
Nancy Leymarie,¹,‡ Eric A. Berg,¹,‡ Mark E. McComb,¹,‡ Peter B. O’Connor,¹,‡ James Grogan,§ Frank G. Oppenheim,²,§ and Catherine E. Costello*,²,³

Mass Spectrometry Resource, Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118, and Department of Oral Biology and Periodontology, Boston University Goldman School of Dental Medicine, Boston, Massachusetts 02118

Proline-rich proteins (PRPs), including collagens, complement 1q, and salivary PRPs, are unusually difficult to sequence by mass spectrometry, due to the high efficiency of cleavage at the amide bond on the N-terminal of proline residues and the consequently low relative abundance of fragment arising from cleavages at other amide bonds. To fully characterize these proteins by mass spectrometry, specialized approaches to fragmentation are needed for the peptides with high proline content. Our work reported herein focused on the analysis of the set of peptides derived by tryptic cleavage of the salivary protein PRP-3. Two methods of fragmentation were compared: Collision-induced dissociation (CID) and electron capture dissociation (ECD). The data obtained demonstrated that ECD spectra of peptides containing more than 30% proline residues are simpler and easier to interpret than are CID spectra of those peptides. Factors that limit the two methods of fragmentation include the complexity of information contained in the CID spectra and the low efficiency of ECD processes. A complementary approach using both decomposition methods provides more complete and interpretable sequence information and yielded >93% sequence coverage for the 11-kDa PRP-3 isolated from human saliva.

Proteins that contain high amounts of proline (>25%), while relatively uncommon, nevertheless constitute several abundant classes of proteins in both plants and animals. These proline-rich proteins (PRPs) include the collagen family,¹ complement 1q (C1q),² and salivary PRPs,³ among others. PRPs have a variety of biological functions, including roles in the extracellular matrix (collagens), immune response (C1q), and oral biology (salivary PRPs). The salivary PRPs are divided into three groups: acidic PRPs, basic PRPs, and glycosylated PRPs. The salivary acidic PRPs are a family of closely related polypeptides,⁴ which together form a major fraction of total salivary protein.⁵ Their characterization is complex due to allelic variation and posttranslational modifications.⁶ Posttranslational modification of these proteins includes phosphorylation within an N-terminal calcium-binding domain and N-terminal pyroglutamidation, as well as limited glucuronidation.⁷

Recognition of the difficulties in obtaining PRP peptide structural information has led to the development of methodologies based on mass spectrometry. Generally, these methodologies utilize enzymatic digestion of the protein of interest followed by the characterization of the resultant proteolytic peptides by MS⁸ or MS/MS techniques.⁹ To date, most tandem mass spectra obtained for proline-rich proteins and peptides have used either high- or low-energy collision-induced dissociation (CID).¹⁰ The abundance of y-type ions associated with proline, the "proline effect", gives rise to difficulties in obtaining complete sequence information for peptides because fragment ions arising by cleavages at amide bonds not adjacent to proline have very low abundances or are missing.¹¹–¹³ While proline-driven cleavages may be useful for sequencing proteins or peptides that have only a few proline residues, PRPs often contain as much as 25–30% proline.

¹ To whom correspondence should be addressed. Mass Spectrometry Resource, Department of Biochemistry, Boston University School of Medicine, 715 Albany St., R806, Boston, MA 02118-2526. E-mail: cec.jms@bu.edu. Phone: (617) 638-6690; Fax: (617) 638-6691.
² Mass Spectrometry Resource, Boston University School of Medicine.
³ Department of Biochemistry, Boston University School of Medicine.
⁴ Boston University Goldman School of Dental Medicine.
¹ Department of Biochemistry, Boston University School of Medicine.
² Department of Biochemistry, Boston University School of Medicine.
³ Department of Biochemistry, Boston University School of Medicine.
⁴ Department of Biochemistry, Boston University School of Medicine.
⁵ Department of Biochemistry, Boston University School of Medicine.
⁶ Department of Biochemistry, Boston University School of Medicine.
⁷ Department of Biochemistry, Boston University School of Medicine.
⁸ Department of Biochemistry, Boston University School of Medicine.
⁹ Department of Biochemistry, Boston University School of Medicine.
¹⁰ Department of Biochemistry, Boston University School of Medicine.
¹¹ Department of Biochemistry, Boston University School of Medicine.
¹² Department of Biochemistry, Boston University School of Medicine.
¹³ Department of Biochemistry, Boston University School of Medicine.

