Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2003

Impact Factor: 5.64 · DOI: 10.1021/ac020422m · Source: PubMed

4 AUTHORS:

Frank Kjeldsen
University of Southern Denmark
71 PUBLICATIONS **2,238** CITATIONS
SEE PROFILE

Kim Haselmann
Novo Nordisk
45 PUBLICATIONS **2,429** CITATIONS
SEE PROFILE

Esben S Sørensen
Aarhus University
93 PUBLICATIONS **2,788** CITATIONS
SEE PROFILE

Roman Zubarev
Karolinska Institutet
251 PUBLICATIONS **10,935** CITATIONS
SEE PROFILE

Anal. Chem. **2003,** *75,* 1267−1274

Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Frank Kjeldsen,[†] Kim F. Haselmann,[†] Esben S. Sørensen,[‡] and Roman A. Zubarev*[,†,§]

Department of Chemistry, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, Department of Molecular and Structural Biology, Aarhus University, Aarhus, Denmark, and Division of Ion Physics, Department of Material Sciences, Uppsala University, Uppsala, Sweden

In hot electron capture dissociation (HECD), multiply protonated polypeptides fragment upon capturing ~11-eV electrons. The excess of energy upon the primary c, z˙ cleavage induces secondary fragmentation in z˙ fragments. The resultant w ions allow one to distinguish between the isomeric Ile and Leu residues. The analytical utility of HECD is evaluated using tryptic peptides from the bovine milk protein PP3 containing totally 135 amino acid residues. Using a formal procedure for Ile/Leu (Xle) residue assignment, the identities of 20 out of 25 Xle residues (80%) were determined. The identity of an additional two residues could be correctly guessed from the absence of the alternative w ions, and only two residues, for which neither expected nor alternative w ions were observed, remained unassigned. Reinspection of conventional ECD spectra also revealed the presence of Xle w ions, although at lower abundances, with 44% of all Xle residues distinguished. Using a dispenser cathode as an electron source, identification of four out of five Xle residues in a 2.7-kDa peptide was possible with one acquisition 2 s long, with identification of all five residues by averaging of five such acquisitions. Unlike the case of high-energy collision-induced dissociation, no d ions were observed in the HECD of tryptic peptides.

Primary structures (sequences) determine higher order structures and ultimately functions of peptides and proteins. Determination of polypeptide sequences is one of the important tasks in biology and an integral part of the proteomics studies. Among the 20 most common naturally occurring amino acids, 2 (leucine (Leu) and isoleucine (Ile)) are constitutional isomers, which complicates their distinction by mass spectrometry. The combined average frequency of the two residues is an impressive 16.4%, with even higher abundance in the important transmembrane proteins. Thus, for a peptide sequence longer than three residues, the probability is higher than 50% that at least one of these residues is either Ile or Leu. The Xle (Ile or Leu) uncertainty reduces the chances of unequivocal protein identification by tandem mass spectrometry and multiplies the number of oligonucleotide primers needed for expression of the protein.

In tandem mass spectrometry, the isomeric side chains $CH(CH_3)(CH_2CH_3)$ and $CH_2CH(CH_3)_2$ of the Ile and Leu residues, respectively, are distinguished by inducing secondary fragmentation in the side chains together with backbone bond dissociation. Until recently, this has only been possible by high-energy collision-activated dissociation (HE CAD) on magnetic sector or hybrid instruments,[1,2] where even-electron d and w fragment ions are formed. If Ile is adjacent to the backbone cleavage site, the mass of the d or w ions is 14.0157 Da (mass of the CH_2 group) larger than in the case of Leu, which distinguishes these two residues. Most modern instruments utilize vibrational (low-energy) excitation techniques. In vibrational excitation of polypeptide cations with CAD,[3,4] SID,[5,6] BIRD,[7,8] or IRMPD,[9] the preferred dissociation channels are the losses of small groups and the backbone C−N (peptide bond) cleavage that yields even-electron b and y′ fragments (using the recently introduced[10] nomenclature).

Using high-energy CAD, Johnson et al.[11] determined 23 out of 36 Ile/Leu/nLeu residues distributed among 33 peptides. And recently, Fernandez-de-Crossio et al.[12] demonstrated that a com-

* Corresponding author. Phone: +46 1847 17209. Fax: +46 1855 5736. E-mail: roman.zubarev@angstrom.uu.se.

[†] University of Southern Denmark.

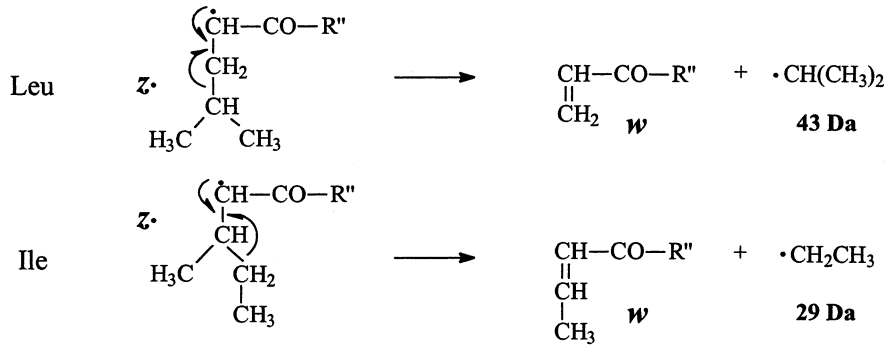
[‡] Aarhus University.

[§] Uppsala University.

(1) Medzihradszky, K. F.; Burlingame, A. L. *Methods: Companion Methods Enzymol.* **1994**, *6*, 284−303.
(2) Schoen, A. E.; Amy, J. W.; Ciupek, J. D.; Cooks, R. G.; Dobberstein, P.; Jung, G. *Int. J. Mass. Spectrom. Ion. Processes* **1985**, *65*, 125−140.
(3) Jennings, K. R. *Int. J. Mass Spectrom. Ion Phys.* **1968**, 227−235.
(4) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233−6237.
(5) Williams, E. R.; Henry, K. D.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. *J. Am. Chem. Soc.* **1990**, *112*, 413−416.
(6) Chorush, R. A.; Little, D. P.; Beu, S. C.; Wood, T. D.; McLafferty, F. W. *Anal. Chem.* **1995**, *67*, 1042−1046.
(7) Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859−866.
(8) Dunbar, R. C.; McMahon, T. B. *Science* **1998**, 194−197.
(9) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Conner, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809−2815.
(10) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201−206.
(11) Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. *Anal. Chem.* **1987**, *59*, 2621−2625.

