JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. **35**, 1399–1406 (2000)

SPECIAL FEATURE: COMMENTARY

Mobile and localized protons: a framework for understanding peptide dissociation

Vicki H. Wysocki,* George Tsaprailis, Lori L. Smith and Linda A. Breci
Department of Chemistry, University of Arizona, Tucson, Arizona 85721-0041, USA

Protein identification and peptide sequencing by tandem mass spectrometry requires knowledge of how peptides fragment in the gas phase, specifically which bonds are broken and where the charge(s) resides in the products. For many peptides, cleavage at the amide bonds dominate, producing a series of ions that are designated b and y. For other peptides, enhanced cleavage occurs at just one or two amino acid residues. Surface-induced dissociation, along with gas-phase collision-induced dissociation performed under a variety of conditions, has been used to refine the general 'mobile proton' model and to determine how and why enhanced cleavages occur at aspartic acid residues and protonated histidine residues. Enhanced cleavage at acidic residues occurs when the charge is unavailable to the peptide backbone or the acidic side-chain. The acidic H of the side-chain then serves to initiate cleavage at the amide bond immediately *C*-terminal to Asp (or Glu), producing an anhydride. In contrast, enhanced cleavage occurs at His when the His side-chain is protonated, turning His into a weak acid that can initiate backbone cleavage by transferring a proton to the backbone. This allows the nucleophilic nitrogen of the His side-chain to attack and form a cyclic structure that is different from the 'typical' backbone cleavage structures.

KEYWORDS: peptide dissociation; surface-induced dissociation; mobile proton; enhanced cleavage; collision-induced dissociation

INTRODUCTION

Gas-phase fragmentation of peptides has been investigated for many years.[1–3] This work has become more important with the recent development of computer algorithms that use protein and DNA databases to identify proteins. Certain programs make use of tandem mass spectrometry data to generate sequence tags (short stretches of amino acid sequence) that can be used to search databases of known sequence.[4–9] A powerful program, Sequest, makes use of uninterpreted tandem mass spectra and compares those spectra with 'predicted' spectra generated from sequence stretches in the database that match the measured molecular mass (M_r) of the 'unknown'.[10–13] It seems logical that programs that are designed to produce sequence information from spectra could be improved if we could more readily predict spectra from sequence. It is the goal of much of our research to explain how peptides fragment and why certain residues or residue combinations lead to unusual cleavage pathways.[14–20] The peptides that fragment in unusual ways generally are not sequenced well by current algorithms and are not useful in protein identification. It is our hypothesis that these peptides could be used in peptide and protein identification if we could understand peptide dissociation in enough detail to predict which peptides will fragment in common vs uncommon ways and to predict the main features of the corresponding spectra.

* *Correspondence to*: V. H. Wysocki, Department of Chemistry, University of Arizona, Tucson, Arizona 85721-0041, USA.
E-mail: vwysocki@u.arizona.edu

Contract/grant sponsor: NIH; *Contract/grant number*: GM R01 51387.

MOBILE PROTON MODEL

Several research groups have contributed to the development of a general model for how peptides fragment in the gas phase; it is the combined knowledge from a large number of studies that effectively weaves together to define and refine the mobile proton model.[2,14,17,21–33] Our group began studies of peptide dissociation by applying the technique of surface-induced dissociation (SID), an activation method for tandem mass spectrometry that involves collisions of polyatomic ions into organic surfaces to cause the projectile ions to fragment[14–20,25,34,35] (G. Tsaprailis, V. H. Wysocki, W Zhong, H. Nair and J. Futrell, submitted for publication). Much of our early work was designed to take advantage of the fact that SID deposits a relatively narrow (few eV) distribution of energies into the ion population and that the average energy deposited can be easily varied, allowing fragmentation to be examined as a function of systematic changes in internal energy. Until recently, all of this research was performed on a tandem quadrupole system, which introduces the limitation that ions must fragment on a time-scale

Received 14 September 2000
Accepted 9 October 2000

of a few microseconds (after collision but before entry to the quadrupole) in order to be mass analyzed and detected.

Several studies have been performed in which small model peptides are dissociated at a variety of collision energies.[14,18,19,35] The peptide structure is systematically changed so that amino acids of varying basicity are included in the peptides. The population of different protonated forms of a peptide depends on the internal energy content of the peptide and the gas-phase basicities of the different protonation sites of the peptide. When the fraction of fragmentation is plotted against the collision energy to produce a fragmentation efficiency curve, there is a correspondence between the basicity of the most basic amino acid and the collision energy required to induce dissociation (on the time-scale of the instrument). These results are consistent with the idea that fragmentation of most protonated peptides requires the involvement of a proton at the cleavage site, i.e. that the cleavages are 'charge-directed.' If an amino acid side-chain tightly binds or 'sequesters' a proton, energy will be required to move that proton from the basic side-chain to the peptide backbone to induce dissociation. That is, addition of energy (e.g. by activation in tandem mass spectrometry, MS/MS) alters the initial population of protonated forms (mobilizes the proton) and increases the population of protonated forms with energies higher than that of the most stable structure. These protonated molecules exist as a population of ions with the proton located at various backbone heteroatoms; the protonation at the backbone sites initiates charge-directed cleavages of the backbone to produce b- and/or y-type sequence ions. It should be noted that, although it is not stated explicitly throughout this paper, a proton located at a particular basic site generally will be solvated by other heteroatoms in the system, i.e. there is a tendency to bury charge via intramolecular solvation.[36]

The energy required for proton 'mobilization' from a basic side-chain or the amino terminus depends on the amino acid composition, with dissociation energy requirements greatest for arginine-containing peptides and decreasing in the order Arg-containing > Lys-containing > non-basic, mimicking the order of decreasing gas-phase basicity.[14] In selected cases, more energy might be required to mobilize the proton than is required to initiate 'charge-remote' fragmentation pathways (i.e. no proton involvement) and in those cases different fragmentation pathways may be followed than is normally the case.

