Side-Chain Losses in Electron Capture Dissociation To Improve Peptide Identification

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2007

Impact Factor: 5.64 · DOI: 10.1021/ac0619332 · Source: PubMed

3 AUTHORS, INCLUDING:

Mikhail M Savitski
Cellzome GMBH
56 PUBLICATIONS **2,892** CITATIONS
SEE PROFILE

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

Anal. Chem. **2007,** *79,* 2296−2302

Side-Chain Losses in Electron Capture Dissociation To Improve Peptide Identification

Mikhail M. Savitski,* Michael L. Nielsen, and Roman A. Zubarev

Laboratory for Biological and Medical Mass Spectrometry, Uppsala University, Box 583, Uppsala S-751 23, Sweden

Analysis of a database of some 20 000 conventional electron-capture dissociation (ECD) mass spectra of doubly charged ions belonging to tryptic peptides revealed widespread appearance of w ions and related u ions that are due to partial side chain losses from radical $z^{\bullet}$ ions. Half of all $z^{\bullet}$ ions that begin with Leu or Ile produce w ions in conventional one-scan ECD mass spectra, which differentiates these isomeric residues with >97% reliability. Other residues exhibiting equally frequent side chain losses are Gln, Glu, Asp, and Met (cysteine was not included in this work). Unexpectedly, Asp lost not a radical group like other amino acids but a molecule CO_2, thus giving rise to a radical $w^{\bullet}$ ion with the possibility of a radical cascade. Losses from amino acids as distant as seven residues away from the cleavage site were detected. The mechanism of such losses seems to be related to radical migration from the original site at the $^{\alpha}C_n$ atom in a $z_n^{\bullet}$ ion to other $^{\alpha}C$ and $^{\beta}C$ atoms. The side chain losses confirm sequence assignment, improve the database matching score, and can be useful in de novo sequencing.

Tandem mass spectrometry (MS/MS) of tryptic peptides is a basis of most proteomics studies.[1−3] By far the most frequently used MS/MS technique is collision-activated dissociation (CAD) that breaks peptide C−N bonds to form N-terminal b and C-terminal y′-ions. Electron-capture dissociation (ECD)[4] is a complementary fragmentation technique that produces c′ and $z^{\bullet}$ ions upon breakage of the N−C_{α} bond. It is also known that ECD induces secondary reactions in radical $z^{\bullet}$ ions, such as intramolecular and intracomplex hydrogen atom transfer[4−11] and partial side chain losses that result in w ions.[12−18] Since losses from Leu are different than those from the isomeric Ile residue, these residues can be differentiated. Because 16% of all amino acids in proteins are either Leu or Ile (Xle), such differentiation is of analytical importance. However, w ions are mostly known to appear in hot-ECD (HECD) that utilizes higher (>10 eV) electron energies,[13−15] while in "conventional" ECD, which utilizes low-energy electrons (<1 eV), these ions appear less frequently. But, because of the higher efficiency, ECD is much more frequently used than HECD and has found application in high-throughput proteomics.[19−22] It has been estimated that HECD produces w ions in ca. 80% of the cases, while ECD in only 44% of the cases.[14] The low yield of side chain losses in ECD is believed to be due to the much lower energy excess in ECD compared to HECD. Yet it is not the energy excess per se but the energy excess per chemical bond that is important for fragmentation. Tryptic peptides are small in size (typical length, 10−12 residues), which means that the recombination energy release per bond can be larger there than in longer polypeptides. To test the extent of side chain losses in "conventional" ECD of tryptic peptides, we performed analysis of some 20 000 ECD mass spectra of doubly charged ions that were accumulated in the course of proteomics work.[19] Analyzing

* To whom correspondence should be addressed. Phone: +46 (0) 18 471 5729. Fax: +46 (0) 18 471 5729. E-mail: Mikhail.Savitski@bmms.uu.se.

(1) Steen, H.; Mann, M. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 699−711.
(2) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198−207.
(3) Mann, M.; Jensen, O. N. *Nat. Biotechnol.* **2003**, *21*, 255−261.
(4) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.
(5) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431−4436.
(6) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857−2862.
(7) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563−573.
(8) Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949−8958.
(9) Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353−3369.
(10) O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 576−585.
(11) Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (1), 113−120.
(12) Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241−249.
(13) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201−206.
(14) Kjeldsen, F.; Haselmann, K. F.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 1267−1274.
(15) Kjeldsen, F.; Zubarev, R. *J. Am. Chem. Soc.* **2003**, *125*, 6628−6629.
(16) Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *Int. J. Mass Spectrom.* **2003**, *228*, 723−728.
(17) Fung, Y. M. E.; Chan, T. W. D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1523−1535.
(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.
(19) Nielsen, M. L.; Savitski, M. M.; Zubarev, R. A. *Mol. Cell. Proteomics* **2005**, *4*, 835−845.
(20) Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. *Mol. Cell. Proteomics* **2005**, *4*, 1180−1188.
(21) Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A. *J. Proteome Res.* **2005**, *4*, 2348−2354.
(22) Savitski, M. M.; Nielsen, M. L.; Zubarev, R. A. *Mol. Cell. Proteomics* **2006**, *5*, 935−948.