10.1021/ac020422m CCC: $25.00
Published on Web 01/09/2003

Scheme 1. w Ion Formation



puter-based sequencing program for interpretation of HE CAD spectra could identify 83 out of 86 Ile/Leu residues from 60 different peptides. Despite these achievements, the trend away from magnetic sector instruments is unlikely to be reversed.

An alternative approach to distinguishing Xle residues employs solution-phase complexation with copper(II) species. Vibrational excitation by low-energy CAD of gas-phase metal−peptide complexes produces radical peptide cations that exhibit characteristic losses from Xle side chains. So far, the utility of the technique has only been demonstrated on small tripeptides containing one Xle residue.[13] It is not clear at the moment whether this approach will work for larger species with several Xle residues.

Recently, we demonstrated that electron capture dissociation with hot (3−13 eV) electrons (HECD) produces abundant w ions.[10] It has been suggested that the hot electrons lose their kinetic energy in inelastic collisions with polycations, with subsequent capture of a thermalized electron. Similar to the conventional ECD, this capture induces fast nonergodic ($<10^{-12}$ s) dissociation of the N−C_α bond to produce c′, z• and to a much lesser extent a•, y′ fragment ions. The 3−13-eV energy excess in HECD induces secondary fragmentation in the side chains. The formation of d and w ions implies characteristic losses from a• and z• fragment ions from Ile and Leu of •CH_2CH_3 (29.0391 Da) and •$CH(CH_3)_2$ (43.0548 Da) groups, respectively. These losses are by ∼1 eV endothermic and are initiated through radical-induced fragmentation of the a• and z• species, respectively. Because of the dominance of z• ions over a• fragments, only w ions have been reported in HECD.[10]

For the SRP peptide dications studied in detail in ref 10, the w fragments appeared in the spectra at the electron kinetic energies exceeding 4−5 eV. Since this publication, we discovered that, for some other peptides, w ions can also be observed in the conventional ECD mode. Apparently, the excess of energy in ECD, however small (exothermicity of the ion−electron recombination is 4−7 eV), can be sufficient to overcome the barrier for dissociation of z• ions. In this paper, we investigate mainly the following three issues. First was the analytical utility of HECD for distinguishing Xle residues in proteolytic fragments taken from one single protein. Solving this task with 100% efficiency would mean a big step toward true de novo protein sequencing by mass spectrometry. Since w ions arise from C-terminal z• fragment ions, HECD was expected to be particularly useful for tryptic peptides that have a basic residue near the C-terminus. The second issue was the relative utility of HECD versus the conventional ECD for Xle identification. Additionally, the question of the HECD efficiency was addressed through determination of the minimal conditions for obtaining spectra sufficient for distinguishing Ile/Leu. As a test molecule, we selected the bovine milk protein PP3. This protein contains 135 residues, of which 25 are Leu or Ile (18.5%). This is somewhat above the average, which is rather typical for biologically important proteins.

EXPERIMENTAL SECTION

Materials and Sample Preparation. PP3 from bovine milk (gi:741536) was digested with Arg-C (Boehringer) and separated by HPLC (Vydac C_{18}). The posttranslational modifications of PP3 (two O-glycosylations, one N-glycosylation, and five phosphorylations) were characterized earlier by Sørensen and Petersen.[14] Confirmation of their sites and types was outside the scope of the current work. To avoid the complication of dealing with posttranslational modifications, some proteolytic peptide sequences were synthesized in-house using an EPS221 automatic peptide synthesizer (Intavis AG) and conventional solid-phase Fmoc strategy. All peptides were dissolved in a water−methanol−acetic acid mixture (49:49:2 v/v) to a total concentration of $\sim10^{-6}$ M and sprayed through gold-coated nano-ESI capillary needles (MDS Proteomics, Odense, Denmark) at a flow of 100−1000 nL/h.

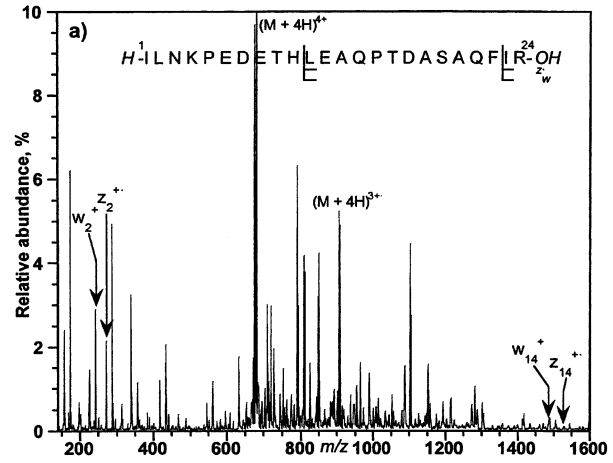
Mass Spectrometry. A 4.7-T Ultima (Ionspec, Irvine, CA) FT mass spectrometer was used in the experiments. Molecular ions were externally accumulated in a hexapole (Analytical of Bradford) for 500−1000 ms and transformed to the cell via a 1-m-long rf-only quadrupole. The ions were captured in the cell using gated trapping followed by isolation of the precursor species having $z \geq 2$ by a stored waveform. The isolated species were subjected to electron irradiation for 50−300 ms. Instead of the standard filament-based electron source, a 20-mm^2 indirectly heated dispenser cathode (STD200, HeatWave) was installed for ECD and HECD experiments.[15] The voltage drop on the filament was 5.2 V. The electron kinetic energies in the center of the ICR cell were <1 and 11 eV in ECD and HECD cases, respectively. Typically, 100−200 acquisitions ("scans") each lasting 2 s were averaged

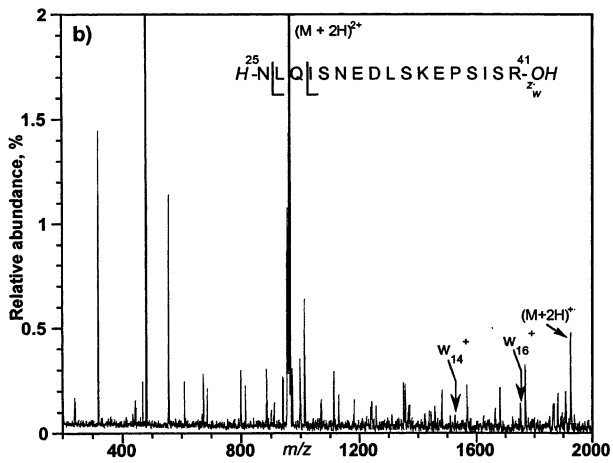
(12) Fernandez-de-Crossio, J.; Gonzales, J.; Betancourt, L.; Besada, V.; Pedron, G.; Shimonishi, Y.; Takao, T. *Rapid commun. Mass. Spectrom.* **1998**, *12*, 1867−1878.

