J|A|C|S
COMMUNICATIONS

Published on Web 02/24/2004

Dissociation Behavior of Doubly-Charged Tryptic Peptides: Correlation of Gas-Phase Cleavage Abundance with Ramachandran Plots

Yingying Huang,[†] Joseph M. Triscari,[‡] Ljiljana Pasa-Tolic,[§] Gordon A. Anderson,[§] Mary S. Lipton,[§] Richard D. Smith,[§] and Vicki H. Wysocki*,[†]

Department of Chemistry, University of Arizona, Tucson, Arizona 85721, Science Application International Corporation, Tucson, Arizona 85711, and Pacific Northwest National Laboratory, Richland, Washington 99352

Received August 20, 2003; E-mail: vwysocki@u.arizona.edu

Large numbers of gas-phase dissociation spectra of protonated peptides are obtained daily and used in protein identification studies. Yet fundamental knowledge of the factors that influence their unimolecular dissociation branching ratios is relatively poor. It is still not possible to predict dissociation branching ratios from a peptide sequence. Clearly, several chemical factors must influence dissociation patterns, including the ψ,ϕ angles determined by the residues involved in an amide bond, the propensities for certain side chains to interact with each other or with the backbone, the tendency for added protons to be intramolecularly solvated, and the stability of the fragment ions once formed.[1−4] Studies of unimolecular dissociation over the past several years have focused on simple model systems[5−9] that were designed to systematically vary structure and determine the influence of the changes on the dissociation pattern. A complementary approach is to acquire a large number of spectra from a great variety of sequences and use a "data mining" approach to glean as much chemical information as possible from the dataset.[10−13] Attempts can then be made to correlate the gleaned information with information known from other studies, e.g., preferred bond angles compiled from the protein databank, as shown below.

Peptide MS/MS spectra from the proteome of the organism *Shewanella oneidensis* were collected using ThermoFinnigan LCQ ion-trap instruments. The SEQUEST algorithm[14] was used to assign peptide sequences to these spectra. Accurate Mass Tag (AMT) measurements[15] by FT-ICR were used to confirm the assigned sequence, and PeptideProphet[16] was used to filter spectra, resulting in 16 738 spectra for which sequences can be assigned with high confidence. The filtered spectra include 5654 doubly charged tryptic peptides (terminating in Lys or Arg) for which charge-directed fragmentation is expected (5 to 38 residues, median and average 16 residues). Analysis of fragmentation of the amide bonds of these 5654 unique doubly charged tryptic peptides are reported here. A more detailed study of the entire dataset is in progress.

Ion types investigated here correspond to cleavage of amide bonds with retention of charge on either the N-terminal fragment (**b** ion) or the C-terminal fragment (**y** ion). Singly charged **b** and **y** ions were first identified from each spectrum according to the assigned sequence and were normalized to the most abundant peak in the same ion series. The normalized abundances were then catalogued by the pair of the amino acid residues at the cleavage site as the relative abundances for such cleavages. After this process was repeated for all Xxx−Zzz amino acid combinations represented in the 5654 spectra, two tables were generated. Figures 1a and 1b illustrate the average bond cleavage abundance at each amino acid combination for **b** and **y** ions, respectively, with the color scheme corresponding to cleavage abundance shown in Figure 1c. Columns and rows for Cys and Trp were purged because of their low occurrence in the data set. Columns for Arg and Lys do not exist for **y** ions because of the low mass cutoff inherent to ion-trap instruments.

The extremely prominent cleavage N-terminal to Pro (column labeled "P") and the extremely weak cleavage C-terminal to Pro and Gly (rows labeled "P" and "G") for both **b** and **y** ions are consistent with previous studies.[12] Proline, the only cyclic amino acid, has long been known to undergo enhanced cleavage at its N-terminus.[8,12] The dominance of enhanced cleavage N-terminal to Pro (between prePro and Pro) and suppressed cleavage C-terminal to Pro and Gly suggests a fundamental dependence of gas-phase peptide fragmentation on conformational constraints. Ramachandran plots, based on protein crystal structure analyses, show that allowed and preferred backbone ψ,ϕ angles for Pro, Gly, and prePro residues differ from those of the general case (all other amino acids).[17] The correlation between those residues that contribute to enhanced/suppressed cleavage in gas-phase peptides and those residues that occupy enhanced or limited ψ,ϕ space in protein structures suggests that the factors that control protein backbone conformation, side chain rotamers, and distortions of CαCβ also contribute to dissociating conformations in gas-phase ions. This is not meant to suggest that a gas-phase protonated peptide has a structure identical to that of a similar sequence in a protein crystal but rather that gas-phase peptide fragmentations are assisted by local sterically favored backbone and side-chain conformations that favor a particular intramolecular nucleophilic attack and/or a particular charge solvation structure (e.g., favorable prePro carbonyl...H+....Pro carbonyl[18]). When plots similar to Figure 1 are prepared for datasets corresponding to peptides with mobile vs nonmobile protons (protons free to transfer intramolecularly, e.g., peptides in Figure 1, vs protons localized at a basic residue), it is clear that cleavage is enhanced at prePro-Pro only when a mobile proton is available to the peptide backbone. In addition, spectra of a pentapeptide with L- vs D-Pro show a different dominant dissociation fragment,[19] and results for the 6-ring analogue of Pro show strong C-terminal cleavage.[8] These results further support a conformational influence on gas-phase prePro-Pro protonation and subsequent fragmentation.

