previously described.[13] Fmoc-Rink amide MBHA resins and Fmoc side chain protected amino acids were purchased from Midwest Biotech (Fischers, IN, USA). Briefly, 20% piperidine in dimethylformamide was used to perform N^{α} deprotections, and 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3*H*)-one was used as the coupling reagent. Peptides were cleaved from the resin with a solution of trifluoroacetic acid/triisopropylsilane/methanol at a 18:1:1 ratio with the addition of 5% 2,2′-(ethylenedioxy)diethanethiol for methionine-containing peptides. Peptides were precipitated and washed with diethyl ether and dried using a vacuum manifold. Peptides were reconstituted in H_2O before dilution into the appropriate water/methanol ESI solution.

Molecular Modeling. Molecular dynamics simulations were performed on doubly protonated CabTRP Ia ions with a simulated annealing procedure in the Insight II software suite (Accelrys, Inc., San Diego, CA, USA) similar to simulations on proline-containing peptides previously described.[45] Simulations were performed on two initial structures of doubly protonated CabTRP Ia that differ by the configuration of the Ala^1−Pro^2 peptide bond. Charges were assigned to the N-terminus and Arg^9 side chain. The Amber force field and a dielectric of 1 were used in the Discover 3 module of the Insight II software package. During simulated annealing cycles, ions are heated from 298 to 500 K over 5 ps, equilibrated at 500 K for 10 ps, cooled to 298 K over 5 ps, and equilibrated at 298 K for 30 ps. One-hundred trial geometries were saved, and collision cross-sections were calculated by the trajectory method in Mobcal.[46]

RESULTS AND DISCUSSION

Statistical Analysis of Penultimate Proline in Neuropeptide Sequences. We analyzed 847 neuropeptide sequences from humans and seven other animal species cataloged in the NeuroPedia[25] database to assess the positioning of proline in neurobiological molecules. To report conservative significance values, we aligned these sequences using the Needleman−Wunsch algorithm[47] with the BLOSUM62 substitution matrix[48] and determined 181 distinct sequences sharing no more than 40% identity. While proline is present throughout human proteins with a frequency of 6.3%,[49] proline occupies the penultimate position with a significantly

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472

greater frequency than any other amino acid in the NeuroPedia database ($P = 0.0002$; binomial test with $p = 0.063$, $n = 181$, $x = 25$). Furthermore, proline occupies the penultimate position of these sequences significantly more often than any other of the first 20 positions ($P = 0.0003$; binomial test with $p = 0.05$, $n = 231$, $x = 25$). We permuted the data set and obtained the same result. Bar plots of proline's occupational frequencies relative to other amino acids and other positions alike are shown in Figure 1. Proline occupies the penultimate position relative to other

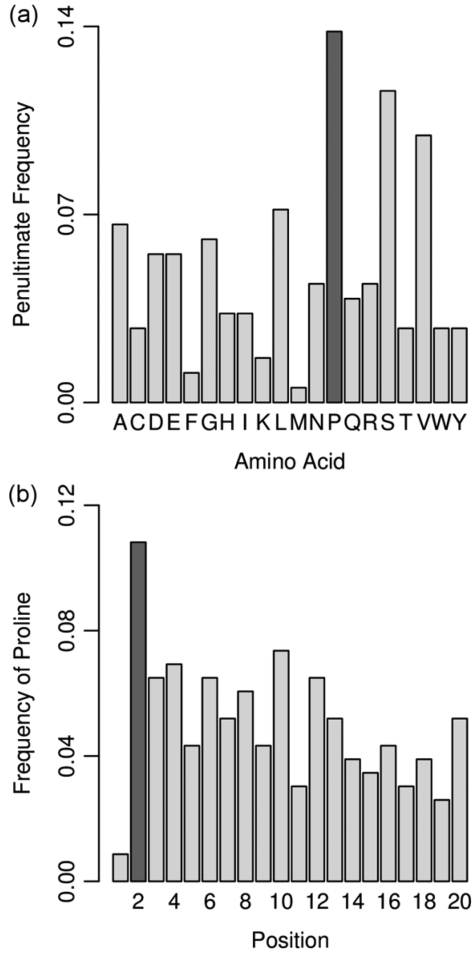


Figure 1. Frequency of occupation for amino acids in the penultimate position of neuropeptide sequences from the NeuroPedia database (a). The frequency of proline occupation in the first 20 positions of neuropeptide sequences (b).