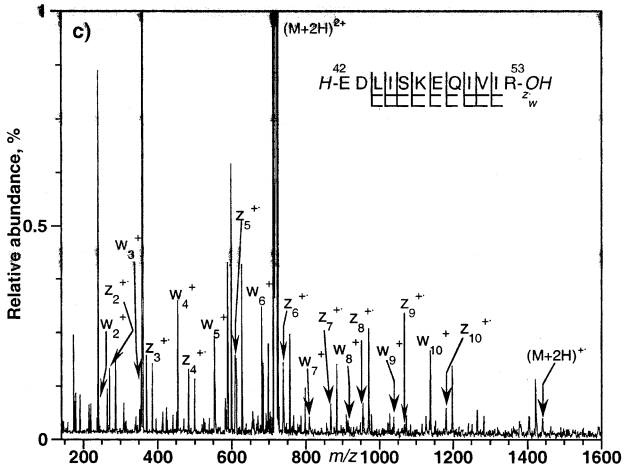
(13) Wee, S.; O'Hair, R. A. J.; McFadyen, W. D. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 884−890.

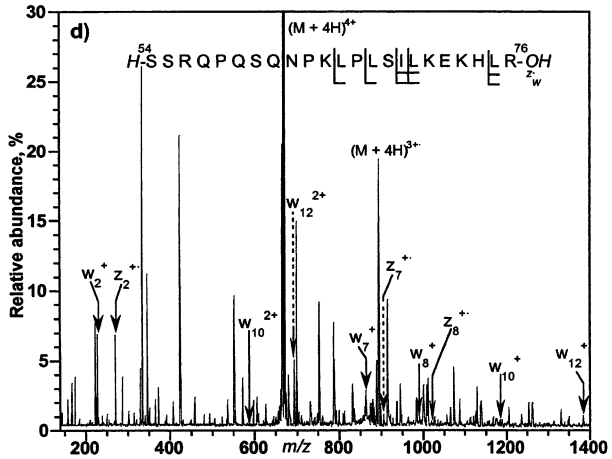
(14) Sørensen, E. S.; Petersen, T. E. *J. Dairy Res.* **1993**, *60*, 535−542.

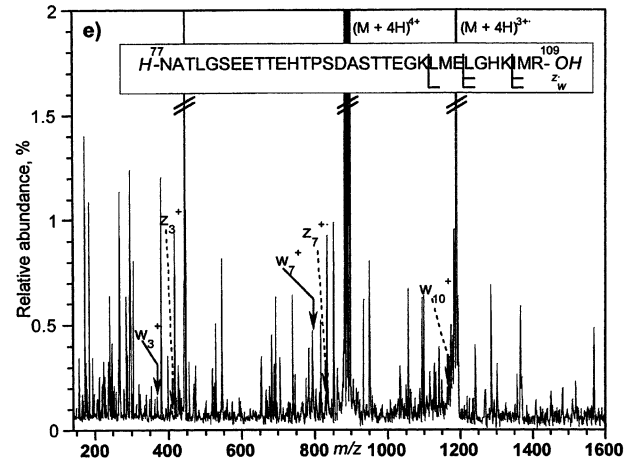
(15) Tsybin, Y. O.; Håkansson, P.; Budnik, B. A.; Haselmann, K. F.; Kjeldsen, F.; Gorshkov, M.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1849−1854.

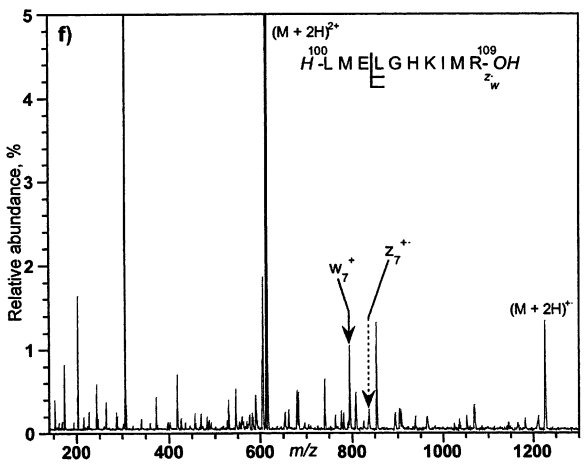

a)
(M + 4H)4+
H-1ILNKPEDETHLEAQPTDASAQFIR-24OH
(M + 4H)3+
Relative abundance, %
m/z


b)
(M + 2H)2+
H-25NLQISNEDLSKEPSISR-41OH
(M+2H)+·
Relative abundance, %
m/z


c)
(M+2H)2+
H-42EDLISKEQIVIR-53OH
(M+2H)+·
Relative abundance, %
m/z


d)
(M + 4H)4+
H-54SSRQPQSQNPKLPLSIILKEKHLR-76OH
(M + 4H)3+
Relative abundance, %
m/z


e)
(M + 4H)4+
(M + 4H)3+
H-77NATLGSEETTEHTPSDASTTEGKLMELGHKIMR-109OH
Relative abundance, %
m/z


