Anal. Chem. **2003,** *75,* 1963−1971

Articles

Cleavage N-Terminal to Proline: Analysis of a Database of Peptide Tandem Mass Spectra

Linda A. Breci,[†] David L. Tabb,[‡] John R. Yates, III,[§] and Vicki H. Wysocki*,[†]

Department of Chemistry, University of Arizona, Tucson, Arizona 85721, Department of Genome Sciences, University of Washington, Seattle, Washington 98195, and SR11 Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Fragmentation at the Xxx-Pro bond was analyzed for a group of peptide mass spectra that were acquired in a Finnigan ion trap mass spectrometer and were generated from proteins digested by enzymes and identified by the Sequest algorithm. Cleavage with formation of a + b + y ions occurred more readily at the Xxx-Pro bond than at other locations in these peptides, and the importance of this cleavage varied by the identity of Xxx. The most abundant Xxx-Pro relative bond cleavage ratios were observed when Xxx was Val, His, Asp, Ile, and Leu, whereas the least abundant cleavage ratios occurred when Xxx was Gly or Pro. Rationalization for these cleavage ratios at Xxx-Pro may include contribution of the Asp or His side chain to enhanced cleavage or the conformation of Pro, Gly, and the aliphatic residues Val, Ile, and Leu at the Xxx location in the Xxx-Pro bond. Although unusual fragmentation behavior has been noted for Pro-containing peptides, this analysis suggests that fragmentation at the Xxx-Pro bond is predictable and that this information may be used to improve the identification of proteins if it is incorporated into peptide sequencing algorithms.

Two general approaches are used to analyze proteins by mass spectrometry: a single stage experiment (MS), providing mass analysis of the peptide products of the enzymatically digested protein,[1] or a tandem experiment (MS/MS), providing the amino acid sequence of the peptides.[2] Analysis of biomolecules such as peptides began with fast-atom bombardment (FAB) ionization techniques,[3] and improved with the introduction of matrix-assisted laser desorption ionization (MALDI)[4,5] and electrospray ionization (ESI).[6,7] Choices of and improvements in mass analyzers, such as time-of-flight (TOF),[8] ion-trap,[9] and quadrupole[10] instruments, provide great flexibility in matching experimental method to need. Experimental mass spectra are coupled to computer search algorithms[1,11−15] for protein identification.

Vast amounts of genome information have become available concurrently with advances in mass spectrometry, and protein and DNA databases are searched for sequences to identify proteins. Computer algorithms have been developed to use peptide masses

* To whom correspondence should be addressed. E-mail: vwysocki@u.arizona.edu.

[†] University of Arizona.

[‡] University of Washington.

[§] The Scripps Research Institute.

(1) Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5011−5015.

(2) Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D. *Anal. Chem.* **1995**, *67*, 1426−1436.

(3) Barber, M.; Bordoli, R. S.; Elliot, G. J.; Sedgewick, R. D.; Tyler, A. *Anal. Chem.* **1982**, *54*, 645−657.

(4) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2301−2303.

(5) Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1991**, *63*, A1193−A1202.

(6) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64−71.

(7) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spec. Rev.* **1990**, *9*, 37−70.

(8) Beavis, R. C.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6873−6877.

(9) Louris, J. N.; Brodbelt Lustig, J. S.; Cooks, R. G.; Glish, G. L.; van Berkel, G. J.; McLuckey, S. A. *Int. J. Mass Spectrom. Ion Processes* **1990**, *88*, 117−137.

(10) Yost, R. A.; Boyd, R. B. *Methods Enzymol.* **1990**, *193*, 154−200.

(11) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.

(12) Pappin, D. D. J.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327−332. Available online at http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse, http://prospector.ucsf.edu/ucsfhtml3.2/msfit.htm.

(13) Yates, J. R.; Speicher, S.; Griffin, P. R.; Hunkapiller, T. *Anal. Biochem.* **1993**, *214*, 397−408.

(14) James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. *Biochem. Biophys. Res. Commun.* **1993**, *195*, 58−64.

(15) Mann, M.; Hojrup, P.; Roepstorff, P. *Bio. Mass Spec.* **1993**, *22*, 338−345.

10.1021/ac026359i CCC: $25.00
Published on Web 03/27/2003

(MS) to search protein or translated genome databases and identify proteins in an approach called "peptide mass mapping,"[12,16,17] in which the observed m/z values detected after digestion of the protein with an enzyme are compared with the peptide masses predicted. The protein can be identified if a sufficient number of peptide ions are observed, if the protein is not heavily modified, and for most programs, if there are not more than two proteins present.[18]

Other algorithms are available for use in identification of proteins from the fragmentation (MS/MS) of individual peptides.[2,11,17,19] The use of MS/MS data addresses the drawbacks of peptide mass mapping with the added advantage that fewer peptides are required; while one identified peptide will limit the possible proteins, a second identified peptide may be enough for absolute identification.[18] The Sequest algorithm uses MS/MS data to identify a peptide. The selected database is searched for all possible contiguous series of amino acids that equal the mass of the precursor ion (with mass tolerance determined by the researcher). Structural modifications (e.g., phosphorylation) may be suggested by the user and allowance is made for those mass differences. A cross-correlation function then provides a measurement of similarity between the m/z of the fragment ions predicted from the amino acid sequence and the fragment ions of the experimental tandem mass spectrum. The strength of the cross-correlation scores and the difference between the first- and second-ranked result determines a successful match.[11] Proteins with sequence information available in the selected database can be identified by 1−5 peptides; however, by using multiple enzymatic cleavage reagents, 97−100% of a protein can be identified in an exceptionally good case.[20]

Judging the significance of a result from a peptide mass mapping or tandem sequencing experiment is important, particularly in the case of automated high-throughput protein analysis used today.[21] Protein identification by tandem mass spectrometry methods does not take into account unusual fragmentation events that can occur as a result of the presence of specific residues. For example, the importance of Arg in a peptide has been noted in a statistical analysis of peptide spectra, because its presence can affect both the appearance of certain ion types (a/b vs y) and the mobility of an ionizing proton on the backbone, because Arg sequesters the proton at its very basic side chain.[22,23] The requirement for a mobile proton is met when an Arg-containing peptide with no other basic residues carries two charges as one is sequestered and one is available to induce charge-directed cleavage.[23] It has been demonstrated that preferential cleavage C-terminal to aspartic acid often occurs,[24−26] and it has been noted and is under investigation that product ion spectra may also be influenced by the presence of histidine in a peptide.[27−29] An improved understanding of peptide dissociation and application of that knowledge to sequencing algorithms is one way to improve those algorithms, because an experimental mass spectrum would better match a more accurately modeled MS/MS spectrum created from sequences obtained from DNA or protein databases.

