Can the (M – X) region in electron capture dissociation provide reliable information on amino acid composition of polypeptides?

Kim F. Haselmann,* Bogdan A. Budnik and Frank Kjeldsen
Department of Chemistry, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. E-mail: kfh@chem.sdu.dk
Nicolas C. Polfer
Department of Chemistry, The University of Edinburgh, King’s Buildings, West Mains Road, Edinburgh EH9 3JJ, Scotland, UK
Roman A. Zubarev
Division of Ion Physics, Department of Materials Science, Uppsala University, Uppsala, Sweden and Department of Chemistry, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

It has been suggested that small losses from reduced peptide molecular species in electron capture dissociation (ECD) could indicate the presence of certain amino acids [H.J. Cooper, R.R. Hudgins, K. Håkansson and A.G. Marshall, J. Am. Soc. Mass Spectrom. 13, 241 (2002)], similarly to immonium ions in high-energy collision-activated dissociation. The diagnostic value in ECD of the (M – X) region (1 Da ≤ X ≤ 130 Da) was tested on several synthetic peptides. The insufficiency of the existing knowledge for making correct conclusions on the amino acid composition is demonstrated and new suggestions of the origin of losses are presented based on the “hot hydrogen atom” ECD mechanism. Generally, it is shown that not only protonation but also charge solvation is responsible for the small losses. The origin of 17 Da and 59 Da losses is revisited and a new mechanism for the 18 Da loss is suggested. The loss of a side chain plus a hydrogen atom is found to be a rather reliable indicator of the presence of histidine, tryptophan, tyrosine and, to a lesser degree, threonine. The overall conclusion is that the (M – X) region does contain information on the amino acid composition, but extraction of this information requires additional studies.

Keywords: Fourier-transform mass spectrometry, electrospray ionisation, hot hydrogen atom rearrangement mechanism, hydrogen atom affinity, gas-phase basicity, proton affinity, hypervalent species

Introduction

Sequencing of polypeptides, partial or total, de novo or library-based, is an important task in biological and biomedical research. Because of the inherent risk of mass overlap between the N- and C-terminal fragments, even the cleavage of all inter-residue bonds may prove insufficient for obtaining a unique sequence candidate. Additional sequence information is, therefore, always useful. For small and medium size peptides (< 20 residues), information on the presence or absence of certain amino acid residues can provide a significant aid in establishing the sequence. In high-energy collision-activated dissociation (CAD), such information is obtained from the presence or absence of immonium ions [H₂N=CHR]+ and other related ions in the region below m/z 160. High-energy CAD is, however, unavailable on many modern instruments. Here we investigate the question whether similar information can be obtained from the analysis of small losses in electron capture dissociation (ECD).³

ECD is a relatively new sequencing technique in Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS). One of the most abundant ECD channels is the loss of small groups, for example, hydrogen, ammonia or guanidino groups, from the precursor molecular ions.¹ Recently, Cooper et al.¹ characterised some of these small losses using fragmentation patterns of ten different peptides. They confirmed the loss of the amino acid side chains from...
histidine, previously reported by Haselmann et al., and found new losses from arginine and lysine. All these losses were one dalton heavier than the mass of the side chain radical, suggesting hydrogen rearrangement and/or incorporation in the side chain prior to cleavage. They also found a loss of a group one dalton lighter than the side chain radical of the methionine residue. Apparently, the hydrogen rearrangement was, in this case, from the side chain. Perhaps the most important finding was that, with the exception of methionine, small losses were only observed from the most basic residues in the gas phase (Arg > His > Lys). Although these findings are undoubtedly potentially important from an analytical point of view, their completeness and generality remains unclear. We decided to test the available knowledge on the synthetic peptide KH with the sequence KIMHASELMANN, i.e., to deduce its partial amino acid composition by analysing the small losses observed in ECD of its dications. In ambiguous situations, we made use of a flexible peptide synthesizer to test different hypotheses on several peptides. To explain these observations, the “hot hydrogen atom” model was used, as originally proposed by McLafferty et al.