f)
(M + 2H)2+
H-100LMELGHKIMR-109OH
(M + 2H)+·
Relative abundance, %
m/z

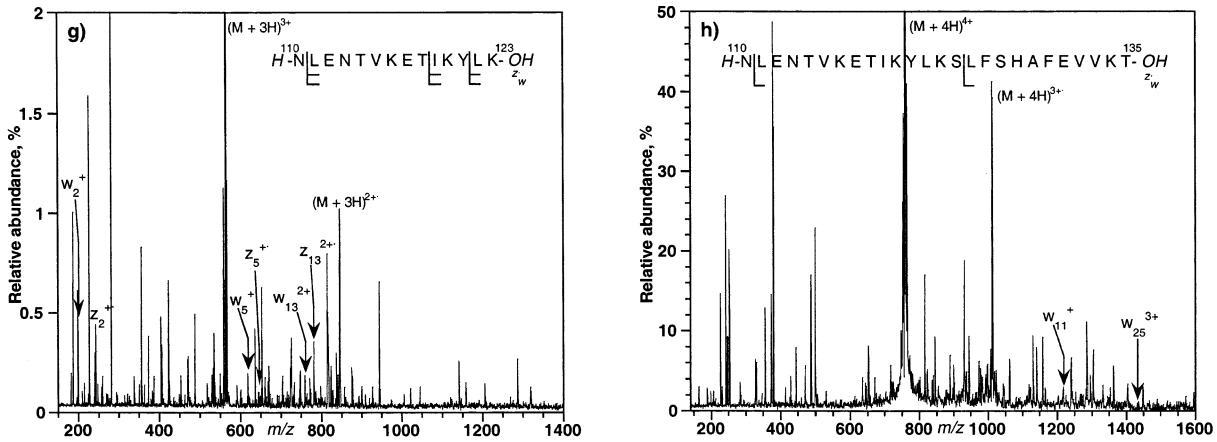


Figure 1. HECD FTICR mass spectra of the eight peptides derived from the protein PP3 showing assignment of secondary w ions for determination of Ile/Leu residues. Corresponding precursor z• ions are marked if present in the spectra. Superscript Arabic numbers in sequences indicate the position of amino acid residues in the intact protein. For peptide 42−53 (c) all w ions in the spectrum are designated. (a) Peptide 1−24, 4+, 125 ms with 11-eV electrons, 500 acquisitions. (b) Peptide 25−41, 2+, 300 ms with 9-eV electrons, 200 acquisitions. (c) Peptide 42−53, 2+, 125 ms with 11-eV electrons, 500 acquisitions. (d) Peptide 54−76, 4+, 125 ms with 11-eV electrons, 100 acquisitions. (e) Peptide 77−109, 4+, 200 ms with 10-eV electrons, 500 acquisitions. (f) Peptide 100−109, 2+, 125 ms with 11-eV electrons, 100 acquisitions. (g) Peptide 110−123, 3+, 125 ms with 11-eV electrons, 100 acquisitions. (h) Peptide 110−135, 4+, 125 ms with 11-eV electrons, 200 acquisitions.

for achieving excellent signal-to-noise ratios. In mass determination of the ECD and HECD fragments, molecular ions were used as internal calibrants.

RESULTS AND DISCUSSION

HECD of PP3 Peptides. Figure 1 shows the HECD mass spectra of the peptides 1−24, 25−41, 42−53, 54−76, 77−109, 100−109, 110−123, and 110−135 (numbers denote the first and the last residue in the protein sequence) corresponding to full sequence coverage with two overlaps. The peptides varied in size from 10 to 33 residues, corresponding to molecular weights between 1123 and 3571. A total of 13 HECD and 10 ECD spectra of different charge states (2+ to 4+) were taken. In all spectra, only w ions were observed, consistent with the previously reported[10] absence of d ions. In Figure 1, the denoted peaks are those of the w ions with N-terminal Xle, as well as of the corresponding z• ions (if present). Additionally for demonstration, w ions originating from all residues in the peptide 42−53 (Figure 1c) were assigned. The large number of peaks present in HECD mass spectra called for establishing a formal procedure for distinguishing Xle.

Such a procedure was created (see Tables 1 and 2) based on the following scoring principles. The Xle residue was considered *positively identified* if a suspected peak of a w ion had a signal-to-noise ratio of ≥1.5, and the deviation from the theoretical mass was ≤30 ppm (the average mass deviation of positively identified ions was 8 ppm). An additional and most important condition was for the magnitude of w ion peak to exceed that of the peak of the "alternative assignment" (i.e., Ile if Leu was assigned, and vice versa) by a factor of ≥1.5. With positive identification of the residue, the corresponding letter code, L or I, was put in the assignment column of Tables 1 and 2. If the last criterion was not fulfilled, the uncertainty of any definite assignment was considered too high, and therefore, a question mark was assigned in these cases. If one out of the first two criteria was not fulfilled, a weaker identification was assigned ("*strong indication*"), denoted as (L) or (I), respectively. If only the last criterion was fulfilled, the assignment was the weakest ("*indication*"), L? or I?. If no peak was present, including the alternative one, or could not be analyzed because of the overlap with other ions, a question mark was given.

Furthermore, when different charge states of the same peptide were subjected to HECD, the final assignment (see the corresponding column of Table 1) was made based on the following rules: two equal assignments gave one of a higher order, e.g., (I/L) + (I/L) = I/L (with the obvious exception of ? + ? = ?), otherwise the highest assignment was used, e.g., I/L + ? = I/L. The same rules were applied to the final assignments if the residue in question was part of both overlapping sequences. For example, I_{107} was part of the peptides 77−109 and 100−109, with the final assignments I and ?, which gave the overall "I" assignment.

On the basis of this formal scoring procedure, 20 out of 25 I/L residues were positively (and correctly!) identified. No residue received a weak final identification grade, which testifies to the adequacy of the selected formal rules. The identities of I_1, L_2, L_{33}, I_{39}, and L_{80} remained unknown by the same rules. Of these residues, the w ion peak for the N-terminal I_1 overlapped with abundant small losses from the reduced molecular species,[16] while w ions for L_2 and L_{33} overlapped with other fragment ions (b ions). Such "bad luck" overlaps are a common phenomenon in polypeptide sequencing.[17] No alternative w ions were present for the two residues L_2 and L_{33}. Therefore, outside the formal identification rules, one could still make correct guesses based on the absence of the alternative peak. Indeed, with 20 out of 25 residues positively identified, the absence of a w ion from a Xle residue means ~80% chance that the opposite assignment is correct. Only in the cases of I_{39} and L_{80}, neither expected nor alternative w ions were detected. These residues represent the only "pure" failure of HECD to produce w ions. Curiously, the I_{39} residue is a part of the partial sequence that reads "SKEPSIS".

(16) Cooper, H. J.; Hudgins, R. R.; Håkansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241−249.
(17) Budnik, B. A.; Nielsen, M. L.; Olsen, J. V.; Haselmann, K. F.; Hörth, P.; Haehnel, W.; Zubarev, R. A. *Int. J. Mass Spectrom.* **2002**, *219*, 283−294.