Proline Fragmentation Behavior. The presence of proline residues in a peptide has a significant effect upon the fragments observed by tandem mass spectrometry. The abundance of intense fragment ions formed at a Pro residue, termed the *proline effect*, results from cleavages at the N-terminal side of Pro.[30−35] In fact, in the fragmentation of whole proteins, fragment ions produced N-terminal to proline are dramatically over-represented in the spectra.[36] Fragmentation of Pro-containing peptides, particularly when poly-Pro stretches exist in the peptide, produce unusual fragment ions in high-energy experiments, some of which are suggested to be produced by charge-remote cleavage.[37,38] Proline is unusual in that it is the only naturally occurring amino acid in which the side chain participates in a ring to the peptide backbone. This creates an N-alkylation at the peptide nitrogen of Pro and the ring formed endows that position with a natural turn, which makes this an important residue in the formation of protein higher order structure.[39−41] Although it has been observed previously that Pro-containing peptides show unusual fragmentation effects, no study has been performed to systematically analyze the effect of Pro in the tandem mass analysis of peptides. The statistical analysis of a peptide spectral database presented here was undertaken to answer two questions. First, are there predictable patterns in cleavages observed at Pro residues that would be useful in the sequence identification of peptides using computer algorithms? Second, do any such patterns suggest why cleavage occurs, or does not occur, adjacent to a Pro residue?

(16) Zhang, W. Z.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 2482−2489. Available online at: http://prowl.rockefeller.edu/cgi-bin/ProFound or http://www.proteometrics.com/browl-cgi/ProFound.exe.

(17) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551−3567. Available online at: http://www.matrixscience.com/cgi/searchform.pl?SEARCH=PMF (fragmentation) http://www.matrixscience.com/cgi/searchform.pl?SEARCH=MIS (mass mapping).

(18) Yates, J. R. *J. Mass Spectrom.* **1998**, *33*, 1−19.

(19) Fenyo, D.; Qin, J.; Chait, B. T. *Electrophoresis* **1998**, *19*, 998−1005.

(20) Gatlin, C. L.; Eng, J. K.; Cross, S. T.; Detter, J. C.; Yates, J. R. *Anal. Chem.* **2000**, *72*, 757−763.

(21) Fenyo, D. *Curr. Opin. Biotechnol.* **2000**, *11*, 391−395.

(22) van Dongen, W. D.; Ruijters, H. F. M.; Luinge, H. J.; Heerma, W.; Haverkamp, J. *J. Mass Spectrom.* **1996**, *31*, 1156−1162.

(23) Dongre, A. R.; Jones, J. J.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365−8374.

(24) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142−5154.

(25) Tsaprailis, G.; Somogyi, A.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195/196*, 467−479.

(26) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72*, 5804−5813.

(27) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399−1406.

(28) Farrugia, J. M.; O'Hair, R. A. J.; Reid, G. E. *Int. J. Mass Spectrom.* **2001**, *210*, 71−87.

(29) Huang, Y. Y.; Wysocki, V. H.; Tabb, D. L.; Yates, J. R. *Int. J. Mass Spectrom.* **2002**, *219*, 233−244.

(30) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233−6237.

(31) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Anal. Chem.* **1993**, *65*, 425−438.

(32) Martin, S. A.; Biemann, K. *Int. J. Mass Spectrom. Ion Processes* **1987**, *78*, 213−228.

(33) Tang, X. J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824−2834.

(34) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185−1187.

(35) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1998**, *33*, 505−524.

(36) Schaaff, T. G.; Cargile, B. J.; Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **2000**, *72*, 899−907.

(37) Wang, Y.; Johansson, J.; Griffiths, W. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 2182−2202.

(38) Heerma, W.; Boon, J. J. L.; Versluis, C.; Kruijtzer, J. A. W.; Hofmeyer, L. J. F.; Liskamp, R. M. J. *J. Mass Spectrom.* **1997**, *32*, 697−704.

(39) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283−437.

(40) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol. Biol.* **1998**, *279*, 449−460.

(41) Pal, D.; Chakrabarti, P. *J. Mol. Biol.* **1999**, *294*, 271−288.

The peptide spectral database presented here contains peptides generated primarily with trypsin. All peptides included in the analysis were doubly charged to represent the most common case, because most peptides produced by tryptic digestion and introduced into a mass spectrometer by ESI carry two charges (thus, 2^+ peptides represent the majority of peptides analyzed). In addition, selecting a single charge state for this analysis eliminates the variable of different cleavage effects due to charge state.

EXPERIMENTAL METHODS

Acquisition of Spectra. The peptide spectra provided for analysis were from a variety of proteins and organisms. They were cleaved initially by enzymatic digestion, primarily by trypsin, with chemical cleavage utilized for some of the proteins. The peptides produced were then roughly separated by high-pressure liquid chromatography (HPLC) or other capillary LC methods and introduced into a Thermo Finnigan ion trap mass spectrometer (LCQ) by electrospray ionization. By utilizing the Finnigan software, the tandem mass spectrometry experiment was automated and data-dependent, so continuous acquisition of mass spectra (MS) was performed as the chromatography run continued, and when ion signals of a certain intensity indicated the presence of a peptide, a tandem mass spectrum (MS/MS) was acquired. Multiple peptides often elute simultaneously under these conditions, and a method of acquisition termed dynamic exclusion provided by the Finnigan software avoids the collection of multiple same-ion spectra over a limited time period and allows the acquisition of MS/MS spectra for lower abundance peptides.

The resulting MS/MS spectra were submitted to the Sequest program for data analysis. This program searches a selected database for candidate peptide sequences (which may include the conversion of a DNA database to theoretical amino acids stretches) that match the m/z of the peptide found in the first stage (MS) of the MS/MS experiment. Theoretical MS/MS spectra are then generated for those matching peptides and are compared to the experimental MS/MS spectrum. The best matches between the spectra generated from database searches and experimental mass spectra are those that have the greatest stretches of contiguous matching a, b, or y fragments. The quality of a match is rated with a cross-correlation score, and a peptide is considered "matched" if that score exceeds a selected threshold (determined by the researcher; e.g., a cross-correlation score of 2.5 is generally accepted for doubly charged peptides) and differs from the next highest score by a preselected amount. The individual, theoretical matches from the database are then compared against each other, and the most likely protein matches are those with the greatest percentage of sequence identified, although the program provides the other top-ranked possibilities in the final report, and so the algorithm is useful for both protein mixtures and isolated proteins.