The mobile proton model was tested in our group in several ways.[14–16,18–20] Examples are shown in Figs 1 and 2. The top panel of Fig. 1 shows fragmentation efficiency curves, plots of fraction of fragmentation vs SID collision energy, for two peptides, RPPGFSPF and PPGFSPFR, that differ in the location of arginine (R). Singly charged des-R1- and des-R9-bradykinin show the same high-energy onset for dissociation because in either case the charge is sequestered at arginine (top right curves labeled +1). When doubly protonated, RPPGFSPF shows a lower onset for dissociation than does PPGFSPFR (Fig. 1A, left). In these peptides, the first proton in each peptide can locate at the highly basic arginine but the second proton is more tightly bound in PPGFSPFR than in RPPGFSPF because the second most basic site in PPGFSPFR is the proline at the

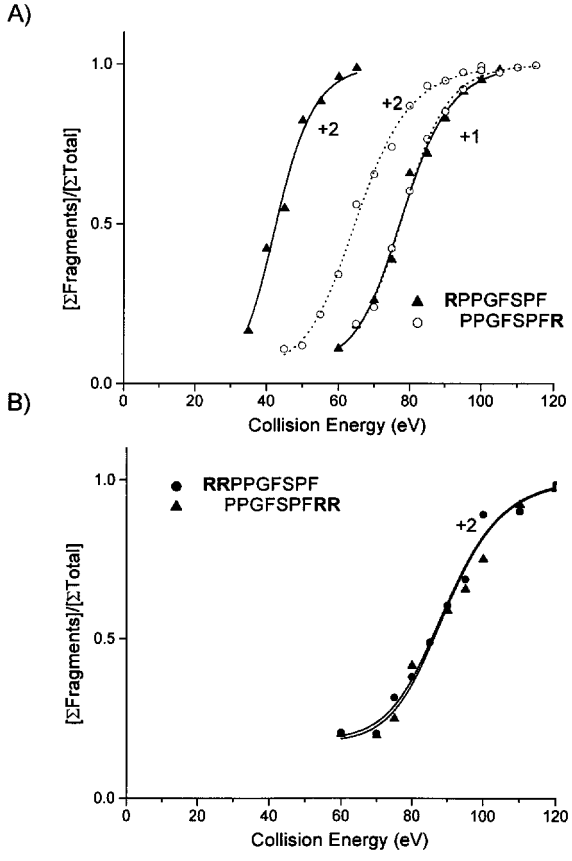


Figure 1. Fragmentation efficiency curves [(sum of fragment % relative abundance)/(total % relative abundance)] for singly and doubly protonated peptides that differ in the number and locations of arginines (**R**). Spectra were acquired by surface-induced dissociation on a tandem quadrupole mass spectrometer. From Ref. 14.

amino terminus, a secondary amine. All sites, other than R, in RPPFGSPF are less basic than the proline free amino terminus in PPGFSPFR. When two arginines and two protons are present (bottom curves), the onsets for dissociation are higher and identical regardless of the terminus on which the two Rs are located, consistent with the two protons being 'sequestered' at the two arginines.

Figure 2 shows fragmentation efficiency curves for peptides that were derivatized to contain a fixed charge. The peptide without Arg (R), $\varphi_3P^+CH_2C(O)LDIFSDF$, shows very different onset energies for dissociation depending on whether a proton is added or not, consistent with the added proton being 'mobile,' i.e. available to induce dissociation, in $[\varphi_3P^+CH_2C(O)LDIFSDF + H^+]$. When the fixed charge, Arg, and an added proton are all present in the fragmenting ion $[\varphi_3P^+CH_2C(O)LDIFSDFR + H^+]$, the dissociation energy is higher than when the fixed charge and an added proton are present in the absence of arginine. This is consistent with sequestration or strong binding of the added proton at the arginine, requiring extra energy to be input to cause dissociation of this ion compared with the fixed-charge ion with an added proton but no Arg. Both Figs 1 and 2 and our published work on multiply protonated dendrimers[34] address the common misconception that doubly charged ions fragment more easily than singly charged ions simply because of Coulombic repulsion. Doubly charged

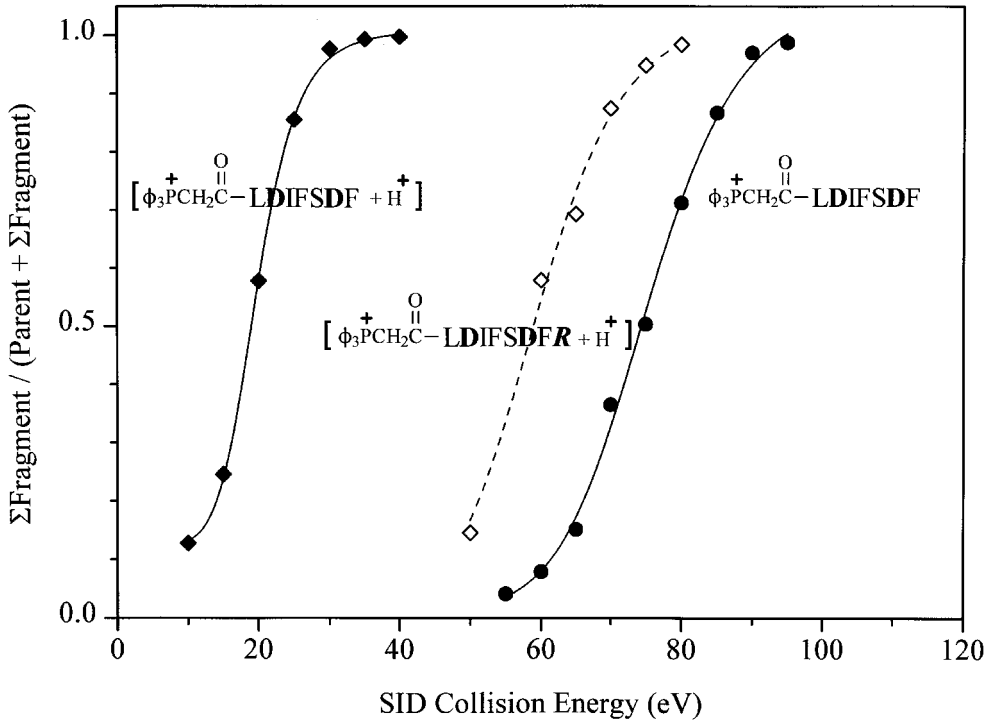


Figure 2. Fragmentation efficiency curves [(sum of fragment % relative abundance)/(total % relative abundance)] for three peptides with a fixed charge at the amino terminus. The peptides differ in the number of charges and in the presence or absence of arginine (**R**) in the sequence. The spectra were acquired by surface-induced dissociation on a tandem quadrupole mass spectrometer. From Ref. 16.

