4

COLLISIONALLY INDUCED DISSOCIATION OF PROTONATED PEPTIDE IONS AND THE INTERPRETATION OF PRODUCT ION SPECTRA

4.1. INTRODUCTION

In mass spectrometric sequencing, the information that describes the amino acid sequence of a peptide is contained in a product ion spectrum. This product ion spectrum is obtained in a tandem mass spectrometry experiment by using collisionally induced dissociation of a protonated or multiply protonated peptide ion. The purpose of this chapter is to describe the fragmentation reactions that lead to the product ion spectrum and a strategy for interpreting these spectra. The strategy used is an iterative process based on an understanding of the structure of protonated peptides, the reaction mechanisms by which they fragment, and the structure of the product ions that are produced.

Routine, complete interpretation of product ion spectra to deduce the entire sequence of a peptide has, to a substantial degree, been replaced by database search programs. These programs can now utilize unprocessed or minimally processed

Protein Sequencing and Identification Using Tandem Mass Spectrometry, by Michael Kinter and Nicholas E. Sherman.
ISBN 0-471-32249-0

product ion spectra to search large databases of essentially theoretical spectra derived from the protein and translated gene sequences to identify peptide sequences in the databases that are consistent with that spectrum. In effect, the product ion spectrum is interpreted and the peptide is sequenced by mathematically matching it to a finite, albeit large, set of possible amino acid sequences. By defining the set of possible amino acid sequences as the experimentally determined protein and gene sequences, the source protein sequence is also identified. Ideally, a protein is identified when a significant number of peptides in a digest can be matched to peptides derived from a particular database amino acid sequence. Because a computer does the searching, a product ion spectrum can be interpreted in periods ranging from a few seconds to perhaps a few minutes at the longest, making it possible to identify dozens of proteins per day. Further, the speed and utility of database searches will continue to grow in parallel to the growth of the databases and the advancement of computing power. The use of these search programs is discussed in Chapter 8 of this volume.

For a number of reasons, however, database searches have not, and should not, eliminate the need for scientists engaged in these experiments to understand peptide fragmentation and to interpret product ion spectra. Firstly, and most important to note, understanding these spectra is necessary to understand and analyze the output from any search program. One should be aware that all search programs are designed to assign a score to each possible match and to use those scores to rank the possible matches and select the best. Therefore, within a given set of possibilities, these scores distinguish the sequence that has the best possibility of matching the spectrum of interest. However, there are few absolute rules for judging the significance of those scores beyond this intra-set ranking, particularly cut-off values for the incorrect matches that may occur for a variety of reasons. Ultimate confirmation of an appropriate match requires some sort of inspection of the spectrum and test of the proposed sequence based on an expected fragmentation pattern. Second, until the databases are complete, computer-based sequencing methods will fail to identify a significant number of proteins for any of a number of reasons. These reasons include low homology of the protein of interest to related database entries, post-translational modifications of the protein of interest, or simply that the protein of interest is a novel protein, unrelated to any database entry. In these cases, some or all of the peptides must be sequenced by interpretation of the product ion spectra to provide information for other types of searches or for other experiments such as cloning or antibody production. These experiments will succeed or fail based on the quality of the product ion spectrum interpretation. Techniques used to aid in the complete, confidant interpretation of the product ion spectra of peptides from novel proteins are described in Chapter 9.

4.2. PEPTIDE FRAGMENTATION CHEMISTRY

Understanding the structure of protonated peptides and their fragmentation pathways plays key roles in one's ability to interpret product ion spectra. Two types of

fragmentation reactions will be discussed in this chapter: the fragmentation of multiply protonated peptide ions induced by low-energy collisions in tandem quadrupole, quadrupole-time of flight, and ion trap mass spectrometers; and the fragmentation of singly protonated peptide ions formed by matrix-assisted laser desorption/ionization in delayed extraction-reflectron-time-of-flight mass spectrometers. Fragmentation reactions seen in tandem instruments that operate in a manner that produces high-energy collisions are not discussed in this volume.

The nomenclature used to describe the different product ions defines two sets of ions that are named based on the peptide terminus retained in the ion (4.1). In this nomenclature system, the a-, b-, and c-ions all contain the N-terminus of the peptide, while the x-, y-, and z-ions all contain the C-terminus. The major N-terminus-containing ion series is the b-ion series and the major C-terminus-containing ion series is the y-ion series. The discussion of peptide fragmentation will focus on the formation of the b- and y-ions because they are the most useful and the most common sequence ions.

4.2.1. Collisionally Induced Dissociation of Peptide Ions Formed by Electrospray Ionization

In positive-ion operating conditions, electrospray ionization produces peptide ions that enter the mass spectrometer with protons attached to all of the strongly basic sites in the peptide. These sites include the N-terminal amine and the side group of any lysine, arginine, or histidine residues. For the purpose of the following discussion, the gas-phase basicity of these sites can be divided into the more basic arginine, histidine, and lysine sites, and the less basic N-terminus site. In the gas phase, a proton associated with the more basic sites is strongly attached and remains associated or fixed at that site even on collisional activation. In contrast, a proton on the less basic N-terminus may move by internal solvation to any of the amide linkages.

As illustrated in Figures 4.1 and 4.2, the result of the migration of the proton originally localized on the N-terminus is that a given protonated peptide formed by electrospray ionization is best viewed as a heterogeneous population of peptide ions, in which different sub-populations of ions have the same amino acid sequence but with a proton associated with each amide linkage. After mass selection, this population of peptides is then accelerated into multiple collisions with gas molecules in the collision cell of tandem mass spectrometers. In the tandem instruments of interest in this volume, those collisions occur with kinetic energies in the 10 eV to 50 eV range and are defined as low-energy collisions. Kinetic energy from these collisions is converted into vibrational energy in the peptide ion, which the peptide ion releases through fragmentation reactions directed by the site of the protonated amide bond. As a result, proton migration by internal solvation is important to the fragmentation chemistry because the variety of protonation sites it produces directs the fragmentation reactions to occur at each of the different amide bonds. This particular view of the manner by which the structure of protonated peptides directs the fragmentation is known as the mobile proton hypothesis (4.2–4.7).

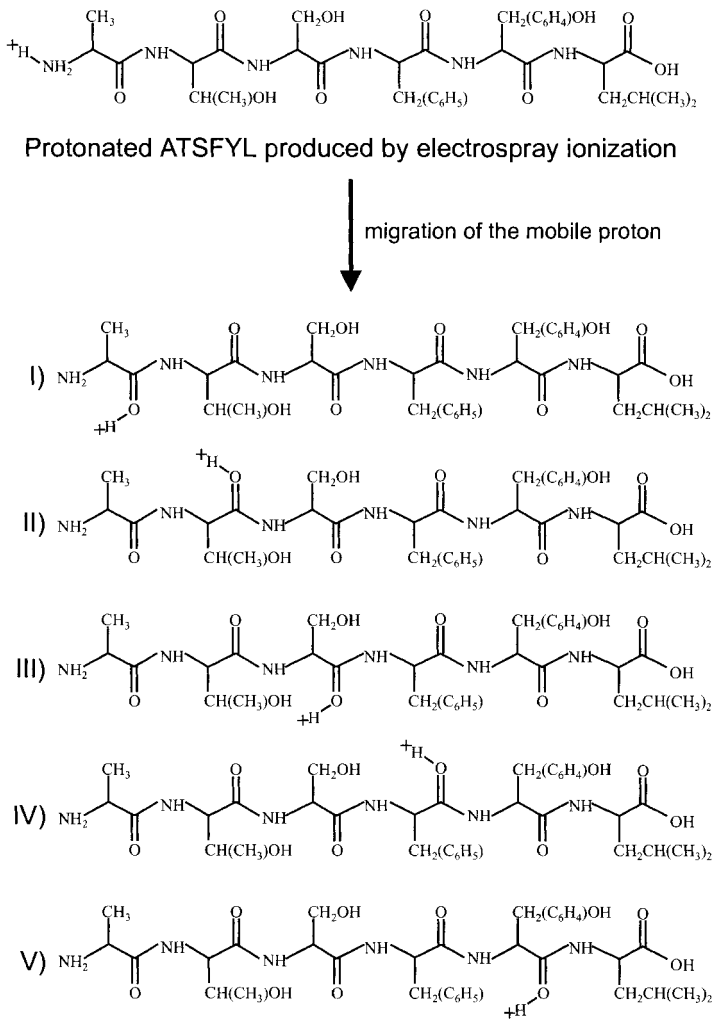

Figure 4.1. The structure of a singly protonated peptide produced by electrospray ionization. A peptide with only one basic site is protonated at that site by electrospray ionization. The proton is subsequently able to migrate by internal solvation, down the length of the peptide, producing a mixed population of protonated peptides with the same amino acid sequence but different sites of proton attachment.

The effect of the mobile proton changes dramatically depending on the charge state of the peptide and the presence of the more basic amino acid residues in the peptide's sequence. The peptide illustrated in Figure 4.1 contains none of the more basic amino acids so that the only site for protonation by electrospray ionization is the N-terminus. This proton can migrate and produce sub-populations of peptide

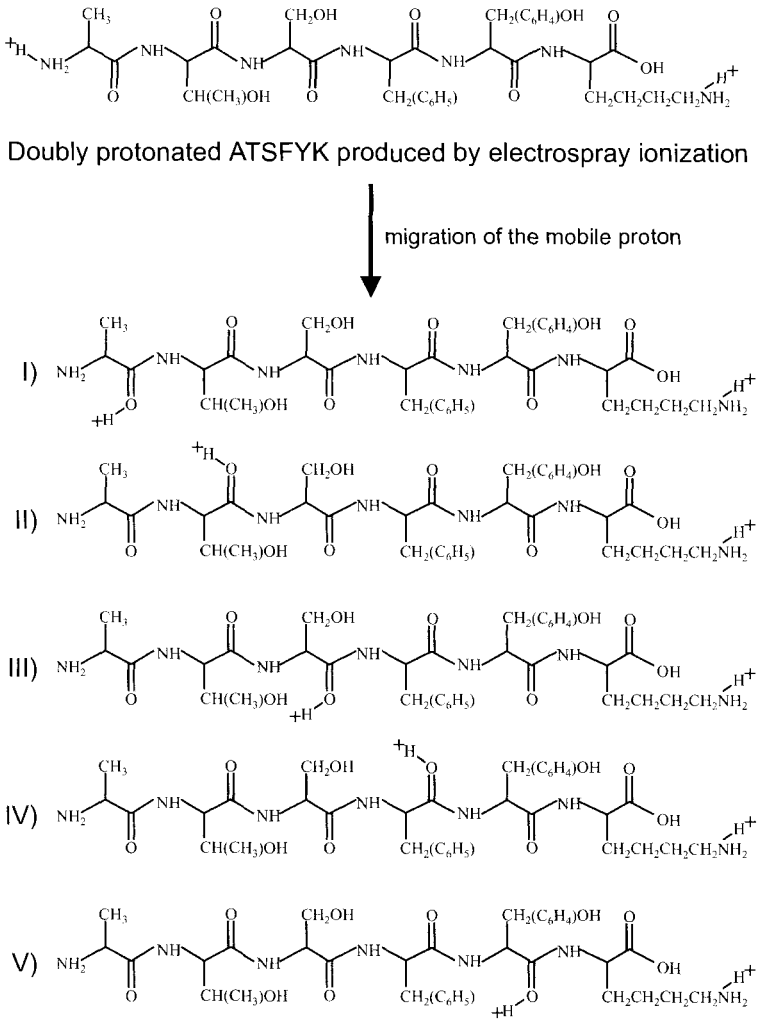

Figure 4.2. The structure of a doubly protonated peptide produced by electrospray ionization. A peptide with two basic sites is protonated at both sites by electrospray ionization. The basicity of the side chain of the C-terminal lysine is such that the proton associated with this site is essentially fixed. The proton attached to the N-terminal amine, however, is not fixed and can migrate down the length of the peptide by internal solvation to produce a mixed population of peptides with the same amino acid sequence but different sites of protonation.

ions with the five different structures shown in the figure. These protonation sites then direct the fragmentation reactions so that each sub-population of peptide ions fragments differently, producing a series of product ions that, taken as a whole, reveal the entire peptide sequence.

In contrast, the peptide illustrated in Figure 4.2 is a tryptic peptide and contains a C-terminal lysine residue. The inclusion of this more basic amino acid residue in the peptide sequence leads to the production of a doubly charged ion in electrospray ionization with one proton attached to the side chain of the C-terminal lysine and the other proton attached to the N-terminus. In this case, the proton attached to the lysine is fixed at that position while the proton at the N-terminus is the mobile proton and would migrate to produce the five structures shown in this figure. As described for the singly charged ion above, one would expect a highly informative series of product ions to be formed because the sub-populations of the doubly protonated peptide that are produced have a proton on each of the amide bonds. In the event that electrospray ionization produced a population of singly charged peptides with this sequence, the proton would be expected to reside on the lysine residue and there would be no mobile proton to direct fragmentation. Without the mobile proton, only limited fragmentation is observed with even the most severe low-energy collision conditions. As a result, the most complete and informative fragmentation is always seen by fragmenting the highest possible charge state of a peptide.