amino acids approximately 13.8% of the time. Proline occupies the penultimate position relative to the other positions approximately 10.8% of the time. In addition to database of sequences analyzed here, we note that several large families of peptides from organisms not included in NeuroPedia frequently possess penultimate proline residues, such as tachykinin-related peptides in invertebrates.[36] These results suggest that proline's occupation of the penultimate position in neuropeptide sequences has been conserved evolutionarily and may serve an important functional role in these molecules. This inspired us to examine the role of penultimate proline residues on the conformational dynamics of neuropeptides.

IMS-MS and IMS-IMS-MS of NPY_{1-9}. Collision cross-section distributions for doubly charged $[NPY_{1-9} + 2H]^{2+}$ ions produced by electrospray ionization from a range of water/methanol solutions are displayed in Figure 2. This study focuses on the doubly charged species, as this is the dominate charge

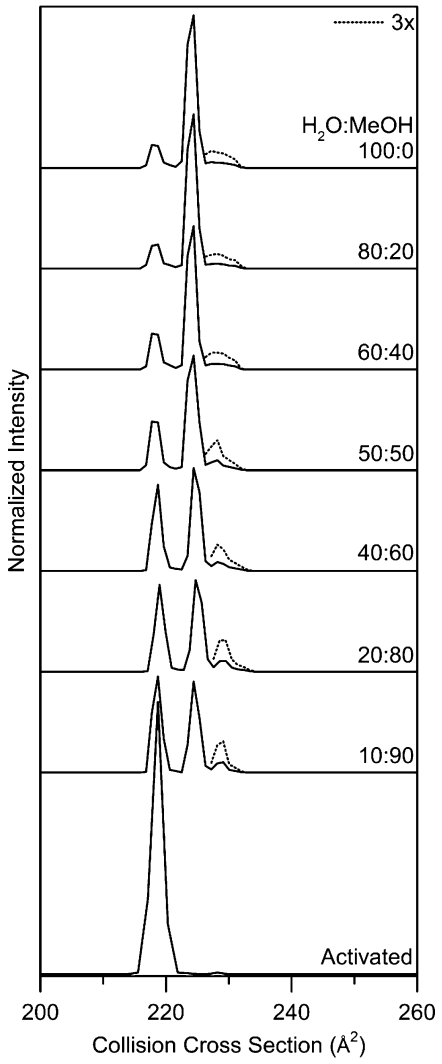


Figure 2. Collison cross-section distributions for $[NPY_{1-9} + 2H]^{2+}$ ions obtained from different water/methanol solutions. The distribution labeled activated is the quasi-equilibrium distribution obtained from collisional activation by IMS-IMS-MS experiments, as explained in the text.

state for all peptides presented here. The distribution of $[NPY_{1-9} + 2H]^{2+}$ ions is composed of at least four partially resolved peaks at 219, 224, 228, and 231 Å^2, referred to as conformers A, B, C, and D, respectively. It is important to note the four conformations are partially resolved and that cross-section distributions do not reach baseline for the entire range of conformations; thus, it is possible that additional conformations of relatively low abundance exist, making this a conservative estimate of the number of conformation types. Although the peaks are not fully resolved, it is interesting that we observe defined peaks in the mobility distribution given previous studies that report that this peptide is unstructured.[30−34] CD and NMR studies of the N-terminal region on NPY suggest that it does not adopt a single well-defined structure; IMS-MS data suggests NPY_{1-9} adopts an ensemble of defined conformer populations that are difficult to monitor in solution but are kinetically trapped during ESI and sampled by

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472

IMS-MS. Although it is unlikely that the structures are identical to those from solution, this provides further evidence that IMS-MS can be used to monitor conformer populations from solution.[12,23]

