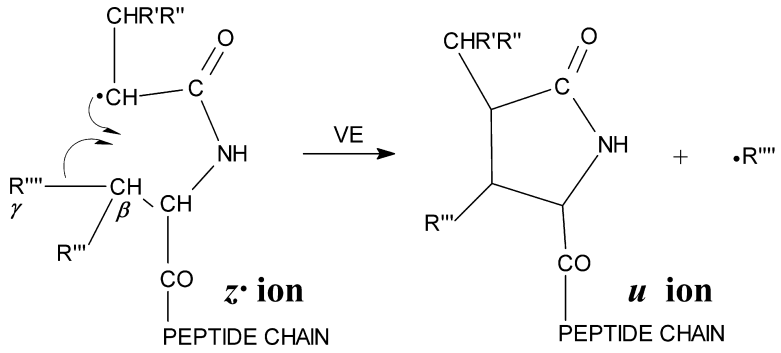
Communication

Secondary Losses via □-Lactam Formation in Hot Electron Capture Dissociation: A Missing Link to Complete *de Novo* Sequencing of Proteins?

Frank Kjeldsen, and Roman Zubarev

J. Am. Chem. Soc., **2003**, 125 (22), 6628-6629• DOI: 10.1021/ja035042c • Publication Date (Web): 08 May 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 05/08/2003

Secondary Losses via γ-Lactam Formation in Hot Electron Capture Dissociation: A Missing Link to Complete *de Novo* Sequencing of Proteins?

Frank Kjeldsen[†,‡] and Roman Zubarev*[,†]

University of Southern Denmark, Department of Chemistry, Campusvej 55, DK-5230 Odense M, Denmark, and Uppsala University, Laboratory for Biological and Medical Mass Spectrometry, Uppsala, Sweden.

Received March 7, 2003; E-mail: Roman.Zubarev@Angstrom.uu.se

De novo sequencing of proteins is an important task when genome sequence information is not available or inaccurate, as in many proteomic studies.[1−4] Complete de novo sequencing with mass spectrometry (MS) remains a challenge, especially distinguishing between Xle (isoleucine and leucine) amino acid residues and identification of the two first N-terminal residues. The latter is mainly due to the absence of fragmentation at the N-terminal amide bond and identical nominal masses of many amino acid pairs. To overcome this problem, condensed-phase approaches using one cycle of Edman degradation followed by mass determination of the truncated peptides have been suggested by Hunt et al.[5] More recently, Gaskell and co-workers suggested tandem MS analysis of N-terminal phenyl isothiocyanate derivatives of tryptic peptides, yielding complementary b_1 and y_{n-1} fragment ions.[6] Reid et al. combined N-terminal derivatization with gas-phase fragmentation of neutral peptides via ion−molecule reactions with acylium ions.[7]

The challenge of determining the Xle residue identities by MS has traditionally been met by high-energy collision-activated dissociation (HE CAD).[8] Recently, a new technique of hot electron capture dissociation (HECD)[9] has been introduced that produces secondary losses in the side chains of N-terminal amino acids of $z^\bullet$ fragment ions. The formed *w* fragments revealed the identities of 20 out of 25 Xle residues in enzymatic peptides of the PP3 bovine milk protein.[10] Of the remaining five Xle residues, two N-terminal ones could be a subject of a two-cycle Edman degradation reaction, as in the method by Hunt et al.

Here we report on the novel reaction in HECD that afforded identification of the remaining three Xle residues. This achievement finalizes MS-based sequence characterization[11] of a 15 kDa protein, which, to the best of our knowledge, has become possible for the first time. The mechanisms of secondary losses yielding *w* fragments and *u* ions (our nomenclature) are shown in Figure 1. Both fragmentations involve a radical-site initiated bond breakage. For w_n ion formation from $z^\bullet_n$ ions, the cleavage occurs between the β- and γ-carbons in the N-terminal side chain forming a double bond between the α- and β-carbons.