In the case of tryptic peptides producing ions that are triply charged and greater, more than one additional strongly basic residue must be present in the sequence. In these cases, a mobile proton is associated with the N-terminus but the other protons are fixed at the C-terminal lysine or arginine and fixed at the internal histidine, lysine, or arginine residues. This internal fixed charge is important because it tends to alter the migration of the mobile proton in a generally unfavorable manner. Specifically, the mobile proton would tend not to localize at nearby amide bonds so that little fragmentation is observed at those positions. Therefore, in peptides that are triply charged and greater, the amount of sequence information seen in the product ion spectrum is highly dependent on the position of the internal charge sites.

Charge-site-directed fragmentation of the amide bonds can occur through a number of pathways but a major low-energy collisionally induced dissociation pathway is illustrated in Figure 4.3 for the singly charged peptide ATSFYL (shown in Figure 4.1). This fragmentation pathway is initiated by the site of protonation and proceeds through a cyclic intermediate that subsequently fragments by one of the two reactions (4.8, 4.9). The first reaction (Reaction I in the figure) retains the charge on the cyclic oxazolone moiety. When the peptide ion being studied is singly charged, as shown in this figure, this reaction forms the ions in the b-ion series because the N-terminus of the peptide is retained in the ionic species that are detected. In the example shown in this figure, formation of the b_3-ion is illustrated and analogous reactions would occur at all of the other protonated amide bonds except the N-terminal-most amide bond as needed to form the b_1-ion. Product ion spectra of protonated peptides do not contain b_1-ions because the cyclic intermediate that is required cannot be formed. Because this fragmentation reaction (Reaction I) is favored, b-ions dominate the product ion spectrum of singly charged peptides with no basic sites. The second reaction (Reaction II in the figure) retains the charge on the former amide nitrogen, forming a y-ion. Relative to Reaction I in the figure, this reaction is less favorable as indicated by the relative lack of y-ions in the product ion spectra of singly charged ions.

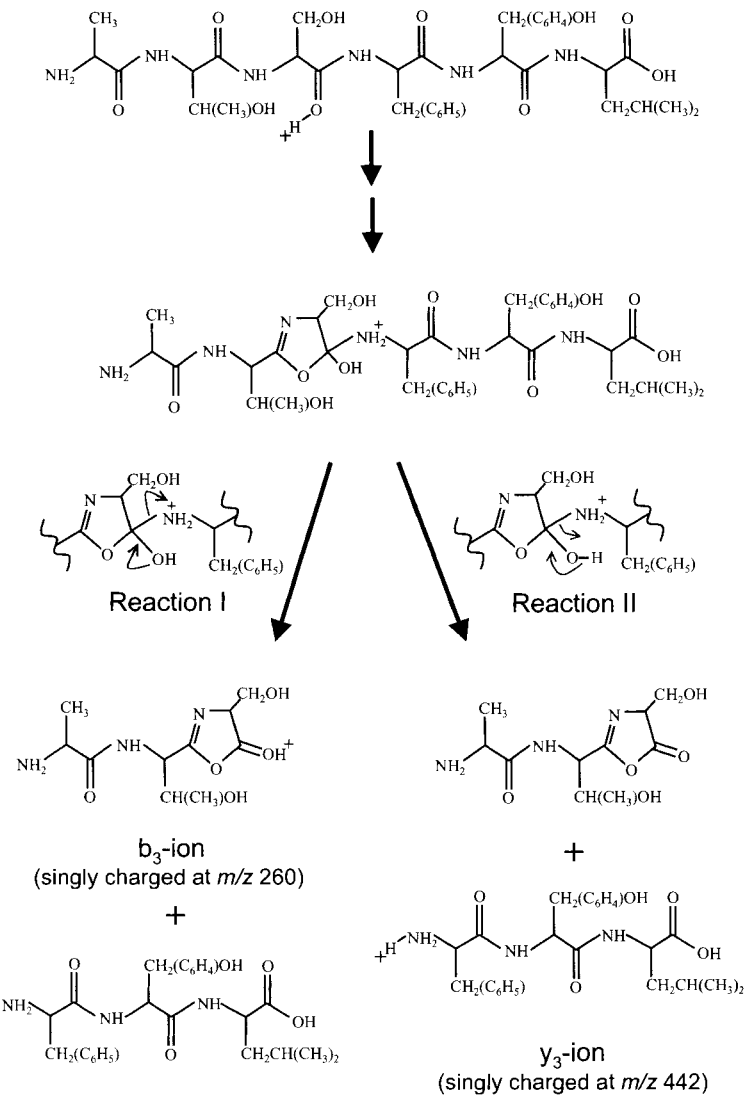

Figure 4.3. The major collisionally induced dissociation pathway occurring in a singly charged peptide ion. Following collisional activation, the site of protonation directs fragmentation reactions that occur as a means of releasing the excess internal energy added to the peptide ion by the collision. Following a cyclization reaction initiated by the site of protonation, two distinct reactions can occur in which the most significant difference is the portion of the molecule that retains the charge and is detected in the product ion spectrum. Based on observations made in the product ion spectra of singly charged ions, Reaction I is the favored pathway and b-ions are most common in these spectra. The reactions are illustrated for the protonated peptide shown in Figure 4.1.III.

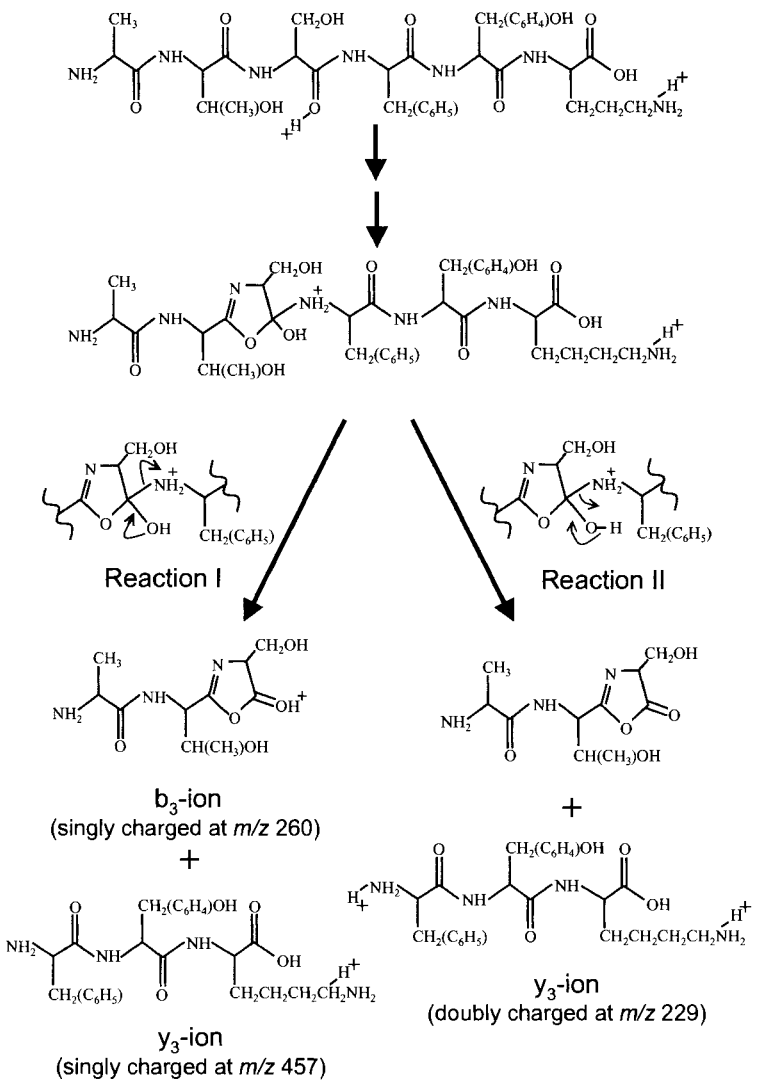

Figure 4.4. The major collisionally induced dissociation pathway occurring in a doubly charged peptide ion. The reaction for a doubly charged ion is directly analogous to that shown in Figure 4.3 for a singly charged ion. A cyclization reaction is initiated by the internal protonation site. Following cyclization, two distinct fragmentation reactions can occur and Reaction I is the favored pathway. In doubly charged ions, this reactions leads to the formation of two product ions, a complementary b-ion and y-ion pair. In the instances where Reaction II occurs, a doubly charged y-ion is formed. The reactions are illustrated for the protonated peptide shown in Figure 4.2.III.

These same fragmentation reactions are illustrated in Figure 4.4 for the tryptic peptide ATSFYK (from Figure 4.2). In the case of Reaction I, the formation of the b-ions takes place as described above. However, a second charged product is also formed in this reaction because of the additional proton, fixed at the C-terminus of the peptide ion. This second ion is a y-ion because it contains the C-terminus of the peptide. As a result, for doubly charged tryptic peptides, the favored low-energy collisionally induced dissociation reaction forms two singly charged product ions, a b-ion, and a complementary y-ion (4.10, 4.11). For doubly charged tryptic peptides, the Reaction II in the figure forms a doubly charged y-ion because both charges are associated with the C-terminal portion of the peptide. Again, this reaction appears to be less favorable as evidenced by the lack of doubly charged y-ions in the product ion spectra of doubly charged peptide ions.

Two aspects of the formation of b-ions should be noted. First, methods have been described that allow the formation of b_1-like ions (4.12–4.14). These methods function by derivatizing the N-terminus of the peptide to add a structure that allows the appropriate cyclic intermediate to form. It is interesting to note that the structures

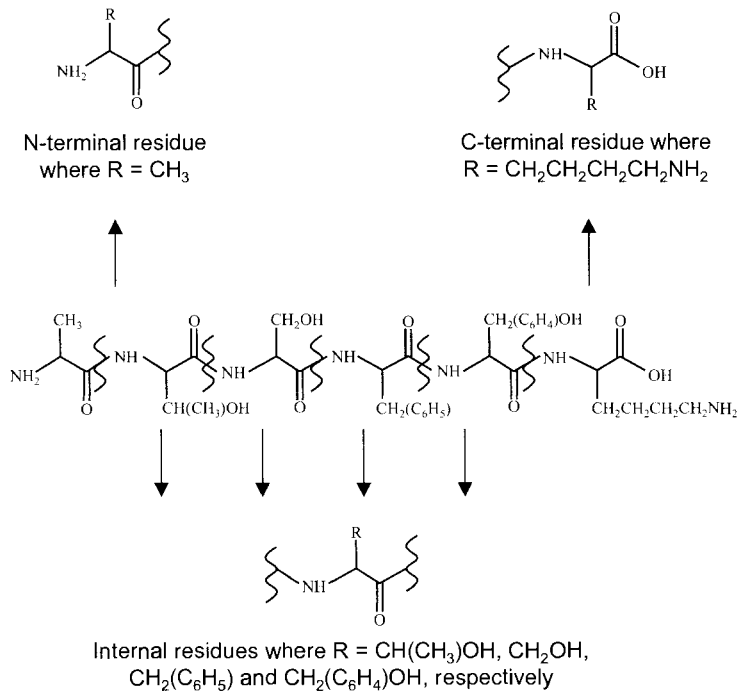

Figure 4.5. An illustration of the origin of the amino acid residue masses. The division of a hypothetical peptide at the amide bonds gives three distinct sets of structures. The most important set is the internal residue structure from which the residue masses tabulated in Table 4.1 are derived. The N-terminal residue structure differs from the internal residue structure by the addition of a -H, whereas the C-terminal residue structure differs from the internal residue structure by the addition of an -OH. These differences allow the differentiation of the N- and C-termini of a peptide.

that are added can be selected to allow mobile protonation and more extensive fragmentation along the peptide sequence (4.12, 4.13) or fix the charge at the N-terminus and give specific but limited fragmentation (4.14). Second, the larger b-ions that are formed may fragment by consecutive reactions to form smaller b-ions or a-ions (4.9, 4.15). The fragmentation to lower b-ions also proceeds through a cyclic intermediate that cannot be formed for the formation of a b_1-ion. The result of these consecutive fragmentation reactions of b-ions is that the lower b-ions can be disproportionally abundant in the product ion spectra of doubly charged peptide ions, as exemplified by the generally abundant b_2-ion. The abundance of this ion is enhanced because it represents the end of the decomposition chain of the higher b-ions. The only additional fragmentation reaction available to the b_2-ion is the loss of CO to form the a_2-ion, which is also seen in good abundance in most product ion spectra of tryptic peptides. The fact that higher a-ions are not generally observed is an indication that, where excess internal energy is contained in the b-type fragment ions, fragmentation to form the lower b-ions is favored over the formation of a corresponding a-ion.

As an illustration of the relationship between the structures of different ions in the b- and y-series, one should consider the peptide shown in Figures 4.2 and 4.4, ATSFYK. Ignoring for the moment any protonation, dividing this peptide at each of the amide bonds gives the three general types of structures shown in Figure 4.5; the N-terminal—most amino acid, the C-terminal—most amino acid, and the series of internal amino acid residues. The structure of the internal amino acids shown in this figure have the form –NH–CH(R)–CO–, where R is the side chain of the amino acid. This structure is referred to as the "residue structure" of an amino acid, and the formula weight of this moiety is defined as its residue mass. Table 4.1 lists the residue masses of the 20 amino acids and a select number of modified amino acids. These residue structures and the masses are central to the interpretation of product ion spectra because they provide the means for distinguishing the different amino acids. Remember, in mass spectrometry, mass is used to characterize structure. The structures of the N-terminal amino acid and the C-terminal amino acid are distinctly different, both from each other and from the residue structure described above. The N-terminal amino acid has the structure NH_2–CH(R)–CO–, which includes an additional –H relative to the residue structure. The C-terminal amino acid has the structure –NH–CH(R)–COOH, which includes an additional –OH relative to the residue structure. These structural differences are important because they produce characteristic mass differences relative to the residue masses in Table 4.1, which allow recognition of the N- and C-terminus of the peptide sequence. These structural differences are more precisely considered with the inclusion of the appropriate protons as shown in Figures 4.6 and 4.7.