Ions Used for Statistical Analysis of Fragmentation Patterns. The a, b, and y fragment ions were chosen for study because these are the main ions identified by the Sequest algorithm and because they are often the most useful sequencing ions, whether identification is by Sequest or other methods. Briefly, b and y ions are generated by fragmentation at the peptide bond between amino acid residues. Cleavage at that position is most often initiated by an ionizing proton, and if the charge remains on the amino terminal fragment, a b ion is formed and can lose CO to form an a ion. If the charge is transferred to the C-terminal fragment, a y ion results. In the case of doubly charged peptides, which are commonly formed in tryptic digests, two fragment ions can result from a single cleavage, b ions (which may form a) in partnership with y ions, that is, a b/y complementary pair.

Data Conversion and Analysis. Two analyses were performed for Pro cleavage. First, a database of 316 peptides was probed for cleavage C-terminal to proline at Pro-Xxx, where Xxx indicates any amino acid. Of this database, 168 peptides contained Pro and were included in the results shown here. The second analysis included 1043 doubly charged peptide spectra that were inspected for cleavage N-terminal to proline at Xxx-Pro. Peptides excluded from the results include 414 that did not contain any Pro residues, 38 that were not used because of poor spectral quality without confirmation of accuracy in identification, 12 duplicates, and 63 that had the m/z of Xxx-Pro cleavage outside of the m/z range of the experiment. This resulted in a total of 516 peptides for which Xxx-Pro cleavage was analyzed.

Peptides were from a variety of proteins and organisms and had been identified by Sequest. All included spectra were inspected manually to confirm correct assignment. Those that were questionable (low signal intensity, low percentage of fragment ions, many ions of high relative intensity not identified) were manually sequenced by using Protein Prospector[12] to generate theoretical sequence ions, which were used to identify confirming fragment ions, such as internal fragments not identified by the Sequest routine. The 516 peptides analyzed contained single or multiple instances of Xxx-Pro. Any positions of Xxx-Pro that would generate one or more a, b, or y fragment ions outside the range of the experiment were not factored into the average cleavage for that particular Xxx-Pro cleavage.

For each peptide analyzed, a ratio of the total relative abundance of a + b + y fragments formed at Xxx-Pro versus all possible a + b + y ions for that peptide is termed the relative bond cleavage ratio and was determined by the formula

$$\frac{\text{relative abundance } (a + b + y)_{Xxx-Pro}}{\text{relative abundance } (a + b + y)_{all}}$$

The most important sequencing fragment ions are the a, b, and y ions, but these are only some of the fragment ions formed by peptide dissociation. For example, the MS/MS spectrum of the peptide LVEGLANDPENKVPLIK shown in Figure 1 has the a, b, and y ions labeled, and it can be seen that there are many additional fragment ions that are not identified. This analysis is concerned only with a, b, and y fragment ions, and their relative abundances were translated into Excel spreadsheet format. For peptides with multiple Xxx-Pro cleavage positions, the ratios at each of the cleavage sites were summed to determine an overall relative bond cleavage ratio for Xxx-Pro cleavage in the peptide. The average ratio at a particular cleavage site (e.g., Ala-Pro) was determined by summing all ratios produced in all peptides at that cleavage position and dividing by the number of ratios. The results presented for cleavage at Pro-Xxx bonds was calculated in a similar manner using the abundance of a + b + y ions formed at that bond.

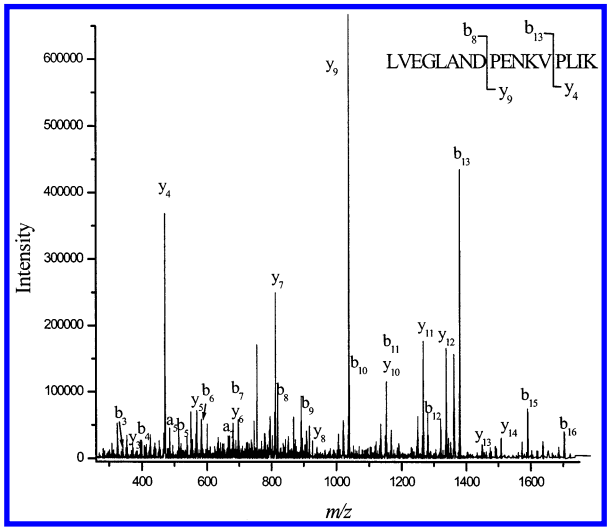

Figure 1. MS/MS spectrum of the peptide [LVEGLANDPENKVPLIK + 2H]$^{2+}$ acquired by CID in an ion trap. Although many peaks are a-, b-, and y-sequence ions, many other peaks are unidentified.

Table 1. Average Pro-Xxx Relative Bond Cleavage Ratios[a] and the Number of Peptides Analyzed

identity of Xxx	av ratio	no. peptides	identity of Xxx	av ratio	no. peptides
Pro	0.045	4	Glu	0.012	30
Gln	0.022	5	Tyr	0.012	3
His	0.020	2	Leu	0.012	24
Gly	0.017	17	Asn	0.012	12
Trp	0.015	1	Phe	0.011	9
Ser	0.015	14	Ile	0.010	16
Asp	0.014	14	Met	0.006	2
Val	0.014	22	Arg	0.004	3
Thr	0.013	9	Lys	0.002	15
Ala	0.013	17	Cys	n/a	0

[a] Ratio of a + b + y ions formed at Pro-Xxx to a + b + y ions formed at all positions.

RESULTS

C-Terminal Fragmentation, Pro-Xxx. Ions formed at positions C-terminal to Pro, where Xxx is any amino acid, were expected to produce low Pro-Xxx relative bond cleavage ratios, in agreement with previous observations in our laboratory and by others.[34] To confirm these previous observations, a small number of peptides (168) were included in this analysis with the intent of expanding the data set if early results suggested an in-depth analysis was necessary. The results of this survey are included here as a general snapshot of the low value of sequencing information generated by cleavage at the Pro-Xxx bond.