Table 1. Assignment of Xle Residues from HECD Spectra[a]

		parent ion, Z							parent ion, $Z+1$							
peptide	residue w(Xle) (exp mass of $(M+H)^+$)	w^{z+} z =	S/N for w^{z+}	$A(w)/A(w_{alt})$	mass devn, ppm for w (for w_{alt})	$A(w_n)/A(z^\bullet_n)$	comment	assgn	w^{z+} z =	S/N for w^{z+}	$A(w)/A(w_{alt})$	mass devn, ppm for w (for w_{alt})	$A(w_n)/A(z^\bullet_n)$	comment	assgn	final assgnment
1−24	I1 (2677.3 Da)						*b, c*	?						*b, c*	?	?
(2.72 kDa)	L2 (2551.2 Da)						*c, d*	?						*c, d*	?	?
$Z=3+$	L11 (1487.7 Da)	1+	5.8	3.3	6.3(23.7)	1.2		L	1+	5.0	≥5.0	0.7 (nd)	2.3	*d*	L	L
	I23 (243.1 Da)	1+	11.1	≥11.1	24.1 (nd)	1.1	*d*	I	1+	72.2	59.1	0.8 (−4.3)	1.3		I	I
25−41	L26 (1756.9 Da)	1+	3.6	≥3.6	9.6 (nd)	≥3.1	*d*	L								L
(1.93 kDa)	I28 (1529.7 Da)	1+	3.1	≥3.1	−3.9 (nd)	≥2.1	*d*	I								I
$Z=2+$	L33 (957.5 Da)	1+	8.5	≥8.5	4.8 (nd)	7.6	*c, d*	?								?
	I39 (330.2 Da)						*e*									?
42−53	L44 (1139.7 Da)	1+	7.5	4.4	0 (−6.1)	2.5		L	2+	4.2	≥4.2	−6.6 (nd)	≥3.7	*d*	L	L
(1.44 kDa)	I45 (1040.6 Da)	1+	2.2	≥2.2	17.3 (nd)	1.3	*d*	I	2+					*e*	?	I
$Z=2+$	I50 (455.3 Da)	1+	11.7	8.5	−11.9 (27.2)	1.9		I	1+	5.6	≥5.6	5.7 (nd)	≥5.5	*d*	I	I
	I52 (243.1 Da)	1+	3.8	≥3.8	−10.6 (nd)	0.6	*d*	I	1+					*e*	?	I
54−76	L65 (1387.8 Da)	1+	5.8	≥5.8	−0.7 (nd)	4.2	*d*	L	1+	4.5	≥4.5	−11.5 (nd)	≥3.5	*d*	L	L
(2.68 kDa)	L65 (1387.8 Da)						*e*	?	2+	11.1	≥11.1	−1.4 (nd)	≥10.7	*d*	L	L
$Z=3+$	L67 (1177.7 Da)	1+	2.2	≥2.2	6.8 (nd)	≥2.0	*d*	L	1+	2.9	≥2.9	15.6 (nd)	≥2.0	*d*	L	L
	L67 (1177.7 Da)						*e*	?	2+	2.1	≥2.1	3.2 (nd)	≥1.5	*d*	L	L
	I69 (991.6 Da)	1+	1.8	≥1.8	5.9 (nd)	0.7	*d*	I	1+	2.2	1.5	−7.0 (−22.6)	1.6		I	I
	L70 (864.5 Da)	1+	4.2	≥4.2	−5.3 (nd)	1.0	*d*	L	1+	5.6	≥5.6	0.9 (nd)	2.5	*d*	L	L
	L75 (229.1 Da)	1+	3.4	≥3.4	−6.1 (nd)	1.0	*d*	L	1+	17.3	15.3	−3.0 (8.1)	1.0		L	L
77−109	L80 (3226.5 Da)						*e*	?								?
(3.57 kDa)	L100 (1168.6 Da)	1+	1.8	≥1.8	−15.4 (nd)	≥1.8	*d*	L								L
$Z=4+$	L103 (795.4 Da)	1+	5.4	≥5.4	−8.1 (nd)	1.6	*d*	L								L
	I107 (374.2 Da)	1+	3.7	≥3.7	−9.8 (nd)	2.2	*d*	I								I
100−109	L100 (1167.6 Da)						*b, c*	?						*b, c*	?	?
(1.23 kDa)	L103 (795.4 Da)	1+	22.7	2.2	−0.8 (12.6)	3.8	*f*	L	1+	8.7	0.9	−1.3 (20.0)	4.8	*g*	?	L
$Z=2+$	I107 (374.2 Da)	1+	11.5	≥11.5	13.6 (nd)	3.5	*c, d*	?	1+	10.6	≥10.6	18.1 (nd)	3.4	*c, d*	?	?
110−123	L111 (1519.8 Da)	2+	14.3	≥14.3	−13.9 (nd)	0.6	*d*	L								L
(1.69 kDa)	I119 (619.4 Da)	1+	4.9	≥4.9	−4.8 (nd)	2.3	*d*	I								I
$Z=3+$	L122 (201.1 Da)	1+	5.8	≥5.8	−9.5 (nd)	1.1	*d*	L								L
110−135	L111 (2865.5 Da)						*e*	?	3+	1.5	≥1.5	−29.9 (nd)	≥1.2	*d*	L	L
(3.04 kDa)	I119 (1965.1 Da)						*e*	?						*e*	?	?
$Z=3+$	L122 (1546.8 Da)	1+	1.6	≥1.6	105 (nd)	1.2	*c, d*	?	2+	2.8	≥2.8	−20.1 (nd)	≥1.3	*c, d*	?	?
	L125 (1218.6 Da)	1+	1.8	1.4	49.8 (−46.7)	0.6	*f*	?	1+	1.5	2.0	−28.7 (−137)	≥1.5		L	L

[a] $A(w)/A(w_{alt})$ is the ratio of abundances of the expected w ion and that for the alternative Xle assignment. Ranking: L/I > (L/I) > (L?/I?) > ?. [b] First amino acid from the N-terminal side. [c] Too close in mass to another ion. [d] No alternative peak found. [e] Not analyzed. [f] Weak signal. [g] The alternative peak is much closer (<4.5 ppm) to b_7^+.

The ECD analysis of PP3 peptides (Table 2) gave 11 positively identified Xle residues (44%). That makes HECD almost twice as efficient for Ile/Leu determination as ECD, which can be attributed to the larger amount of excessive energy. Compared to HECD spectra, ECD data gave no new Xle information.