proline, giving rise to peptides containing 6–10 or more proline residues. The prevalence of internal fragment ions or multiple terminal proline losses, generated when many proline residues are present in one peptide, often makes it difficult to obtain interpretable CID MS/MS data. The high proline content, however, could have an interesting and useful result when a different technique of fragmentation such as electron capture dissociation (ECD) is applied. The novel ECD method of fragmentation is based on recombination of multiply protonated polypeptide and protein ions with thermal electrons. It has proven to be a valuable tool for peptide sequencing and allows the localization of posttranslational modifications. The advantage of this technique lies in the fact that the series of fragment ions are generally more complete and more specific (N-terminal c, and C-terminal z-) than are the corresponding ion series produced in low-energy CID fragmentation (N-terminal b, C-terminal y, and internal fragments). However, proline sites represent an exception to this trend, as the cyclic structure of proline precludes formation of c/z- ions, with the result that ECD is predicted to generate many fewer fragments for the proline-rich peptides, while still maintaining the same sequence information. Awareness of this peculiarity of proline-containing peptides led us to apply the ECD ion fragmentation method to characterize PRP.

The work summarized in this report was centered on a purified human proline-rich protein, PRP-3, collected from volunteers at the Boston University School of Dental Medicine. The samples were first screened on the quadrupole/orthogonal acceleration time-of-flight (Q-o-TOF) mass spectrometer and a first methodology based on the traditional CID method of fragmentation was investigated. With the goal to improve the PRP sequencing methodology, we investigated the ion fragmentation method ECD combined with Fourier transform mass spectrometry (FTM S). The main purpose of this investigation was to compare the different effects of the presence of proline on the fragmentation patterns of tryptic proline-rich peptides under CID or ECD conditions and compare the advantages of those methods of fragmentation. In the case of peptides containing more than 25% of proline, the resultant ECD data proved to be simpler and easier to interpret than data obtained with the CID experiments. However, PRP-3 also includes a region that is not proline-rich. The advantage of the complementary approach, using both ECD and CID to determine the full structure of an acidic proline-rich protein (PRP-3), including complete sequence information as well as localization of posttranslational modifications, was demonstrated.

EXPERIMENTAL SECTION
Isolation and Tryptic Digestion of PRP-3. Human parotid saliva was collected from donors in the Boston University School of Dental Medicine at Boston Medical Center, using techniques based on methods described by Jensen et al. A 10-mL dialyzed sample containing 150 mg of protein was loaded on a DEAE column (3.0 × 30 cm; DEAE “fast flow”; Pharmacia, Piscataway, NJ) and equilibrated with 20 mM Tris buffer, pH 8.0 (buffer A). The sequential gradients employed were 0–6, 6–18, and 18–54% buffer B (buffer A containing 1 M NaCl) at a flow rate of 2 mL/min at 4 °C. A single late-eluting acidic proline-rich protein fraction was collected, dialyzed extensively against water, and lyophilized. This fraction was subjected to high-resolution anion exchange chromatography using an HPLC system (Pharmacia) equipped with a Mono Q column (5 mm × 10 cm; Pharmacia). The leading edge of the eluting protein peak was collected, dialyzed extensively against water, and lyophilized. The protein was desalted using reversed-phase chromatography with an HPLC system (Gilson Co., Inc., Lewis Center, OH) equipped with a C18 column (4.6 mm × 25 cm; TosohHaas). The protein was eluted using a shallow linear gradient from 8 to 32% acetonitrile and dried. This protein, termed PRP-3, was used to prepare the tryptic digests.

Four proline-rich peptides (T1, residues 1–30; T2, 31–74; T3, 75–91; and T4, 92–106) were then generated by treatment with trypsin. Tryptic digestion of PRP-3 was made in 1 mL of buffer prepared according to the enzyme supplier (Promega, Madison, WI). Reactions were incubated at 37 °C for 24 h, and the products were separated immediately using C18 chromatography on a Gilson HPLC system using a 0–80%ACN/H2O with 0.1% TFA continuous gradient. Collected fractions were dried and stored at −20 °C.