The results of this analysis are listed in Table 1, which includes the average ratio and the number of peptides included for each occurrence of Pro-Xxx, where the identity of Xxx is listed in the first column. Low ion abundance results from cleavage at Pro-Xxx, ranging from 4.5% (Pro-Pro) to 0.2% (Pro-Lys). Due to the small data set, some Pro-Xxx positions are poorly represented, although there are 12 or more examples for 10 of the 20 amino acids shown in Table 1. The results of this abbreviated survey were in agreement with previous observations that occurrences of Pro-Xxx cleavage are observed, but those are rare and it may be productive to consider low ion abundance at Pro-Xxx for peptide identification purposes.

Table 2. Average Xxx-Pro Relative Bond Cleavage Ratios,[a] the Standard Deviation, and the Number of Peptides Analyzed

identity of Xxx	av	SD	no. peptides
Val (V)	0.383	0.223	55
His (H)	0.377	0.209	42
Asp (D)	0.367	0.235	35
Ile (I)	0.364	0.202	42
Leu (L)	0.331	0.183	71
Lys (K)	0.291	0.225	67
Glu (E)	0.288	0.143	43
Phe (F)	0.282	0.199	26
Tyr (Y)	0.279	0.160	28
Ala (A)	0.263	0.175	60
Gln (Q)	0.259	0.172	20
Thr (T)	0.209	0.166	25
Asn (N)	0.204	0.160	32
Arg (R)	0.166	0.141	55
Trp (W)	0.114	0.140	7
Ser (S)	0.111	0.076	33
Gly (G)	0.068	0.071	37
Pro (P)	0.056	0.064	27
Poorly Represented Residues:			
Cys (C)	0.613	0.154	2
Met (M)	n/a	n/a	0

[a] Ratio of a + b + y ions formed at Xxx-Pro to a + b + y ions formed at all positions.

N-Terminal Fragmentation, Xxx-Pro. A comprehensive analysis was made of cleavages at Xxx-Pro positions in 516 peptides. On average, it was found that 36.3% of a + b +y ion intensity observed in these 516 peptides was due to cleavages at Xxx-Pro bonds. This result is surprisingly large, and a closer inspection of these results is in order. The average relative intensity based upon the identity of Xxx is listed in Table 2, and this table also lists the standard deviation and numbers of instances analyzed. Of the 20 amino acids, 18 were represented by enough data to provide meaningful results (≥20), and although the number of representative cleavage sites for Trp (W) is low (7), it must be considered that the average abundance of Trp in proteins is one of the lowest of all amino acids at 1.1%.[42] Although Cys and Met also occur at low abundance in proteins (2.8% and 1.7% respectively),[42] there were only 2 (Cys) and 0 (Met) instances of these amino acids, and they are considered under-represented in this analysis.

These results are graphically represented in Figure 2, and it can be seen that the confidence intervals shown in this figure are too large to make definitive statements about the relative ordering of cleavage intensities based upon the identity of Xxx, and that was not the goal of the analysis. The results show that, first, there is a good deal of cleavage N-terminal to proline residues, and second, certain clues to Xxx-Pro cleavage propensity may be discovered in those individual results shown in Table 2 and Figure 2. In Figure 2, the amino acid representing Xxx is plotted on the *x* axis. The *y* axis is the average relative bond cleavage ratio for that Xxx-Pro bond. No simple relationship between fragment ion abundance and the identity of Xxx based upon the amino acid's physical properties, such as ionization energy,[43] proton affinity,[44]

(42) Klapper, M. H. *Biochem. Biophys. Res. Commun.* **1977**, *78*, 1018−1024.

Table 3. His-Pro Relative Bond Cleavage Ratios for Peptides That Contain a His-Pro Bond, with Ratios of a + b + y Ion Intensities Formed at That Position vs All a + b + y Ion Intensity in the Peptide

IAGVVY**HP**SNNELVR	0.654	GSIDEQ**HP**R	0.271
DKY**HP**GYFGK	0.617	VLEH**HP**R	0.258
IGAW**HP**AHVM	0.601	NFNSQGVQISSGAL**HP**LNK	0.233
KY**HP**DELAK	0.596	GKGSIDEQ**HP**R	0.218
DKY**HP**GYFGKVGM	0.557	YVLEH**HP**R	0.200
VCHA**HP**TLSEAFK	0.515	SKNMHL**HP**L	0.190
MTVSLI**HP**IAMDQGLR	0.470	ENVWFS**HP**R	0.146
AGNFAL**HP**EVVR	0.457	MGKGAIDEQ**HP**R	0.069
IIGLDYH**HP**DFEQESK	0.402	HENVWFS**HP**R	0.062
GGQAIVV**HP**DYLYGAITEDR	0.375	AHENVWFS**HP**R	0.060
Y**HP**GYFGKVGM	0.304	FLTG**HP**K	0.033

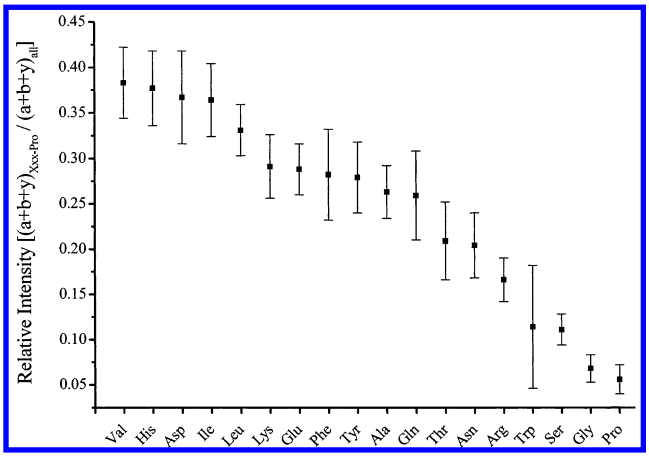

Figure 2. Average Xxx-Pro relative bond cleavage ratios. The *x* axis is the amino acid represented by Xxx in the Xxx-Pro bond, plotted versus the ratio of ion intensity formed at a + b + y ions at Xxx-Pro vs other positions in the peptides and shown with error bars (80% confidence interval).

or size,[45] is evident in these data. Instead, a combination of factors determines how and when fragmentation occurs N-terminal to Pro.