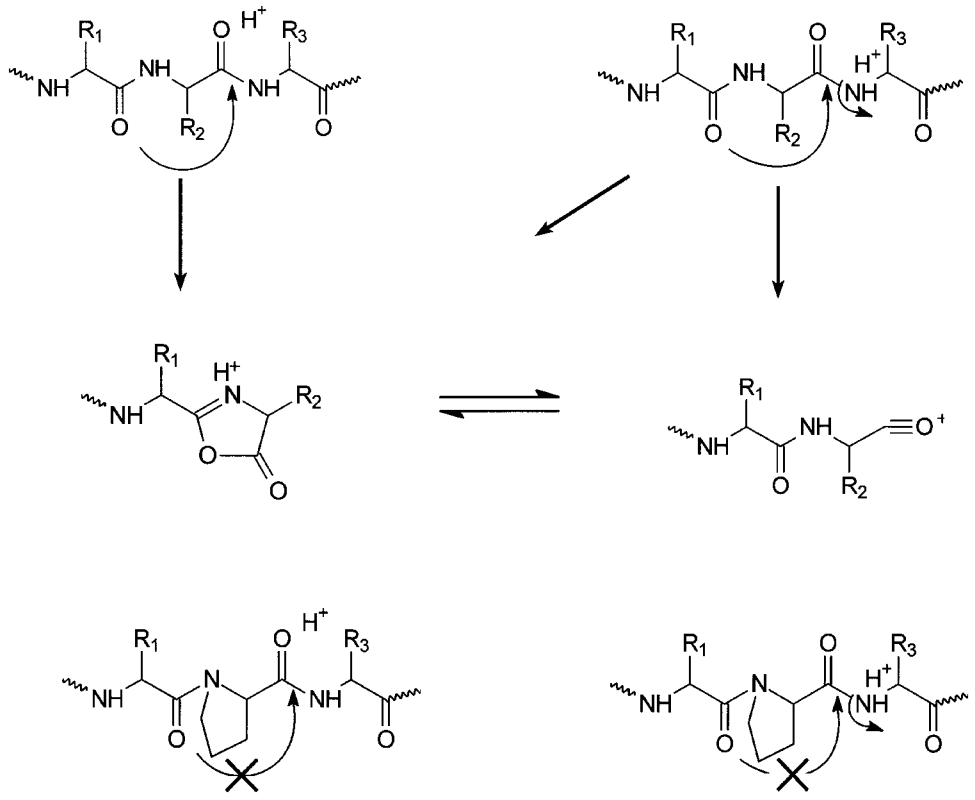
RPPGFSPF fragments at a much lower collision energy than singly charged **R**PPGFSPF, while doubly-charged **RR**PPGFSPF fragments at a higher energy than either singly or doubly charged **R**PPFGSPF (Fig. 1). It would be difficult to explain these results with a Coulombic repulsion argument but the results are readily explained by the mobile proton model (see above). Likewise, doubly charged $[\varphi_3P^+CH_2C(O)LDIFSDF + H^+]$ has a significantly lower onset for dissociation than does $[\varphi_3P^+CH_2C(O)LDIFSDF\mathbf{R} + H^+]$. Both are doubly charged and differ in size by only one amino acid residue. Strong binding of the charge at Arg, with intramolecular solvation, readily explains the difference in dissociation onsets.

Several papers have appeared that support the general mobile proton model for peptide dissociation. Examples come from the work of Harrison, Gaskell, Vaisar, Glish, Russell, Vékey, Paisz and co-workers.[21,24,28,32,33,37–39] The utility of the 'mobile proton model' is that it provides a general framework for understanding and predicting how peptides will dissociate in the gas phase. The ability to predict dissociation is useful in the development of peptide sequencing and protein identification algorithms that make use of mass spectrometric data.

Where must the proton be located to initiate backbone cleavages?

In the above discussion of the mobile proton model, it was stated that protons are transferred intramolecularly from basic side-chains to the heteroatoms along the backbone in order to initiate backbone cleavages to produce, e.g. b_n and y_n ions. The location of the proton after transfer to the backbone was not stated explicitly because the location of the proton on the backbone in the fragmenting structures has not been clearly established. Scheme 1 shows two possibilities. Protonation at carbonyl oxygens is more logical because carbonyl oxygens of amide bonds are more basic than amide nitrogens. With a proton at a carbonyl oxygen, a nearby carbonyl (*N*-terminal to the protonated carbonyl oxygen) can serve as a nucleophile to attack the electropositive carbon of the protonated carbonyl,[40] resulting in a protonated oxazolone b ion or, via an ion–molecule complex, a neutral *N*-terminal fragment and a corresponding y ion.[41] The accompanying contributions by Schlosser and Lehmann,[40] Wesdemiotis and co-workers[41] and O'Hair[42] show mechanisms that could involve protonation at amide nitrogen or carbonyl oxygen. Although protonation on amide nitrogen is less likely because of the lower gas-phase basicity of the amide N, protonation at this site has been shown by calculations to decrease significantly the C(O)—N bond order suggesting facile cleavage of the amide bond.[14,43] This cleavage would produce an acylium ion that can close to a protonated oxazolone or a protonated oxazolone directly if the C(O)—N cleavage occurs concomitant with attack of the adjacent nucleophilic carbonyl oxygen (Scheme 1).