Experimental

Synthesis

The peptides used in this study were, unless stated otherwise, synthesised in-house by automated solid-phase Fmoc (9-fluorenylmethoxycarbonyl) synthesis using a research-scale (< 25 μmol) ResPep peptide synthesizer (Intavis AG, Gladbach, Germany). The coupling time was chosen to be 25 min. After completion of the synthesis, the resin was washed three times with dichloromethane (Merck, Darmstadt, Germany) and dried in a dry nitrogen flow for 30 min. Simultaneous cleavage of the peptide from the resin and side chain deprotection was performed at 21°C in a mixture (95 : 2.5 : 2.5 v/v/v) of trifluoroacetic acid (Fluka, Steinheim, Germany), triisopropylsilane (Fluka) and MiliQ-water (Millipore, CA, USA). After two hours of deprotection, the peptide was precipitated with cold (0°C) t-butyl methyl ether (Fluka) and then redissolved in 70% acetonitrile and lyophilised. The identity of the product and the presence of contaminants were checked by Fourier transform (FT) mass spectrometry. The LH-RH peptide (L9761) was obtained from Sigma (St. Louis, MO, USA) and used without further purification.

Mass spectrometry

Nano-electrospray ionisation, using a hexapole-based interface (Analytica of Branford, MA, USA) modified with a heated metal capillary, was performed on a 4.7 T Ultima FT mass spectrometer from IonSpec (Irvine, CA, USA). Peptide samples were dissolved in a standard electrospray mixture of water, methanol and acetic acid (49 : 49 : 2 v/v/v) to a concentration of 10⁻⁵ M. A 3–5 μL aliquot was loaded into a metallised pulled-glass capillary (Protana Engineering, Odense, Denmark). The electrospray-produced multiply-charged ions were externally accumulated in the hexapole for 0.5 s and then transmitted to the FT cell by RF-only quadrupole ion guides. The capture of ions in an open-ended cylindrical FT cell with extra trapping plates was achieved by gated trapping. For ECD, the ions of interest were isolated by application to the excitation electrodes of a pre-programmed waveform. An indirectly-heated dispenser cathode, 3 mm in diameter and operated at 5 V and 1.18 A, was employed. For ECD, the electron-emitting surface was biased to –1 V for 100–200 ms. Between 100 and 200 acqui-
sitions were accumulated and averaged. In Fourier transforma-
tion of the time domain, one zero-filling without apodisation was performed. Mass spectra were internally
 calibrated with respect to the parent ion peak. The average
 mass accuracy of the reduced species obtained in the spectra
 shown was 4.1 ppm with a standard deviation less than
 3 ppm.

Computational method

In some cases, gas-phase proton affinities and con-
former energies were calculated by a single-point energy
technique optimised at the B3LYP/6-311++G(d,p) level of theory
using the Gaussian 98 software on Unix based machines.
Structures were first optimised by the PM3 semi-empirical
method.

Results and discussion

For clarity, we only considered in this study losses aris-
ing from the capture of just one electron by the precursor ion
\([M + nH]^+\). The mass region of interest was 1 to 130 Da
below the singly-reduced molecular ion \([M + nH]^+\). Here, we shall call it the \((M - X)_k^+\) region, where \(k = n - 1\).
The \((M - X)_k^+\) region of the KH peptide is shown in Figure 1.
The smallest loss was observed to be the loss of 1 Da; this
 corresponds to hydrogen atom desorption, a commonly
 observed loss for peptides less than ~ 5 kDa. The observed
 ratio \([M + H]^+ / (M + 2H)^+\) = 0.9 is typical for peptides of
 this mass range (< 1.5 kDa). Although even this hydrogen
 atom loss might potentially provide structural information,
 exploring this opportunity lay outside the scope of this study.