DISCUSSION

The most abundant Xxx-Pro relative bond cleavage ratios are observed when Xxx is Val, His, Asp, Ile, and Leu, whereas the least abundant cleavages occur when Xxx is Gly or Pro. These residues may be divided into three categories for further discussion: acidic/basic residues (His, Asp), Pro (Pro-Pro cleavage), and aliphatic residues (Val, Ile, Leu, Gly). It must be considered that results of cleavage of particular Xxx-Pro bonds can become quite complicated when multiple Xxx-Pro sites exist in a peptide or when residues that are implicated in enhanced fragmentation effects, such as His and Asp (without Pro), are located in a particular peptide. Thus, the error bars shown in the graphical result chart of Figure 2 are large; however, some general trends can still be observed.

Acidic/Basic Residues. Two of the amino acids that result in intense cleavage when positioned N-terminal to Pro are His (0.377) and Asp (0.367). It has been shown that both the basic residue His and the acidic residue Asp may be implicated in preferential ion formation at bonds C-terminal to these residues whether the next residue is Pro or another amino acid. Enhanced cleavage has been noted at protonated His in peptides,[27] with protonation expected on the His side chain in doubly charged peptides if there is one or fewer other basic site and that basic site is sufficiently separated from the His residue so that the need for separation of charge is not a factor.[46] Thus, it is not surprising that His-Pro cleavage scores high in this analysis. In the 14 peptides in which His-Pro was the only Xxx-Pro bond in the peptide and His was not located as the third-to-last residue (C-terminal), intensities at His-Pro were at least 23.3% of a + b + y ions formed (Table 3). When His was the 3rd to last residue, ion intensity at the His-Pro bond was suppressed, likely because protonation at His does not occur when His is two residues away from Arg or Lys in these tryptic peptides and the requirement for charge separation is not met. When multiple examples of Xxx-Pro existed in a peptide, then strong cleavage was observed at His-Pro unless the identity of one of the other positions was Val-Pro, Lys-Pro, or Asp-Pro. In those cases, preferred cleavage occurred at the other sites (and His-Pro cleavage was suppressed), or cleavage intensities were split between the multiple sites.

Preferential cleavage has been noted C-terminal to Asp in peptides, and studies in the Wysocki laboratory have shown that preferential cleavage occurs at Asp primarily when the number of basic residues (Arg) equals or exceeds the number of protons located on the peptide.[24−26,29] This suggests that when the proton is sequestered by the basic residue, it is unavailable to initiate cleavage at other sites in the peptide.[47] Under those circumstances, the acidic hydrogen of Asp may initiate cleavage, resulting in anhydride formation at the Asp C-terminus.[26] It is expected that Asp-Xxx relative bond cleavage ratios will be large regardless of the identity of Xxx when the conditions of basic sites and charge state are met; thus, it is not surprising that Asp-Pro cleavage scores high in this analysis. Three factors lead to a high probability that the number of charges will equal the number of basic residues for the peptides examined in this study: (1) the peptides were tryptic (and so contain at least one Arg or Lys); (2) tryptic digests are not always complete and may contain a Lys/Arg, Lys/Lys, or Arg/Arg combination and His residues may be present; and (3) fragmentation for peptides carrying a 2^+ charge provide good

(43) Campbell, S.; Marzluff, E. M.; Rodgers, M. T.; Beauchamp, J. L.; Rempe, M. E.; Schwinck, K. F.; Lichtenberger, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 5257−5264.
(44) Campbell, S.; Beauchamp, J. L.; Rempe, M.; Lichtenberger, D. L. *Int. J. Mass Spectrom. Ion Processes* **1992**, *117*, 83−99.
(45) Counterman, A. E.; Clemmer, D. E. *J. Am. Chem. Soc.* **1999**, *121*, 4031−4039.
(46) Sullards, M. C.; Reiter, J. A. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 40−53.
(47) Gu, C.; Somogyi, A.; Wysocki, V. H.; Medzihradszky, K. F. *Anal. Chem. Acta* **1999**, *397*, 247−256.

Table 4. Asp-Pro Relative Bond Cleavage Ratios for Peptides That Contain an Asp-Pro Bond, with Ratios of a + b + y Ion Intensities Formed at That Position vs All a + b + y Ion Intensity in the Peptide

RVV**DP**FTR	0.945	KTDLFADL**DP**SQYFETR	0.462
TLR**DP**VIYR	0.745	KNVEVVALN**DP**FISNDYSAY	0.372
KVWL**DP**NETSEIAQANSR	0.659	VWL**DP**NETSEIAQANSR	0.268
ILYGHLD**DP**HGQDIQR	0.613	GLVWEGSVL**DP**EEGIR	0.246
AFHDTEWATQ**DP**R	0.588	VNVGQL**DP**IGDK	0.183
GHLD**DP**HGQDIQR	0.560	AIGSYEQAYYVGT**DP**K	0.072
LITSHLVDT**DP**EVDSIIKDEIER	0.560	NAEFIIAN**DP**DADR	0.072
GDLGVEIG**DP**ELVGVQKK	0.476	FVSG**DP**LLHTTAWK	0.044

performance in Sequest, so a "self-selection" of peptides with two basic residues occurs when a minimum Sequest score is used to generate a data set.

Peptides that contain an Asp-Pro bond (average ratio 0.369) but no other Xxx-Pro combinations are listed with their a + b + y cleavage ratios in Table 4. It can be seen in the table that when Arg is the C-terminal residue and the peptide contains one additional basic residue (Arg, Lys or His), as in eight of the first nine peptides, that a + b + y ion intensity ratios for cleavage at Asp-Pro equal or exceed 0.462. The seven peptides of intensity lower than 0.462 did not contain a second basic residue (when Arg was C-terminal) or did not contain a C-terminal Arg. As suggested by previous studies, the acidic hydrogen of the Asp residue could initiate cleavage at the Asp-Xxx bond, greatly increasing observed fragment ion intensity at that position. The fragment ions resulting from such a cleavage and containing the original N-terminus would be neutralized at the now C-terminal Asp, forming an anhydride-terminated b ion and forming a truncated peptide from the fragment containing the original C-terminus. The charge held at each basic site would provide the single charge to each of the two fragment ions formed. Intense Asp-Pro bond cleavage is observed when the number of basic residues in a peptide equal or exceed the number of protons added, but it is particularly pronounced when the terminating residue of the intact peptide is Arg. One possible difference in Arg-terminating peptides that must be considered is formation of a salt bridge, which has been observed in peptides.[48−50] A salt bridge formed between the protonated N-terminus (+), deprotonated C-terminus (−), and protonated Arg side chain (+) would "tie up" the proton, preventing localization to a position to initiate charge-directed cleavage.