Based on work in progress in our group on enhanced cleavage at proline, we speculate here that protonation on the carbonyl oxygen, and fragmentation from that structure, is more consistent with experimental data for b_n and y_n formation. Cleavage at proline commonly occurs at proline's *N*-terminus to produce either b_n or y_n ions. Cleavage of a peptide at the *C*-terminal side of proline does not often occur and this is used as evidence that b_n ions are protonated oxazolones.[40] If the *N*-protonated



Scheme 1

structure of Scheme 1 (bottom right) is considered, there is no reason to expect that cleavage at proline should not occur to produce acylium b_n ions (which could further fragment to produce a_n ions). The low bond order of the C(O)—N bond and the lack of an activation barrier for this cleavage suggest that facile cleavage should occur (the molecule would not sample the subsequent ring closure and 'decide' not to cleave the C(O)—N bond). For the oxygen-protonated structure, however, attack of the adjacent carbonyl oxygen at the electropositive carbon of the protonated carbonyl is necessary to initiate the H-transfer that must occur prior to loss of the *C*-terminal fragment. This attack is hindered at proline (Scheme 1, bottom left). The fact that cleavage, and b_n or y_n formation, generally does not occur *C*-terminal to proline is suggested here to be evidence that b_n formation occurs from the *O*-protonated intermediate and not the *N*-protonated intermediate. In a future publication, this idea will be further supported by the different ratios of cleavage that occur at the *N*-terminal side vs the *C*-terminal side of D and L stereoisomers of *N*-alkylated residues (*N*-methylalanine and piperidine) (L. Breci, V. Wysocki and T. Vaisar, in preparation). Additional evidence for involvement of *O*-protonation in dominant cleavage pathways of protonated peptides comes from the work of O'Hair, Reid and coworkers. They showed that dominant water loss from pentaglycine and tetraglycine and from GGC-OMe and GCG-OMe involves loss of one of the backbone oxygens via a dissociation that must be initiated by *O*-protonation rather than *N*-protonation.[44,45]

There is not yet enough data in the literature to state conclusively that peptide backbone fragmentations occur mainly from *O*-protonated vs *N*-protonated forms. Both possibilities should be considered when attempting to reconcile bodies of peptide fragmentation data.

RESIDUE-SPECIFIC CLEAVAGES

A recent paper that appeared in *Electrophoresis* examined tandem mass spectra of 319 peptides and showed that even a powerful program such as SEQUEST does not identify a relatively high percentage of the peptides of a digest of a known protein.[46] Other approaches identify even fewer peptides. This is often irrelevant in the identification of a known protein, because partial sequence information and molecular mass information are usually sufficient to identify the protein unambiguously. It is relevant, however, in attempts to sequence peptides and unknown proteins. The inability to sequence the complete subset of the peptides of a protein by their tandem mass spectra is not surprising if one considers the great variability that exists in the gas-phase dissociation patterns of peptides. Several authors have pointed out that certain peptides show unusual spectra that involve enhanced cleavage at particular amino acid residues or residue combinations and that these peptides lack a complete pattern of backbone cleavages that would provide complete sequencing of the peptide.[19,20,47–53] For unknown proteins or for peptides, these spectra are a problem for mass tagging or *de novo* interpretation programs. The spectra do not provide enough information to generate a mass tag (a short sequence stretch that can be used for database searching) and do not provide enough sequence ions for *de novo* interpretation programs. These spectra could be matched to 'theoretical' tandem mass spectra if we knew how to predict these 'unusual' spectra

from sequence strings. The spectra are 'unusual' only in the sense that they do not produce contiguous series of backbone cleavage sequence ions (e.g. a/b, y). In spite of their current unpredictable appearance, spectra with only a few major peaks are common as evidenced by the above-mentioned work of Simpson and co-workers.

We present below a summary of our efforts to explain fragmentation mechanisms for two cases of selective cleavage of gas-phase peptides: enhanced cleavage at acidic residues[16,20] and enhanced cleavage at protonated histidine (G. Tsaprailis, V. H. Wysocki, W Zhong, H. Nair and J. Futrell, submitted for publication). Additional similar studies are ongoing, including a refinement of the current knowledge of enhanced cleavage at proline. For this research, a variety of different multi-stage mass spectrometers are used to follow the dissociation patterns of the peptides. By applying different activation methods in a variety of instrument types (eV SID, eV CID, keV CID; trapping vs non-trapping instruments), the molecules are activated by different energy deposition mechanisms and dissociate over different time frames. If the dissociation chemistry persists and dominates throughout these conditions, the true unimolecular chemistry of the molecule is being examined and not some instrument-dependent enhancement of particular ion types. In addition to the application of different MS/MS tools, traditional means of studying the dissociation mechanisms are applied in these gas-phase studies, e.g. isotopic labeling, blocking of proposed reaction sites and amino acid substitutions.