10.1021/ac0619332 CCC: $37.00
Published on Web 02/03/2007

Table 1. Masses of Radicals Lost from Side Chains of Amino Acids in ECD Monitored in This Work

amino acid residue	side chain loss, Da	chemical formula
L	43.0542	$^{\bullet}C_3H_7$
E	59.0128	$^{\bullet}C_2H_3O_2$
M	61.0107	$^{\bullet}C_2H_5S$
I	29.0397	$^{\bullet}C_2H_5$
Q	58.0287	$^{\bullet}C_2H_4NO$
V	15.0240	$^{\bullet}CH_3$
N	44.0131	$^{\bullet}CH_2NO$
T	15.0229, 17.0022	$^{\bullet}CH_3$, OH
S	17.0022	$^{\bullet}OH$
W	116.049	$^{\bullet}C_8H_6N$
D	44.9971	$^{\bullet}CHO_2$
H	67.0291	$^{\bullet}C_3H_3N_2$
F	77.0386	$^{\bullet}C_6H_5$
Y	93.0335	$^{\bullet}C_6H_5O$

the data, we found w ions and related u ions[15,23] to be much more widespread than was previously thought. Below we report on this study and its potential for improving proteomics analysis.

EXPERIMENTAL SECTION

MS/MS data were collected from the proteomics analysis of lysates of human cell lines and *Escherichia coli* (*E. coli*; all samples from Sigma) performed on a 7 T LTQ FT mass spectrometer (Thermo) using consecutive ECD and computer-assisted design (CAD) fragmentation of peptides eluting from the analytical column of a nano-LC system (Agilent 1100). The peptides were identified by the Mascot search engine (Matrix Sciences). The details of the experimental procedure and data processing are described in ref 19.

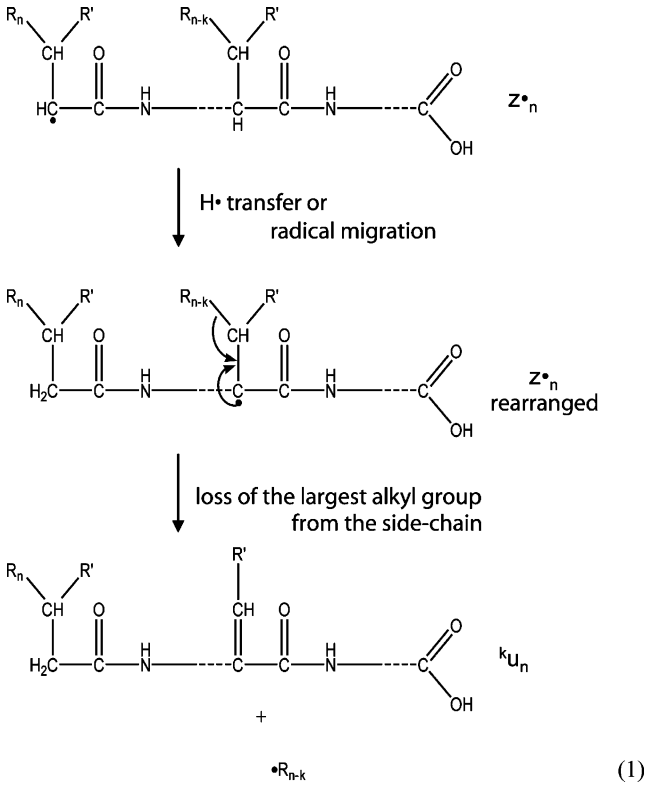
RESULTS AND DISCUSSION

Dependence upon Amino Acid Type and Position. Upon N−C_α bond cleavage, the C-terminal $z_n^\bullet$ fragment (here and throughout the text the indexation is from the C-terminus) has a radical centered on its α-carbon, $^{\alpha}C_n$. The radical site can then migrate to the $^{\beta}C_n$ atom and further to the side chain with formation of a double bond between $^{\alpha}C_n$ and $^{\beta}C_n$ atoms and a loss of an alkyl radical R_n from $^{\beta}C_n$.[14]

As a rule ("loss of the largest alkyl"),[24] the lost radical R_n is larger than the remaining R′, which can be either H, OH, or CH_3. For instance, $z^\bullet$ ion with Ile_n loses $R_n = C_2H_5$ preferentially over the loss of R′ = CH_3.