Mass Spectrometry. Q-o-TOF Mass Spectrometer. CID-tandem MS experiments were performed on an QStar Pulsari Q-o-TOF mass spectrometer (MD-Sciex/Applied Biosystem Inc, Framingham, MA), equipped with a nanoelectrospray (ESI) ion source in the positive ion mode. After calibration, the instrument was capable of achieving better than 10 ppm mass accuracy with a minimum resolution of 19000 (fwhm). T3 and T4 samples were diluted to a final concentration of ~500 fmol/µL in 49.5/49.5/1 methanol/water/formic acid, and T1 and T2 samples were diluted to the same concentration in 50/50 methanol/20 mM ammonium acetate. The ESI capillary potential was increased slowly from 0 to 1.2 kV. The declustering potential was held at 35 V. For tandem MS experiments, the parent ions were isolated using the first quadrupole (Q1) and fragmented via CID inside the collision cell (Q3) with an N2 pressure of 3.5 × 10−5 Torr. The quadrupole potential was adjusted from 15 to a maximum 55 V to reduce the abundance of the precursor ion to about one-third of its original intensity. The fragment ions were analyzed in the TOF region.

FTICR Mass Spectrometer. ECD-tandem MS experiments were performed on a modified FTICR MS system (IonSpec, Irvine, CA) with a 7-T active shielded superconducting electromagnet (Cryomagnetics, Oak Ridge, TN) using a capacitively coupled closed cylindrical Penning trap and an ESI ion source (Analytica of Branford, Branford, CT) modified with a home-built nano-ESI interface. The peptides were diluted with the same solvents used for CID experiments to a final concentration of 5–10 pmol/µL. The pulse valve backed by 15 mbar N2 was pulsed for 4 ms during a 0.5-s ion accumulation, and then ions were allowed to cool for 3 s. For ECD, an electron filament gun was mounted on the rear trap plate of the cell, and the filament termini were held at +1.3 V.
and \(-0.3\) V. The duration of the electron irradiation varied from 5 to 20 s, as needed. Typical broadband resolving power was \(1:50,000\). Data analysis was done using the Boston University Data Analysis (BUDA) software developed by the Mass Spectrometry Resource. In early experiments, the efficiency of electron capture varied significantly from day to day; this phenomenon was correlated with instability on the filament current/voltage. The problem was partially solved by replacing the IonSpec filament power supply with a commercial constant-current supply (BK Precision, model 1620).

RESULTS AND DISCUSSION

Tryptic Peptides T4 and T3. Our investigation began with the study of the shorter tryptic peptides, T3 (Figure 1) and T4 (Figure 2), which contain 6 and 7 proline residues out of 17 and 15 amino acids, respectively, and have in common the N-terminal sequence PQGPPQQGGHP. The fragmentation of the triply protonated peptides at \(m/z 578\) (T3) and 536 (T4) was investigated for both types of experiments. Although 100% of the sequence is covered with each of the fragmentation techniques used, the utility of the data for elucidation of the structure of an unknown is distinctly different.