Enhanced cleavage at acidic residues

Several authors have noted enhanced cleavage at acidic residues (aspartic and glutamic acid).[22,23,38,47–49,54] Gaskell and co-workers suggested the involvement of arginine in peptides that cleave selectively at acidic residues and proposed that side-chain interactions between the arginine and the acidic side-chain might explain the results.[21,24,31,37] This explanation has a lot of appeal and is consistent with much of the experimental data. There are certain aspects of the data, however, that cannot be explained by this model. A series of papers have been published that show that enhanced cleavage occurs *C*-terminal to aspartic acid if arginine is present in a peptide that also contains aspartic acid and if the number of added protons is equal to or less than the number of arginines present.[15,16,19,20] This is illustrated in Fig. 3. A peptide containing two Asp, one Arg and one added proton, [LDIFSDFR + H^+], shows enhanced cleavage at both Asp—Xxx bonds [y_2 and y_6, Fig. 3(A)] while the corresponding compound with two Asp, no Arg and one added proton, [LDIFSDF + H^+], cleaves non-selectively [Fig. 3(B)]. Our explanation for these and related results is that the arginine(s) tightly binds the proton(s) allowing the acidic hydrogen of the aspartic acid side chain to be the 'reactive' proton that initiates cleavage as shown in Scheme 2. This is supported by data for related peptides that have been derivatized to add a fixed charge to them instead of a proton; even these peptides that contain no added proton fragment selectively at the aspartic acid [b_2 and b_6, Fig. 3(C)].[16] The product of the selective cleavage at Asp is an anhydride. The anhydride structure is consistent with MS/MS/MS data that show facile loss of CO_2 plus CO to produce a d ion. The selective cleavage

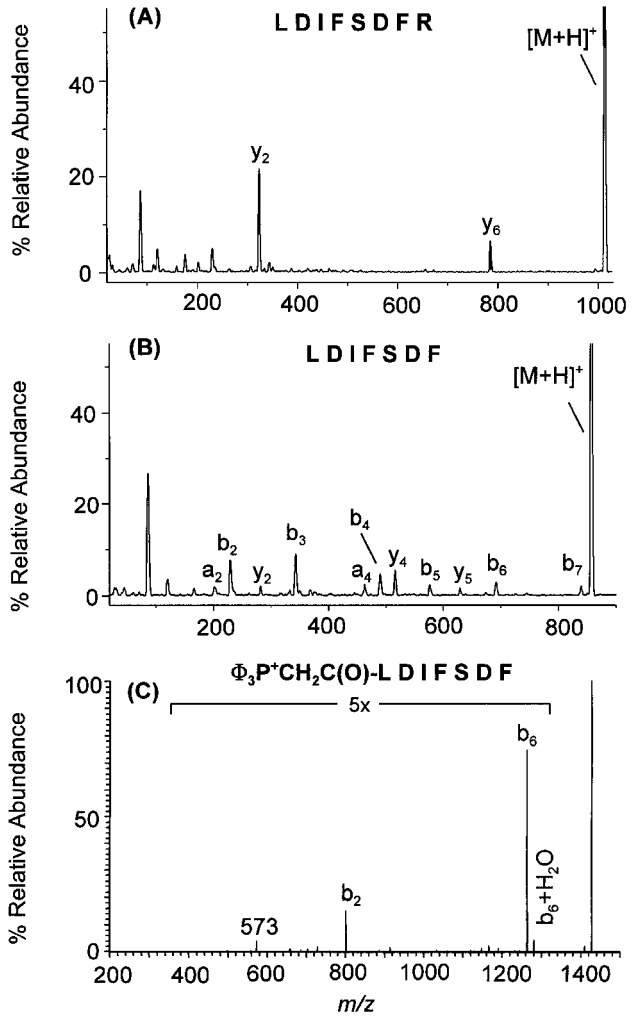
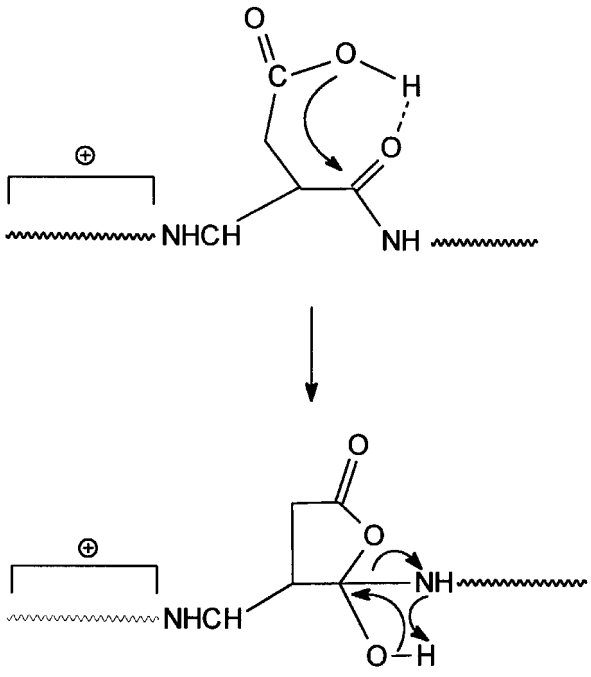


Figure 3. Surface-induced dissociation spectra of singly-protonated (A) LDIFSDFR and (B) LDIFSDF. (C) Collision-induced dissociation spectrum of ϕ_3P^+LDIFSDF; spectrum acquired with a quadrupole ion trap instrument (Finnigan LCQ). Data from Ref. 16.

chemistry at Asp/Glu has been shown to persist even in the case of direct dissociation of larger peptides and proteins. The 11^+ charge state of lysozyme, which contains 11 arginines, fragments selectively at acidic residues whereas the 14^+ charge state (which contains protons in excess of the number of arginines, i.e., mobile protons) fragments non-selectively.[55]

In order to test the robustness of the Arg–Asp sequencing 'rule,' a search and inspection of a small database of peptide spectra acquired on a commercial quadrupole ion trap instrument was performed. The spectra were provided by the Yates group (Scripps Institute). Inspection of the singly charged spectra for 10 tryptic peptides that contain Arg plus Asp and/or Glu showed that *all* fragment selectively at the acidic residues (see Table 1), in agreement with the rule that enhanced cleavage will occur at acidic residues if any added protons are bound to an Arg. We have shown previously that peptides containing lysine do not show the same trend.[15] In agreement with our earlier work, inspection of spectra of singly charged peptides that contain Lys and Asp and/or Glu showed no enhancement of cleavage at acidic residues. From the same database, spectra of 42 doubly charged



Scheme 2

Table 1. Peptide sequences and the most abundant fragment ions produced upon dissociation of the singly protonated peptides in a quadrupole ion trap mass spectrometer