Ammonia loss

Another matter is the loss of 17 Da, corresponding to
 ammonia desorption, that is also prominent in the spectrum.
Of the amino acids that are suggested to be associated with
 this loss, the KH peptide contains N-terminal lysine but no
 arginine. This suggestion has been supported by the studies
 of fragmentation of hypervalent species of primary amines,
 showing that both N–H (corresponding to a hydrogen atom
 loss) and N–C (ammonia loss) bond cleavages occur, with
 ammonia loss being a minor pathway. Cooper et al. observed no loss of ammonia from a synthetic peptide con-
taining C-terminal lysine (Ac-CAAAAAK), and suggested
 that either ammonia is lost from a residue other than lysine or
 the ammonia loss from lysine must necessarily involve
 another amino acid. We suggest an alternative explanation
 based on the intramolecular charge solvation effect. The
 synthetic peptide used in the cited study was tailor-made for
 folding in the gas phase to an α-helix stabilised by the solva-
tion of the charged lysine side chain on backbone carbonyl
groups. Hudgins and Jarrold have done extensive ion mobili-
 ty and molecular dynamics simulations of the peptide in

Scheme 1.
question. \(^{11}\) They have shown how the terminal ε-amino group in the lysine participates in hydrogen bonding with backbone carbonyl groups. The hypervalent species produced upon electron capture by such a structure would most likely lose a hydrogen atom to the carbonyl possessing higher hydrogen atom affinity than the amino group. Consistent with that explanation, no hydrogen atom loss (desorption) was observed from that peptide.

Besides the lysine, there is another site in the KH peptide that can lose ammonia, namely the N-terminal amino group. This site has a slightly lower gas-phase basicity than the lysine side chain. Rodriguez et al. \(^{12}\) by ab initio calculations that the most stable conformation in protonated triglycine (GGG) is a five-membered ring formed by the protonated N-terminal amine and the backbone carbonyl group of the same amino acid. Electron capture by such a structure would likely give \(c\) and \(z\) neutral fragments. Glycine is the most flexible amino acid and can, therefore, more readily fold to produce a five-membered ring structure that is less likely to lose ammonia. This was observed in the ECD spectrum of the synthetic peptide with the sequence (GAV)_6 P-NH₂, that did not lose ammonia from the reduced molecular species (spectra not shown). In other peptides exhibiting more bulky side chain groups and/or gas-phase folding, such a five-membered ring can also exist, albeit in a distorted and, thus, less stable form and could, therefore, be lost upon neutralisation. The mechanism of the ammonia loss from the N-terminal amino group is shown in Scheme 1. The observations of Cooper et al. \(^{4}\) are also consistent with this mechanism, since no ammonia loss was expected from the N-terminus.

**18 Da loss**

The peak at –18 Da, corresponding to a water loss (18.011 Da), has only been reported once by Cooper et al. who proposed no mechanism for this loss. \(^{4}\) We rationalise it in Scheme 2 by the partial proton solvation of a hydroxyl group leading to a radical-site-initiated, direct bond cleavage to form a water molecule and a secondary alkyl radical upon electron capture. Even though the proton affinities of alcohols are, in general, ~ 100 kJ mol\(^{-1}\) less than those of amines (but ~ 200 kJ mol\(^{-1}\) higher than those of hydrocarbons), the
participation of hydroxyl groups in charge solvation cannot be excluded.\(^1\)

To test the mechanism, we investigated the \((\text{M} - \text{X})\) region of the synthetic peptide TTTDSTTPAPPTK (Figure 2). The 18 Da loss is, here, rather abundant, consistent with the presence in the molecule of eight hydroxyl groups besides the carboxylic acid functionality. With 62% of all amino acids in this peptide containing a hydroxyl group, the probability of this group participating in proton solvation is very high. Additionally, losses corresponding to 35 and 36 Da are also visible. Cooper et al.\(^4\) suggested that the 36 and 34 Da losses were two water losses (34.053 Da) and two ammonia losses (34.053 Da), respectively. Direct ECD losses upon one electron capture usually include just one group. Therefore, we suggest that the second group is lost much later (through an ergodic mechanism), as a result of deposition of excessive energy by hot electrons that are always present in an electron beam.\(^1\)