Figure 4.6 illustrates the structure of the b-series ions from fragmentation of the doubly protonated ATSFYK peptide. Of note is the structure of the b_1 ion, which cannot be drawn with the proper cyclic oxazalone structure. However, by using the acylium structure for this ion one can see that this structure differs from the structure defined for a residue mass, NH_2–CH(R)–CO–$^+$ versus –NH–CH(R)–CO–, by an –H. Therefore, the m/z of b_1 equals the residue mass of that amino acid plus 1 Da,

Table 4.1. Residue masses of the amino acids. The residue masses of the 20 genetically encoded amino acids and selected modified amino acids.*

Amino acid	One-letter code	Residue mass (Da)	Immonium ion (m/z)
Glycine	G	57.02	30
Alanine	A	71.04	44
Serine	S	87.03	60
Proline	P	97.05	70
Valine	V	99.07	72
Threonine	T	101.05	74
Cystine	C	103.01	76
Leucine	L	113.08	86
Isoleucine	I	113.08	86
Asparagine	N	114.04	87
Aspartate	D	115.03	88
Glutamine	Q	128.06	101
Lysine	K	128.09	101
Glutamate	E	129.04	102
Methionine	M	131.04	104
Histidine	H	137.06	110
Oxidized Methionine	Mo	147.04	120
Phenylalanine	F	147.07	120
Arginine	R	156.10	129
Carbamidomethylcysteine	C*	160.03	133
Tyrosine	Y	163.06	136
Acrylocysteine	C[a]	174.04	147
Tryptophan	W	186.08	159

* This table also includes the one-letter abbreviations commonly used when writing peptide sequences and the m/z of the immonium ions with the form $NH_2{=}CHR^+$.

assuming that the fragment ions are singly charged. For the alanine shown in the figure this value is 72 Da. Subsequent b-ions then add the respective amino acid residue structures, –NH–CH(R_n)–CO–, where R_n is the side chain of the amino acid at position n relative to N-terminus, according to the amino acid sequence of the peptide. For singly charged fragment ions, the m/z of the b_n-ions show the sequential addition of the residue mass of the amino acid at that position. As a result, the b_2-ion in this example is seen at m/z 173, which is 101 Da greater than the b_1-ion at m/z 72. The 101 Da difference corresponds to the residue mass of a threonine. The series extends to m/z 260 for the b_3-ion, m/z 407 for the b_4-ion, and m/z 570 for the b_5-ion. Again, each of the mass differences corresponds to the residue mass of the amino acid at that position; $b_2 \rightarrow b_3 = 87$ Da for serine; $b_3 \rightarrow b_4 = 147$ Da for phenylalanine; and $b_4 \rightarrow b_5 = 163$ Da for tyrosine. Completing the structure of the doubly protonated peptide ion requires adding the C-terminal amino acid and the two protons. As discussed above, the structure of this moiety differs from the residue structure by addition of an –OH, such that the mass difference equals the residue mass of the C-terminal amino acid plus 17 Da plus 2 Da for the protons. In this case,

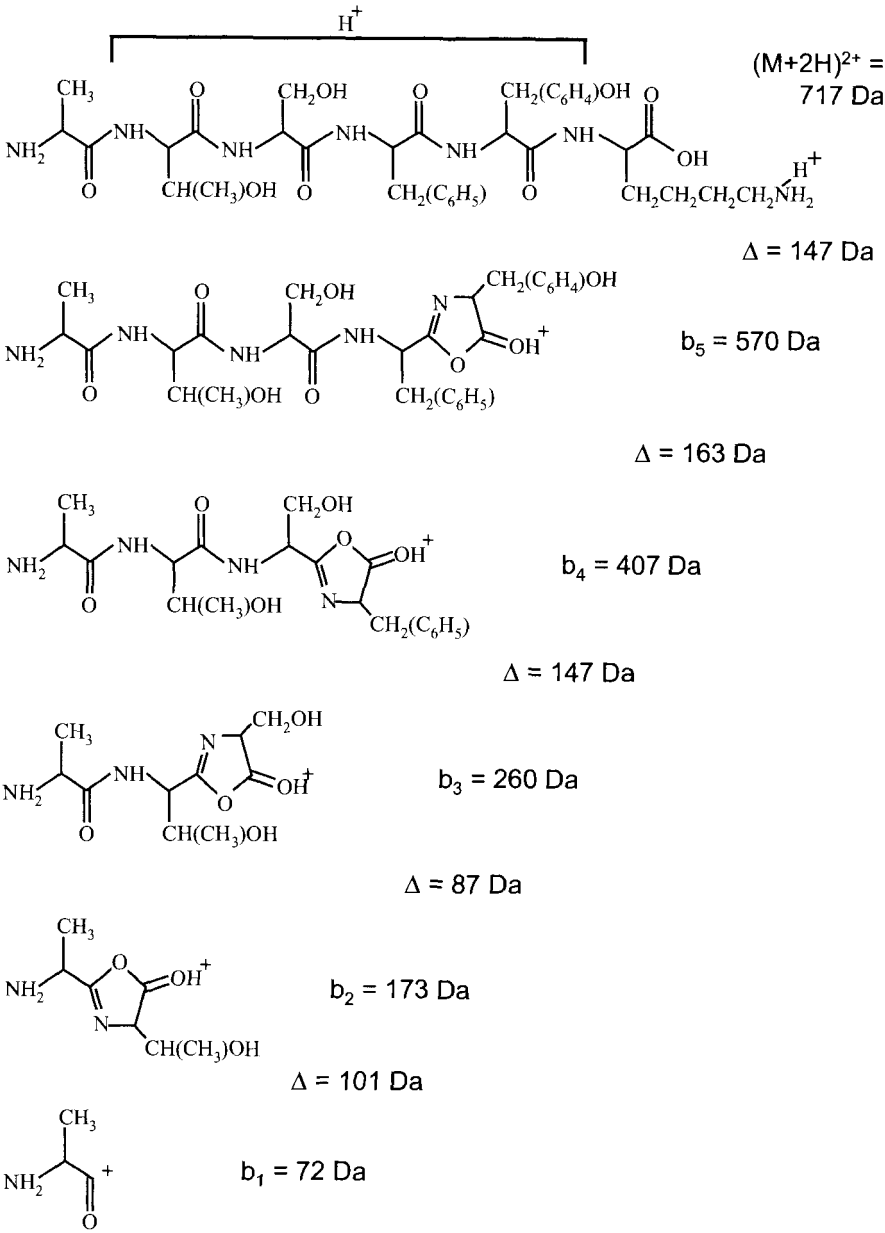

Figure 4.6. The structures of the b-ion series formed from a hypothetical doubly charged peptide. The b_2-, b_3-, b_4-, and b_5-ions are shown with the proper oxazalone structure. The b_1-ion is shown as an acylium ion for the purpose of this illustration because it cannot be drawn in the correct oxazalone structure. From this series of ions, the mass of each ion can be calculated from the mass of the previous ion and the residue mass of the amino acid at that position. Alternatively, the residue mass of the amino acid at each position can be calculated from the mass of each bracketing pair of product ions.

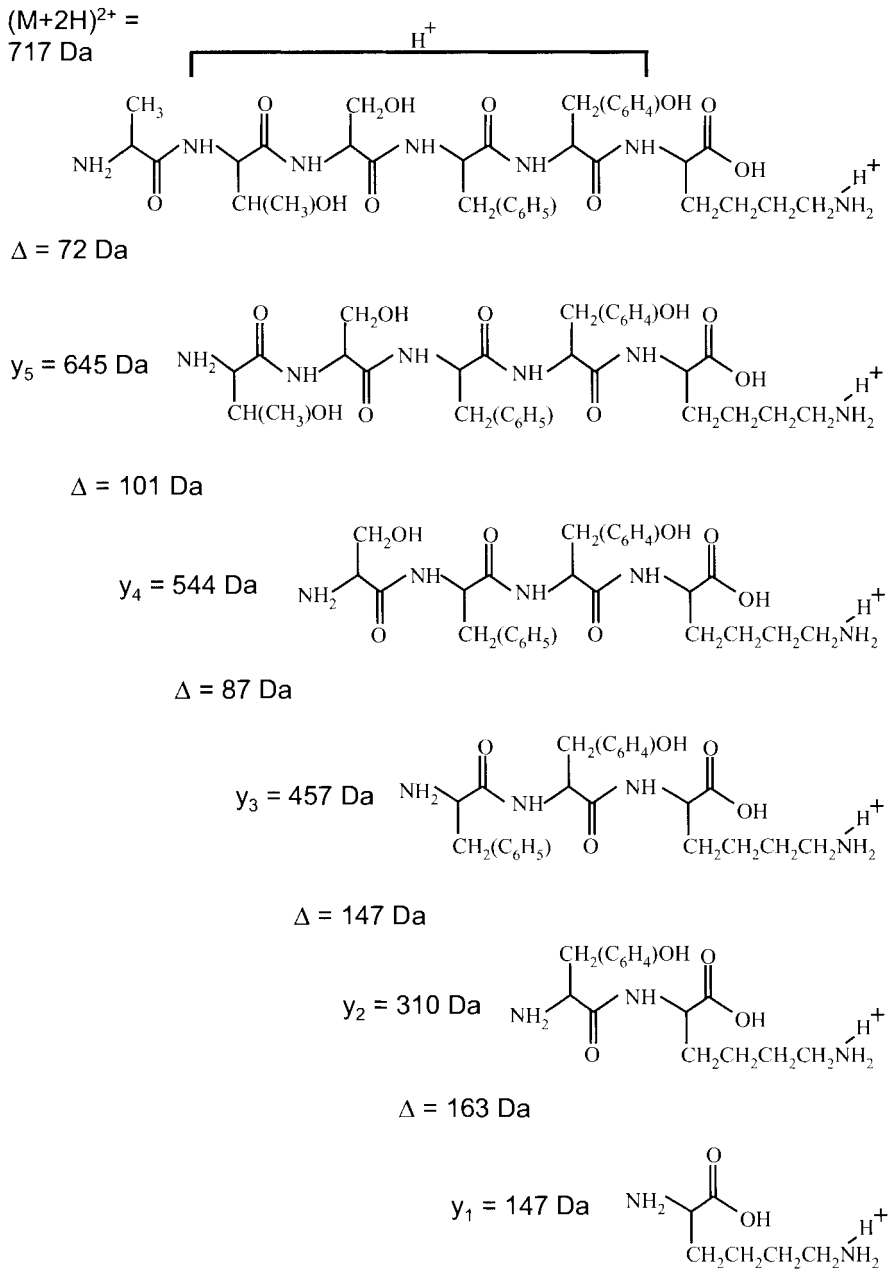

Figure 4.7. The structures of the y-ion series formed from a hypothetical doubly charged peptide. From this series of ions, the mass of each ion can be calculated from the mass of the previous ion and the residue mass of the amino acid at that position. Alternatively, the residue mass of the amino acid at each position can be calculated from the mass of each bracketing pair of product ions.

the addition is 147 Da for lysine to give the formula weight of the doubly protonated species, 717 Da.

Figure 4.7 shows the expected y-ions resulting from fragmentation of this same doubly protonated ATFSYK peptide. The y_1-ion has the structure $H_2N-CH(R_1)(H+)-COOH$ where R_1 is the side chain of the C-terminal amino acid with the proton fixed on the basic side chain of the lysine residue. As described above, the structure of this moiety is different from the residue structure of an amino acid such that the m/z of y_1 = the residue mass of the C-terminal amino acid, +1 Da(–H) + 17 Da(–OH) + 1 Da (proton), assuming that it is singly charged. As was seen with the b-ion series, subsequent y-ions then add the residue structure of the respective amino acids –NH–CH(R_n)–CO–, where R_n is the side chain of the amino acid at position n relative to C-terminus according to the amino acid sequence of the peptide. For singly charged fragment ions, the m/z of the y_n-ions shows the sequential addition of the residue mass of the amino acid at that position: $y_1 \rightarrow y_2 = 163$ Da for tyrosine; $y_2 \rightarrow y_3 = 147$ Da for phenylalanine; $y_3 \rightarrow y_4 = 87$ Da for serine; $y_4 \rightarrow y_5 = 101$ Da for threonine. Completing the structure of the protonated peptide ion requires adding the N-terminal amino acid. As discussed above, the structure of this moiety differs from the residue structure so that the added mass is equal to the residue mass of the C-terminal amino acid plus 1 Da for the extra –H. The final addition in this example is, therefore, 72 Da for alanine to give the 717 Da formula weight of the doubly protonated peptide ion.

One should keep in mind that both the b-ion series shown in Figure 4.6 and the y-ions series shown in Figure 4.7 are derived from single fragmentation reactions occurring among members of the population of ATSFYK ions protonated at the different amide bonds. As seen in these figures, the result is two distinct sets of product ions where the m/z difference between adjacent members of that set is equal to the residue mass of the amino acid at that position. Recognition of the members of these two series of product ions and calculation of the residue masses is the fundamental process of amino acid sequence interpretation. For doubly charged tryptic peptides, the b- and y-ion series are complementary in that cleavage of any of the protonated amide bonds generates both a b-ion and a y-ion. As described in Section 4.3 of this chapter, the complementary nature of the b- and y-series greatly facilitates the interpretation process.