The radical initially located at $^{\alpha}C_n$ can migrate not to $^{\beta}C_n$ but to carbon atoms of the neighboring residue instead.[15,17] There it can initiate analogous reactions but leading to the loss of the R_{n-1} radical. In this case, the formed even-electron species are called u_n ions.[16] The radical migration is believed to preferentially occur "through space" from $^{\alpha}C_n$ to $^{\beta}C_{n-1}$ atoms and involve formation of a five-member lactam ring.[15,23] In general, u_n ions are less abundant in mass spectra than w_n ions.[15]

In eq 1 the general case is shown where the $z^\bullet$ ion is rearranged by means of either H$^\bullet$ transfer or radical migration. This is followed by the double bond formation and the loss of the $^{\bullet}R_{n-k}$ group (in the case $k = 0$, a w_n ion is formed).



(1)

Table 1 gives masses of the radical neutral fragment losses which were monitored from $z^\bullet$ ions with the structure $^{\bullet}X_nAY_m$, where X is any amino acid but A, Y is any amino acid, and $n + m + 1$ is the number of residues in $z^\bullet$. Only amino acids frequently present in internal sequences of tryptic peptides were included in the analysis. Thus Lys, Arg, Cys, and Trp were excluded because of their rarity. The frequency of cases when a specific mass loss (determined within ±0.015 Da accuracy) from the monoisotopic peak of the $z^\bullet$ ion was present in the ECD mass spectrum is presented in Figure 1 as a function of the distance from the original radical location. On average almost half the $z^\bullet$ ions that begin with Leu or Ile produce w ions. Other residues exhibiting frequent w ions are Gln, Glu, and Met. Frequencies of all losses decrease with the distance from the cleavage site, but abundant side chain losses were found occurring as far as from the fourth residue, and for methionine from even more distant sites (see below).

Xle Identification in High-Throughput Analysis. The most valuable analytical feature of w and u ions is the distinguishing between the isomeric Ile and Leu (Xle) residues. As an example, the top panel of Figure 2 shows the ECD mass spectrum of a peptide identified by Mascot as either EFLLIFR or EFILLFR. Since the difference between the two variants is only in the identity of Xle residues, both assignments received the same score of 18, well below the threshold value of 34. The probable reason for the low scoring was the standard ECD settings that included c, y, and z type ions as possible fragments. Abundant w and u ions present in the spectrum must have upset the scoring. Manual analysis of the spectrum supported the EFLLIFR sequence by three w and three u ions. To distinguish u ions arising from the

(23) Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. *Anal. Chem.* **1987**, *59*, 2621−2625.
(24) McLafferty, F. W.; Turecek, F. *Interpretation of Mass Spectra*; University Science Books: California, 1993.

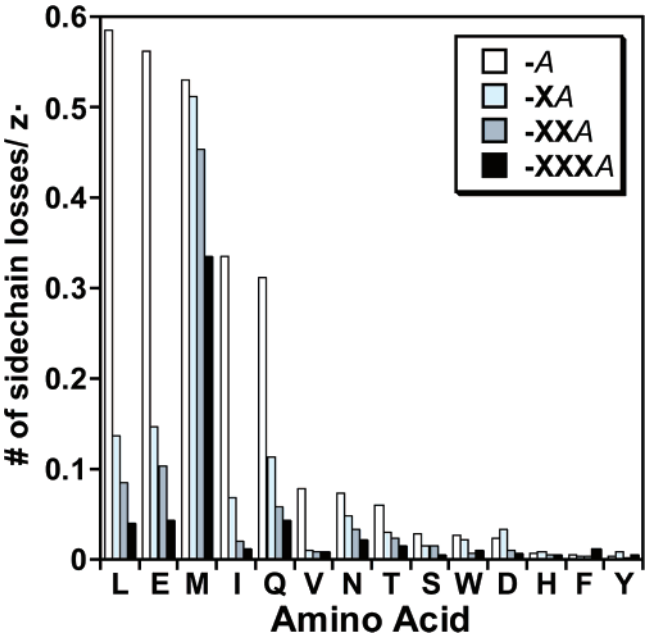


Figure 1. Frequencies of w_n ion detection relative to the presence of z_n· ions determined for amino acids in different positions to the cleavage site. A is the amino acid from which the partial side chain loss occurs; X is any other amino acid except A.

side chain losses in amino acids at different distances k from the N-terminus of the $z^{\bullet}$ ion, we introduce for such ions the notation ^{k}u. By definition, w = ^{0}u and u = ^{1}u. Including w ions as possible fragments in Mascot search resulted in $M = 21$ for EFLLIFR, a very modest increase given that w_4 is the most abundant fragment in the mass spectrum. Most likely, Mascot gives to w ions a lower score than to other types of fragments. The alternative sequence, EFILLFR, received $M = 19$. Manual removing of masses corresponding to w and u ions increased the score for both sequences to 40, way above the identification threshold of $M = 34$. Thus merging CAD and ECD data with automatic removal of w and u ions from the fragment list submitted to Mascot[19] may be the optimal way for Mascot identification using ECD.