Changes in conformer populations are observed as a function of the water/methanol ratio from which NPY_{1-9} is electrosprayed (Figure 2). From 100% H_2O, conformer B is 80%, conformer A is 14%, and the unresolved conformers C and D are 6% of the collision cross-section distribution. However, as the methanol fraction is increased, conformer A increases in abundance and comprises 48% of the distribution from the 10:90 water/methanol solution. Although the combined abundance of conformers C and D remains constant at ∼7% with increasing methanol content, conformer C increases relative to conformer D. The observance of changes in peptide conformation populations as a function of solution composition is similar to IMS-MS studies of triply protonated bradykinin (Arg^1-Pro^2-Pro^3-Gly^4-Phe^5-Ser^6-Pro^7-Phe^8-Arg^9) electrosprayed from different water/methanol solutions.[12] It is worth pointing out that bradykinin is similar to NPY_{1-9} in that both are proline-rich sequences, possess a penultimate proline, and lack rigid secondary structure in solution.[12]

Multidimensional IMS-IMS-MS experiments were conducted to further probe the role of solvent in the formation of specific NPY_{1-9} conformations. Upon collisional activation, we observe the formation of the compact conformation, A, regardless of the initial conformation selected (Figure 2). Previous studies describe collisionally activated distributions of gas-phase conformations obtained by IMS-IMS-MS experiments as the quasi-equilibrium distribution.[50] A characteristic of the quasi-equilibrium distribution is that ions collisionally activated above the energetic barriers between states form the same distribution of annealed ions regardless of what conformation is selected and activated. This provides insight into the conformations that are preferred from solution versus those that are stable in the absence of solvent−molecule interactions. This agrees with the water/methanol solution scans that conformation B is favored in aqueous environments and conformation A is favored in nonaqueous environments such as the gas phase and high fractions of methanol.

IMS-MS of Pro → Ala Substituted and N-Terminally Truncated NPY. In order to provide insight into the specific proline residues that are important for the conformational heterogeneity of $[NPY_{1-9} + 2H]^{2+}$ ions, Pro → Ala substituted peptides were analyzed by IMS-MS. Pro → Ala substitutions allow for the assignment of specific proline residues as either *trans* or *cis* because Xaa−Ala bonds are almost exclusively *trans*. Thus, we assign conformers that are present in both the natural and Ala substituted sequences as *trans* and those that are eliminated upon Pro → Ala substitution as *cis*. Furthermore, we monitor the number and relative populations of conformers upon substitution to provide insight into the specific residues important for conformational heterogeneity.

Figure 3 shows collision cross-section distributions for Pro^2 → Ala, Pro^5 → Ala, Pro^8 → Ala, and $Pro^{2,5,8}$ → Ala substituted analogues of NPY_{1-9} electrosprayed from 50:50 water/methanol. Throughout this article, collision cross-section distributions of Ala substituted peptides are adjusted for the intrinsic size difference between Pro and Ala residues as explained above. The distribution for the Pro^2 → Ala peptide is dominated by a single conformation at 219 Å^2 that is 97% of the total distribution. This suggests that the Pro^2 residue is likely in the *trans* configuration for conformer A of $[NPY_{1-9} + 2H]^{2+}$ ions. The distribution for the Pro^5 → Ala is broad, with the most abundant cross-section at 224 Å^2 and peak shoulders at ∼220 and ∼227 Å^2. From this distribution, we assign Pro^5 as *trans* for conformer B of NPY_{1-9}, as this is the major peak from the Pro^5 → Ala distribution. The Pro^8 → Ala distribution has two conformations of similar abundance at 219 and 224 Å^2 and a trailing shoulder of relatively low abundance at 228 Å^2. This distribution is strikingly similar to that of $[NPY_{1-9} + 2H]^{2+}$ ions, suggesting that Pro^8 is in the *trans* configuration for the two major conformers, A and B, observed for $[NPY_{1-9} + 2H]^{2+}$. Furthermore, this suggests that Pro^8 contributes the least significantly to the conformational heterogeneity of NPY_{1-9}.