The production of C-terminal secondary u_n ions can be explained by the creation of a five-membered lactam (γ-lactam) via bond formation between the N-terminal α-carbon atom and the β-carbon in the adjacent amino acid residue of the $z^\bullet_n$ ion. This type of fragmentation, first reported by Biemann and co-workers[12] for HE CAD, is now found in HECD spectra of many polypeptides.

Figure 2 presents the HECD mass spectra obtained with Fourier transform mass spectrometry. Electrospray-produced dications and tetracations of PP3 peptides 25−41 (2+, *m*/*z* 965) and 77−109 (4+, *m*/*z* 894) were isolated and irradiated for 300 and 200 ms by 9 and 10 eV electrons, respectively. In these spectra, secondary *d* and *w* ions from totally three Xle residues were either absent or

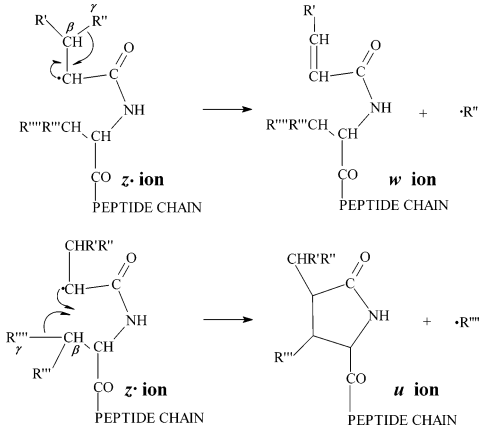


Figure 1. The proposed mechanisms for *w* and *u* ion formation from the precursor $z^\bullet$ radical ions.

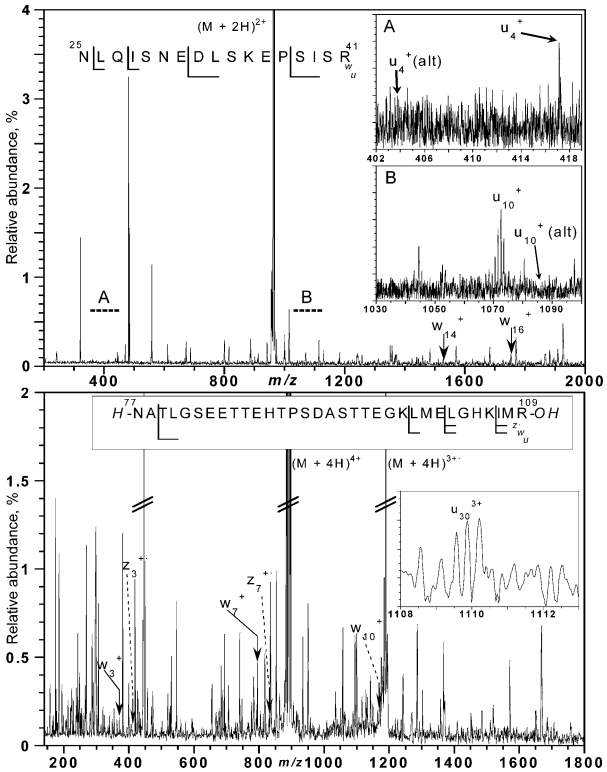


Figure 2. HECD spectra of PP3 peptides 25−41 (2+, *m*/*z* 965) and 77−109 (4+, *m*/*z* 894). The u_4, u_{10}, and u_{30} ions determine the residues as Leu_{33}, Ile_{39}, and Leu_{80}, respectively. Evidence for H-atom losses was observed for the u_{10} and u_{30} ions.

overlapping with other backbone fragment ions. The presence of the u_4 and u_{10} ions for peptide 25−41 and u_{30} ions for peptide 77−109 identifies these residues as Leu_{33}, Ile_{39}, and Leu_{80}. In all cases,

[†] University of Southern Denmark.
[‡] Uppsala University.