CID Spectra. The CID product ion spectra of the T3 and T4 peptides are similar to CID spectra reported from other proline-rich peptides.\(^{19}\) The CID spectra are dominated by singly charged fragment ions, but some doubly and triply charged fragment ions are also present. The proline immonium ion at \(m/z 70\) is particularly abundant in both CID spectra. The N-terminal fragment ions series (b) is present in the CID spectra at a low relative intensity, except for the \(b_2\) and \(b_3\) ions, due to the presence of an arginine and histidine residue at the C-terminus.\(^{20}\) The CID of the T3 peptide is characterized by a complete C-terminal ion (y) series, and only the \(y_4\) fragment ion is missing in the C-terminal ion series of the T4 peptide. Certain y ions are particularly abundant (>15%) in the spectra: \(y_5, y_6, y_7, y_{13}\), and \(y_{14}\) for the T3 peptide and \(y_3, y_{11}\), and \(y_{12}\) for the T4 peptide, showing clearly that the fragmentation at each peptide bond is not equally facile. This phenomenon is related to the "proline effect" reported in the literature by several different groups.\(^{19,21,22}\) The proline effect refers to preferential fragmentation of the more labile bond at the N-terminal side of Pro, as compared to less preferential fragmentation of the more stable bond on the Pro C-terminal. The high basicity of the Pro residue\(^{14}\) facilitates the protonation of this peptide...
residue and leads to a high yield of y ions having a Pro residue on the N-terminus. The first direct consequence of this proline effect is to form a predominantly C-terminal fragment ion series (y), to the detriment of the corresponding N-terminal fragment ions series (b). An indirect effect of Pro in the sequence was noted in the CID spectra: the b2 and y13 fragment ions of T4 and T3 peptides, which correspond to cleavage between Gln 2 and Gly 3, are the most abundant peaks in both CID spectra. Jonsson et al.10 have reported that the Gln-Gly cleavage is a dominant dissociation site in the sequence Pro-Gln-Gly-Pro. While the mechanism of the Gly-Gln cleavage has not been definitively elucidated, the interaction of proline with the side chain of the Gln residue seems to be involved.23,24 The most abundant y fragment ions (y5, y6, y7, y11, and y13 for T3 and y5, y11, y12, and y13 for T4) reveal the partial sequence (QGPP and PPP) of the T3 peptide and the partial sequence (QGPP) of the T4 peptide. The other y or b ions, which reveal the remaining sequence of the peptide, are much less abundant (1–10%). If the sequence of the peptide were unknown, the much lower signal intensity of these y and b fragment ions would complicate the interpretation of the CID spectrum.

The second consequence of the presence of the proline residue at a high ratio in the sequence is to promote the formation of internal fragment ions where the N-terminal residue of the internal fragment is usually a Pro residue. Their relative abundances vary from 0.1 to 10% in the range m/z 70–1000. Almost all of these internal fragment ions have the structure type yb, but some ya and yb, with ammonia loss are observed as well.12 The internal fragment ions present in the spectra are mostly singly charged, but a few are doubly charged (for example y16b11 for the T3 peptide and y14b11-NH3 for the T4 peptide). While this type of information is useful in sequencing peptides containing only a few proline residues, the increasing number of fragment ions generated when many proline residues are present in one peptide makes sequencing difficult. In this case, even though the sequence is known, some fragments may be attributed to more than one possible structure: the ions m/z 226 and 283 for both peptides could be


Figure 2. (a) CID spectrum (ESI-Q-o-TOF MS/MS) obtained from the triply charged ion of proline-rich peptide T4. (b) ECD spectrum (FTMS, 10 scans) obtained from the triply charged ion of the peptide T4.
interpreted as ions b₂ (m/z 226.1186) and b₃ (m/z 283.1401) or as the internal fragment ions PQ (m/z 226.1192) and QGP (m/z 283.1406). Another example to illustrate the ambiguity of the interpretation of some ions is the m/z 380.202 ion of the peptide T4. The m/z 380.202 ion could be interpreted as y₂⁺ (m/z 380.2223) or the y₅b₅+ (m/z 380.1808) fragment ion. A third and probably more realistic interpretation is that the observed peak is made up of approximately equal contribution from both species and has the average m/z 380.2016. This overlap is explained by the low S/N ratio of this peak, which limits the resolution in the detected signal. Furthermore, some fragment ions are interpretable only because the sequence of the studied peptides is known, while others are ambiguous because of sequence redundancy. For example, the ions y₃b₅ and y₅b₁₂ of the T3 peptide represent the sequence PQP. Thus, rather than giving real structural information on the studied peptide, such internal fragment ions further complicate the interpretation of the CID spectra.

Both spectra contain a number of unidentified peaks with low relative abundances (1–10%). We also noted the presence of b-NH₃ ions and the almost complete series of (y-NH₃) fragments, which are characteristic of tryptic peptides with a C-terminal Arg residue. The presence of these ion series further complicates interpretation of the CID spectra without providing more structural information.