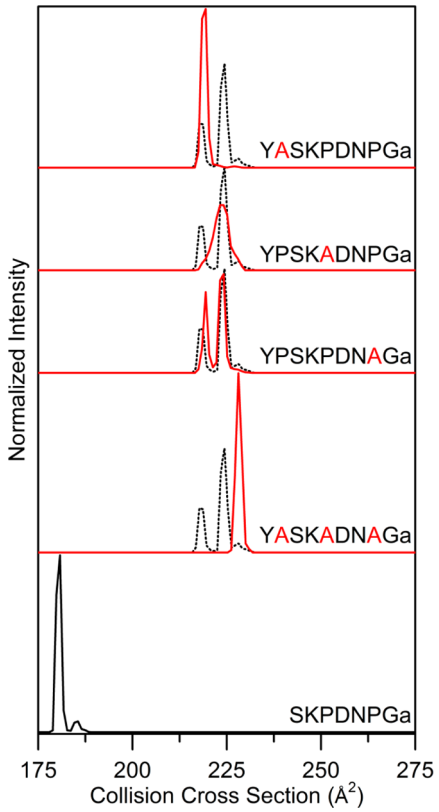


Figure 3. Collison cross-section distributions for Pro^2 → Ala, Pro^5 → Ala, Pro^8 → Ala, and $Pro^{2,5,8}$ → Ala analogues of $[NPY_{1-9} + 2H]^{2+}$ ions and N-terminally truncated $[NPY_{3-9} + 2H]^{2+}$ ions. Distributions are obtained from 50:50 water/methanol solutions. Ala substituted distributions are shown in red and overlaid on the dotted $[NPY_{1-9} + 2H]^{2+}$ distribution for ease of comparison. Distributions for Ala substituted peptides are shifted for size difference between Pro and Ala residues, as explained in the text.

The collision cross-section distribution for triply substituted $Pro^{2,5,8}$ → Ala peptide has a single peak at 228 Å^2 (Figure 3). This peptide allows for the all *trans* form of NPY_{1-9} to be assigned as conformer C. The presence of a single peak in the IMS distribution for triply substituted $Pro^{2,5,8}$ → Ala peptide confirms that the conformational heterogeneity observed for $[NPY_{1-9} + 2H]^{2+}$ arises from the *cis*−*trans* isomerization of proline residues.

Considering all Pro → Ala substituted peptides, we assign conformer A as *trans*-Pro^2, *cis*-Pro^5, and *trans*-Pro^8, as this peak dominates the mobility distribution for the Pro^2 → Ala substitution and is in relatively high abundance for the Pro^8 → Ala substitution. Conformer B is assigned as *cis*-Pro^2, *trans*-Pro^5, and *trans*-Pro^8 because this peak is the most abundant

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472

conformation for Pro^5 → Ala and Pro^8 → Ala substitutions and is not observed for the Pro^2 → Ala or $Pro^{2,5,8}$ → Ala peptides. As mentioned above, conformer C is assigned as *trans*-Pro^2, *trans*-Pro^5, and *trans*-Pro^8 because the conformer is observed for the triply substituted peptide. We are unable to definitely assign the least abundant conformation, D, as we do not observe a substituted peptide with significant population near this peak.

In order to provide further evidence that the penultimate proline residue is important for the conformational heterogeneity of $[NPY_{1-9} + 2H]^{2+}$ ions, we have analyzed the peptide with the Tyr^1—Pro^2 residues truncated from the N-terminus (NPY_{3-9}). The collision cross-section distribution for $[NPY_{3-9} + 2H]^{2+}$ has a major peak at 181 Å^2 and a minor peak at 185 Å^2, with relative abundances of 92 and 8%, respectively (Figure 3). We do not observe a single peak that is >80% total abundance for any of the penultimate proline-containing peptides. However, both NPY_{3-9} and the Pro^2 → Ala analogue have distributions that are dominated by a single conformation at >90% relative abundance, providing further evidence that penultimate proline is important for the conformational flexibility of the N-terminal region of neuropeptide Y. We point out that NPY_{1-9} contains a proline that is penultimate from the C-terminus. However, the peptide bond that adopts both *cis* and *trans* configurations is N-terminal to Pro residues (e.g., Asn^7—Pro^8). The analogous peptide bond to N-terminal penultimate proline occurs when proline is located at the C-terminal position.