Figure 4.8 shows an idealized product ion spectrum expected for the ATSFYK peptide. This idealized spectrum contains the complete set of b- and y-ions noted in the previous figures and, therefore, contains all of the information needed to deduce the structure of the peptide. Further, all ions in the spectrum are shown with significant abundance. In practice, however, this idealized mass spectrum is complicated by two factors, variations in the abundance of the different ions in the spectrum and the formation of other types of product ions.

In a mass spectrum, the relative abundance of an ion reflects the relative frequency of the reaction that produces that ion, including the effects of any subsequent reactions of the ion. In the case of a protonated peptide, the equal relative abundances shown in Figure 4.8 imply that proton migration to and retention

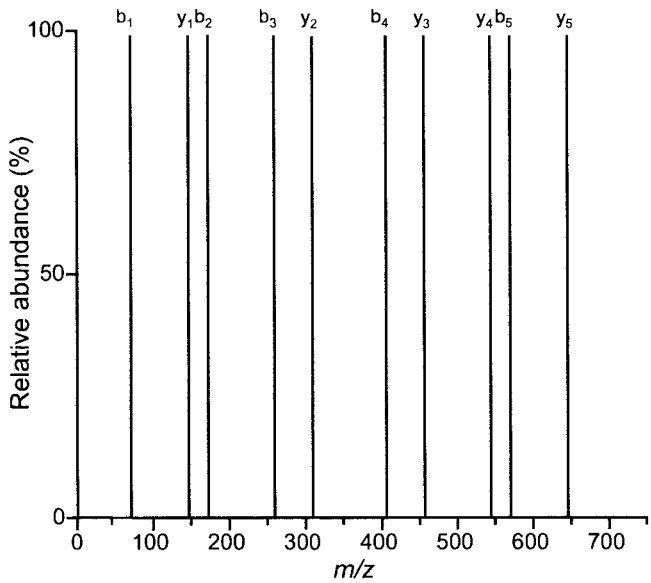

Figure 4.8. An idealized product ion spectrum. A product ion spectrum showing a complete b- and y-ion series. The spectrum is derived from the fragment ions shown in Figure 4.6 and Figure 4.7. All ions are given a relative abundance of 100 %, reflecting the assumption that all bonds will fragment with equal efficiency.

on all amide bonds is equal, thereby giving all amide bonds an equal opportunity to cleave. As will be evident when considering specific product ion spectra later in this chapter, the relative abundance of the different product ions varies widely, including the fact that some product ions are not even observed. This variation reflects the simple fact that subtle differences in the amide bonds, dependant on the nature of amino acid side chain and the position in the peptide sequence, lead to both favored and disfavored fragmentation sites. Because the population of protonated peptide ions is finite, favored fragmentation at one amide bond must be accompanied by a reduction in the fragmentation at other amide bonds.

Low-energy collisionally induced dissociation of protonated peptides also leads to other fragmentation reactions and other types of product ions, although these reactions occur with less frequency than the reactions that give the b- and y-ions. These other types of product ions include those that are formed by the loss of small neutral molecules from the b- and y-ions, and the by-ions. In the product ion spectra of tryptic peptides, lower abundance ions resulting from the loss of ammonia and water from the b- and y-ions are commonly observed. These fragmentation reactions lead to ions that may be designated as b-17 Da or y-17 Da ions, when the loss of ammonia occurs, and b-18 Da or y-18 Da ions, when the loss of water occurs. Similarly, the a-ions result from the loss of CO from b-ions and are seen as b-28 Da ions. In the product ion spectra of tryptic peptides, only the b_2-ion routinely generates a detectable amount of a-ion (the a_2-ion). However, the presence of the

b_2-ion/a_2-ion pair in the low mass end of the spectrum, recognizable by the characteristic 28 Da mass difference, is an extremely useful feature of product ion spectra of tryptic peptides. Finally, the by-ions are product ions formed by consecutive fragmentation reactions of y-ions to produce b-type ions. These ions are generally seen only in spectra where a y-ion with unusually high relative abundance is observed and are often called internal cleavage products. Overall, the formation of any of these other ion types is governed by the specific type and position of the amino acids present or possibly by interaction of multiple amino acids within the peptide chain. For completeness, it should also be mentioned that a number of re-arrangement products can also be observed. These other ions are generally, but not always, low abundance and are not useful for sequence interpretation.

A final type of ion observed in the product ion spectra of tryptic peptides is the immonium ions. Immonium ions have the structure $H_2N{=}CHR^+$, and the masses of the immonium ions (summarized in Table 4.1 with the residue masses) are such that they are observed in the low-mass end of the product ion spectrum. The immonium ions are important because their observation can be used as an indicator of the presence of that amino acid in the peptide sequence.

In summary, protonated peptides from electrospray ionization fragment on collisional activation via one major and a number of minor reaction pathways. The results of these fragmentation reactions are the product ions that are reflected in the product ion spectrum and are used to determine the sequence of the peptide. The primary ions used to deduce the sequence are the b- and y-ion series, but other ions will also contribute to the spectrum. The interpretation strategy described below is fundamentally a method to aid the recognition of the b- and y-ion series so that appropriate residue masses can be calculated to determine the amino acid at each position.

4.2.2. Fragmentation of Protonated Peptide Ions Formed by Matrix-Assisted Laser Desorption/Ionization

The fragmentation of peptide ions formed by matrix-assisted laser desorption/ionization differs significantly from the fragmentation of peptide ions formed by electrospray ionization. Product ion spectra can be obtained in matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry by either of two methods: recording the m/z of ions formed in the ion source during a delay in the extraction of ions from that source, or recording the m/z of ions produced by fragmentation reactions that occur in the field-free region between the ion source and the reflectron.

The less common method for recording the product ion spectrum of peptide ions formed by matrix-assisted laser desorption/ionization is to use the ion extraction delay on a two-stage ion source (a delayed extraction source) to allow sufficient time for the fragmentation reactions to occur (4.16, 4.17). The fragmentation reactions must occur in less than the $\sim$ 500-nsec time scale of the extraction delay so that the products can be accelerated out of the ion source for mass analysis by conventional

time-of-flight techniques. The primary drawback to this approach is that no type of precursor ion selection is possible so the analysis of mixtures of peptides, particularly complex mixtures, is not practical.

Recording the m/z of ions produced by fragmentation reactions that occur in the field-free region between the ion source and the reflectron is the more common of these two methods and yields what are referred to as post-source decay spectra (4.18, 4.19). These fragmentation reactions are the result of metastable decay reactions of ions formed in the source in a manner that produces sufficient excess internal energy to drive the fragmentation reactions. The origin of this excess internal energy is not totally understood but is most likely the result of collisional activation in the plume of neutral species generated by the laser desorption process. The m/z analysis of the product ions is carried out based on the portioning of the kinetic energy of the precursor ion in the fragmentation reactions. Because these reactions occur outside of the source, mass selection with an ion gate can be used to isolate a specific precursor ion, as described in Chapter 3. This precursor ion selection makes it possible to record the post-source decay spectrum of specific peptide ions detected in the analysis of a protein digest by matrix-assisted laser desorption/ionization. As a result, these analyses can and are applied to the types of protein sequencing experiments of interest to proteomic research.

The fundamental aspects of the fragmentation reactions that produce post-source decay spectra are not as extensively studied as the collisionally induced dissociation reactions described in the preceding section. Several observations, however, can be made concerning the nature of the amino acid sequence information in these product ion spectra. First, the peptide ions that are being fragmented are almost exclusively singly charged so that the proton producing the charge is likely located at a strongly basic site and less able or unable to direct the site of fragmentation. As a result, a significant portion of the fragmentation reactions appear to occur by so-called "charge remote pathways" (4.20, 4.21). Second, the energetics of the collision process would appear significantly different from low-energy collisional activation. The peptide ions are formed in a plume of ions and neutrals produced by the laser pulse and accelerated through this plume by the accelerating voltage of the time-of-flight mass analyzer in a manner that makes high-energy collisions likely (4.22). If the extraction is delayed, then the degree of fragmentation is diminished, presumably because the neutrals are no longer present for the activating collisions (4.23). Finally, and most important to note, the systematic and preferential fragmentation of the amide bonds that allow collisionally induced dissociation spectra to be interpreted are generally not observed in post-source decay spectra. This situation does not mean that post-source decay spectra do not contain sequence-specific information because quite the opposite is true; post-source decay spectra contain a great deal of sequence specific information. That information, however, is in the form of a number of immonium ions, internal fragment ions, and neutral loss ions that are observed in addition to a generally incomplete and less-abundant set of b- and y-ions. This information can be matched to database sequences by the search programs described in Chapter 8 (4.24) but generally cannot be interpreted by the strategy described in the following section of this chapter.

4.3. INTERPRETATION OF THE PRODUCT ION SPECTRA OF TRYPTIC PEPTIDES

As described above, product ion spectra are interpreted to determine the amino acid sequence of a peptide by recognizing the b- and y-ion series and calculating the residue masses of amino acids from adjacent members of those ion series. This interpretation is an iterative process that produces a mathematical solution of the product ion spectrum being considered. A proper solution explains all of the ions in the spectrum, including low abundance ions, and is consistent with other observations of the peptide and the experiment such as the use of the enzyme trypsin for the proteolytic digestion, the charge state of the peptide and product ions, and the observation of high relative abundance ions in the product ion spectrum. The interpretation process is aided by information presented in Tables 4.1, 4.2, 4.3, and 4.4, making it important to understand the information presented in those tables.

4.3.1. Tabulated Values Used in the Interpretation

Table 4.1 lists the 20 common amino acids, the one-letter abbreviations, and the monoisotopic residue mass given to two decimal places. These residue mass values are used when interpreting product ion spectra, although it is easiest to use the simple nominal mass when working through a spectrum interpretation. The difference between nominal and monoisotopic mass becomes significant, as a series of values are totaled and will produce what is essentially a rounding error. In all of the spectra presented in this chapter, the m/z measured for the product ions are the monoisotopic values. As a result, this rounding error can produce calculated m/z values that differ from those observed in the spectrum, but this difference does not affect the interpretation process. One should always base the specific amino acid assignments on the residue masses calculated from the measured m/z of the ions in the spectrum and use the calculated nominal masses as an aide to track the interpretation process.

In addition to the 20 common amino acids, several modified amino acids are also included in Table 4.1. Oxidized methionine is included because it is commonly observed. Also, carboxyamidomethyl- and acrylo-modifications of cysteine are included because these are the products of the alkylation reactions used to block the reformation of disulfides during protein digestion. In the digestion methods described in Chapter 6, iodoacetamide is used for this reaction and would produce carboxyamidomethylcysteine. However, acrylocysteine may still be observed due to adventitious alkylation during electrophoresis by acrylamide contained in the polyacrylamide gels.

One may note that, in general, the residue masses of the amino acids vary considerably, from 57 Da for glycine to 186 Da for typtophan. There is, however, some overlap among the residue masses of several amino acids. First, leucine and isoleucine have identical residue masses and under low-energy collisionally induced dissociation conditions these two amino acids cannot be distinguished. For the

Table 4.2. Look-up table for the m/z of the b_2-ion.*

	G	A	S	P	V	T	C	L/I	N	D	Q/K	E	M	H	Mo/F	R	*C**	Y	C^a	W
G	115																			
A	129	143																		
S	145	159	175																	
P	155	169	185	195																
V	157	171	187	197	199															
T	159	173	189	199	201	203														
C	161	175	191	201	203	205	207													
L/I	171	185	201	211	213	215	217	227												
N	172	186	202	212	214	216	218	228	229											
D	173	187	203	213	215	217	219	229	230	231										
Q/K	186	200	216	226	228	230	232	242	243	244	257									
E	187	201	217	227	229	231	233	243	244	245	258	259								
M	189	203	219	229	231	233	235	245	246	247	260	261	263							
H	195	209	225	235	237	239	241	251	252	253	266	267	269	275						
Mo/F	205	219	235	245	247	249	251	261	262	263	276	277	279	285	295					
R	214	228	244	254	256	258	260	270	271	272	285	286	288	294	304	313				
*C**	218	232	248	258	260	262	264	274	275	276	289	290	292	298	308	317	321			
Y	221	235	251	261	263	265	267	277	278	279	292	293	295	301	311	320	324	327		
C^a	232	246	262	272	274	276	278	288	289	290	303	304	306	312	322	331	335	338	349	
W	244	258	274	284	286	288	290	300	301	302	315	316	318	324	334	343	347	350	361	373

* The *m/z* of all possible b_2-ions for combinations of the amino acids residue masses shown in Table 4.1. One-letter codes are used to designate the different amino acids with the addition of Mo to designate oxidized methionine, C* to designate carbamidomethylcysteine, and C^a to designate acrylocysteine. Single entries are made for the isobaric amino acid pairs L and I, Q and K, and Mo and F.

Table 4.3. Amino acids combinations that are equal to a single amino acid residue mass.*

Amino acid combination	Residue mass (Da)	Equivalent amino acid
GG	114	N
GA	128	Q, K
GV	156	R
GE	186	W
AD	186	W
SV	186	W
SS	174	C[a]

* The single-letter amino acid codes shown in Table 4.1 are used in this table.