Sequence	Most abundant ions, ion trap data
AGENIGALLR	y_7, H_2O loss
AGENVGVLLR	y_7, H_2O loss
AGFAGDDAPR	y_3, y_4, H_2O loss
ALLEEQPR	y_3, y_4, H_2O loss
EHILLGR	y_6, H_2O loss
INEGILQR	y_5
LAQEDPSFR	y_4
VAAEGVILAR	y_6
DLGEALR	y_3, y_6, H_2O loss
IPDIDLIVIR	y_5, y_7, y_8, H_2O loss

peptides containing one Arg and Asp and/or Glu were inspected. The 18 doubly charged peptides that contain Arg and Asp and/or Glu (but not His or Pro, other residues that can cause selective cleavage), were found not to show enhanced cleavage at acidic residues. This result is consistent with the 'rule' that an available mobile proton (a proton in excess of the number of Arg) will allow charge-directed cleavage at various sites along the peptide backbone. When other residues that can cause selective cleavage are present in the same molecules that contain Arg and Asp and/or Glu (22 additional database spectra of doubly charged peptides), selective cleavages at these other residues often dominate, presumably because of the charge directed nature of the cleavages (e.g. see histidine discussion below).

Enhanced cleavage at protonated histidine

Cleavage near histidine is enhanced for many peptides. A common example is encountered by mass spectrometrists who use angiotensin peptides for tuning or calibration purposes; these peptides contain an H—P bond that is the dominant cleavage site under certain conditions. In order to explain this type of cleavage and to develop that understanding into a robust fragmentation 'rule,' so that it might be incorporated into sequencing algorithms, a series of peptides were examined under a variety of dissociation conditions (G. Tsaprailis, V. H. Wysocki, W Zhong, H. Nair and J. Futrell, submitted for publication). The peptide structures were varied and changes included the presence or absence of arginine, the charge state, the type of charge (proton vs fixed-charge derivative), the identity of the residue adjacent to histidine and alkylation of the histidine. While all singly protonated peptide ions containing both histidine and arginine fragment non-selectively, the doubly protonated peptide ions with arginine and histidine, or the singly protonated peptides containing histidine but not arginine, cleave in a selective manner. In particular, dominant complementary b^+/y^+ product ions resulting from cleavage between the HP amide bond are observed for doubly protonated DRVYIHPFHL, DRVYIHPF and RVYIHPF [Fig. 4(A)]. For the fixed-charge derivative, non-selective cleavage occurs if charge is provided only by the charge on the derivative and selective cleavage [Fig. 4(B)] occurs if a proton is added to produce a doubly charged parent ion. In contrast to the results for acidic residues, where selective cleavage occurs at Asp if there is no added proton available along the backbone or at the acidic side-chain, enhanced cleavage at His occurs when an added proton is present on the His side-chain.

The role of histidine in directing selective cleavage was probed further by systematically changing the residue *C*-terminal to histidine (from proline to sarcosine and then to alanine). The results indicate that while b^+/y^+ complementary ion pairs dominate in doubly protonated RVYIHPF [Fig. 4(A)], b_5^{2+} and b_6^{2+} product ions dominate the spectra of doubly protonated RVYIHAF [Fig. 4(C)]. Based on all of the results, a fragmentation mechanism for enhanced cleavage at histidine was proposed (Scheme 3). The proposed selective fragmentation mechanism involves a common intermediate that produces doubly charged b^{2+} ions or a singly charged b^+/y^+ complementary ion pair. The branching ratio between these two pathways is determined by the identity of the residue *C*-terminal to histidine and by the ability of protonated histidine to transfer a proton to the *C*-terminal leaving fragment. The latter contributing factor is based on strong experimental MS/MS spectral evidence which reveals that dominant b_5^{2+} product ions are observed by SID and CID when the histidine side-chain is alkylated in doubly protonated RVYIH*PF [i.e. when there is no proton available for transfer b_5^{2+} is formed instead of b_5^+; compare Fig. 4(A) and (D)]. The b ion structures terminating at histidine residues (b_5^+ and b_5^{2+} as illustrated in Scheme 3) are different from the typical b ions that are thought to be protonated oxazolones and/or acylium ions (Scheme 1). Evidence for this difference comes from MS/MS/MS and resonance ejection experiments. In addition to enhanced cleavage of the His—Xxx bond, it has also been determined that enhanced cleavage can occur at the amide bond one removed (His—Xxx—Yyy, cleavage of the Xxx—Yyy bond, e.g. b_6^{2+} formation from RVYIHPF). This is consistent with molecular modeling results that show strong hydrogen bonding between the His side-chain and the Xxx—Yyy bond.