To assess the utility of the CID data, multiple database searches (Prospector, Mascot, Profound) were performed, using as input the entire CID spectrum or the most abundant fragment ions (8 or 10) of the CID spectra of the T3, T4, and T2 peptides and expanding the search parameter for mass accuracy from 10 to 200 ppm. Only one database (Profound) was able to match the experimental data from any proline-rich peptides (only T4) to a corresponding PRP-3 in the databases. The presence of at least 35% proline in the sequence of T3 and T4 peptides induces a characteristic pattern of fragmentation. While the CID spectra may be rich in information and may contain the complete amino acid sequence, the complexity of the spectra is such that extracting structural information is difficult and very often is not useful for determination of an unknown sequence.

ECD Spectra. The ECD spectra are described by the nomenclature adapted in Table 1. The ECD spectrum of T3 (Figure 1b) is composed primarily of 10 c and z fragment ions, representing 10 possible cleavages. The ECD spectrum of the T4 peptide contains 8 out of the 9 possible c and z fragment ions (among 13 ion signals). The presence of different basic amino acids such as glutamine, proline, and histidine along the sequence allows the formation of c and z fragment ions with similar probability. The mechanism of ECD is far less specific than CID and offers relatively nonspecific cleavages at all the bonds facilitating sequence determination of peptides. The proline residue is the only exception to the low specificity of ECD. Due to the cyclic proline structure, the formation of c and z fragment ions is not possible (reaction 1 below).

The opened cyclic (R₁⁺⁺⁺CON(HCH₂)₂CONH-R₂⁺⁺⁺) ion issued from reaction 1 has very little probability of capturing another electron in the FTMS cell. An additional H⁺ transfer in this region, permitting the formation of the fragment ions, is improbable: the proton affinity of the carbonyl is comparable for almost all amino acid residues (except for the tryptophan residue). Therefore, by ECD, the proline residues can be located in the sequence without ambiguity and fragmentation of all other amide bonds provides sequence information of the rest of the peptide.

The internal fragments present in the CID spectra, which complicate interpretation, are completely absent in the ECD spectra. In comparison to the CID spectra, the ECD spectra have much less complexity and are consequently easier to interpret. Some neutral losses are observed in the ECD spectra of the T3 and T4 peptides. The neutral loss of NH₃ or H provides no sequence information, while the neutral loss of 45.028 Da (CH₂NO), the neutral loss of 59.045 Da (CNH₂), or the neutral losses of 82.045 Da (CH₂N₂) are characteristic of the presence of glutamine, arginine and histidine residues.

Both T3 and T4 ECD spectra show the presence of anomalous fragment ions denoted c and z (Table 1). These correspond to 1-Da differences, lower and higher, respectively, in masses of the expected ions c and z, and have been observed previously. The symbol C corresponds to the following structure \([\text{Nterm} + \Sigma (\text{Aa}) + \text{NH}_2] (\text{Nterm} = H \Sigma (\text{Aa}) \text{represents the summation of the amino acids present in the fragment product). The symbol Z corresponds to the following structure [\text{Cterm} + \Sigma (\text{Aa}) - \text{NH}_2] (\text{Cterm} = OH).}

### Table 1. Nomenclature Adopted To Label the Fragment Ions Formed during the ECD Experiments

<table>
<thead>
<tr>
<th>fragment label</th>
<th>corresponding ion</th>
</tr>
</thead>
<tbody>
<tr>
<td>c&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(C + (n − 1)H)&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>c&lt;sup⇐&lt;/sup&gt;</td>
<td>(C + nH)&lt;sup&gt;⇐&lt;/sup&gt;</td>
</tr>
<tr>
<td>z&lt;sup&gt;+&lt;/sup&gt;</td>
<td>(Z + (n + 1)H)&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>z&lt;sup⇐&lt;/sup&gt;</td>
<td>(Z + nH)&lt;sup⇐&lt;/sup&gt;</td>
</tr>
</tbody>
</table>


The presence of these types of ions could potentially complicate the interpretation of the spectra, but comparison of isotope distribution patterns can be used to distinguish possible ambiguities. An unusual x fragment ion (x7 and x5 for the T3 and T4 peptides) corresponding to the cleavage between His and Pro residues was observed (Figure 2). This type of x fragment ion has not been reported in the ECD literature except when energetic electrons (>5–6 eV) are used during the ion–electron collisions.