10.1021/ja035042c CCC: $25.00

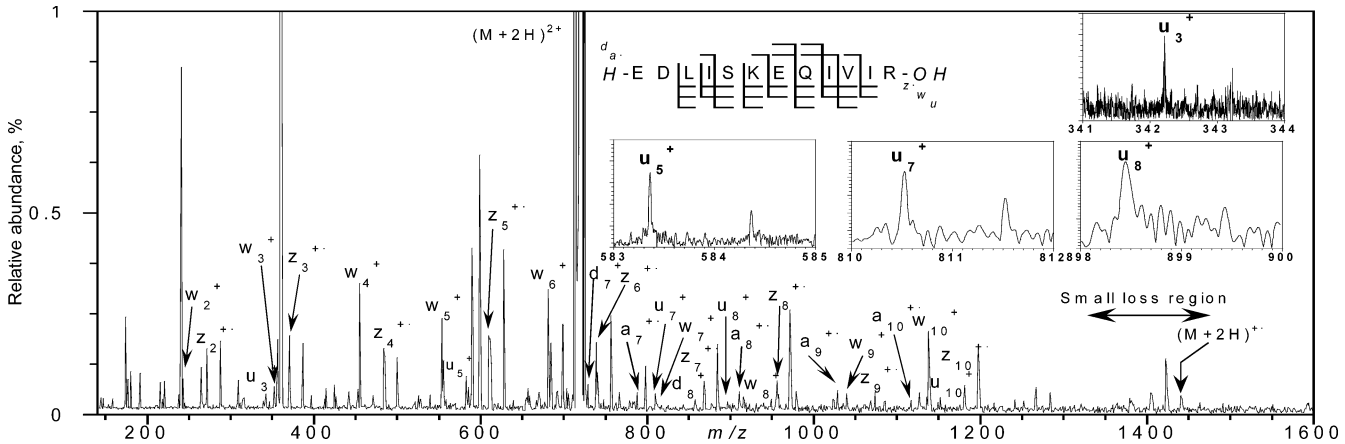


Figure 3. Broadband HECD mass spectrum of 2+ ions of the peptide EDLISKEQIVIR with assignments of secondary fragment ions and their precursors.

the *u* ions corresponding to alternative assignments were absent. The u_{10} and u_{30} ions showed evidence of H-atom losses. This was an isolated observation, which is currently under further investigation.

While Biemann et al. observed one *u* ion (from Leu) in their study,[12] the origin of the *u* ions reported here is much more general. As an example of this, Figure 3 shows a HECD spectrum of the dodecapeptide EDLISKEQIVIR. This peptide contains only residues that are expected to undergo secondary fragmentation since they all carry nonaromatic side chains on their β-carbon, as suggested by Biemann and co-workers.[12] The spectrum exhibits two *d*, nine *w*, and five *u* ions. *w* ions from direct side-chain loss were observed for all but three residues (E, D, and R). The *u* ions denoted u_3, u_5, u_7, u_8, and u_{10} originated from secondary side-chain losses in the Ile_2, Ile_4, Glu_6, Lys_7, and Ile_9 residues, respectively. All five *u* ions were positively identified with mass deviations $\leq$6 ppm and signal-to-noise ratios (S/N) of $\geq$2.0 except for u_8 (S/N = 1.3). Three out of four Xle residues could be identified via *u* ions alone.

Formation of *u* ions is a minor channel of $z^{\bullet}$ ion fragmentation compared to that giving *w* ions. Hence, from the nine $z^{\bullet}$ ions in the spectrum in Figure 3, nine *w* ions and only five secondary *u* ions were formed; the total abundance of *u* ions was 24% of that of *w* ions. The only exception was the u_7 ion (S/N = 2.3 vs S/N = 1.5 for w_7). The preferential secondary loss from the Glu side chain seems to be a general trend.

To understand the *u*/*w* branching ratio, enthalpies of formation of *w* and *u* ions were calculated by the semiempirical PM3 method for peptides IL, KK, and VVR. The *w* ions turned out to be by 60−90 kJ/mol *less* stable than *u* ions. The higher rates of *w* ion formation in HECD are believed to be due to the entropy factors that favor radical-initiated losses on shorter distances.