A tyrosine-containing peptide was also tested for water loss (Figure 3), but a 19.005 Da loss was observed instead. This loss is not typical and we explain it by secondary fragmentation, i.e. water loss, from the protonated molecular species \([\text{M} + \text{H}]^+\) that has lost a hydrogen atom, i.e. \(\text{H}_2\text{O}\) (19.018 Da). A loss of 108.054 Da from the reduced peak can also be seen in the spectrum. This corresponds to the tyrosine side chain plus a hydrogen atom, i.e. o-cresol (108.058 Da). Phenolic compounds behave like acids whereas alcohols act more often as bases. Phenolic compounds lose water whereas alcohols lose ammonia. Phenol is more acidic in the gas phase (1470 kJ mol\(^{-1}\)) than alcohols (~ 1560–1590 kJ mol\(^{-1}\)). Phenolic compounds behave like acids whereas alcohols act more often as bases. The proton affinities increase when going from water (724 kJ mol\(^{-1}\)) towards more and more substituted alcohols [for example, \(\text{PA}(\text{t-BuOH}) = 816.1\text{ kJ mol}^{-1}\)] because the energetically most favourable site of the proton attachment is para to the OH substitution on the ring, which creates a resonance-stabilised product.\(^1\)

This product ion would lead, not to water loss upon electron capture, but to the side chain loss upon rearrangement. It must be pointed out that in all of the above molecules exhibiting water losses in ECD, carboxylic acid moieties were also present. Therefore, a water loss originating from, for example, a water-bound acylium ion structure cannot be excluded (see below).

45 Da and 59 Da losses

The 45 Da loss observed in the spectrum in Figure 1 (45.026 Da) has been attributed by Cooper et al.\(^4\) to the loss of \(\text{CH}_3\text{N} = \text{O}\) (45.022 Da) from the amide groups of the asparagine/glutamine side chains. The KH peptide does contain two asparagines. On the other hand, the 59 Da loss has so far only been associated with arginine and attributed to \(\text{CH}_5\text{N} = \text{O}\) (59.037 Da). Besides the fact that no arginine is present in the peptide, the closest fit to the observed mass loss of 59.041 Da is \(\text{C}_2\text{H}_5\text{N} = \text{O}\) (59.037 Da). We suggest the following mechanism of this loss (Scheme 3). The most likely site of protonation of an amide is the carbonyl oxygen in both the solution phase and in the gas phase, but the probability of N-protonation increases with increasing substitution (R) of the amide. Lin et al.\(^16\) have investigated the protonated formamide, both experimentally and theoretically, with \(\text{ab initio}\) calculations. They found the O-protonated form to be the most stable isomer that can lose both water and ammonia, whereas the less stable N-protonated isomer is the precursor for CO loss. Intuitively, the less favourable protonation at the N-site can be explained by the delocalisation of the nitrogen lone-pair towards the carbonyl group, which makes it less accessible for protonation. In polypeptides, the carbonyl oxygen of the amide group in the asparagine/glutamine side chain could participate as much in charge solvation as the backbone carbonyl groups. Upon electron capture by such a charge-solvated amide, an intermediate, unstable radical can be formed (Scheme 3). This

![Figure 3. ECD spectrum of the small loss region (M+ – X) of the synthetic peptide with the sequence PYYFYVYH-OH.](image-url)
radical can further proceed either via radical-site-initiated, direct bond cleavage (the primary alkyl radical initially formed at the peptide could migrate to form a more stable tertiary alkyl radical at the alpha-carbon, also shown) or via a radical-site-initiated α-cleavage, which gives rise to the losses of 45 Da and 59 Da, respectively. The loss of ammonia from the amide group is only possible if it is protonated at the nitrogen atom, which is less likely unless no other basic sites are available. For example, in the ECD spectrum of dications of the synthetic peptide PNNFNVNH-OH, the observed losses were 1 Da (major), 18 Da (minor), 45 Da, 59 Da (major) and 82 Da (Figure 4). The 45 and 59 Da losses are in agreement with the above discussion. Apparently, no ammonia loss was present. The water loss here seems to come from the carboxylic acid functionalities.