Table 4.4. Neutral losses observed from ions with different amino acid compositions.*

Amino acid	Neutral loss
A	–
G	–
S	18
P	–
V	–
T	18
C	34
L/I	–
N	17
D	18
Q	17
K	17
E	18
M	48
H	–
Mo	64
F	–
R	17
C*	92
Y	–
C[a]	106
W	–

The consecutive loss of small neutral molecules is an energetically favored process in collisionally induced dissociation. The nature of the neutral that is lost is dependent on the amino acid composition of the product ion. In this table, the one-letter amino acid codes given in Table 4.1 are used. The – designates that no neutral losses occur for that amino acid.

purposes of spectrum interpretation, these two amino acids are treated as one and given the one-letter abbreviation X to denote that either amino acid is possible at that position. One should reiterate that this abbreviation does not mean "unknown", but rather it designates one of two specific possibilities, leucine or isoleucine. Second, glutamine and lysine both have nominal residue masses of 128 Da, and phenylalanine and oxidized methionine both have nominal residue masses of 147 Da. The residue masses of these amino acids cannot be distinguished except when exceptionally high mass resolution is used. Unlike leucine and isoleucine, however, these amino acid pairs can often be differentiated and assigned by using other mass spectrometric data. For example, glutamine and lysine can be distinguished by an acetylation reaction that specifically modifies lysine residues. It is also common to assume that, in tryptic peptides, lysine should be found only at the C-terminus; and if an internal lysine were present, then the peptide ion would add another proton and exist as a triply charged ion. The difference between phenylalanine and oxidized methionine can be less clear. If only a single phenylalanine or oxidized methionine residue mass is observed, one may be able to distinguish between the two based on the presence or absence of ions resulting from the loss of the small neutral molecule $HSOCH_3$ (64 Da) from oxidized methionine or from the generally detectable immonium ion for phenylalanine at m/z 120. If, however, more than one phenylalanine or oxidized methionine residue mass is observed, then a specific assignment may be based on very subtle observations of which ions are losing $HSOCH_3$ such that some uncertainty may remain. Any uncertainty about this assignment should not deter the continuing interpretation of the spectrum.

The other residue masses of the amino acids all differ by at least 1 Da and should be distinguishable when unit mass resolution or better is used. One must remember, however, that the residue masses are calculated by subtraction of two experimentally measured m/z values and the accuracy of those values directly affects the accuracy of the calculated residue mass. For example, if the mass resolution of the instrument system being used produces values that vary by greater than ± 0.5 Da, then the uncertainty of the calculated residue mass will be greater than ± 1 Da due to the propagation of errors. Under these conditions, it can be difficult to distinguish between the amino acids with residue weights that differ by only 1 Da, such as leucine and isoleucine at 113 Da from asparagine at 114 Da, asparagine at 114 Da from aspartate at 115 Da, or glutamine at 128 Da from glutamate at 129 Da. This situation is commonly encountered with tandem quadrupole instruments in which the second mass analyzer is intentionally detuned for poorer resolution but better sensitivity. This uncertainty should not a problem with either ion trap or quadrupole-time-of-flight mass spectrometer systems. It is possible, however, in any instrument system that centroiding artifacts in the data system introduce similar uncertainties that affect the m/z assignment for a given ion. These artifacts are usually the result of what is essentially chemical noise from unresolved ions in the product ion spectrum and have the effect of producing problems in the m/z assignment that increase the uncertainty in the calculated residue mass. In most instances, however, careful inspection of both the b- and y-ion series along with careful consideration of the measured molecular weight of the peptide will assist in resolving these problem assignments.

Table 4.1 also lists the masses of the immonium ions ($NH_2=CHR^+$). Although these ions are informative about amino acid content, they provide no information about the position of that amino acid in the sequence. Also, lack of an immonium ion is generally not conclusive. For example, the presence of m/z 110 indicates a high probability of a histidine somewhere in the peptide chain but the absence of m/z 110 does not preclude the presence of histidine in a peptide.

Table 4.2 lists all possible b_2-ions and their m/z assuming a single charge. This information is useful because although the b_1 ion is rarely seen in a product ion spectrum, the b_2 ion is almost always present provided an appropriate m/z range can be used. Further, because of the facile loss of CO, the b_2-ion is generally observed paired with the a_2-ion and is recognizable by the 28 Da difference between the two ions. Whereas observation of the b_2-ion and consultation of Table 4.2 provide a limited set of possible combinations for the two N-terminal amino acids, precise assignment of identity and order of these two amino acids depends on observing the y_{n-1}-ion. This ion is generally not of high relative abundance and in some cases may not be present. The amino acid pairs shown in this table range from the smallest combination GG, at 115 Da, to the largest combination WW, at 373 Da. For example, if a b_2-ion is seen at m/z 145 Da then the only possible combination of amino acids is SG. The product ion spectrum would then be inspected to find a y-ion corresponding to the loss of 87 Da, to assign an S, or the loss of 57 Da, to assign a G, from the protonated peptide. Finding either would confirm the order of these two amino acids, but at the worst the possibilities would be narrowed to two—SG– or GS–. Other b_2-ions produce more amino acid combinations. For example, a b_2-ion at m/z 228 could be LN, IN, VQ, VK, or RA, which limits the total number of possible combinations to ten. This table can also be used to generate lists of two-amino acid-combinations anywhere there is a suspected two amino acid gap in an amino acid sequence. Remember, however, that the values in this table are reported as b_2-ion combinations and the corresponding internal two-amino acid-combinations would have values 1 Da lower.

Table 4.3 highlights one other aspect to remember when considering combinations of two amino acids. In a limited number of cases, the combined residue mass of two amino acids may equal the residue mass of a single amino acid. For example, the combined residue mass of GE, AD, and SV all equal the 186 Da residue mass of a tryptophan residue. When the residue masses listed in this table are encountered, one should consider any unexplained ions observed between two fragment ions producing the assignment as a possible indication of two amino acids in this position.

Table 4.4 summarizes a final set of information that may be useful when interpreting product ion spectra, mass differences produced by the loss of a small neutral molecule. These losses are observed for both the b- and y-ion series if that ion contains a particular amino acid. As a result, observation of these losses can be used as a general indicator of amino acid content of the peptide or product ion. As described above, the most common loses are water (18 Da) and ammonia (17 Da). Loss of water typically occurs from the alcohol- (serine and threonine) or carboxylic acid- (aspartate and glutamate) containing amino acids but may also be lost from the C-terminus of y-ions. Ammonia losses are generally due to the presence of the

amide- (asparagine and glutamine) or amine-containing amino acids (lysine and arginine) but may also occur from the N-terminus. Observing these types of ions are most useful when following a given ion series because the losses will be seen only as long as the amino acid producing that neutral loss is retained in the product ions. For example, b-ions of serine or threonine-containing peptides will generally be accompanied by ions 18 Da (assuming singly charged) less than the particular b-ion for as long as the b-series contains the S or T. Seeing a change in the pattern of neutral loss within a series often indicates that one of the above amino acids is the next amino acid in that series. Another common neutral loss is 64 Da (assuming singly charged) from oxidized methionine-containing peptides or product ions. The utility of the 64 Da loss is the same as that described for water and ammonia losses, namely, that it shows an ion that contains oxidized methionine.

4.3.2. A Strategy for the Interpretation of Product Ion Spectra of Tryptic Peptides

The strategy used by the authors to interpret the product ion spectra of tryptic peptides has nine components. This strategy is based on methods developed and taught by others (4.25) and is effective for the complete interpretation of product ion spectra of tryptic peptides obtained in electrospray ionization experiments, as evidenced by published reports of the sequencing of novel proteins in the authors' laboratories (4.26, 4.27). Tryptic peptides ionized by electrospray ionization are particularly amenable to tandem mass spectrometric sequencing because of the highly informative fragmentation directed by the mobile proton placed on the N-terminal amine, and the complementary b- and y-ions series produced by the fixed charge on the side chains of the C-terminal lysine or arginine residue. This approach is also fundamentally useful for the interpretation of post- source decay spectra obtained in matrix-assisted laser desorption ionization-time-of-flight mass spectrometry experiments. The success of the interpretation of post-source decay spectra, however, is highly variable because of the more complex fragmentation observed for these singly charged ions, as described in Section 4.2.2 of this chapter.

One can begin by considering an exemplary product ion spectrum derived from a tryptic peptide and presented in Figure 4.9. This spectrum can be viewed as having three different regions, a low-mass region, a high-mass region, and a middle region, as illustrated in the figure. The low-mass region of a product ion spectrum contains the immonium ions, recognizable with use of Table 4.1; the b_2-ion as tabulated in Table 4.2 and generally recognizable because of the b_2-ion/a_2-ion pair differing by 28 Da; and the y_1-ion, recognizable in the product ion spectra of tryptic peptides at either m/z 147 (for lysine-containing peptides) or m/z 175 (for arginine-containing peptides). The high-mass region of the product ion spectrum is generally dominated by y-ions although b-ions will also be present. Because the numbering system for each ion series begins at the peptide terminus that is retained in the product ion, these y-ions will be referred to as the high-end y-ions and their position noted as y_{n-1}, y_{n-2}, or y_{n-3}, where n = the length of the peptide, for the purpose of this discussion. The middle region of the product ion spectrum is an area where the b-ion

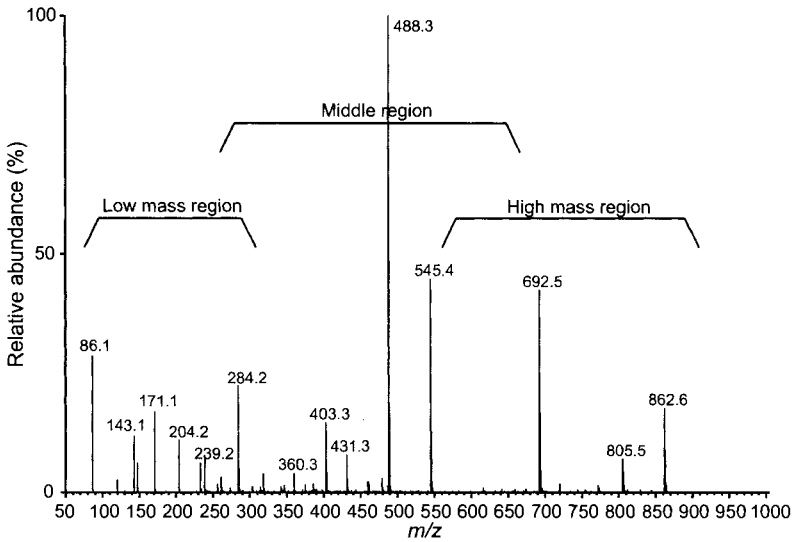

Figure 4.9. An exemplary product ion spectrum. The product ion spectrum of a doubly charged peptide ion acquired by collisionally induced dissociation with a quadrupole-time-of-flight instrument. The m/z of the precursor ion was 488.3 Th. An in-gel tryptic digestion of a Coomassie blue-stained protein band 2D electrophoresis gel produced the peptide. The different regions of the spectrum are defined to allow a generalization of the type of information that is found in each part of the spectrum. The sequence of the peptide is IGLFNAGVGK and the interpretation of the spectrum is given in Figure 4.10.

series and the y-ion series begin to overlap, and care must be taken when identifying ions in each series. In general, the abundance of the higher y-ions will fade toward lower abundance in this region, and some y-ions may not be observed. This region may also contain the precursor ion and, in some cases, doubly charged ions that are either the high-end y-ions or ions formed by losses of small neutral molecules from the precursor ion.

The b-ions and y-ions for the interpretation of the product ion spectrum in Figure 4.9 are shown in Figure 4.10. One may note that not all ions were observed, but in this case a complete interpretation was still possible. More important to note, however, is the simple mathematical relationship between the m/z of the b-ions and the m/z of the y-ions. Specifically, assuming that all product ions are singly charged, the m/z of the b_x-ion plus the m/z of the corresponding y_{n-x}-ion (where $n =$ the length of the peptide and $x =$ the position in the respective ion series) equals the protonated molecular weight of the peptide plus 1 Da. This relationship is due to the generation of these two ions by breaking a specific amide bond in the doubly protonated peptide ion. The critical practical effect of this relationship is that as an ion in one ion series is recognized, its m/z is used to calculate the m/z of the corresponding ion in the other ion series and identify that ion in the spectrum. (It is especially important to remember the rounding error produced by the use of nominal versus exact residue masses when considering these calculations.) With this system,

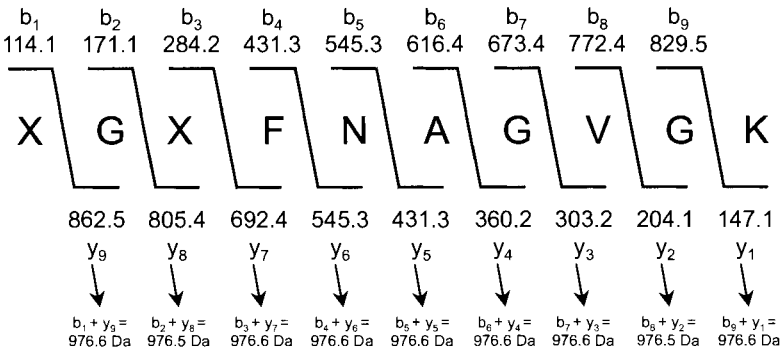

Figure 4.10. A summary of the interpretation of the product ion spectrum shown in Figure 4.9. This interpretation is given in the general format that is used throughout this chapter. Specifically, the calculated m/z of the b-ions are shown above the amino acid sequence and the calculated m/z of the y-ions are shown below the amino acid sequence. All values refer to the monoisotopic m/z of the product ion in its singly protonated form. In this interpretation the calculated m/z are shown to one decimal place, but in many of the other interpretations the calculated m/z will be rounded to the nearest nominal mass. Also noted in the figure is the fact that, because of the complementary nature of the b- and y-ion series, as the m/z of any fragment ion is assigned, the m/z of the complementary fragment ion can be calculated.

each amino acid assignment is effectively checked by observation of the two complementary ions. This system also gives the feeling of classifying a given ion in a spectrum as either a b- or y-ion. Although this type of classification is generally useful for clarifying which ions need to be explained by the on-going interpretation, it is possible that a given m/z is composed of ions from both ion series. One may note that this condition is the case for peptide sequence given in Figure 4.10, where the m/z 545 and 431 ions are a part of both the b-ion and y-ion series.