Related trends are observed that were previously suggested[12] for a more narrowly defined set of peptides. These trends might also be explained by correlations to Ramachandran plots: preferential cleavage C-terminal to β-branched aliphatic residues, Ile and Val, was observed for both **b** and **y** ions and may result from conformational restrictions imposed by the β-branched side-chain; Ile and Val are known to have allowed and preferred areas of Ramachandran plots that occupy slightly lower percentages of total

[†] University of Arizona.
[‡] Science Application International Corporation.
[§] Pacific Northwest National Laboratory.

10.1021/ja038041t CCC: $27.50

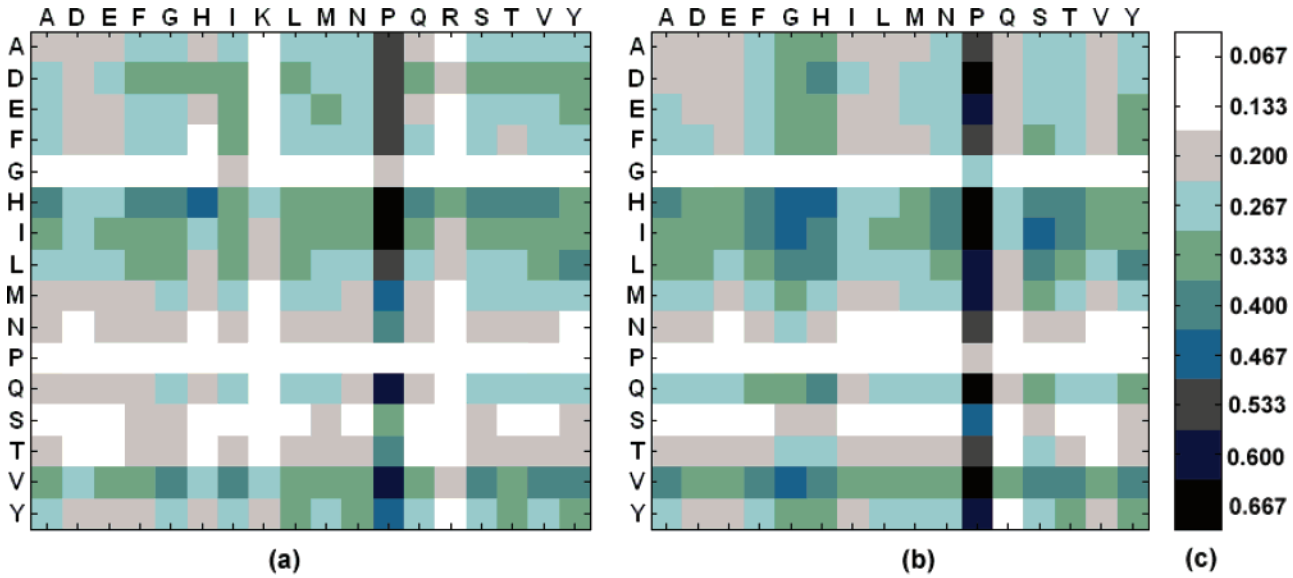

Figure 1. Average relative abundances of bond cleavages at specific amino acid combinations for (a) **b** ions and (b) **y** ions. The single-letter code of the amino acids listed in the leftmost column correspond to the N-terminal amino acid (Xxx) in an Xxx−Zzz pair, while those listed along the topmost row correspond to the C-terminal amino acid (Zzz). Each number in panel c represents the minimum value of the average abundance for the corresponding color.

ψ,ϕ space than the other "general" residues.[20] Leu, with a γ-branched side-chain, shows less enhanced cleavage compared to Ile and Val.