Proline Effect at Pro-Pro. Cleavage at Pro commonly occurs N-terminal to the Pro residue, often providing abundant b or y fragment ions from cleavage at that bond (the proline effect).[30−35] The brief investigation into Pro-Xxx cleavage provided evidence that fragmentation at the C-terminal side of Pro is a far more unusual event, and this is also demonstrated by the very low average of a + b + y ions formed (0.056) at the Pro-Pro bond. It may be mechanistically difficult for the five-member proline ring to participate in necessary cyclic ion intermediate formation or in the formation of a bicyclic 5−5 ring neutral if Pro is the C-terminal residue of an N-terminal neutral fragment. The fragmentation of peptides containing stretches of Pro residues produce unusual results containing some b ions of unknown structure and various ions thought to be generated by charge remote processes.[37] Unusual fragmentation at Pro-Pro bonds may be the result of steric constraints. NMR studies of cis/trans isomerization in oligopeptides found the Pro-Pro bond to have one of the slowest isomerization rates in the peptide series, even though that bond has a high cis abundance in proteins; thus, the authors hypothesize that an accumulation of Pro residues leads to a decelerated reorganization of local structure because of increased steric constraints.[40]

Aliphatic Residues. Perhaps the most unexpected result of this statistical analysis of peptide fragmentation was the strong cleavage ratios observed for Val-Pro, Ile-Pro, and Leu-Pro. Val, Ile, and Leu have no functional groups to perform chemistry, such as the acid/base effects of Asp or His, and yet they clearly provide an enhancement to Xxx-Pro cleavage when positioned N-terminal to a Pro residue. What Val, Ile, and Leu have in common is that they have branched aliphatic side chains. The detailed results for these residues and discussion of the unusual fragmentation effects observed follows.

Cleavage at Val-Pro. Intense cleavage was observed at the Val-Pro bond, and peptides that contain a Val-Pro bond (average relative bond cleavage ratio 0.383) but no other Xxx-Pro combinations are listed with their a + b + y cleavage ratios in Table 5. Strong fragment ions were almost always produced at Val-Pro, and reduction in cleavage ratio score was often due to a distribution of ratios among many bond sites (not just Xxx-Pro), which reduced the individual Val-Pro score even when cleavage at Val-Pro produced relatively intense spectral peaks. Some exceptions include two related peptides FR**VP**TVDVSVVDLTVK (0.074) and R**VP**TVDVSVVDLTVK (0.009), both of which produced intense peaks C-terminal to Asp, demonstrating a likely dominance of Asp-Xxx cleavage when the number of basic residues equals or exceeds the number of protons in a peptide. This is not always the case, however. For example, the Asp-containing peptide KDQYVPEVSALDLSR cleaved preferentially at VP (VP cleavage = 0.759), even though it contained two basic residues and two protons. One possible explanation for this result is charge solvation of one proton by Lys, Arg, and the C-terminal carbonyl oxygen, with the second proton available to initiate backbone cleavage.

Cleavage at Ile-Pro. Intense cleavage was also often observed at the Ile-Pro bond. Peptides that contain an Ile-Pro bond (average relative bond cleavage ratio 0.364) but no other Xxx-Pro combinations are listed with their a + b + y cleavage ratios in Table 6. Similar to Val-Pro, strong fragment ions were often produced at

(48) Wyttenbach, T.; von Helden, G.; Bowers, M. T. *J. Am. Chem. Soc.* **1996**, *118*, 8355−8364.
(49) Freitas, M. A.; Marshall, A. G. *Int. J. Mass Spectrom.* **1999**, *182/183*, 221−231.
(50) Schnier, P. D.; Price, W. D.; Jockusch, R. A.; Williams, E. A. *J. Am. Chem. Soc.* **1996**, *118*, 7178−7189.

Table 5. Val-Pro Relative Bond Cleavage Ratios for Peptides That Contain a Val-Pro Bond, with Ratios of a + b + y Ion Intensities Formed at That Position vs All a + b + y Ion Intensity in the Peptide

EAGL**VP**STSEAIR	0.826	IDMLV**VP**LRDE	0.377
TVH**VP**GTTVTHTVR	0.780	SVAV**PV**DILDHDNNYELK	0.365
KDQY**VP**EVSALDLSR	0.759	EGT**VP**TDLDQETGLAR	0.352
AEHVGVLHVVG**VP**SISAQ	0.754	ENFALIHVSLV**PV**IHGEQK	0.330
STYH**VP**LLLLEQK	0.699	HGIIF**VP**LGYK	0.246
KGH**VP**AHIIEQR	0.677	VGVLHVVG**VP**SISAQAK	0.232
LK**VP**AINVNDSVTK	0.590	AEHVGVLHVVG**VP**SISAQAK	0.229
MKSHW**VP**NTI	0.535	KGEQELEGLTDTT**VP**K	0.158
IIEHD**VP**EHFFGELAK	0.511	PANLVDLN**VP**AK	0.106
EYG**VP**VILHTDHAAK	0.499	FR**VP**TVDVSVVDLTVK	0.074
FGWDTHG**VP**IEHIIDK	0.493	R**VP**TVDVSVVDLTVK	0.009
L**VP**DELTVALVK	0.481	WFGQAGN**VP**HDDILR	0.000
HVVG**VP**SISAQAK	0.444		

Table 6. Ile-Pro Relative Bond Cleavage Ratios for Peptides That Contain an Ile-Pro Bond, with Ratios of a + b + y Ion Intensities Formed at that Position vs All a + b + y Ion Intensity in the Peptide

LLNG**IP**LDEEETR	0.843	NQILVSGE**IP**STLNEESK	0.502
HSITL**IP**SENFTSK	0.658	GAHQF**IP**IGFGIK	0.454
VSIVEI**IP**FEEEQNIR	0.639	KDEA**IP**ELVEGQTFDADVE	0.448
GASQNI**IP**SSTGAAK	0.557	HINVG**IP**VK	0.367
TASGNI**IP**SSTGAAK	0.549	INKDIEEEGDS**IP**LEQSTNGK	0.356
HLGELLG**IP**HDLVWR	0.517	DIQ**IP**VYDTFDGSDLR	0.276
LGAEEITAD**IP**NVGESALSK	0.516	FG**IP**QISTGDMFR	0.155