IMS-MS of CabTRP Ia and Analogue Sequences. Collision cross-section distributions for doubly protonated $[M + 2H]^{2+}$ ions of CabTRP Ia (Ala^1-Pro^2-Ser^3-Gly^4-Phe^5-Leu^6-Gly^7-Met^8-Arg^9-NH_2), Pro^2 → Ala substituted CabTRP Ia, and CabTRP Ib (Ser^1-Gly^2-Phe^3-Leu^4-Gly^5-Met^6-Arg^7-NH_2) electrosprayed from 50:50 water/methanol are displayed in Figure 4. Two peaks are observed in the distribution for CabTRP Ia at 229 and 235 Å^2, with relative abundances of 86 and 14%, respectively. Unlike NPY_{1-9}, we do not observe major changes in the populations of CabTRP Ia when electrosprayed from different ratios of water/methanol. The contrast in solution dependence between NPY_{1-9} and CabTRP Ia may be due to differences in amino acids surrounding the Xaa^1—Pro^2 peptide bond. Several studies have illustrated that both local and nonlocal sequence composition influence proline isomerization.[51,52] For example, the residue N-terminally adjacent to proline is crucial for stabilizing different peptide bond configurations.[1,51] The influence of solvent on specific intramolecular interactions is likely different for each unique peptide sequence.

A major change in the distribution of conformations upon collisional activation by IMS-IMS occurs with the conformation at 235 Å^2 comprising 100% of the activated quasi-equilibrium distribution (Figure 4). It is interesting that both peptides analyzed in this study populate multiple conformations from solution but prefer a single gas-phase conformation upon activation. This suggests that the configuration of Xaa^1—Pro^2 peptide bonds is sensitive to the local environment.

We analyzed the Pro^2 → Ala substituted analogue of CabTRP Ia by IMS-MS to assign the *cis* and *trans* isomers of $[CabTRP\ Ia + 2H]^{2+}$ ions. The collision cross-section distribution has a single peak at 230 Å^2 (Figure 4). Therefore, we assign the peaks of $[CabTRP\ Ia + 2H]^{2+}$ ions at 229 and 235 Å^2 as *trans*- and *cis*-Pro^2, respectively. This differs from the penultimate proline in NPY_{1-9} that adopts the *trans* configuration in the quasi-equilibrium distribution. One potential explanation for this is differences in protonation sites. Both sequences have two basic sites: the N-terminus and Lys or Arg residue. However, the Lys^4 residue of NPY_{1-9} is located closer to the Pro^2 residue than the Arg^9 residue of CabTRP Ia. This likely influences the intramolecular interactions that stabilize different *cis* and *trans* configurations.

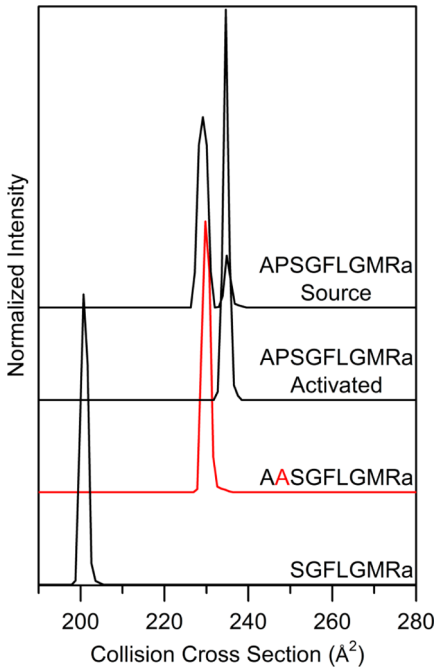


Figure 4. Collison cross-section distributions for $[CabTRP\ Ia + 2H]^{2+}$ ions, Pro^2 → Ala substituted analogue, and N-terminally truncated $[CabTRP\ Ib + 2H]^{2+}$ ions. Distributions are obtained from 50:50 water/methanol solutions. The distribution labeled activated is the quasi-equilibrium distribution for $[CabTRP\ Ia + 2H]^{2+}$ ions obtained from collisional activation by IMS-IMS-MS experiments, as explained in the text. The Ala substituted distribution is shown in red and shifted for size difference between Pro and Ala residues, as explained in the text.

We have analyzed the N-terminally truncated peptide CabTRP Ib by IMS-MS. The collision cross-section distribution for $[CabTRP\ Ib + 2H]^{2+}$ has a single peak at 201 Å^2. This confirms that the *cis*—*trans* isomerization of Ala^1—Pro^2 is important for the multiple conformations observed for CabTRP Ia. From IMS-MS studies of CabTRP Ia and its analogues, we determined that CabTRP Ia samples two environment-dependent conformations that arise from the *cis*—*trans* isomerization of the penultimate proline residue and differ in population from solution versus after collisional activation, whereas the N-terminally truncated CabTRP Ib has a single conformation.