What remains to be incorporated in the automatic procedure is verification of the Mascot-suggested sequence by side chain losses present in the original ECD data file. But, before recommending such a procedure for routine proteomics analysis, such verification itself needs to be validated, as any procedure in proteomics analysis has to be at least 95% reliable. To determine the reliability of Xle identification based on w and u ions, these identifications were checked against Mascot sequence assignments. In the rare cases where both 43 and 29 Da losses were present, the most abundant of these two losses was used to determine the Xle identity.[15] Of 11 303 cases where Xle was the N-terminal amino acid in a z_n· ion and the w_n ion was present, in all but 276 cases (2.5%) the Xle assignment agreed with the Mascot-suggested sequence. The 2.5% figure includes the Mascot miss-assignment rate, so the actual reliability of w ion analysis is >98%. For u ions, only motifs of the ·X(Xle)XX type were considered, where X was not Xle. Out of 1464 cases where a judgment could be passed, only 67 cases (4.5%) disagreed with Mascot-suggested sequences. Thus ^{1}u ions are also a reliable identifier of Xle residues.

Losses from Other Residues. Besides isomer differentiation, partial side chain losses can also be used to confirm tentative sequence assignments. For amino acids Leu, Met, and Glu that are without β-carbon branching, w ions are isobaric with z ions ($z = z^{\bullet} - H^{\bullet}$) with N-terminal Ala. Potentially, this poses a problem for Gln, for which the lost group, $^{\bullet}CH_2C(O)NH_2$, is isobaric with (Gly + H·). Thus, the situations ·QY.Y can in principle be confused with ·GAY...Y (Gly + Ala are isobaric with Gln), as the triplet z_n·, w_n, z_{n-1}· for the former sequence can be confused with the triplet z_n·, z_{n-1}·, z_{n-2}· for the latter sequence. Fortunately, analysis revealed that the confusion only arises in 0.8% of the cases because even-electron w ions unlike radical $z^{\bullet}$ ions are rarely undergoing hydrogen rearrangement with the complementary fragment (see below).

Distant Side Chain Losses. Interaction of the radical site in peptides with nearby groups has been much discussed in literature,[4−18] but little is known about more distant interactions. Methionine is obviously a special case, as in Figure 1 it shows the longest side chain loss propagation. An example of mass spectrum with the Met loss six residues away from the N-terminal residues (^{6}u ion) is shown in the lower panel of Figure 2. The two isotopic peaks and high mass accuracy of FTICR are reliable indicators of the identity of that ion. A less distant loss four residues away (^{4}u) is also present there.

To determine how far away from the initial radical position side chains can be lost from residues other than Met, we employed variance analysis.[11] This analysis (Figure 3) was similar to that in Figure 1, but the reporting quantity was not the w/$z^{\bullet}$ ratio itself but its variance (average deviation from the average value) for different amino acids in the same position relative to the cleavage site. Thus, Figure 3 shows "global" decay values determined for all amino acids in Figure 1 except Met. The maximum variance was expectedly found for the -A position (losses corresponding to w ion formation). The variance monotonously decreases with the distance from the radical site (note the logarithmic scale on the vertical axis). To determine how deep into the peptide sequence the side chain losses propagate, one needs to establish the background level of the variance. Since the positions A- and AX- cannot be implicated in the side chain loss, the variance level for these positions must be due to random data fluctuations and thus determines the background level. Against this background, the variance for the position -A is elevated more than 1000 times, and it drops with each additional residue by a factor of 7 on average, reaching the background level after four or five residues.

Correlation analysis (Figure 3, inset) was employed to verify these findings. Here, linear correlation between propensities for different amino acids to form w ions in different positions and the propensities for the same amino acids in the neighboring position (e.g., -A vs -XA) was measured. Unity was taken as a default value for -A. The correlation drops below the statistically significant value ±0.56 [25] when there are more than four residues between the radical site and the amino acid that loses the side chain. The small or negative correlation values of the positions left to the cleavage site confirm that these positions do not influence side chain losses.

The absence of other features in Figure 3 except the monotonic decline in variance and correlation means that migration of the

(25) Bewick, V.; Cheek, L.; Ball, J. *Crit. Care* **2003**, *7*, 451−459.