Table 7. Leu-Pro Relative Bond Cleavage Ratios for Peptides That Contain a Leu-Pro Bond, with Ratios of a + b + y Ion Intensities Formed at That Position vs All a + b + y Ion Intensity in the Peptide

DFFE**LP**LTDEQVEK	0.677	YG**LP**QLSEEAFDELSDK	0.259
SLGLK**LP**LSVINVSAQR	0.618	IHEASFV**LP**TWAAK	0.254
HFASYAN**LP**GTITHGM	0.594	DVIIDEAE**LP**IK	0.211
VGYT**LP**SHIISTSDVTR	0.479	FHATNE**LP**FLFK	0.191
AQYNEIQGWDHLSL**LP**TFGAK	0.355	RLA**LP**DNVVIR	0.151
GVEVV**LP**VDFIIADAFSADANTK	0.352	TIEEIFLHS**LP**VK	0.149
YNEIQGWDHLSL**LP**TFGAK	0.323	HGHLGF**LP**R	0.147
KIFVLAEEN**LP**SLGYK	0.292	AL**LP**HLTNAIVETNK	0.131
ISLG**LP**VGAIMNCADNSGAR	0.289	FGGNAQQTAQ**LP**R	0.128
DTGNAYNSLAD**LP**ANLIK	0.276	ITTIEEIFLHS**LP**VK	0.106
VGVQVAAQHNTSVFYG**LP**QEK	0.273	YG**LP**QLSEEAFDELSDK	0.046

Ile-Pro. In those cases for which cleavage ratios were low, such as DIQ**IP**VYDTFDGSDLR (0.276) and FG**IP**QISTGDMFR (0.155), the Ile-Pro cleavage was intense but was only one of several enhanced cleavage positions in the peptide.

Cleavage at Leu-Pro. Leu-Pro provided the greatest number of examples of the residues studied and was also an important location of cleavage. Peptides that contain a Leu-Pro bond (average relative bond cleavage ratio 0.331) but no other Xxx-Pro combinations are listed with their a + b + y cleavage ratios in Table 7. Similar to Val-Pro and Ile-Pro, strong fragment ions were often produced at Leu-Pro, and reduction in relative bond cleavage ratio score was often due to a distribution of cleavage throughout the peptide, which reduced the individual Leu-Pro score. An exception to this was cleavage for the lowest scoring peptide YG**LP**QLSEEAFDELSDK, for which dominant fragment ions were formed at the Leu-Ser bond in addition to the distribution of cleavage throughout the peptide.

Cleavage at Xxx-Pro, when Xxx = Val, Ile, or Leu, continued to produce intense fragments even in the presence of other Xxx-Pro bonds. The data do not demonstrate any trends ordering Val, Ile, or Leu by increasing fragmentation efficiency, because many particular examples show both high and low cleavages in the presence of Pro C-terminal to one of these aliphatic residues. For example, Ile-Pro has peptide ratios both lower and higher than the Val-Pro cleavage ratio (0.095:0.415 and 0.517:0.256) when both are present in the same peptide.

Cleavage at Gly-Pro. In opposition to what was observed for Val, Ile, and Leu, cleavage intensities are suppressed at Gly-Pro bonds, with an average relative bond cleavage ratio of 0.068. This effect has a few exceptions, such as in the peptide TIVWN**GP-P**GVFEFEK (0.308), in which intense y ion formation at Gly-Pro may have been promoted by the Gly-Pro-Pro combination at that position.

In the peptides of this study, Val-Pro, Ile-Pro, and Leu-Pro bonds produced intense a + b + y ions with very few exceptions, which suggests that for these residues, intense cleavage ions may be expected. An exception to enhanced cleavage at Xxx-Pro occurred in the presence of an Asp residue when the number of basic residues equaled the number of protons with enhanced cleavage at Asp.[24−26] Conversely, Gly-Pro bonds did not produce abundant fragment ions or did so with little peak intensity. This suggests that a valid prediction is suppressed ion intensity at Gly-Pro positions in peptides.

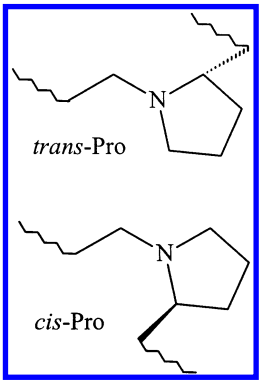

Figure 3. The Xxx-Pro bonds shown as trans (top) and cis (bottom).

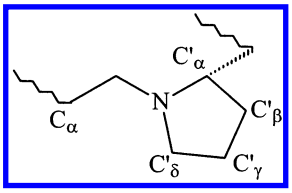

Figure 4. Steric conflict between the C′$_\delta$ of Pro and the side-chain group on the C$_\alpha$ of the preceding residue determine the cis/trans relationship of the Xxx-Pro bond.

Conformational Effects of Proline. In an effort to understand enhanced cleavages that occur N-terminal to Pro residues without chemical effects of the preceding residue, we must consider the unusual conformation of Pro in a protein. X-ray and NMR spectroscopy methods are used to elucidate the tertiary structure of proteins. Cis/trans isomerization at the Xxx-Pro bond of oligopeptides in solution has been shown to be a local effect that correlates to the crystalline structure of proteins.[40] Ion mobility studies have shown that this effect may also be observed in the gas phase, because two conformations of peptides that contain a Pro residue have been observed, and these forms were attributed to the cis/trans forms of Pro.[51] For amino acid residues in general, the double bond character of the peptide bond results in a trans dihedral angle of 180° or a cis dihedral angle of 0°, with trans highly favored,[39] and cis bonds occur in only 0.03[52]−0.05%[53] of the cases. This low occurrence of cis is due to the energy difference of 2.5 kcal/mol between trans and cis and also due to a rotational barrier of ∼20 kcal/mol.[54,55] However, for a bond to Pro (Xxx-Pro), the trans isomer (Figure 3, top) is favored over the cis (Figure 3, bottom) by only 0.5 kcal/mol,[56] and the energy barrier for rotation is reduced to 13 kcal/mol,[41] resulting in higher abundance of cis bonds N-terminal to Pro of 5.2%-6.5%.[41,53,57] The driving force determining this unique cis/trans relationship at Xxx-Pro bonds is the steric conflict between the side-chain group on the C$_\alpha$ of the Xxx residue and the C′$_\delta$ of proline (Figure 4). That relationship has been shown to vary on the basis of the identity of Xxx in the Xxx-Pro bond.[40,41]