An additional complication is that the 45 Da and 59 Da losses can be confused with 46 Da and 60 Da losses observed in ECD of the 2+ ion of TTTDSTTPAPTTK (Figure 2). The 46 Da loss corresponds to the side chain of threonine plus one hydrogen atom, ethanol (46.042 Da). Threonine is an abundant amino acid in the sequence and is shown above to participate in proton solvation. The carbonyl group of a carboxylic acid can also participate in charge solvation. The 46 loss can also correspond to formic acid (HCOOH, 46.006 Da). This could be due to the loss of a side chain carboxyl group plus one hydrogen atom, such a loss could originate from direct bond cleavage of the neutralised carbonyl group participating in charge solvation, similarly to the amide case. The 60 Da loss can come from aspartic acid (side chain of Asp + 1 Da = 60.021) contained in the sequence. Consistent with this idea, the 60 loss was observed when asparagines were replaced with aspartic acids, giving the sequence PDDFDVDH-OH (data not shown).

Loss of 82 Da

This loss has been attributed by Cooper et al. to the loss of a histidine side chain (82.053 Da), previously reported by Haselmann et al. in the ECD spectra of b+ ions. Since the histidine residue there was situated close to the C-terminal, this loss was rationalised through neutralisation of the nearby acylium site. Here, as well as in Reference 4, no acylium ion structure is assumed, meaning that other mechanisms must be used to explain the phenomenon. Nguyen and Turecek have studied the protonation sites in gaseous pyrrole and imidazole. They found the following proton affinities of pyrrole sites: 866, 845 and 786 kJ mol⁻¹ for C2, C3 and C5, respectively.
C3 and N, respectively. The most likely site for protonation in pyrrole is C2 and not N, again indicating that the nitrogen lone-pair is made unavailable by delocalisation. For imidazole, the picture changes slightly, with the proton affinities calculated to be 941, 804, 791, 791 and 741 kJ mol\(^{-1}\) for the N3, C4, C2, C5 and N1 positions, respectively. The protonation of the imine nitrogen, the favoured site, yields a stable aromatic cation. Neutralisation by electron capture at this site would lead to hydrogen desorption or the stable secondary carbon radical, originally proposed by McLafferty et al.\(^3\) The charge in the aromatic imidazole ring is, however, not localised to a single (e.g. N3) atom but delocalised across the whole five-membered ring. Other possible, but less likely, structures include the tertiary radical species at C5, from which the side chain loss could readily proceed by a radical-site-initiated $\alpha$-cleavage. This gives a resonance-stabilised radical product and the methylimidazole (Scheme 4), which is 74 kJ mol\(^{-1}\) more stable than the aromatic secondary carbon radical (the lowest energy conformer among the different tautomers). From the ECD cleavage pattern (c\(_z\) to c\(_{z'}\) and z\(_{y'}\) to z\(_{y}\) ions are observed in Figure 1), His4 in the KH peptide can be identified as a protonation site. Therefore, this side chain loss is not surprising; it is also observed in spectra of other histidine-containing peptides (Figures 3, 4 and 5).

**Unobserved 74 Da loss**

Cooper et al.\(^4\) rationalised a 74 Da loss observed in one out of two methionine-containing peptides through the loss of the methionine side chain. We note that it is the only full side chain loss involving hydrogen transfer from the side chain. This is despite the high hydrogen atom affinity of the side chain sulphur atom. In ECD of the KH peptide, no 74 Da loss is observed, even though the sequence contains two methionines. The loss of a methionine side chain incorporating the “hot hydrogen atom” would correspond to –76.035 Da; this loss is not observed either. The absent losses from methionines are probably due to the fact that these amino acids are not participating in charge solvation in the KH peptide.