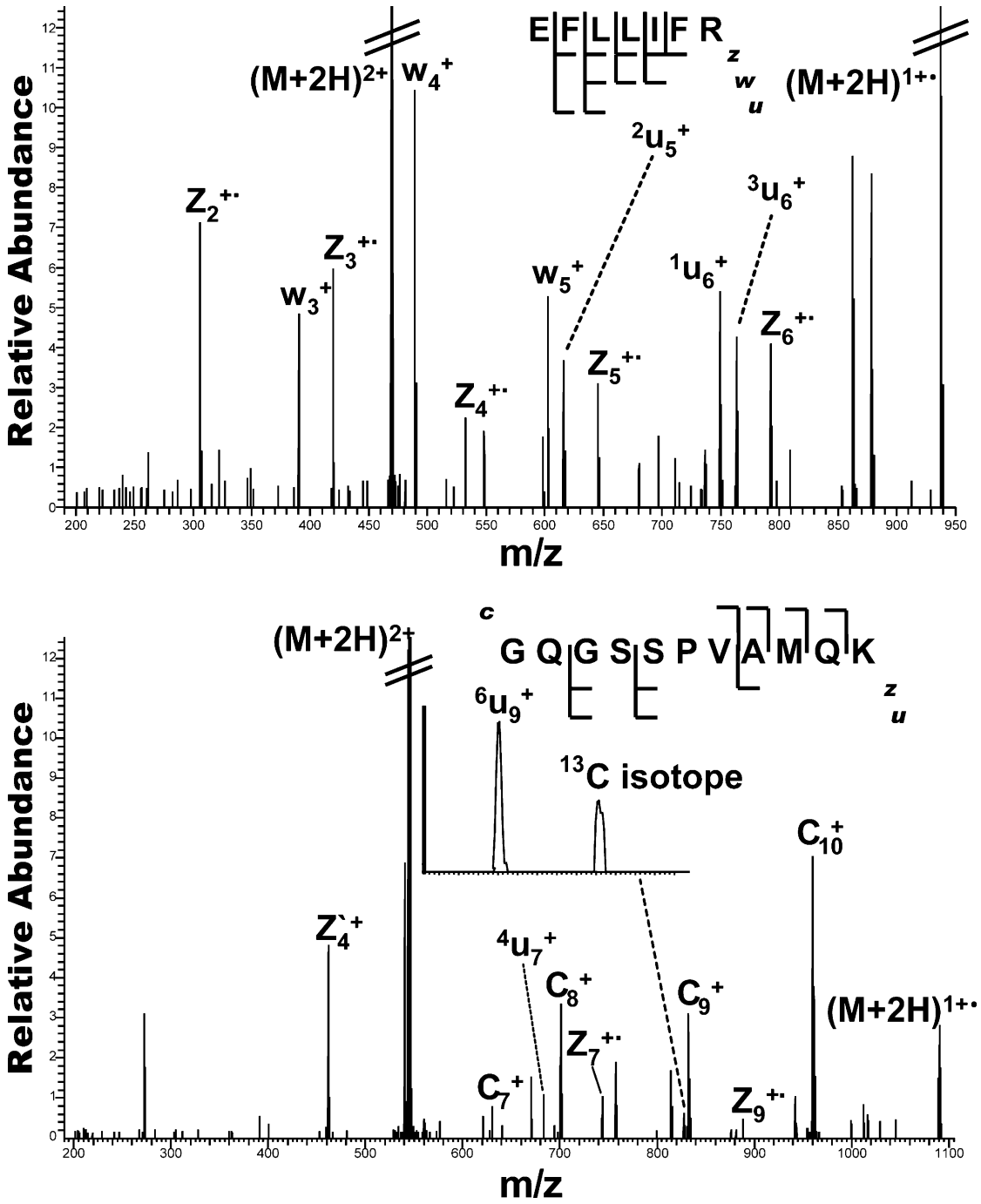


Figure 2. Top panel: Single-scan ECD mass spectrum of 2+ ions of the peptide EFLLIFR. w and u ions were manually assigned. Lower panel: Example of distant Met side chain losses in ECD of a Met-containing peptide. See text for explanation of the u ion nomenclature.

radical site along the peptide backbone is likely to be a universal process behind both the local and the distant side chain losses. The variance histogram in Figure 3 reflects the kinetics of this process, which is influenced by two factors: the probability for the radical to reach ${}^{\alpha}C_{n-m}$ or ${}^{\beta}C_{n-m}$ atom starting from ${}^{\alpha}C_n$ and the probability for the side chain to be lost. The latter probability involves in turn different energy barriers for different residues.