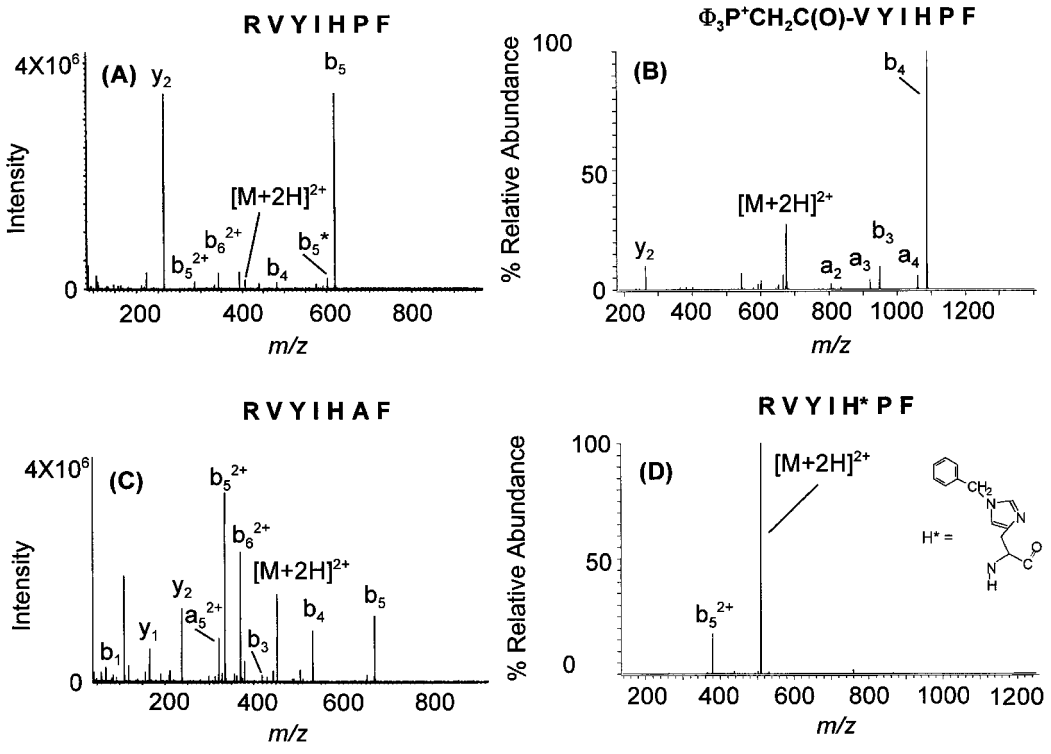
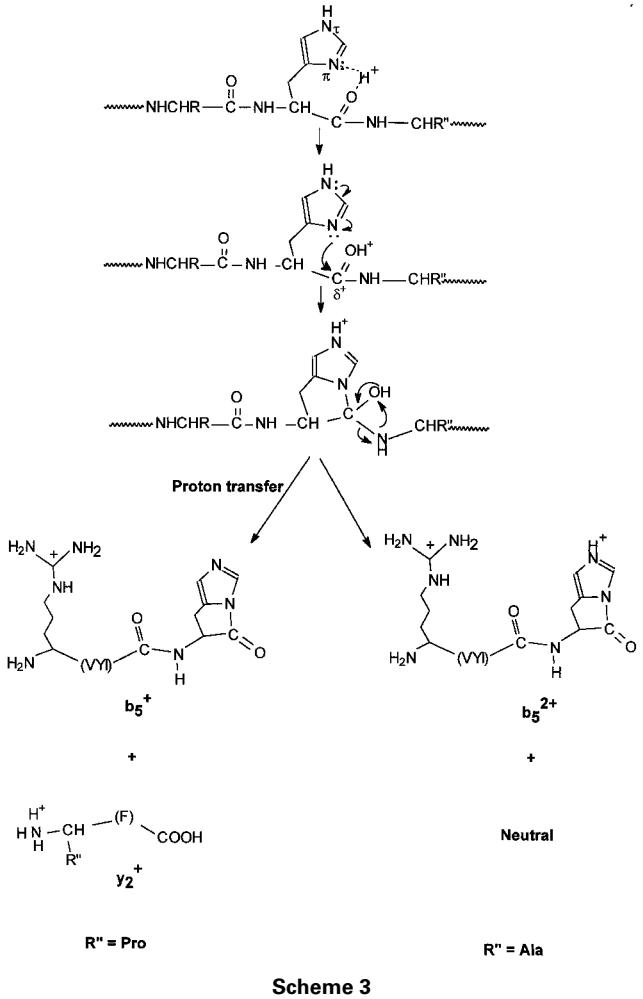


Figure 4. SORI-CID spectra for a series of peptides containing histidine (H) or alkylated histidine (H*). The peptides cleave selectively at H or H*. The charge state of the product ion depends on the charge on His in the forming products and the residue carboxy-terminal to His. Data from G. Tsaprailis, V. H. Wysocki, W. Zhong, H. Nair and J. Futrell, submitted for publication.



Scheme 3

CONCLUSIONS

A detailed knowledge of gas-phase peptide dissociation mechanisms is important to the improvement of algorithms that are used for the automated sequencing of peptides or identification of proteins. Most current sequencing algorithms are designed to use a contiguous ion series present in the MS/MS spectrum. Frequently, enhanced cleavages at specific amino acids occur and the algorithms produce poor results for these MS/MS spectra (even though the MS/MS spectra have a high signal-to-noise ratio). In this paper, SID data that support the general 'mobile proton model' for peptide fragmentation are summarized. Two examples of selective cleavage are presented: enhanced cleavage at acidic residues in the absence of an added proton at the cleavage site and enhanced cleavage at protonated histidine. The understanding developed from such studies of dissociation mechanisms is used to generate fragmentation rules that are then tested against databases of spectra. Once a fragmentation rule has been established and found to apply in a general, predictive way to a large percentage of peptides containing a particular residue or combination of residues, it should be possible to incorporate the fragmentation rule into programs that are designed to predict sequence from spectra.

Acknowledgements

This research was supported by NIH grant GM R01 51387. George Tsaprailis acknowledges a fellowship from the Fonds pour la Formation de Chercheurs et l'Aide á la Recherche. We thank John Yates and David Tabb, Scripps Institute, for providing the spectral database used to test and verify sequencing rules.