The interpretation of a product ion spectrum begins with the assumption that a mass spectrum has also been recorded and that the monoisotopic molecular weight of the peptide in its singly protonated form, $(M + H)^+$, is known. It is also useful to determine the most abundant charge state of the peptide from the mass spectrum. The charge state is particularly useful when considering tryptic peptides because peptides where the most abundant charge state is triply charged or greater usually contain internal histidine, lysine, or arginine residues.

The Nine-Step Strategy for Interpretation

1. *Inspect the low-mass region for immonium ions.* The first step in the interpretation is to inspect the low-mass region of the spectrum, noting the presence of any immonium ions and the amino acid composition that they indicate.
2. *Inspect the low-mass region for the* b_2*-ion.* In the second step of the interpretation, the low-mass region of the spectrum is inspected to identify the b_2-ion, generally recognizable by the b_2-ion/a_2-ion pair separated by

28 Da. By using Table 4.2, the possible two-amino acid-combinations indicated by the b_2-ion are noted. The m/z of the b_2-ion is then used to calculate the m/z of the corresponding y_{n-2}-ion, and the high-mass region of the product ion spectrum is inspected to identify this ion.

3. *Inspect the low-mass region for the* y_1*-ion.* The third step of the interpretation is to assign the C-terminal amino acid. The low-mass region of the spectrum is inspected to identify the y_1-ion at either m/z 147, for C-terminal lysine peptides, or m/z 175, for C-terminal arginine peptides. The m/z of the y_1-ion is then used to calculate the m/z of the b_{n-1}-ion, and the high-mass region of the product ion spectrum is inspected to identify that ion, if present.
4. *Inspect the high-mass region to identify the* y_{n-1}*-ion.* The fourth step of the interpretation is to attempt to assign the N-terminal amino acids from combinations indicated by the b_2-ion. The high-mass region of the spectrum is scrutinized to identify the y_{n-1}-ion, if present. The list of possible amino acid combinations derived from the b_2-ion limits the possible residue masses to consider. If an ion is identified, the m/z of that ion is used to calculate the residue masses of the first two amino acids and to assign those peptides.
5. *Extend the y-ion series toward lower m/z.* Working with the residue masses listed in Table 4.1, begin to extend the y-ion series backwards (toward lower m/z) from the y_{n-2}-ion. As a y-ion is identified calculate the m/z of the corresponding b-ion and identify that ion in the spectrum. Work towards extending the y-ion series from the y_{n-2} ion to the y_1-ion
6. *Extend the b-ion series toward higher m/z.* If progress extending the y-ion series falters, use the residue masses listed in Table 4.1 to extend the b-ion series from the last identified b-ion. As any b-ions are identified, use the m/z of that ion to calculate the m/z of the corresponding y-ion and identify that ion in the spectrum.
7. *Calculate the mass of the peptide.* When the interpretation of the spectrum is complete, calculate the mass of the proposed peptide sequence and check its agreement with the measured mass.
8. *Reconcile the amino acid content with spectrum data.* Check that the amino acid content agrees with the immonium ions observed. Also consider the charge state of the peptide in terms of the presence of histidine, and internal lysine or arginine residues.
9. *Attempt to identify all ions in the spectrum.* Work to identify the other ions in the spectrum based on the proposed peptide sequence and pay particular attention to the ions from the loss of H_2O, NH_3, and $HSOCH_3$; any doubly charged ions; and any ions due to internal cleavages.

4.3.3. Sample Interpretation Problem Number One

Figure 4.11 contains the product ion spectrum of a tryptic peptide with a monoisotopic molecular mass of 778.6 Da for the protonated peptide, $M + H^+$, obtained by fragmenting a doubly charged ion at m/z 389.8 with no triply charged ion

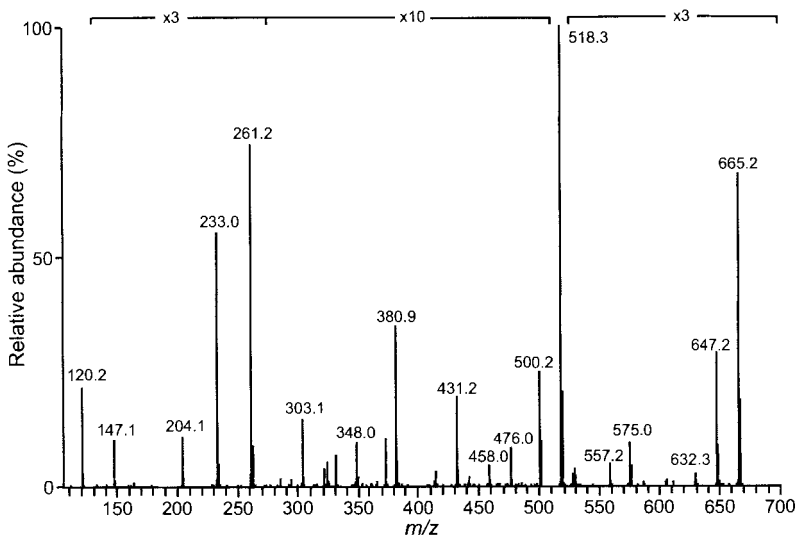

Figure 4.11. Interpretation problem number one. The product ion spectrum was acquired by using an ion trap mass spectrometer by collisionally induced dissociation of a doubly charged ion, m/z 389.9. The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 2D electrophoresis gel. All of the ions in the spectrum are normalized to the most abundant ion in the spectrum. However, selected magnification has been used over various portions of the spectrum to enhance the clarity of those regions of the spectrum. No additional ions were observed outside of the displayed m/z range.

observed in the mass spectrum. The lack of a triply charged ion makes it unlikely that there is a histidine or internal lysine or arginine contained in the peptide sequence. The spectrum was recorded using an ion trap instrument operated with unit resolution so that the product ions' m/z are monoisotopic masses. This product ion spectrum will be used to illustrate the principles of the interpretation strategy. The steps are numbered to correspond to those listed above.

1. *Inspect the low-mass region for immonium ions.* Inspection of the low-mass region of the spectrum reveals a prominent ion at m/z 120, indicating the presence of a phenylalanine in the sequence. No other immonium ions are present, in part because of the low-mass cut-off of the ion trap mass analyzer.
2. *Inspect the low-mass region for the* b_2*-ion.* Inspection of the low-mass region also reveals the b_2-ion at m/z 261. This identification is aided by the observation of the a_2-ion at m/z 233. As seen in Table 4.2, the amino acid combinations YP-, VC-, FX-, and EM- are consistent with this b_2-ion. Because observation of an ion in one series allows calculation of the corresponding ion in the other series, one can calculate the y_{n-2}-ion as m/z 518 and identify this ion in the spectrum.
3. *Inspect the low-mass region for the* y_1*-ion.* Further inspection of the low-mass region reveals an ion at m/z 147, the y_1-ion for lysine-containing peptides.

This assignment is strengthened by the fact that no m/z 175, the y_1-ion for arginine, is present. Identifying the y_1-ion allows calculation and identification of the b_{n-1}-ion at m/z 632. The interpretation, up to this point, is summarized in Figure 4.12. As may be noted in the figure, no real sequence data are present in the interpretation at this point. It is, however, an important stage in the process because the interpretation has now oriented the N- and C-terminal ends of the peptide, establishing the beginning of each ion series.

4. *Inspect the high-mass region to identify the* y_{n-1}*-ion.* With the b_{n-1}-ion identified, one can then inspect the high-mass region for the possible observation of the y_{n-1}-ion that would be at an m/z greater than the y_{n-2}-ion. Two prominent possibilities are seen at m/z 647 and m/z 665. If the ion at m/z 647 were the y_{n-1}, then the calculated residue mass would be 129 Da (relative to the m/z 518 ion) and would assign a glutamate consistent with our possible amino acid pairs determined from the b_2-ion. There would be, however, no explanation for the ion at m/z 665. If the ion at m/z 665 were assigned as the y_{n-1}-ion, then the calculated residue mass would be 147 Da (relative to the m/z 518 ion) and a phenylalanine would be assigned. This assignment is also consistent with the possible amino acid pairs determined from the b_2-ion and the immonium ion. Further, if m/z 665 is the y_{n-1}-ion then the m/z 647 ion, which is 18 Da less, could be the result of a water loss. This possible explanation is supported by the similar loss of water seen for the m/z 518 ion, producing the ion at m/z 500. Finally, the assignment of the phenylalanine allows one to assign a leucine or isoleucine as the N-terminal amino acid. Figure 4.13 summarizes the interpretation to this point.
5. *Extend the y-ion series toward lower m/z.* The next step would be to use the residue masses shown in Table 4.1 to extend the high y-ion series. It is reasonable to test the more abundant ions first. On this assumption, the m/z 431 ion or the m/z 381 ion would be used to attempt to extend the y-ion series from the m/z 518 ion and indicate residue masses of 87 Da, assigned a serine, and 137 Da, assigned a histidine, respectively. The serine assignment is preferred because an internal histidine would lead to the triply charged peptide. Also, identification of m/z 431 Da as the y_{n-3}-ion allows calculation of m/z of the b_3-ion as 348 Da, which is observed in the spectrum. If histidine

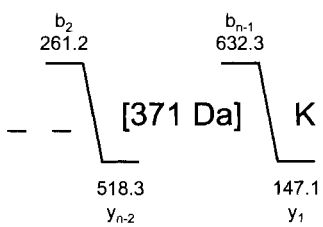

Figure 4.12. A summary of the on-going interpretation of the production ion spectrum in Figure 4.11. The value in brackets denotes the remaining mass difference between the assigned ions. The – designates an unknown amino acid.

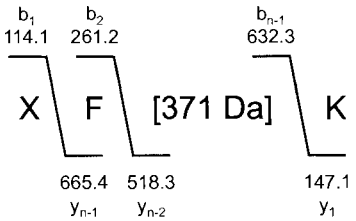

Figure 4.13. A summary of the on-going interpretation of the production ion spectrum in Figure 4.11. The value in brackets denotes the remaining mass difference between the assigned ions.

were assigned, the calculated m/z of the b_3-ion would be 398 Da, which is not observed in the spectrum. In this case, one might note that the ion at m/z 381 is 9 Da less than the precursor ion at m/z 389.9. As a result, the m/z 381 ion appears to be a doubly charged ion resulting from the loss of water and/or ammonia from the precursor ion. One might also note that the ions at m/z 476 and 458 cannot be y-ions because they would represent non-existent residue masses of 42 Da and 60 Da, respectively. The fact that these ions cannot be y-ions makes it likely that they will be b-ions to be explained when that ion series is extended.

With the ions at m/z 381 Da and m/z 348 Da already assigned, further extension of the y-series should focus on the m/z 372 and m/z 303 ions. If m/z 372 were the y_{n-4}-ion then the residue weight would be 59 Da, which is not an amino acid. Alternatively, if m/z 303 were the y_{n-4}-ion then the residue mass would be 128 Da, a glutamine, and the corresponding b_4-ion would be calculated as m/z 476, which is observed in the spectrum. This amino acid would not be assigned as a lysine because an internal lysine would produce a triply charged ion for this peptide. The interpretation of the spectrum to this point is summarized in Figure 4.14.

Further extension of the y-ion series now has few choices. Assigning the m/z 204 ion as the next y-ion gives a residue mass of 99 Da, a valine, with a calculated b_5-ion at m/z 575, which is observed in the spectrum. In addition, this assignment leaves a final residue mass of 57 Da, a glycine, to extend the y-ion series to the y_1-ion at m/z 147 Da. As a result, this assignment completes

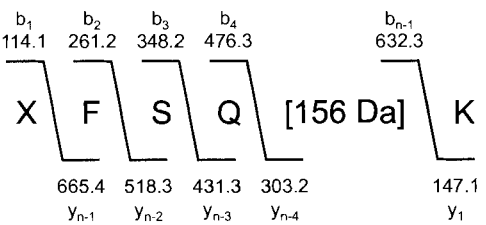

Figure 4.14. A summary of the on-going interpretation of the production ion spectrum in Figure 4.11. The value in brackets denotes the remaining mass difference between the assigned ions.

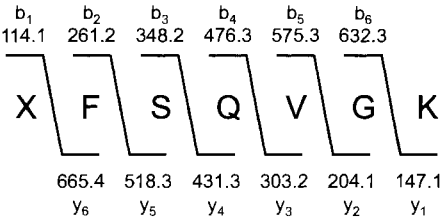

Figure 4.15. A summary of the complete interpretation of the product ion spectrum shown in Figure 4.11. Note that because the length of the peptide is now known, the b- and y-ion notation has changed to reflect the position in the peptide sequence.

the amino acid sequence of this peptide and the interpretation is summarized in Figure 4.15.