In these experiments, the energy of the electrons used during the ECD experiments was 0.5 eV. The origin of the x7 and x5 ions is not clear.

The fragments ions present in the ECD spectrum, almost c and z ions, are fewer in number than the fragments observed in the CID spectrum but provide equal and complete structural information of the studied proline-rich peptide, without presenting any ambiguity.

**Tryptic Peptide T2.** We extended our investigation with the comparison of the CID and ECD spectra of the T2 peptide (Figure 3a and b, respectively). The T2 peptide contains 10 proline residues out of 44 amino acids. The (M + 4H)4+ ion was not perfect: some sodium and potassium adduct ions are observed.

The CID spectrum is primarily composed of singly, doubly, and some triply charged fragment, and the C-terminal y series dominates the CID spectrum. The y5, y6, y7, y18, y19, and y23 ions, corresponding to cleavages on the N-terminal side of a proline residue (proline effect), are the most abundant fragment ions. Many different internal fragment ions, whose interpretation could be complex, are also present in the CID spectra. A number of unidentified low-abundance fragment ion species are observed. The assignment of fragment ions in the CID spectra covers only 70% of the total sequence.

The ECD spectrum of the T2 peptide is dominated by 20 c and 18 z fragment ions, which yield 73% coverage of the sequence. As for the T3 and T4 ECD spectra, some fragment ions are present, as well as side-chain fragment ions. The internal fragments are absent.

---

Although the CID and ECD spectra cover nearly the same percentage of the sequence (70–73%), the structural information contained in the two spectra is different and they complement each other. For example, ECD spectra allowed determination of the sequences PL (residues 4–5), QP (residues 11–12), and GDGN (residues 15–18) of the T2 peptide, while the CID clarified the NDG (residues 20–22), QGPPQQGG (residues 25–32) sequence. In this case, the complementarity of the two techniques of fragmentation affords greater than 91% sequence coverage. The utility of analyzing the same peptide by both ECD and CID fragmentation techniques is clearly demonstrated.

**Tryptic Peptide T1.** Finally, we compared the CID and ECD spectra obtained for the triply charged ion of the T1 peptide, m/z 1174 (Figure 4a and b, respectively). The T1 peptide contains 12 acidic acid residues (out of 30 amino acids) and has only one proline residue. The T1 peptide is also characterized by the presence of two phosphorylation sites at Ser 8 and Ser 22 and by cyclization of the N-terminal glutamine residue.

**CID Spectrum.** The CID spectrum covers 80% of the sequence and contains singly and doubly charged b and y ions. The complementary and the b_{12}^{+} and y_{18}^{2+} ion pair corresponding to the cleavage of the amide bond between the Val12 and Pro13 are abundant ions. The only proline residue present in the sequence of the T1 peptide induces the proline effect. Additionally, more than 30% of the residues of the T1 peptide are acidic residues and thus contribute to a further specific pattern of fragmentation. The mobile proton of an acidic residue is particularly likely to be involved in cleavage on the C-terminal side of this amino acid if all basic sites are protonated.37,38 This tendency is illustrated by the high abundance of the y_{2}, y_{3}, y_{4}, b_{6}, b_{7}, and b_{8} ions, corresponding to cleavages on the C-terminal side of the aspartic acid residues.

The peak (M + 3H – HPO_{4})^{3+} indicates clearly the presence of at least one site of phosphorylation. The most abundant b and y ions were observed with phosphate attached. Few b and y ions

---

Figure 4. (a) CID spectrum (ESI-Q-o-TOF MS/MS) obtained from the triply charged ion of peptide T1. (b) ECD spectrum (FTMS, 100 scans) obtained from the triply charged ion of the peptide T1. The lower case s symbolizes phosphoserine and q indicates a pyroglutamic acid.
with phosphate loss are observed, and they have low abundance. It has been reported that the extent of loss of phosphate in MS/MS experiments depends on different parameters such as the sequence of peptides, the collision energy, the pressure, and the type of gas used in the collision cell. In this experiment, the collision energy and pressure are relatively low. Furthermore, as discussed above, the proline effect and the acidic effect preclude and circumvent the situation that the CID spectrum is dominated by the loss of phosphate. The positions of the two phosphate groups (at Ser8 and Ser22) are identified without ambiguity. The CID spectrum shows clearly that Ser17 is not phosphorylated.