![Scheme 4.](image)

**Figure 4.** ECD spectrum of the small loss region (M$^+$ – X) of the synthetic peptide with the sequence PNNFNVNH-OH.
In principle, the hot hydrogen atom mechanism does not require the presence of charge for the cleavage of neutral fragments; high hydrogen atom affinity is enough. Our experience is, however, that protonation or charge solvation is essential for the initiating losses. To illustrate the latter point, consider the (M – X) region of a peptide containing a tryptophan residue. McLafferty et al. calculated the hydrogen atom affinity of the tryptophan side chain residue to be 18 kJ mol\(^{-1}\) higher than that of the disulfide bond, which, in turn, is much higher than that of the carbonyl group. Backbone cleavages near tryptophan are often observed in ECD spectra; we found that its side chain is lost as well. For example, luteinising hormone releasing hormone (LH-RH), with the sequence pEHWSYWRP-G-NH\(_2\), exhibits an abundant 131 Da loss (Figure 5) which corresponds to the tryptophan side chain plus 1 Da. ECD of synthetic peptides with the sequence FWHGMLHNFTNW-NH\(_2\) and AFAAWARE-OH also gave abundant tryptophan side chain losses (data not shown). Is it possible for the tryptophan side chain to participate in proton solvation? We calculated the proton affinities for atoms of 3-methylindole by DF methods (B3LYP/cc-pVTZ); they turned out to be 872, 892 and 841 kJ mol\(^{-1}\) for N1, C2 and C3, respectively. Therefore, upon protonation or charge solvation, the C2 position is favoured and would give rise to the protonated species shown in Scheme 5. Note that our suggestion regarding the

![Scheme 5.](image)

Figure 5. ECD spectrum of the small loss region (M+ – X) of the natural peptide luteinising hormone releasing hormone (LH-RH) with the sequence pEHWSYWRP-G-NH\(_2\).
role of charge solvation does not apply for disulfides, for which an alternative cleavage mechanism has been suggested.  

Conclusions

For the KIMHASELMANN peptide, based on previous knowledge from the (M–X) region of the ECD spectrum of dications, one would have predicted the presence in the sequence of histidine, asparagine/glutamine and arginine, and an absence, or at least a low abundance, of methionine residues.Whilst the first two predictions would be perfectly correct, the other two would miss the target completely. The erroneous identification of arginine would have been due to the presence of the 59 Da loss that, as shown here, arises from asparagine/glutamine. To the postulated losses from charged groups we add the possibility of losses due to charge solvation. The additional knowledge of the origin of ammonia and water losses detects the presence of serine/threonine and either N-terminal amine or lysine.

It appears that losses of side chains plus one Da can give direct identification of the presence of histidine, tryptophan, tyrosine and, perhaps, threonine, whereas only a combination of 17, 44, 59 and 101 Da losses is indicative of arginine, 45 and 59 Da for asparagine and 46 and 60 Da for aspartic acid. The ammonia loss and water loss are, by themselves, of lesser value. It must be emphasised that the absence of a loss is not a direct indication of the absence of that amino acid; it only indicates that this amino acid is not participating in charging or charge solvation. This can be considered as a warning against possible over interpretation of the ECD data, including the data presented in this paper.

With the possible exception of the methionine residue, all full side chain losses were accompanied by hydrogen atom transfer to the side chain. This, as well as the fact that the presented mechanisms succeeded in explaining the observed effects, supports the hot hydrogen rearrangement mechanism.

Summarising, the experienced difficulty in reaching an unequivocal conclusion on amino acid composition testifies to the inadequacy of the current knowledge of the (M–X) region. However, we expect that deriving valid information from that region will be possible in the future. Obviously, more work needs to be done in this direction, both in terms of experiments as well as theoretical studies.

Acknowledgements

Heinrich Gausepohl is gratefully acknowledged for help in setting up the peptide synthesis facility. Frank Jensen is acknowledged for help during the theoretical calculations. The FT MS instrument was funded by the Instrument Centre Programme (grant number NSF 9700471 and STVF 00001242). The peptide synthesis programme in our laboratory is supported by the Danish National Research Council, grants SNF 51-00-0338 and SNF 51-00-0238.

References


Received: 11 August 2001
Revised: 9 October 2002
Accepted: 11 October 2002
Web Publication: 29 October 2002