6. *Extend the b-ion series toward higher m/z.* In this example, work with the y-ions series has produced a complete interpretation with a suitable b-ion series. The logic used in working with the b-ions series, however, would be the same as that used in step 5 with the y-ion series.
7. *Calculate the mass of the peptide.* The calculated monoisotopic mass for this sequence is 778.4 Da, $(M + H)^+$, which compares well with the measured mass.
8. *Reconcile the amino acid content with spectrum data.* The proposed sequence contains a phenylalanine and is, therefore, consistent with the observed immonium ion. Such consistency is especially important in this case because the 147 Da residue mass could also be an oxidized methionine. Further, no ions for a neutral loss of 64 Da, indicative of the presence of oxidized methionine, are observed.
9. *Attempt to identify all ions in the spectrum.* The final step is to inspect the spectrum for prominent, unexplained ions. The ions at m/z 500 and m/z 647 have been explained as water losses, an explanation that is consistent with the serine residue in the peptide. Further, the losses of water are no longer observed at the point in the sequence that the serine is no longer contained in the y-ions. As a result, the proposed amino acid sequence, XFSQVGK, is completely consistent with the product ion spectrum with the only ambiguity being whether the X is a leucine or an isoleucine.

One must remember that, while this amino acid sequence is an appropriate solution for the product ion spectrum, the interpretation of a spectrum is not an absolute result. The interpretation, however, could be tested by methods described in Chapter 9 such as analysis of a synthetic peptide giving the same product ion spectrum (described in Section 9.3.2). Alternatively, a protein sequence in the databases matching this sequence would also support this interpretation, particularly if the product ion spectra of other peptides in the digest match other portions of that sequence. These types of matches are the basis of protein identification by database searching as described in Chapter 8.

Figure 4.16 shows one last noteworthy feature of this product ion spectrum, the prominence of the ion at m/z 518.2. In the presentation of this spectrum shown in Figure 4.11 portions of the spectrum were magnified for ease of interpretation. In Figure 4.16 these ions are all normalized to the most abundant fragment ion, the base peak. This pattern of abundance is often seen in product ion spectra of tryptic peptides and corresponds to the y-ion for the facile loss of the two N-terminal-most amino acids.

4.3.4. Sample Interpretation Problem Number Two

Figure 4.17 shows the next product ion spectrum to be interpreted. The peptide studied is a tryptic peptide in which the most abundant charge state in the electrospray mass spectrum is the doubly charged ion and the measured protonated molecular weight is 1295.0 Da. The experiment was carried out using an ion trap mass spectrometer system with unit m/z resolution. The interpretation will follow the nine-step strategy described above.

1. *Inspect the low-mass region for immonium ions.* In this spectrum, no immonium ions were recorded because the low-mass cut-off of the ion trap system used for the analysis prevents detection of any ions below m/z 200.
2. *Inspect the low-mass region for the b_2-ion.* Inspection of the low-mass portion of the spectrum does recognize an apparent b_2-ion at m/z 262, with the a_2-ion at m/z 234. Calculation of corresponding y_{n-2}-ion allows identification of the m/z 1033 ion as this y-ion. As seen in Table 2, it is possible that the low-mass cut-off seen in this spectrum precludes recording a number of possible b_2-ions.

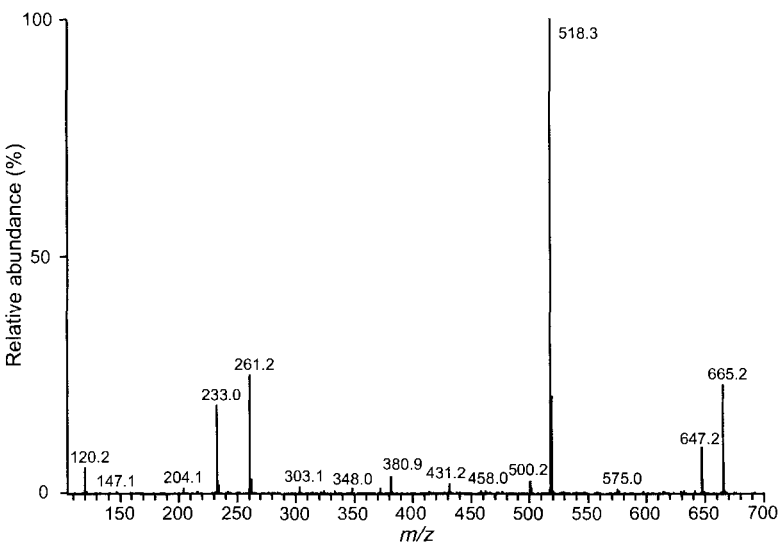

Figure 4.16. The product ion spectrum for interpretation problem one. This spectrum is identical to the spectrum shown in Figure 4.11 except no magnification has been used in the display. In this figure, the high abundance of the b_2- and y_5-ion pair can be seen.

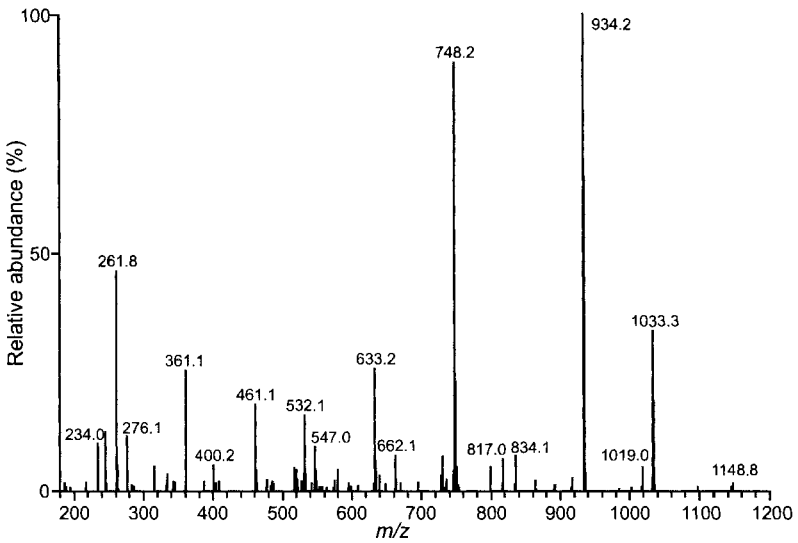

Figure 4.17. Interpretation problem number two. The product ion spectrum was acquired using an ion trap mass spectrometer by collisionally induced dissociation of a doubly charged ion, m/z 648.0 The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 2D electrophoresis gel. All of the ions in the spectrum are normalized to the most abundant ion in the spectrum. No additional ions were observed outside of the displayed m/z range.

However, the observation of the b_2-a_2 pair and the corresponding y_{n-2}-ion supports the selection of this b_2-ion.

3. *Inspect the low-mass region for the* y_1*-ion.* The low-mass cutoff does, however, prevent inspection for the y_1-ion at m/z 147 for lysine or m/z 175 for arginine. As an alternative, one can calculate the corresponding b_{n-1}-ions and attempt to observe those ions in the spectrum. For this peptide, if the C-terminal amino acid were lysine then the b_{n-1}-ion would be seen at m/z 1148, and if the C-terminal amino acid were arginine then the b_{n-1}-ion would be seen at m/z 1120. In this spectrum, an ion is seen at m/z 1148 corresponding to the b_{n-1}-ion for a C-terminal lysine. One should remember, however, that 147 Da could also represent an N-terminal phenylalanine. In fact, if one uses Table 2, the b_2-ion at m/z 262 could be either FN or C^aS so caution at this point would appear warranted. However, the appearance of the b_{n-1} ion for C-terminal lysine or arginine is often seen in spectra obtained by using an ion trap system and, because no b_{n-1}-ion for arginine is apparent, the assignment of lysine is well-founded. One should reiterate that this assumption of the possible C-terminal amino acids is made possible by the specificity of the proteolytic enzyme trypsin and is useful to begin the process of recognizing the b- and y-ion series.
4. *Inspect the high-mass region to identify the* y_{n-1}*-ion.* Inspecting the high-mass region for a y_{n-1}-ion that might distinguish between the choices of amino acid

pairs for the b_2-ion also turns to the m/z 1148 ion. This ion has been used as a b-ion to assign the C-terminus amino acid as a lysine. It is also the only ion in this region of the spectrum that provides a rational distinction between the possible b_2-ion combinations. As shown in Figure 4.18, which summarizes the interpretation to this point, there are two possible ways to handle this situation. One approach would be to utilize the data as completely as possible to produce a sequence. This approach leads to the FN- assignment shown in part A of this figure. The other approach would be to reflect the uncertainty at this point, which is reflected in the _ _ notation used in part B of this figure to indicate the place of two amino acids but shows that the identities of those residues are not known. The authors' general approach has been to use the data in the spectrum as completely as possible, such that the FN- assignment will be used for the remainder of the discussion.

5. *Extend the y-ion series toward lower* m/z. The interpretation can now begin by using the amino acid residue masses summarized in Table 4.1 to extend the y-ion series towards the y_1-ion. With the ion at m/z 1033 noted as the y_{n-2}-ion, the next abundant ion in the high-mass region is at m/z 934, a residue mass of 99 Da that assigns a valine residue. The corresponding b_3-ion is calculated as m/z 361, which is observed in the spectrum. The next abundant ion in the high-mass region is m/z 748 and corresponds to a tryptophan (residue mass 186 Da); a confirming b_4 is observed at m/z 547. At this point, however, one must remember (as shown in Table 4.3) that three amino acid

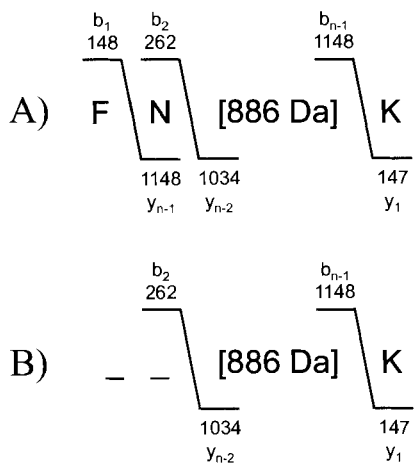

Figure 4.18. A summary of the on-going interpretation of the production ion spectrum in Figure 4.17. Two possible interpretations are shown. (A) The assignment of the most N-terminal amino acids based on the data observed in the spectrum. (B) A more conservative interpretation reflecting the uncertainty associated with the assignment of the ion at m/z 1148 as both a b-ion and a y-ion. The value in brackets denotes the remaining mass difference between the assigned ions. The – designates an unknown amino acid.

pairs have combined residue weights of 186 Da. In evaluating these choices, only –VS– appears possible, giving a calculated y-ion at m/z 835. However, not only is the ion in the spectrum at m/z 834, but the corresponding calculated b-ion would be expected at m/z 460 while the ion in the spectrum is at m/z 461. As a result, although worth noting as the interpretation progresses, this choice is less appealing than the tryptophan assignment. In addition, the relative abundance of the m/z 934 and m/z 748 ions is similar whereas the m/z 835 ion is an order-of-magnitude less abundant. While placing too much emphasis on the relative abundance of the different ions can be misleading, in this spectrum the pattern of the relative abundances supports a link between the m/z 934 and m/z 748 ions making tryptophan the better choice. The interpretation of the spectrum to this point is summarized in Figure 4.19.

Proceeding to the next y-ion, the ion at m/z 633 represents a residue mass of 115 Da, indicative of an aspartate residue with the calculated b_5-ion at m/z 662. The next y-ion would be at m/z 532, a residue mass of 101 Da corresponding to a threonine. The corresponding b_6-ion is calculated to be m/z 763 Da but is not present. This lack of a corresponding b-ion is notable because it can indicate a problem with the interpretation. In general, however, it is common for the abundance of the b-ion series to fade in the middle region of the spectrum, and such a trend in abundance is seen through this series of b-ions. The assignment of the threonine is further supported by the ability to assign the next y-ion as the m/z 461 ion. This assignment indicates an alanine, with a residue mass of 71 Da, at the next position with the calculated b_7-ion is seen at m/z 835 Da. The interpretation of this spectrum to this point is summarized in Figure 4.20.

6. *Extend the b-ion series toward higher m/z.* At this point, the next y-ion is not readily apparent because the ions at m/z 400 and 361 do not give appropriate amino acid residue masses. Therefore, as progress extending the y-ions series has halted, one can begin to use the residue masses in Table 4.1 to extend the b-ion series either working up from the b_7-ion or working down from the b_{n-1}-ion. In this case, the ion at m/z 1019 is 129 Da less than the b_{n-1}-ion at m/z 1148. This residue mass corresponds to a glutamate and the calculated y_2-ion is seen in the spectrum at m/z 276. One might also note that this ion is 14 Da

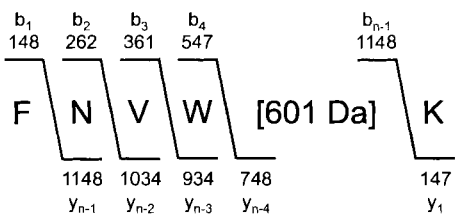

Figure 4.19. A summary of the on-going interpretation of the production ion spectrum in Figure 4.17. The value in brackets denotes the remaining mass difference between the assigned ions.