Detection of the Presence of Met. Of all amino acids studied, methionine showed the most distant losses. To test whether specific losses of $\cdot CH_2SCH_3$ from methionine-containing $z^{\bullet}$ ions can be used as an indicator of the presence of this amino acid in the molecule, ECD mass spectra all Met-containing peptides were analyzed. Of 2668 such mass spectra, the loss of 61.0107 ± 0.0015 Da from $z^{\bullet}$ ions was found in 804 spectra, or 28%. The same "loss" was also detected in 94 peptides that did not contain methionine, which corresponds to a 12% false positive rate. Such a rate is above the acceptable level of 5%, and therefore the presence of methionine should not be concluded from the distant side chain losses alone. Note that no loss of 61.011 Da was detected from the charge-reduced molecular species $MH_2^{+\bullet}$. The latter species are hydrogen-abundant radicals and show different chemistry than hydrogen-deficient $z^{\bullet}$ ions.[26]

(26) Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57−77.

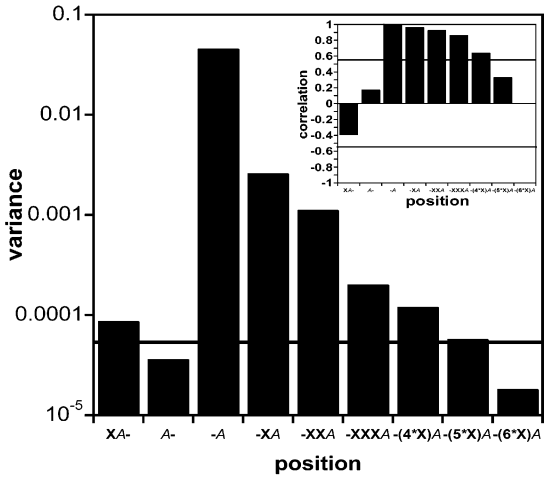


Figure 3. Variance analysis of the influence on side chain losses of amino acids in different positions relative to the cleavage site. Inset: Linear correlation between the patterns of amino acid preferences for side chain losses in neighboring positions. The value at -A is 1.000; the value at -XA is the correlation between the preferences for -A and -XA, and so on.

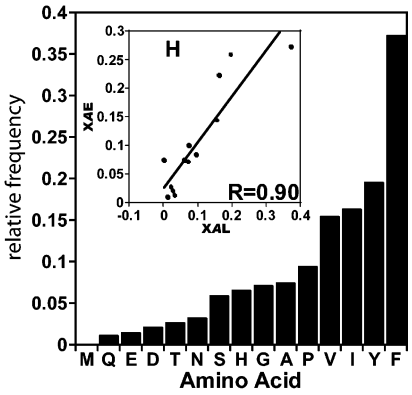


Figure 4. Relative influence of amino acids in the intermediate position X on the side chain losses from Leu_{n-2} in z_n• ions, where the *n*th amino acid is not Leu. Inset: Correlation between Leu_{n-2} and Glu_{n-2} data. The outlying amino acid is His; excluding it gives a correlation of 0.90.

Radical Migration. To determine the exact mechanism of radical migration to distant side chains was outside the scope of the current analytical work. Yet, even from the analytical point of view it was important to determine the main features of radical propagation that induces distant losses. To this end, the rate at which the loss frequency drops with the distance to the initial radical site was determined for Leu, Ile, Gln, and Glu. All these amino acids showed near-exponential decay behavior with similar exponential factors: 0.9 for Ile, 1.2 for Leu, 1.3 for Glu, and 1.5 for Gln. The near-exponential decay is known to occur in multistep random-walk propagation phenomena, e.g., radical transfer in DNAs.[27] To reveal the role of intermediate residues in radical conductance to Leu residue, the frequency of losses of 43.0542 ± 0.015 Da was studied when Leu was in the 3d position from the N-terminus ($m = 2$) in a z_n• ion, while none of the preceding amino acids was Leu. The values determined for each of the 15 amino acids in the second position ($m = 1$) are shown in Figure 4. It follows that the aromatic amino acids Phe and Tyr are the best radical "conductors". This may be due to stabilization of the radical at α-carbon by these aromatic side chains which are believed to be involved in radical transfer in enzymes.[28] Furthermore, the order of stability of aliphatic residues due to the increasing size of their chains, Gly, Ala, Val, Ile, is reflecting their conductive properties in Figure 4. Note that the inflexible Pro residue should present an obstacle for through-space transfer, while glycine due to its exceptional flexibility should promote it. In Figure 4, Pro has higher radical conductance than Gly. Thus Figure 4 is more consistent with multistep, random-walk-type H• migration, possible involving the backbone.

Of the amino acids acting as poor conductors, Met, Gln, Glu, and Asp (see below) tend themselves to lose side chains, which might prevent u ion formation from the 3d amino acid. The presence of Asp among poor conductors is supported by H/D scrambling data,[10] where the scrambling rate of Asp α-carbon hydrogens was found to be reduced.