**ECD Spectrum.** The ECD spectrum of the T1 peptide shows 13 c and z fragmentation ions, which represent 12 cleavages of the 28 possible cleavages and cover 42% of the sequence. We note the presence in the spectra of b fragment ions (b10, b11, b12, b26), and y fragment ions (y18, y19, y23, y33), which are likely the result of the collision process between the isolated ions and residual gas molecules. The b ions observed in the ECD spectrum are consistent with both the proline effect and the “basic residue” effect described above. The complementary ion pair (b12, y18) is particularly abundant and can be explained in this case by the low efficiency of the ECD processes.

The ECD spectrum is furthermore characterized by side-chain fragmentation, primarily involving aspartic acid or glutamic acid (neutral loss of C2H4O2, 60.016 Da) and glutamine (neutral loss of CH4NO, 45.028 Da). No side-chain fragmentation of the two phosphoserines present in the T1 peptide is observed. The low proton affinity of the phosphate group, in comparison with the amide bond, explains this observation. The presence of the c4 and z21 fragmentation ions locates one of the phosphorylation sites at Ser8. The second site of phosphorylation (shown in the CID spectrum at Ser22) cannot be unambiguously identified in the ECD spectrum, due to the lack of fragmentation between Ser17 and Ser22.

In this case, the CID spectrum is richer in structural information (the CID spectrum covers 80% of the sequence of the T1 peptide and localizes both phosphorylation sites). However, the information contained in the ECD spectrum is complementary to the CID spectrum. It confirms the phosphorylation site at the Ser8 and identifies the sequences qDL (residues 1–3) and sE (residues 22–23) of the peptide T1. Combination of data obtained from ECD and CID spectra results in coverage for 90% of the T1 peptide sequence.

Finally, the two spectra of the larger peptides T1 and T2 (Figures 3 and 4) confirm the previously reported tendency for ECD: the provision of extensive sequence information on the terminal regions of a peptide and limited information on the middle region. This tendency was reduced by Horn et al. using “activated Ion ECD” whereby the tertiary structure of the peptide was unfolded in the gas phase, yielding more fragments in the middle of the sequence.

**CONCLUSION**

We demonstrated that the results of tandem mass spectrometry experiments employing ECD and CID techniques, depending on the nature of the peptides sequenced, provide for complementary information. For PRPs, due to the proline effect, the CID spectra containing the b, b-NH3, y, and y-NH3, as well as internal fragment ion series, are more difficult to interpret. Conversely, ECD spectra of the T3 and T4 peptides are simple, containing primarily c and z fragmentation ions, and permit complete sequence coverage without ambiguity. When the backbone amine bond X-Pro is cleaved, no c and z fragmentation ions are formed because the side chain of proline still links the peptide backbone. By ECD, the proline residues

---

can be located in the sequence without ambiguity and fragmentation of all other amide bonds provides sequence information of the rest of the peptide. For the 44-residue T2 peptide, by combination of the information obtained with the two fragmentation methods, coverage is increased from 70 to 91% clearly demonstrating the complementary nature of ECD and CID. For the acidic peptide T1, the ECD efficiency is relatively low but the CID spectrum contains more structural information and is able to localize both phosphorylation sites (Ser8 and 23). Upon combining information from the two techniques, more than 93% of the full length of the PRP-3 protein is covered (Figure 5). The approach used to analyze the PRP-3 will be applied to investigate the structure of uncharacterized proline-rich proteins.

ACKNOWLEDGMENT
This work was supported by National Institutes of Health Grants P41 RR 10888 (to C.E.C.) and R01 DE 05672 and R01 DE 07652 (to F.G.O.).

Received for review February 17, 2002. Accepted June 18, 2002.

AC0255835