Several trends in the data of Figure 1 appear to be related to the ability of the side chains to act as proton acceptors/donors. Relatively strong cleavage C-terminal to histidine in **b** ion formation, as reported for model peptides and smaller sets of spectra,[1,11,21] is also observed here. This is presumably the result of the localization of the charge on the histidine side chain and initiation of a charge-directed cleavage. Serine, in both **b** and **y** ions, shows weak cleavage at its C-terminal amide bond but relatively strong cleavage at its N-terminal amide bond. H-bonding intermediates may explain this observation. The oxygen on the Ser side chain may be involved in solvating a proton attached to the neighboring carbonyl oxygen, providing a structure with an electropositive carbon that can be attacked by the adjacent N-side carbonyl to form a **b** ion (or a **y** if a proton subsequently transfers). A seven-member ring is formed if such H-bonding takes place N-terminal to the Ser side-chain and is more stable than the six-member ring formed[22] if such H-bonding is C-terminal to the Ser side-chain. A lack of strongly enhanced cleavage at Asp, Glu is anticipated for the 2+ ions investigated here; cleavage is initiated by the added "mobile" proton.[1,5,6,10]

The results show that there exists great variability in relative cleavage efficiencies between different amino acid residue combinations in doubly protonated gas-phase peptides. This variability is dominated by enhanced or suppressed cleavage at those residues (Gly, Pro, prePro) that have Ramachandran ψ,ϕ plots that differ from the general case. For nonaliphatic side chains, enhancement or suppression of cleavage provides insight on possible intramolecular chemical interactions. Using results of this type to guide both experimental studies of fragmentation of model peptides and computational modeling of fragmentation transition states will allow a more complete knowledge of unimolecular dissociation and its correlation to conformational constraints to be obtained.

Acknowledgment. We thank Dr. Ken J. Auberry for assistance in data transfer. This research was supported by NIH grant GM R0151387 to V.H.W. and a 2003 Pfizer Graduate Research Fellowship in Analytical Chemistry to Y.H. Research at Pacific Northwest National Laboratory (PNNL) was supported by the U.S. Department of Energy, Office of Biological and Environmental Research. PNNL is operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830.

Supporting Information Available: Count for each of the amino acid residue combinations analyzed. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. *J. Mass Spectrom.* **2000**, *35*, 1399−1406.
(2) O'Hair, R. A. J. *J. Mass Spectrom.* **2000**, *35*, 1377−1381.
(3) Schlosser, A.; Lehmann, W. D. *J. Mass Spectrom.* **2000**, *35*, 1382−1390.
(4) Polce, M. J.; Ren, D.; Wesdemiotis, C. *J. Mass Spectrom.* **2000**, *35*, 1391−1398.
(5) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W. Q.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142−5154.
(6) Gu, C. G.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000**, *72*, 5804−5813.
(7) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365−8374.
(8) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185−1187.
(9) vanDongen, W. D.; Ruijters, H. F. M.; Luinge, H. J.; Heerma, W.; Haverkamp, J. *J. Mass Spectrom.* **1996**, *31*, 1156−1162.
(10) Huang, Y.; Wysocki, V. H.; Tabb, D. L.; Yates, J. R. *Int. J. Mass Spectrom.* **2002**, *219*, 233−244.
(11) Tabb, D. L.; Smith, L. L.; Breci, L. A.; Wysocki, V. H.; Lin, D.; Yates, J. R. *Anal. Chem.* **2003**, *75*, 1155−1163.
(12) Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H. *Anal. Chem.* **2003**, *75*, 1963−1971.
(13) Kapp, E. A.; Schutz, F.; Reid, G. E.; S., E. J.; Moritz, R. L.; O'Hair, R. A. J.; Speed, T. P.; Simpson, R. J. *Anal. Chem.* **2003**, *75*, 6251−6264.
(14) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(15) Lipton, M. S. et. al. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11049−11054.
(16) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. *Anal. Chem.* **2002**, *74*, 5383−5392.
(17) Lovell, S. C.; Davis, I. W.; Adrendall, W. B.; de Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct. Funct. Genet.* **2003**, *50*, 437−450.
(18) Addario, V.; Guo, Y. Z.; Chu, I. K.; Ling, Y.; Ruggerio, G.; Rodriquez, C. F.; Hopkinson, A. C.; Siu, K. W. M. *Int. J. Mass Spectrom.* **2002**, *219*, 101−114.
(19) Breci, L.; Kuppannan, K.; Herrmann, K.; Vaisar, T.; Laskin, J.; Futrell, J. H.; Wysocki, V. H. Manuscript in preparation.
(20) Kleywegt, G. J.; Jones, T. A. *Structure* **1996**, *4*, 1395−1400.
(21) Tsaprailis, G.; Nair, H.; Zhong, W.; Kuppannan, K.; Futrell, J. H.; Wysocki, V. H. *Anal. Chem.* Submitted.
(22) Meot-Ner, M. *Int. J. Mass Spectrom.* **2003**, *227*, 525−554.

JA038041T