For illustration, ECD and HECD spectra of 2+ of the 1.4-kDa peptide 42−53 (EDLISKEQIVIR) are shown in Figure 2. Characteristic for the HECD spectrum is the larger number of product ions compared to the ECD spectrum. Besides that, HECD allowed positive identification of all four Xle residues in the peptide, whereas the ECD spectrum gave two. The ratios of the abundances of w ions and the corresponding $z^\bullet$ ions were also monitored (see Tables 1 and 2). These ratios were generally much higher in HECD than ECD, consistent with the larger extent of secondary fragmentation in the former technique. Overall, 83% of w ions in the HECD spectra had higher abundances than the corresponding $z^\bullet$ ions, compared to just 35% in ECD. This is consistent with the proposed mechanism[10] in which the w ions arise from secondary fragmentation of $z^\bullet$ species.

The presence of secondary fragmentation in some of the ECD spectra points toward vibrational excitation of the ions either before or after electron capture. In principle, the excess of energy in ion−electron recombination can be sufficient for inducing secondary fragmentation in some peptides. Additional internal energy increment could arise from electrons with energies above 0.2 eV. Although dispenser cathodes operate at much lower temperatures than filaments, the higher current density leads to the kinetic energy spread of the emitted electrons of ∼0.5−0.7 eV,[15] with a fraction of electrons tailing far above that value.

Size and Sequence of Peptides. With a fixed amount of energy deposited by 11-eV electrons in the precursor polycations, one would expect HECD to decrease in efficiency as the size of the cations increases. However, from the body of the HECD data obtained in this study, it appears that peptide size is not the only critical parameter for successful identification of Xle residues. This can be seen from Table 1 and Figure 1 for the example of two similar size peptides, 42−53 (EDLISKEQIVIR) and 100−109 (LMELGHKIMR). While dications of the peptide 42−53 produced many fragment ions including all (four) possible w ions for Xle residues, the dications of the peptide 100−109 fragmented very poorly and allowed only one determination out of two possible (besides the N-terminal L_{100} that could not be identified). This is not the first case when peptides with methionine do not readily form w ions: previously, HECD failed to distinguish two Xle pairs

Table 2. Assignment of Xle Residues from ECD Spectra[a]

		parent ion, Z							parent ion, $Z+1$							
peptide	residue w(Xle) (exp mass of $(M+H)^+$)	w^{z+} z =	S/N for w^{z+}	$A(w)/A(w_{alt})$	mass devn, ppm for w (for w_{alt})	$A(w_n)/A(z^{\bullet}_n)$	comment	assgn	w^{z+} z =	S/N for w^{z+}	$A(w)/A(w_{alt})$	mass devn, ppm for w (for w_{alt})	$A(w_n)/A(z^{\bullet}_n)$	comment	assgn	final assgnment
1−24	I1 (2677.3 Da)													*b, c*	?	?
(2.72 kDa)	L2 (2551.2 Da)													*c, d*	?	?
$Z=3+$	L11 (1487.7 Da)								1+	2.9	≥2.9	6.0 (nd)	0.7	*d*	L	L
	I23 (243.1 Da)													*e*	?	?
25−41	L26 (1756.9 Da)						*e*	?								?
(1.93 kDa)	I28 (1529.7 Da)						*e*	?								?
$Z=2+$	L33 (957.5 Da)						*e*	?								?
	I39 (330.2 Da)						*e*	?								?
42−53	L44 (1139.7 Da)	1+	5.3	≥5.3	−1.8 (nd)	1.0	*d*	L								L
(1.44 kDa)	I45 (1040.6 Da)						*e*	?								?
$Z=2+$	I50 (455.3 Da)	1+	6.4	≥6.4	−6.1 (nd)	0.6	*d*	I								I
	I52 (243.1 Da)						*e*	?								?
54−76	L65 (1387.8 Da)								1+	21.8	≥21.8	1.4 (nd)	8.9	*d*	L	L
(2.68 kDa)	L65 (1387.8 Da)								2+	24.1	≥24.1	−19.0 (nd)	30.9	*d*	L	L
$Z=3+$	L67 (1177.7 Da)								1+	4.1	≥4.1	−0.9 (nd)	3.7	*d*	L	L
	I69 (991.6 Da)								1+	2.6	0.8	−7.4 (−7.1)	0.4		?	?
	L70 (864.5 Da)								1+	16.9	≥16.9	−2.2 (nd)	1.3	*d*	L	L
	L75 (229.1 Da)								1+	12.3	≥12.3	−17.9 (nd)	0.6	*d*	L	L
	L80 (3226.5 Da)						*e*	?								?
77−109	L100 (1168.6 Da)						*e*	?								?
(3.57 kDa)	L103 (795.4 Da)	1+	2.3	≥2.3	−0.6 (nd)	0.5	*d*	L								L
$Z=4+$	I107 (374.2 Da)						*e*	?								?
	L100 (1167.6 Da)						*b, c*	?						*b, c*	?	?
100−109	L103 (795.4 Da)	1+	3.0	1.0	−3.6 (21.0)	0.7	*g*	?	1+	5.7	1.1	5.3 (25.6)	5.1	*g*	?	?
(1.23 kDa)	I107 (374.2 Da)						*e*	?	1+	3.1	≥3.1	38.0 (nd)	2.2	*c, d*	?	?
$Z=2+$	L111 (1519.8 Da)	2+	12.4	≥12.4	−15.4 (nd)	0.3	*d*	L								L
110−123	I119 (619.4 Da)	1+	7.3	6.1	−5.8 (−3.8)	0.8		I								I
(1.69 kDa)	L122 (201.1 Da)	1+	6.7	≥6.7	−16.4 (nd)	0.4	*d*	L								L
$Z=3+$	L111 (2865.5 Da)						*e*	?	3+	5.7	≥5.7	−18.3 (nd)	0.8	*d*	L	L
110−135	I119 (1965.1 Da)						*e*	?						*e*	?	?
(3.04 kDa)	L122 (1546.8 Da)						*e*	?	2+	4.5	≥4.5	−14.2 (nd)		*c, d*	?	?
$Z=3+$	L125 (1218.6 Da)						*e*	?						*e*	?	?