Studies of X-ray crystal data have provided details of the percent cis found by identity of Xxx in the Xxx-Pro bond, with values ranging from <3% for Asp-Pro to 11.2% for Pro-Pro,[41] and from 1.8% for Ile-Pro to 12.4% for Trp-Pro.[40] Gly-Pro was found to produce cis bonds in one of the higher abundances (>7%,[41] 7.4%[40]), whereas for the branched aliphatic residues Val, Ile, and Leu cis bonds (for the Xxx-Pro bond) were found to be of the lowest frequency (1.8−3.4%.)[40,41] In fact, the slowest rates of cis/trans isomerization of Xxx-Pro bonds in oligopeptides was observed by NMR to be the branched aliphatic residues Val and Ile.[40] Thus, the Gly-Pro and Pro-Pro bonds, which show limited propensity to cleave have higher cis abundance, while Val-Pro, Ile-Pro, and Leu-Pro which provide intense fragment ions, have lower cis abundance. Although the propensity to form a relatively low abundance of cis-isomers does not explain why Val-Pro, Ile-Pro, and Leu-Pro cleavages are enhanced, it suggests that the steric arrangement of the Val, Ile, and Leu side chains might play a role in the formation of the trans isomer state for cleavage of the Xxx-Pro bond.

The branched aliphatic residues (Val, Ile, and Leu) impart conformational restrictions in the formation of secondary structure,[41,58] while the amino acid Gly imparts little restriction to secondary structure. Although gas-phase volumes of amino acids have been shown to be somewhat smaller than in the core of proteins, Val, Ile, and Leu pack less efficiently than most amino acids and occupy a larger space than would be ascribed if their van der Waals radii alone were considered.[45] Cross sections determined by ion mobility and molecular modeling methods for oligomers Val$_n$, Ile$_n$, and Leu$_n$ (n = 4−9) have been reported to be larger than for more polar residues, with the side chains extending radially away from the solvated core of the peptides.[59] It may be considered that greater fragmentation occurs at Xxx-Pro bonds when Xxx is Val, Ile, or Leu, because the bulky branched aliphatic side chains of these residues restrict conformation, preventing rotation of the bond to the cis isomer and perhaps producing a "reactive" conformation that directly leads to product ions. This suggests that enhanced cleavage may occur C-terminal to Val, Ile, and Leu (at the Val-Xxx, Ile-Xxx, and Leu-Xxx bonds), regardless of the identity of the Xxx residue. In contrast, the flexibility of the Gly-Pro bond as demonstrated by its high cis abundance in proteins leads to a reduction in bond fragmentation/ion formation at Gly-Pro because the lack of a side chain means that there is no conformational restriction that "locks" the molecule into a "reactive" conformation. In other words, the high propensity of Gly to form cis isomers is an indication of its conformational mobility, which does not favor a specific reactive conformation, whereas the low propensity of cis isomerization for Val, Ile, and Leu is an indication that these molecules are sterically locked into a reactive conformation.

CONCLUSIONS

The statistical analysis of a spectral database to investigate cleavage N-terminal to Pro residues in peptides provided unique evidence of the predictive nature of ions formed from dissociation of that bond. Strong ions formed at Xxx-Pro when Xxx was His, Asp, Val, Ile, Leu, and weak or no ions formed when Xxx was Gly

(51) Counterman, A. E.; Clemmer, D. E. *Anal. Chem.* **2002**, *74*, 1946−1951.
(52) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291−304.
(53) Stewart, D. E.; Sarkar, Q.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*.
(54) LaPlanche, L. A.; Rogers, M. T. *J. Am. Chem. Soc.* **1964**, *86*, 337−341.
(55) Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212−4216.
(56) Maigret, B.; Perahia, D.; Pullman, B. *J. Theor. Biol.* **1970**, *29*, 275−291.
(57) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397−412.
(58) Kleywegt, G. J.; Jones, T. A. *Structure* **1996**, *4*, 1395−1400. Supporting materials, available online at: http://xray.bmc.uu.se/gerard/rama
(59) Henderson, S. C.; Li, J. L.; Counterman, A. E.; Clemmer, D. E. *J. Phys. Chem. B* **1999**, *103*, 8780−8785.

or Pro. Dissociation at the Xxx-Pro bond is clearly important because these ions provided 36.3% of the total a + b + y ion intensity formed for the 516 peptide spectra analyzed. Cleavage results at Xxx-Pro when Xxx was His or Asp was not surprising, because strong cleavage is expected C-terminal to His and Asp in peptides, and this effect is amplified when they are located N-terminal to Pro. Cleavage behavior when an aliphatic residue was N-terminal to Pro was surprising and provided important insight into the proline effect. Residue specific fragment ion intensity (strong cleavage at Val, Ile, and Leu; weak cleavage at Gly and Pro) suggests that the conformation of the peptide influences cleavage N-terminal to Pro. Greater fragmentation occurs at Xxx-Pro bonds when Xxx is Val, Ile, or Leu because the bulky aliphatic side chains of these residues restrict conformation "locking the peptide into a reactive structure", whereas the flexibility of the Gly-Pro bond leads to a reduction in bond fragmentation/ion formation at that position.

The identification of proteins by mass spectrometry might be improved if sequencing algorithms used in tandem mass spectrometry experiments included chemical effects in the fragmentation models that predict spectra from sequence strings. Unusual fragmentation behavior has been noted for Pro-containing peptides, and we have shown here that the relative abundance for cleavage of Xxx-Pro bonds may be predicted. This predictive behavior will be tested in a sequencing algorithm to determine the benefits to peptide and protein identification.[60]

ACKNOWLEDGMENT

We gratefully acknowledge funding of this work by NIH through a grant to V.W. (GM R0151387).

Received for review November 27, 2002. Accepted February 21, 2003.

AC026359I

(60) Huang, Y. Y.; Triscari, J. M.; Wysocki, V. H. Poster presented at the 2002 ASMS Conference on Mass Spectrometry and Allied Topics, Orlando, FL.