The H•-conducting properties of intermediate residues with Glu as the 3d residue ($m = 2$; the first two amino acids are not Glu) are very similar to those in Figure 4. The only major difference was detected for His, which was found to be the best radical conductor for Glu_{n-2} and a mediocre one for Leu_{n-2} (inset in Figure 4). The explanation can be that the hydrogen bond that Glu readily forms with the side chain of the neighboring His stabilizes the radical at $^{\alpha}$C of His. Without His, the data sets for Glu_{n-2} and Leu_{n-2} correlate with $R = 0.90$. Data sets for Gln_{n-2} and Ile_{n-2} also mutually correlated with themselves as well as with Glu_{n-2} and Leu_{n-2}, supporting the hypothesis that the radical-transfer mechanism between the first and the third residues has common ground for all these amino acids.

Intramolecular versus Intracomplex H• Transfer. The reason why w ions are not appearing more frequently in conventional ECD is the competition of the side chain losses with other radical-terminating processes. One of such processes is intracomplex H• transfer.[11] If after N−C_{α} bond cleavage and before separation of c′ and *z*• fragments H• is transferred from c′ to *z*•, the latter species becomes an even-electron molecule, and the side chain loss through mechanism (1) is no longer possible. As an experimental verification of the above statement, Figure 5 presents an integral mass spectrum for all z_n• ions with Ile_n. The peak at 0 Da corresponds to the monoisotopic mass, while the peak at 1 Da is a doublet with the lighter component being due to the ^{13}C isotope (+1.003 Da) and the second one due to H• addition (+1.008 Da). The peak at −29.0396 Da is the monoisotopic mass of the w ion, while the peak to the right is a singlet with only a ^{13}C component (inset in Figure 5). Thus H• transfer to *z*• ions and w ion formation from *z*• ions are two competing processes: w ions are only formed from *z*• species, and once formed, they no longer engage in hydrogen rearrangement.

The same message follows from Figure 6, where the ratio of occurrence frequencies of w and *z*• or *z*′ ions is presented as a function of the cleavage position in the peptide sequence. Note that *n* here is the total length of the tryptic peptide and not of the *z* ion. Contrary to the behavior of intracomplex H• transfer that preferentially occurs when the N−C_{α} cleavage happens in internal

(27) Giese, B.; Wessely, S.; Spormann, M.; Lindemann, U.; Meggers, E.; Michel-Beyerle, M. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 996−998.
(28) Stubbe, J.; Nocera, D. G.; Yee, C. S.; Chang, M. C. Y. *Chem. Rev.* **2003**, *103*, 2167−2201.

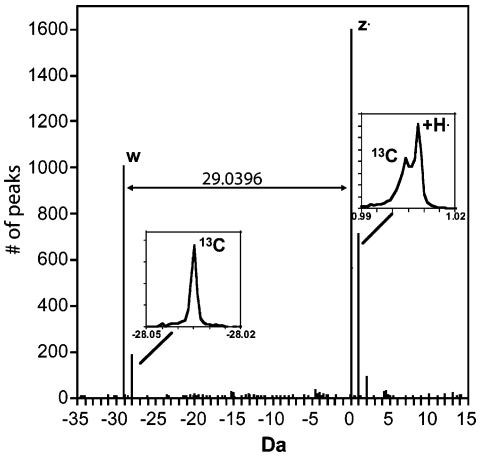


Figure 5. Summed mass spectra for all z_n, ions with Ile_n. The peak at 0 Da corresponds to the monoisotopic mass of z_n, while the peak at −29 Da is due to the monoisotopic mass of w_n. Note that the satellite peak at z• is a doublet, while that at w is a singlet, indicating that w ions formation is an alternative to H• transfer to z• ions.

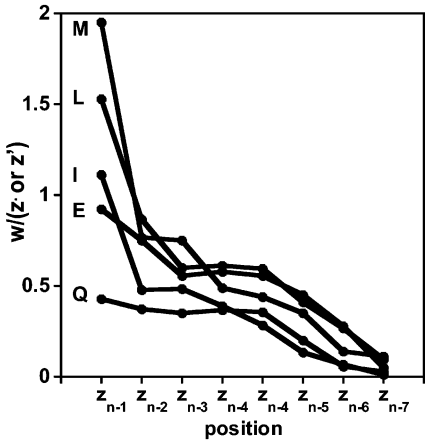


Figure 6. Decay of the ratios of w ion and z ion frequencies (z = z• or z′) with the distance of the lost side chain from the N-terminus (*n* is the total length of the tryptic peptide). The general decaying trend holds for all amino acids.