[a] $A(w)/A(w_{alt})$ is the ratio of abundances of the expected w ion and that for the alternative Xle assignment. Ranking: (L/I) > (L?/I?) > ?. [b] First amino acid from the N-terminal side. [c] Too close in mass to another ion. [d] No alternative peak found. [e] Not analyzed. [f] Weak signal. [g] The alternative peak is much closer (<10.7 ppm) to b_7^+.

in the synthetic peptide KIMHASELMANN.[18] However, 4+ of the peptide 77−109 (NATLGSEETTEHTPSDASTTEGKLMELGHKIMR) gave in HECD three out of four possible Xle w ions despite the presence of two methionines. At the same time, the 4+ of the peptide of similar size but without methionines, 110−135 (NLENTVKETIKYLKSLFSHAFEVVKT), yielded only two out of four residues (two additional w ions were of very low abundances). It appears that the closeness of an Xle residue to the C-terminus increases its chance to be identified: when a shorter (110−123) version of the last peptide was subjected to HECD, the residues L_{111}, I_{119}, and L_{122} were easily identified (Figure 1g). The effect can also be attributed to the reduced size.

HECD of Different Charge States. HECD spectra of five peptides in Table 1 (1-24, 42−53, 54−76, 100−109, 110−135) were recorded in different charge states for comparison. For four of these peptides (42−53, 54−76, 100−109, 110−135), new Xle information was obtained; in one other case (peptide 1−24), the same w ions were obtained. For two of the four peptides (54−76, 110−135), higher charge states were more informative than the lower, with seven identified w ions for 4+ of the peptide 54−76 compared to five for 3+. For peptide 110−135, 4+ gave two Xle w ions, while 3+ of the same peptide gave no such ions. On the contrary, the 2+ and 3+ of the peptide 42−53 (Figure 2 and Table 1) provided identification of four and two Xle w ions, respectively. And similarly, 2+ of the peptide 100−109 gave determination of one Xle residue, while for the 3+, no w ions where observed. It still remains unclear which charge states (higher or lower) tend to give more sequence information. From the general point of view, higher charge states are less stable and thus should give more ionic fragments. However, since the latter will come in a larger variety of charge states, the fragment peaks will overlap with each other with higher probabilities.

Fast Xle Identification. Generally, obtaining HECD spectra with a good signal-to-noise ratio requires integration of more data compared to ECD, typically 100 scans. However, for at least two of the eight peptides in this study, less than 10 scans were sufficient. We investigated for one of these peptides, 54−76 (SSRQPQSQNPKLPLSILKEKHLR), the minimal conditions for distinguishing Xle according to the formal criteria used. The 4+ ions of this peptide were irradiated for 125 ms with 11-eV electrons. With one 2-s-long scan, four out of five possible Xle residues were positively identified (Figure 3a), and only five scans

(18) Haselmann, K. F.; Budnik, B. A.; Kjeldsen, F.; Nielsen, M. L.; Olsen, J. V.; Zubarev, R. A. *Eur. J. Mass. Spectrom.* **2002**, *8*, 117−121.

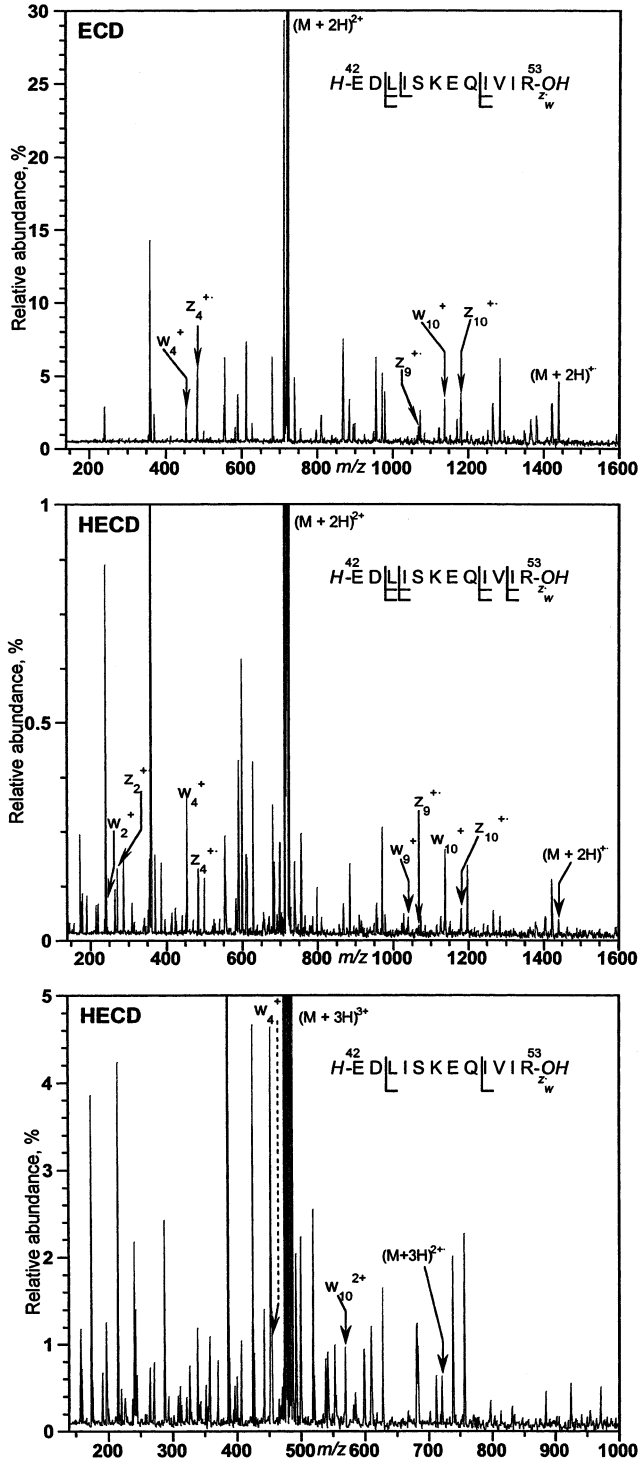


Figure 2. ECD (top) versus HECD spectra of the 2+ (middle) and the 3+ (bottom) of peptide 42−53 (EDLISKEQIVIR) irradiated with 0.2 and 11 eV, respectively. Four w ions were obtained with HECD of 2+, while ECD gave two w ions. HECD of 3+ gave two w ions.