Insight into Peptide Structure through Molecular Dynamics Simulations. In order to gain insight into structural differences that arise from the rotation of the Ala^1—Pro^2 peptide bond, molecular modeling was performed on doubly protonated CabTRP Ia. CabTRP Ia was selected for simulations because *cis* and *trans* isomers for only a single proline residue need to be considered, making the simulations relatively straightforward. Figure 5 shows two representative low-energy structures for the *trans* and *cis* isomers of CabTRP Ia. Calculated collision cross-section values are within 1% of the experimental; the *trans* (Figure 5a) and *cis* (Figure 5b)

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472

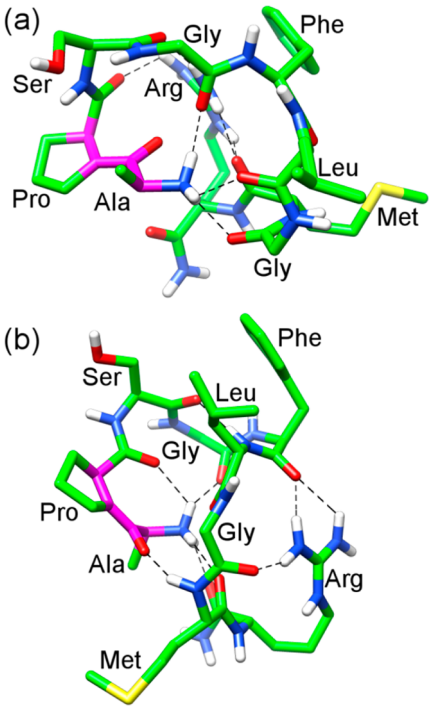


Figure 5. Representative structures of $[CabTRP\ Ia + 2H]^{2+}$ ions obtained by molecular modeling simulations. Ala^1–Pro^2 bonds in the *trans* (a) and *cis* (b) configurations are highlighted in magenta. H-bonds determined with recommended relaxed constraints in Chimera are shown as dashed lines. Carbon-bonded hydrogens are hidden for ease of peptide backbone and amino acid side chain comparison.

calculated collision cross-sections are 230 and 235 $Å^2$, respectively. In both structures, hydrogen bonds formed by the protonated N-terminus and guanidine group of Arg^9 with the peptide backbone are the main intramolecular interactions. Hydrogen bonds were determined using the default relaxed constraints in Chimera.[53] It appears that these interactions a markedly different for each respective *cis* and *trans* isomers. For example, in the *trans*-Pro^2 conformation, the N-terminus does not form a hydrogen bond with the carbonyl oxygen of Pro^2. However, in the *cis*-Pro^2 conformation, the N-terminus forms a hydrogen bond with the carbonyl oxygen of Pro^2. This suggests that the configuration of the Ala^1-Pro^2 is crucial for dictating the hydrogen bond network formed by the protonated N-terminus with the peptide backbone. We point out that the interaction between the protonated N-terminus and the carbonyl oxygen of the Pro^2 residue was also observed for the *cis* conformation of doubly protonated SPQAEK and singly protonated GPGG peptides studied by IMS-MS and IMS-cold ion spectroscopy, respectively, suggesting that this interaction may be important for stabilizing the *cis*-Pro^2 conformation in a variety of peptide sequences.[17,54] These simulations illustrate that the *trans* (Figure 5a) and *cis* (Figure 5b) configurations of the Ala^1–Pro^2 bond in CabTRP Ia favor different intramolecular interactions around the protonated N-terminus that lead to changes in global conformation and side chain orientations.

CONCLUSIONS

A statistical analysis of 847 neuropeptide sequences found in an online database shows that proline residues are frequently located at the penultimate position from the N-terminus. Studies of NPY_{1-9} and CabTRP Ia by IMS-MS demonstrated that these peptides adopt multiple conformations that arise from the *cis*–*trans* isomerization of Xaa^1–Pro^2 peptide bonds, suggesting that penultimate proline residues are important for the conformational heterogeneity of neuropeptide sequences. Upon substitution or truncation of penultimate proline residues, we observe a decrease in the conformational heterogeneity of these systems. Furthermore, conformer populations related to different *cis*- and *trans*-Pro residues are sensitive to the local environment. That is, we observe changes in conformer populations as a function of solution composition or upon collisional activation (i.e., the gas-phase distribution). Molecular modeling simulations show that the configuration of the Ala^1–Pro^2 peptide bond of CabTRP Ia determines the network of intramolecular hydrogen bonds formed by the protonated N-terminus and therefore the orientation of the amino acid side chains and global conformation.