parts of the sequence,[11] w ions are readily formed after cleavages close to the N-terminal and more seldom after cleavages close to the center of the molecule. This is because for cleavages close to the N-terminus the lifetime of the (c′ + z•) complex is shorter,[11] which limits the probability of intracomplex H• transfer and increases the chances of w ion formation. On the other hand, cleavages distant from the N-terminus create larger c′ species that are more tightly bound in a complex with z• ions than small c′ species, and more likely to donate H• to z•.[11] Thus, for enhanced formation of w ions, the lifetime of (c′ + z•) complexes should be shortened. This can be achieved by increasing the energy of the electrons, as in HECD, or preheating the precursor ions with IR radiation.[29]

Radical w• Ions Formed by N-Terminal Asp. Unlike homologues Glu, Asp does not exhibit abundant w ion formation. Instead of 44.997 Da loss (w ions), Asp tends to lose 43.989 Da, i.e., to form w• ions (w• = w + H•). Figure 7 demonstrates that w• ions are formed from Asp 50 times more often than w ions. It

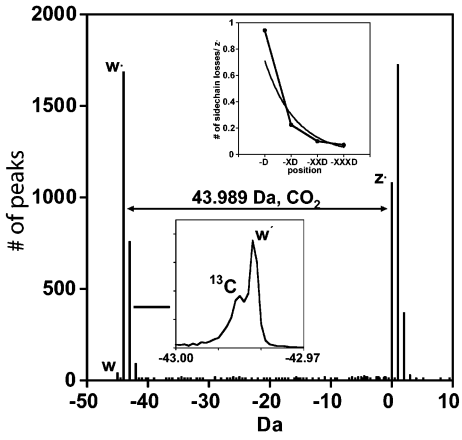


Figure 7. Same as Figure 5, but for Asp. Note that the main loss is not of •COOH (w) but of CO_2 (w•). Since the latter species is a radical, hydrogen rearrangement from the complementary fragment is possible, leading to w′ (zoom-in). Inset shows exponential decay of the w•/z• ratio.

follows from the inset in Figure 7 that such losses corresponding to decarboxylation (CO_2 ejection) are relatively more frequent than any other partial side chain loss, including Met side chain. The Asp losses are however as rapidly decaying with distance to the original radical site as losses from most other amino acids. The dominance of the even-electron loss is apparently due to the much larger stability of the CO_2 molecule compared to the •COOH radical (•CH_2−COOH lost from Glu is more stable). Since the radical remains on the peptide fragment, decarboxylation opens the way for further radical reactions, i.e., a radical cascade.[8,10] For instance, hydrogen rearrangement from the complementary fragment is possible, which results in w′ ions (compare zoom-ins in Figure 5 and 7).

CONCLUSIONS

Partial side chain losses in "conventional" ECD leading to w and u ions were found to be more widespread than previously thought. The losses are particularly abundant from amino acids Leu, Ile, Glu, Asp, Gln, and Met. In analysis of dications of tryptic peptides, the identity of roughly half of all Xle residues can be assigned on the basis of w ions in single-scan ECD data. The low rate of disagreement with Mascot-predicted sequence indicates high reliability of such assignment and its potential for high-throughput, proteomics-grade de novo sequencing.[21]

The side chain losses were found not only from the N-terminal amino acid (w ions) and the adjacent residue (u ions), but also from amino acids up to four and more residues away from the cleavage site, thus confirming and extending previous findings that were based on analysis of individual mass spectra.[17] Statistical analysis of the data confirmed that the most likely explanation for such distant losses is a multistep radical migration driven by the desire to minimize the potential energy of the radical site. In practice, all losses more distant than from the third residue can be ignored, as they appear with the <5% probability. The exceptions are the losses from methionine and possibly from cysteine that were previously studied[30] but fell outside the scope of current investigation because of the low statistics.

(29) Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759−1768.

(30) Chalkley, R. J.; Brinkworth, C. S.; Burlingame, A. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1271−1274.

The losses from Asp are unique, since they leave the radical on the peptide fragment. Such losses open way for a radical cascade. Knowledge of this and other details of radical-driven processes in peptide fragments should increase our knowledge of radical propagation in polypeptides and improve analytical utility of radical-based fragmentation techniques, such as ECD and ETD.[31]

ACKNOWLEDGMENT

This work was supported by the Knut and Alice Wallenberg Foundation and Wallenberg Consortium North (Grant WCN2003-UU/SLU-009 to R.A.Z. and an instrumental grant to R.A.Z. and Carol Nilsson) as well as the Swedish research council (Grants 621-2004-4897, 621-2002-5025, and 621-2003-4877 to R.A.Z.). Frank Kjeldsen and Christopher Adams are acknowledged for insightful discussion.

Received for review October 12, 2006. Accepted December 21, 2006.

AC0619332

(31) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9528−9533.