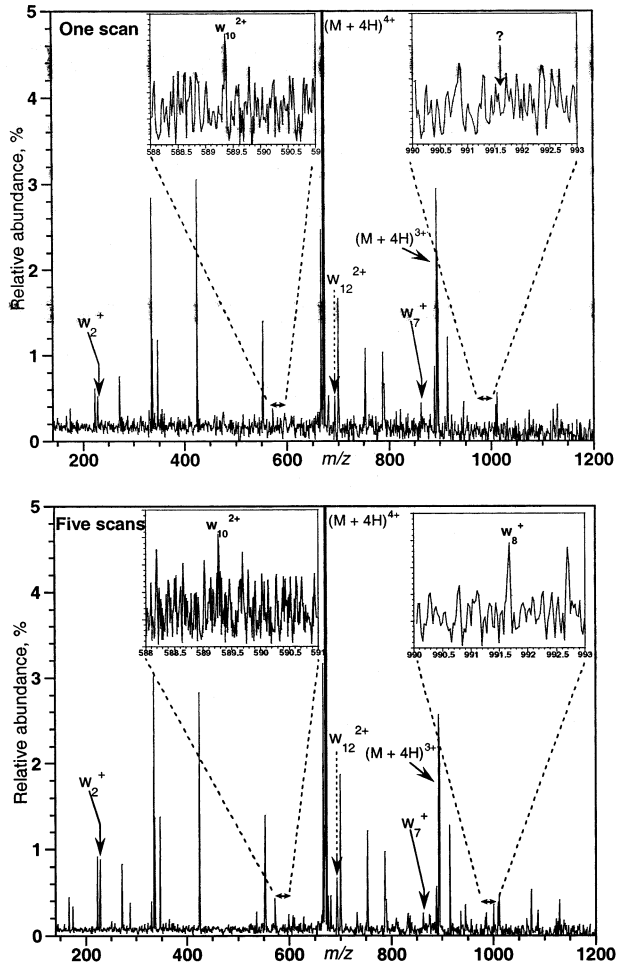


Figure 3. Fast determination of Xle residues in the peptide 54−76 (SSRQPQSQNPKLPLSILKEKHLR) by HECD: (a) with one scan (acquisition), four of five Xle residues were identified, and (b) with only five acquisitions, the last w_8^+ was observed, identifying Xle_8 as isoleucine.

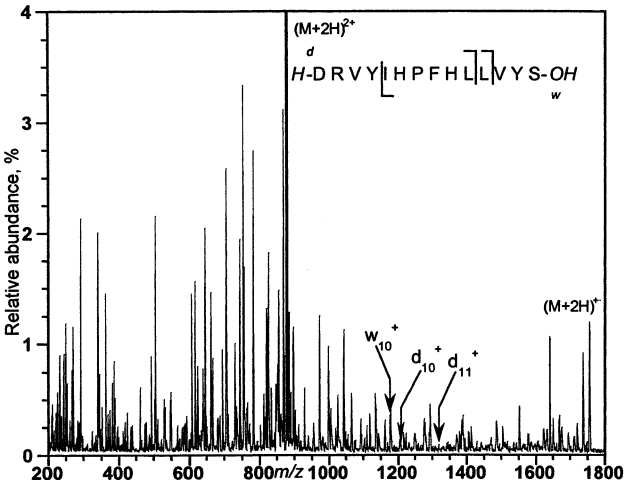


Figure 4. HECD spectrum of doubly protonated renin substrate (DRVYIHPFHILLVYS), obtained by irradiation for 150 ms with 11-eV electrons. Identification of all three Xle residues by d_{10}^+, d_{11}^+, and w_{10}^+ fragment ions.

were needed for identification of all five residues (Figure 3b). This means that the rate of HECD Xle determination is approaching the level at which the technique can be used on-line with HPLC.

Rare Observation of d Ions. Neither in the first HECD publication[10] nor in the HECD spectra of the eight peptides used in this study were ions detected that could be unequivocally assigned to d species. However, HECD of 2+ of renin substrate (DRVYIHPFHLLVYS) did produce two d ions (d_{10}^+ and d_{11}^+), which, together with one w ion (w_{10}^+), identified all three Xle residues in this nontryptic peptide (Figure 4). This observation remains so far unique and will be a subject of further studies. A

possible explanation for it is that HECD of renin substrate, a peptide with a basic N-terminus, gives more abundant a• ions than are usually observed for tryptic peptides charged at the C-terminus.

CONCLUSIONS

Formal rule-based identification of 20 of 25 Ile/Leu residues in PP3 by HECD can be considered very satisfactory, given the fact that two of the remaining five residues could be correctly guessed by the absence of the alternative peaks. These results illuminate both the progress and the obstacles on the way to complete de novo sequencing of proteins. The latter requires 100% efficiency in determination of all residues in all peptides—a goal that mass spectrometry has yet to achieve. With HECD, one step closer is made. Yet, distinguishing the Xle is imperfect, mainly due to the rare appearance of d fragments. Little is known about the formation mechanism of their precursors—the a• ions. Additional research in ion−electron interactions is needed to shed light on that mechanism and perhaps even to find a way of increasing the d ion abundance.

The discovery, catalyzed by the successful HECD results, that mass spectra of conventional ECD also contain w ions, illustrates the need to improve understanding of the mechanism of ion−electron reactions. The close relation between the two techniques has already been discussed.[10] HECD appears to be more technically challenging than ECD, although the principal difference amounts only to the electron energy. However, higher kinetic energies of HECD fragments together with the ion cloud diffusion due to energetic ion−electron collisions reduce the FT signal and make the question of sensitivity more acute. This is also viewed as a motivation for future studies and improvements. Higher sensitivity in HECD would make this technique even more attractive; e.g., the reduced time frame of analysis would make it compatible with on-line LC systems. Additionally, less extensive sample purification for contaminated samples and mixtures would be required if even low-abundant species were possible to isolate and subject to HECD.

The HECD spectra are more complex compared to CAD or conventional ECD spectra. Making sense of them for de novo sequencing is difficult without sophisticated data processing.[19] Comparison with ECD and low-energy CAD spectra of the same species can reduce the problem, but again requires implementation of adequate software. For the limited task of Xle determination, the formal procedure used in this work has shown its practical utility. Its implementation into the software should be rather straightforward.

ACKNOWLEDGMENT

The work was supported by the Danish National Research Councils (Grants SNF 9891448, 51-00-0358, and STFV 0001242 to RZ, STFV 26-01-0058 to R.A.Z. and F.K.).

NOTE ADDED AFTER ASAP POSTING

This article was inadvertently posted ASAP before final corrections were made. The corrected version was posted on February 13, 2003.

Received for review June 27, 2002. Accepted October 23, 2002.

AC020422M

(19) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 320−332.