It is interesting to consider possible biological implications of the ability of proline-containing peptides to adopt multiple conformations. We propose that, in addition to the previously known role of protecting bioactive peptides from enzymatic degradation,[55] penultimate proline residues may serve as a functional switch that allows bioactive peptides to sample multiple conformations that may have different functions or binding affinities for their receptors. The ability of proline residues and, specifically, penultimate prolines to allow neuropeptides to sample different conformer populations upon changes to the local environment is likely an important structural consideration as peptides experience similar environmental changes, such as when approaching or binding receptors.

AUTHOR INFORMATION

Corresponding Author

*E-mail: clemmer@indiana.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Matthew Acton, David Smiley, and the DiMarchi research group at Indiana University for assistance with peptide synthesis. This work is supported by a grant from the NIH (R01 GM103725).

REFERENCES

(1) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*, 253–260.
(2) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397–412.
(3) Fischer, G. *Chem. Soc. Rev.* **2000**, *29*, 119–127.
(4) Brazin, K. N.; Mallis, R. J.; Fulton, D. B.; Andreotti, A. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1899–1904.
(5) Andreotti, A. H. *Biochemistry* **2003**, *42*, 9515–9524.
(6) Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. *Nature* **2005**, *438*, 248–252.
(7) Sarkar, P.; Reichman, C.; Saleh, T.; Birge, R. B.; Kalodimos, C. G. *Mol. Cell* **2007**, *25*, 413–426.
(8) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. *Nat. Chem. Biol.* **2007**, *3*, 619–629.
(9) Theillet, F.-X.; Kalmar, L.; Tompa, P.; Han, K.-H.; Selenko, P.; Dunker, A. K.; Daughdrill, G. W.; Uversky, V. N. *Intrinsically Disord. Proteins* **2013**, *1*, e24360.
(10) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466–8472