Interestingly, the conventional low-energy ECD[13] of the same peptide as in Figure 3 gave two *u* ions. The abundance of u_5 was 60% lower and u_7 30% higher than in HECD; u_3, u_8, and u_{10} were absent. This was not surprising since *w* ions have also been reported in ECD at lower abundances than in HECD.[10]

When two neighboring amino acids in a sequence are identical, w_n and u_n ions become isomeric. A complication can arise in the case where two different Xle residues are adjacent. Then the u_n ion can be confused with the w_n ion of the alternative assignment.[10] Such an example is found for the isomeric analogue SDREYP**I**LIR of the tryptic decapeptide SRP from *saccharomyces cerevisiae*. Here, HECD gave rise to an abundant (S/N = 12.4) w_4 ion at mass-to-charge ratio (*m*/*z*) 469 as expected for Ile_7, and a less abundant (S/N = 3.1) u_4 ion at *m*/*z* 455 as expected for Leu_8. The same *m*/*z* corresponds to an w_4 ion from a Leu_7 residue. However, the factor of 4 difference in the abundances between these two ions enabled the distinguishing.[9] Additional complication exits for the identification of N-terminal Xle residues. For the first Xle, *u* ions do not exist, while its *w* ions overlap with small losses from the reduced molecular species.[14,15] This was verified for two isomeric peptides ISDRIELYPIALR and LSDRLEIYPLAIR, HECD spectra of which looked identical in terms of Xle_1 identification. PP3 sequence is the most unfavorable case with two N-terminal Xle residues. The identification of the second Xle hinges on the presence of its *w* ion, which in the case of the N-terminal peptide 1−24 (ILNKPE-DETHLEAQPTDASAQFIR) was overlapping with other ions. In more favorable cases of the N-terminal sequence, full MS-only sequence characterization should be possible.

Acknowledgment. The work was supported by the Danish National Research Councils (Grants SNF 9891448, 51-00-0358, and STFV 0001242 to R.Z., STFV 26-01-0058 to R.Z. and F.K.). We thank Esben S. Sørensen, Aarhus University, for the PP3 protein.

References

(1) Chang, Y.; Hong, H.; Kim, B.; Hall, S.; Settineri, T. *Int. J. Mass. Spectrom.* **2001**, *209*, 47.
(2) Uttenweiler-Joseph, S.; Neubauer, G.; Christoforidis, S.; Zerial, M.; Wilm, M. *Proteomics* **2001**, *1*, 668.
(3) Cagney, G.; Emili, A. *Nature* **2002**, *20*, 163.
(4) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.; Begley, T. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 672.
(5) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. **1986**, *83*, 6233.
(6) Summerfield, S. G.; Steen, H.; O'Malley, M.; Gaskell, S. J. *Int. J. Mass Spectrom.* **1999**, *188*, 95.
(7) Reid, G. E.; Tichy, S. E.; Pérez, J.; O'Hair, R. A. J.; Simpson, R. J.; Kenttämaa, H. I. *J. Am. Chem. Soc.* **2001**, *123*, 1184.
(8) Medzihradszky, K. F.; Burlingame, A. L. *Methods: A Companion to Methods in Enzymology*; Academic Press: San Diego, 1994; Vol. 6, p 284.
(9) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201.
(10) Kjeldsen, F.; Haselmann, K. F.; Sørensen, E. S.; Zuberav, R. A. *Anal. Chem.* **2003**, *75*, 1267.
(11) The full sequence except the identities of the two N-terminal Xle residues was obtained as follows. Nine ECD and three low-energy CAD mass spectra of five Arg-C peptides and seven Lys-C peptides gave cleavages of all 134 interresidue bonds. The total sequence obtained by this "shot-gun" approach was aligned using seven overlaps between the peptides. As a final step, HECD was performed for Xle identification.
(12) Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. *Anal. Chem.* **1987**, *59*, 2621.
(13) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265.
(14) Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241.
(15) Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Polfer, N. C.; Zubarev, R. A. *Eur. J. Mass Spectrom.* **2002**, *8*, 461.

JA035042C