REFERENCES

1. Biemann K. *Anal. Chem.* 1986; **58**: 1288.
2. Biemann K, Martin SA. *Mass Spectrom. Rev.* 1987; **6**: 1.
3. Hunt DF, Yates JR, Shabanowitz J, Winston S, Hauer CR. *Proc. Natl. Acad. Sci. USA* 1986; **83**: 6233.
4. Johnson RS, Biemann K. *Biomed. Environ. Mass Spectrom.* 1989; **18**: 945.
5. Mann M, Wilm M. *Anal. Chem.* 1994; **66**: 4390.
6. http://prospector.ucsf.edu.
7. http://www.mann.embl-heidelberg.de.
8. http://www.matrixscience.com.
9. http://prowl.rockefeller.edu.
10. Eng JK, McCormack AL, Yates JR. *J. Am. Soc. Mass Spectrom.* 1994; **5**: 976.
11. Yates JR, Eng JK, McCormack AL. *Anal. Chem.* 1995; **67**: 3202.
12. Yates JR, Eng JK, McCormack AL, Schieltz D. *Anal. Chem.* 1995; **67**: 1426.
13. http://www.thermoquest.com.
14. Dongré AR, Jones JL, Somogyi Á, Wysocki VH. *J. Am. Chem. Soc.* 1996; **118**: 8365.
15. Gu CG, Somogyi A, Wysocki VH, Medzihradszky KF. *Anal. Chim. Acta* 1999; **397**: 247.
16. Gu C, Tsaprailis G, Breci L, Wysocki VH. *Anal. Chem.* 2000; 72.
17. McCormack AL, Somogyi Á, Dongré AR, Wysocki VH. *Anal. Chem.* 1993; **65**: 2859.
18. Nair H, Wysocki VH. *Int. J. Mass Spectrom.* 1998; **174**: 95.
19. Tsaprailis G, Nair H, Somogyi A, Wysocki VH, Zhong WQ, Futrell JH, Summerfield SG, Gaskell SJ. *J. Am. Chem. Soc.* 1999; **121**: 5142.
20. Tsaprailis G, Somogyi A, Nikolaev EN, Wysocki VH. *Int. J. Mass Spectrom.* 2000; **196**: 467.
21. Summerfield SG, Whiting A, Gaskell SJ. *Int. J. Mass Spectrom. Ion Processes* 1997; **162**: 149.
22. Burlet O, Orkiszewski RS, Ballard KD, Gaskell SJ. *Rapid Commun. Mass Spectrom.* 1992; **6**: 658.
23. Burlet O, Yang CY, Gaskell SJ. *J. Am. Soc. Mass Spectrom.* 1992; **3**: 337.
24. Summerfield SG, Cox KA, Gaskell SJ. *J. Am. Soc. Mass Spectrom.* 1997; **8**: 25.
25. Dongre AR, Somogyi A, Wysocki VH. *J. Mass Spectrom.* 1996; **31**: 339.
26. Poulter L, Taylor LCE. *Int. J. Mass Spectrom. Ion Processes* 1989; **91**: 183.
27. Nold MJ, Wesdemiotis C, Yalcin T, Harrison AG. *Int. J. Mass Spectrom. Ion Processes* 1997; **164**: 137.
28. Harrison AG, Yalcin T. *Int. J. Mass Spectrom.* 1997; **165**: 339.
29. Mueller DR, Eckersley M, Richter WJ. *Org. Mass Spectrom.* **1988**: 217.
30. Johnson RS, Krylov D, Walsh KA. *J. Mass Spectrom.* 1995; **30**: 386.
31. Cox KA, Gaskell SJ, Morris M, Whiting A. *J. Am. Soc. Mass Spectrom.* 1996; **7**: 522.
32. Vaisar T, Urban J. *Eur. Mass Spectrom.* 1998; **4**: 359.
33. Vaisar T, Urban J. *J. Mass Spectrom.* 1998; **33**: 505.
34. de Maaijer-Gielbert J, Gu CG, Somogyi A, Wysocki VH, Kistemaker PG, Weeding TL. *J. Am. Soc. Mass Spectrom.* 1999; **10**: 414.
35. Jones JL, Dongre AR, Somogyi A, Wysocki VH. *J. Am. Chem. Soc.* 1994; **116**: 8368.
36. Wyttenbach T, von Helden G, Bowers M. *J. Am. Chem. Soc.* 1996; **118**: 8355.
37. Summerfield SG, Gaskell SJ. *Int. J. Mass Spectrom.* 1997; **165**: 509.
38. Harrison AG, Tu YP. *J. Mass Spectrom.* 1998; **33**: 532.
39. Paisz B,Vekey K. *Rapid Commun. Mass Spectrom.* 1999; **13**: 525.
40. Schlosser A, Lehmann WD. *J. Mass Spectrom.* 2000; **35**: 1382.
41. Polce MJ, Ren D, Wesdemiotis C. *J. Mass Spectrom.* 2000; **35**: 1391.
42. O'Hair RAJ. *J. Mass Spectrom.* 2000; **35**: 1377.
43. Somogyi A, Wysocki V, Mayer I. *J. Am. Soc. Mass Spectrom.* 1994; **5**: 704.
44. Reid GE, Simpson RJ, O'Hair RAJ. *J. Am. Soc. Mass Spectrom.* 1998; **9**: 945.
45. Reid G, Simpson R, O'Hair R. *J. Am. Soc. Mass Spectrom.* 1998; **9**: 945.
46. Simpson RJ, Connolly LM, Eddes JS, Pereira JJ, Moritz RL, Reid GE. *Electrophoresis* 2000; **21**: 1707.
47. Yu W, Vath JE, Huberty MC, Martin SA. *Anal. Chem.* 1993; **65**: 3015.
48. Bakhtiar R, Wu Q, Hofsradler SA, Smith RD. *Biol. Mass Spectrom.* 1994; **65**: 707.
49. Qin J, Chait BT. *J. Am. Chem. Soc.* 1995; **117**: 5411.
50. Jockusch RA, Schnier PD, Price WD, Strittmatter EF, Demirev PA, Williams ER. *Anal. Chem.* 1997; **69**: 1119.
51. Lee SW, Kim HS, Beauchamp JL. *J. Am. Chem. Soc.* 1998; **120**: 3188.
52. Loo JA, Edmonds CG, Smith RD. *Anal. Chem.* 1993; **65**: 425.
53. Vaisar T, Urban J. *J. Mass Spectrom.* 1996; **31**: 1185.
54. Price WD, Schnier PD, Jockusch RA, Strittmatter EF, Williams ER. *J. Am. Chem. Soc.* 1996; **118**: 10640.
55. Stephenson Jr. J, Cargile BJ, McLuckey SA. *Rapid Commun. Mass Spectrom.* 1999; **13**: 2040.