(11) Csizmok, V.; Felli, I. C.; Tompa, P.; Banci, L.; Bertini, I. *J. Am. Chem. Soc.* **2008**, *130*, 16873−16879.
(12) Pierson, N. A.; Chen, L.; Valentine, S. J.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2011**, *133*, 13810−13813.
(13) Pierson, N. A.; Chen, L.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2013**, *135*, 3186−3192.
(14) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136*, 10308−10314.
(15) Shi, L.; Holliday, A. E.; Shi, H.; Zhu, F.; Ewing, M. A.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2014**, *136*, 12702−12711.
(16) Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2014**, *65*, 175−196.
(17) Glover, M. S.; Shi, L.; Fuller, D. R.; Arnold, R. J.; Radivojac, P.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 444−452.
(18) Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577−592.
(19) Bohrer, B. C.; Merenbloom, S. I.; Koeniger, S. L.; Hilderbrand, A. E.; Clemmer, D. E. *Annu. Rev. Anal. Chem.* **2008**, *1*, 293−327.
(20) Ruotolo, B. T.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2006**, *10*, 402−408.
(21) Wyttenbach, T.; Bowers, M. T. *J. Phys. Chem. B* **2011**, *115*, 12266−12275.
(22) Lee, S.-W.; Freivogel, P.; Schindler, T.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 11758−11765.
(23) Silveira, J. A.; Fort, K. L.; Kim, D.; Servage, K. A.; Pierson, N. A.; Clemmer, D. E.; Russell, D. H. *J. Am. Chem. Soc.* **2013**, *135*, 19147−19153.
(24) Fort, K. L.; Silveira, J. A.; Pierson, N. A.; Servage, K. A.; Clemmer, D. E.; Russell, D. H. J. *J. Phys. Chem. B* **2014**, *118*, 14336−14344.
(25) Kim, Y.; Bark, S.; Hook, V.; Bandeira, N. *Bioinformatics* **2011**, *27*, 2772−2773.
(26) Beck-Sickinger, A. G.; Weland, H. A.; Wittneben, H.; Willim, K.-D.; Rudolf, K.; Jung, G. *Eur. J. Biochem.* **1994**, *225*, 947−958.
(27) Grandt, D.; Schimiczek, M.; Rascher, W.; Feth, F.; Shively, J.; Lee, T. D.; Davis, M. T.; Reeve, J. R., Jr.; Michel, M. C. *Regul. Pept.* **1996**, *67*, 33−37.
(28) Christie, A. E.; Lundquist, C. T.; Nässel, D. R.; Nusbaum, M. P. *J. Exp. Biol.* **1997**, *200*, 2279−2294.
(29) Tatemoto, K.; Carlquist, M.; Mutt, V. *Nature* **1982**, *296*, 659−660.
(30) Saudek, V.; Pelton, J. T. *Biochemistry* **1990**, *29*, 4509−4515.
(31) Darbon, H.; Bernassau, J.-M.; Deleuze, C.; Chenu, J.; Roussel, A.; Cambillau, C. *Eur. J. Biochem.* **1992**, *209*, 765−771.
(32) Chu, S. S.; Velde, D. V.; Shobe, D.; Balse, P.; Doughty, M. B. *Biopolymers* **1995**, *35*, 583−593.
(33) Monks, S. A.; Karagianis, G.; Howlett, G. J.; Norton, R. S. *J. Biomol. NMR* **1996**, *8*, 379−390.
(34) Bader, R.; Bettio, A.; Beck-Sickinger, A. G.; Zerbe, O. *J. Mol. Biol.* **2001**, *305*, 307−329.
(35) Beveridge, R.; Chappuis, Q.; Macphee, C.; Barran, P. *Analyst* **2013**, *138*, 32−42.
(36) Nässel, D. R. *Peptides* **1999**, *20*, 141−158.
(37) Severini, C.; Improta, G.; Falconieri-Erspamer, G.; Salvadori, S.; Erspamer, V. *Pharmacol. Rev.* **2002**, *54*, 285−322.
(38) Koeniger, S. L.; Merenbloom, S. I.; Valentine, S. J.; Jarrold, M. F.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E. *Anal. Chem.* **2006**, *78*, 4161−4174.
(39) Merenbloom, S. I.; Koeniger, S. L.; Valentine, S. J.; Plasencia, M. D.; Clemmer, D. E. *Anal. Chem.* **2006**, *78*, 2802−2809.
(40) Tang, K.; Shvartsburg, A. A.; Lee, H.-N.; Prior, D. C.; Buschbach, M. A.; Li, F.; Tolmachev, A. V.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2005**, *77*, 3330−3339.
(41) Hoaglund, C. S.; Valentine, S. J.; Sporleder, C. R.; Reilly, J. P.; Clemmer, D. E. *Anal. Chem.* **1998**, *70*, 2236−2242.
(42) Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York, 1988.
(43) Valentine, S. J.; Counterman, A. E.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1188−1211.
(44) Srebalus-Barnes, C. A.; Clemmer, D. E. *J. Phys. Chem. A* **2003**, *107*, 10566−10579.
(45) Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **2002**, *74*, 1946−1951.
(46) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082−16086.
(47) Needleman, S. B.; Wunsch, C. D. *J. Mol. Biol.* **1970**, *48*, 443−453.
(48) Henikoff, S.; Henikoff, J. G. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 10915−10919.
(49) Morgan, A. A.; Rubenstein, E. *PLoS One* **2013**, *8*, e53785.
(50) Pierson, N. A.; Valentine, S. J.; Clemmer, D. E. *J. Phys. Chem. B* **2010**, *114*, 7777−7783.
(51) Pal, D.; Chakrabarti, P. *J. Mol. Biol.* **1999**, *294*, 271−288.
(52) Wathen, B.; Jia, Z. *J. Proteome Res.* **2008**, *7*, 145−153.
(53) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(54) Masson, A.; Kamrath, M. Z.; Perez, M. A. S.; Glover, M. S.; Rothlisberger, U.; Clemmer, D. E.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2015**, DOI: 10.1007/s13361-015-1172-4.
(55) Vanhoof, G.; Goossens, F.; De Meester, I.; Hendriks, D.; Scharpé, S. *FASEB J.* **1995**, *9*, 736−744.

DOI: 10.1021/acs.analchem.5b01889
Anal. Chem. 2015, 87, 8466−8472