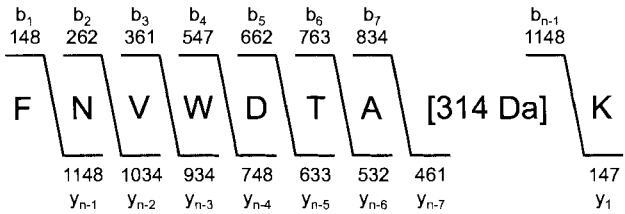

Figure 4.20. A summary of the on-going interpretation of the production ion spectrum in Figure 4.17. The value in brackets denotes the remaining mass difference between the assigned ions.

less than the m/z 1033 ion and therefore is not a neutral loss from the y-ion. The updated interpretation is shown in Figure 4.21.

As seen in Figure 4.21, the remaining mass difference is 185 Da. Table 4.2 can be used to compile a list of two-amino acid-combinations that can produce a given mass difference. One must remember, however, that the tabulated values are for b_2-ions, and that the combined residue mass of internal two amino acid pairs is 1 Da less than the values in the table. From Table 4.2, the two amino acid pairs of –NA– and –QG– are found for the 185 Da gap, giving possible sequence combinations of –NA–, –AN–, –QG–, and –GQ–. No b- or y-ions are present for the first three possible combinations. For example, if –NA– were the correct assignment then the b_8-ion would be expected at m/z 948 and the y_{n-8}-ion would be expected at m/z 347, with neither ion being observed in the spectrum. In fact, the only possible assignment is –GQ– utilizing the y-ion at m/z 404 and the b-ion at m/z 891. These low abundance ions are actually not even labeled in the spectrum. They are, nonetheless, truly detected (not noise) and the low abundance of ions seen for fragmentation at a glysine residue is a common observation. This completes the interpretation as shown in Figure 4.22.

7. *Calculate the mass of the peptide.* Completion of the sequence allows calculation of the mass of the putative sequence. The calculated protonated mass for this peptide is 1294.6 Da, a good match for the measured mass of 1295.0 Da.
8. *Reconcile the amino acid content with spectrum data.* Because no immonium ions were observed, we cannot evaluate the agreement of the interpreted sequence with observation of any specific ions. In this case, the lack of low-

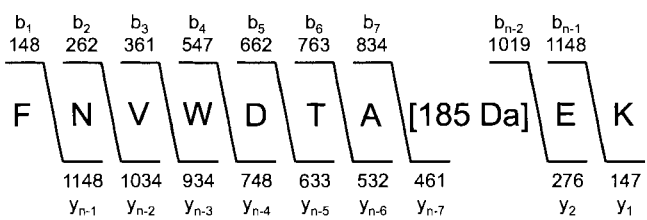

Figure 4.21. A summary of the on-going interpretation of the production ion spectrum in Figure 4.17. The value in brackets denotes the remaining mass difference between the assigned ions.

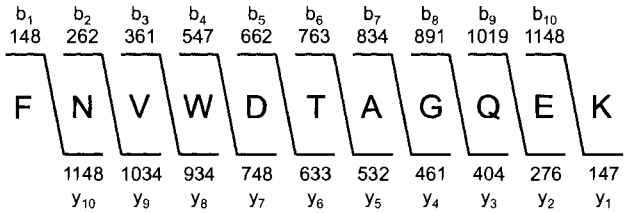

Figure 4.22. A summary of the complete interpretation of the product ion spectrum shown in Figure 4.17. Note that because the length of the peptide is now known, the b- and y-ion notation has changed to reflect the position in the peptide sequence.

mass data is significant because phenylalanine frequently produces detectable immonium ions at m/z 120. Had such an ion been observed, the FN– assignment would have been strengthened. It should be pointed out, however, that were this information critical, additional experiments (described in Chapter 9) could be carried out to assist this assignment.

9. *Attempt to identify all ions in the spectrum.* The only weak points of the amino acid sequence that has been deduced are the possible inclusion of oxidized methionine instead of the phenylalanine and the fact that the –GQ– assignment is based upon low relative-abundance ions. As alluded to above, the distinction between the oxidized methionine and phenylalanine could have been made through the observation of immonium ions. A second diagnostic point for these two amino acids would be the neutral loss of $HSOCH_3$ (64 Da) commonly observed for peptide ions containing oxidized methionine. Because of the position of the putative phenylalanine in this peptide sequence, only the precursor ion contains the residue of interest. The precursor ion can certainly lose the $HSOCH_3$ moiety, and often this loss is observed as a doubly charged ion. However, no ion is observed at m/z 616 so no evidence of an oxidized methionine is found in this spectrum. The use of low relative-abundance ions to make the –GQ– assignment again illustrates the full use of the data present in the spectrum. It is quite common that fragmentation at a glycine residue gives low relative-abundance product ions. The actual question concerning ions such as this should not be the relative abundance so much as the simple question of whether the ion of interest is a genuine part of the spectrum. This question is best answered based on one's experience with a given instrument system.

4.3.5. A Summary of Interpretation Problems One and Two

The two interpretation examples presented above were selected because they illustrate the complete use of the interpretation process to produce a peptide amino acid sequence. One can define three general types of results from the interpretation of product ion spectra: a complete interpretation to give an entire peptide amino acid sequence as was seen in these examples; a partial interpretation to give varying length segments of amino acid sequence of the peptide; or little or no

interpretation of the spectrum. It is the authors' experience that ~ 30 % of product ion spectra recorded for tryptic peptides can be completely interpreted. The majority of product ion spectra, ~ 50 %, are only partially interpretable and contain significant regions over which no amino acid assignments are made. Most often, the missing assignment is the first two amino acids because of a lack of the b_1- and y_{n-1}-ion. It is important to realize, however, that these spectra still contain a great deal of sequence information that may be useful in the overall characterization of the protein of interest and, given a possible protein sequence, may be confidently reconciled with that protein sequence. The remaining ~ 20 % of spectra are uninterpretable. As may be seen below, most uninterpretable spectra are triply charged or greater and therefore do not fragment in a manner that produces a uniform series of b- and y-ions, most likely due to the deleterious effects of the internal, fixed protonation site. One should remember, however, that like the partially interpretable spectra, these uninterpretable spectra are generally still informative. As a result, it is often possible to use even difficult, uninterpretable product ion spectra to test a specific sequence that may be derived from another source, such as a database sequence match.

4.3.6. Examples of More Difficult Product Ion Spectra That Cannot Be Completely Interpreted

The next series of sample interpretations intends to illustrate some spectral features that may confound the interpretation process and limit the completeness of the derived sequence. These features include but are not limited to preferential fragmentation at a proline, poor fragmentation at a glycine, and the fragmentation of a triply charged peptide ions. As noted above, it is the authors' experience that these types of spectra constitute the majority of the product ion spectra obtained from tryptic peptides, so it is important to appreciate that significant information can still be derived from an interpretation, or at least a partial interpretation, of these spectra.

Sample Interpretation Problem Number Three The next example, shown in Figure 4.23, illustrates the effects of the facile fragmentation that is generally seen at the N-terminal side of proline residues. In the electrospray mass spectrum of the peptide in this example, the doubly charged ion is the most abundant but a significant amount of the triply charged species is also observed. The protonated molecular weight measured in this spectrum is 1568.0 Da, $(M + H)^+$. The experiment was carried out by using an ion trap mass spectrometer system with unit m/z resolution fragmenting a doubly charged ion with an m/z of 784.5.

1. *Inspect the low-mass region for immonium ions.* The low-mass cut-off of the ion trap mass spectrometer limited the mass range that could be recorded. As a result, no immonium ions are observed.
2. *Inspect the low-mass region for the* b_2*-ion.* Inspection of the low-mass region does detect a possible b_2-ion/a_2-ion pair at m/z 277 and m/z 249, respectively. Using this b_2-ion allows the calculation of the corresponding y_{n-2}-ion as m/z

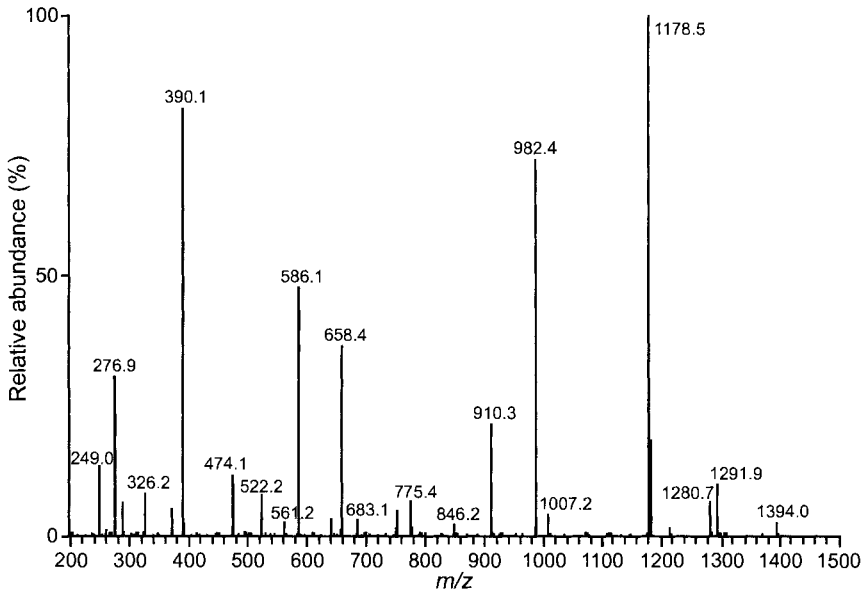

Figure 4.23. Interpretation problem number three. The product ion spectrum was acquired using an ion trap mass spectrometer by collisionally induced dissociation of a doubly charged ion, m/z 784.5 The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 2D electrophoresis gel. All of the ions in the spectrum are normalized to the most abundant ion in the spectrum. No additional ions were observed outside of the displayed m/z range.

1292, which is observed in the spectrum. The possible amino acid combinations for this b_2-ion, taken from Table 4.2, are MoE, FE, and XY. One should be aware that many other two-amino acid-combinations exist with b_2-ions at m/z below the mass range of this particular spectrum. However, the clear observation of the corresponding y_{n-2}-ion strongly supports this assignment.

3. *Inspect the low-mass region for the* y_1*-ion.* The low-mass cut-off also precludes observation of a possible y_1-ion at m/z 147 or 175. However, using the assumption that, due to the tryptic digestion, the peptide must have a C-terminal lysine or arginine residue allows one to assign the ion at m/z 1394 as b_{n-1}-ion for a C-terminal arginine.
4. *Inspect the high-mass region to identify the* y_{n-1}*-ion.* With the y_{n-2}-ion and the b_{n-1}-ion assigned, the high-mass region can be inspected to attempt to identify an ion that would allow the assignment of the two N-terminal amino acids. Regrettably, no other higher mass ions are observed in this spectrum. Even focused searches based only on the possible ions produced by the limited number of combinations given by the b_2-ion cannot clarify this assignment. As a result, no data exist in the spectrum that allows any determination to be made among the possible amino acid combinations given above. Figure 4.24 summarizes the beginning of this interpretation.

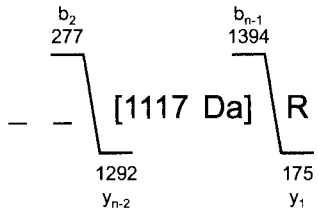

Figure 4.24. A summary of the on-going interpretation of the production ion spectrum in Figure 4.23. The value in brackets denotes the remaining mass difference between the assigned ions. The – indicates and unknown amino acid.

5. *Extend the y-ion series toward lower m/z.* Inspection of the high-mass region of the spectrum to extend the y-ion series should note the m/z 1281 ion. This ion cannot be the result of a small neutral loss from m/z 1292 y-ion, nor can it be a y-ion. This ion is, however, consistent with the b_{n-2}-ion giving a residue mass of 113 Da for a leucine or isoleucine. The calculated y_2-ion at m/z 288, although not labeled in the spectrum, is present at low relative abundance. The next high-mass ion, m/z 1179, is the most abundant ion in the spectrum. As a b-ion it would indicate a residue mass, 102.2 Da, which does not match any amino acid. As a y-ion it indicates a residue mass, 113.4 Da, which identifies a leucine or isoleucine at this position. Continuing from the m/z 1179 ion, no ion fitting the loss of a single amino acid exists. The high relative-abundance ion at m/z 982 represents a difference of 196 Da that must be due to two amino acids. Table 4.2 reveals that the only possible combination for this mass difference is the –PV– combination but no y-ion or b-ion exists in this spectrum to differentiate between –PV– and –VP–. However, the correct order of –PV– is supported by the high relative abundance of the m/z 1179 ion. This type of fragmentation is common in peptides containing proline where y-ions with an N-terminal proline are often the most abundant ions in the spectrum. This enhanced abundance is derived from the ability of the proline residue to preferentially capture the mobile proton and direct fragmentation at that amide bond. The enhanced basicity of the proline residue is also reflected in the observation of a significant amount of triply charged peptide ion in the electrospray mass spectrum with the implication that the basicity is not sufficient to fix the charge at that site as seen with internal arginine, histidine, or lysine residues. Further, the preferential fragmentation at the N-terminal side of a proline residue generally comes at the expense of the next y-ion in the series, which in this spectrum is the m/z 1082 ion. The deduced sequence to this point is summarized in Figure 4.25.

In attempting to continue the y-ion series from the m/z 982 ion, one can rule out the ions at m/z 910 and m/z 846 because they give non-existent residue masses of 72 Da and 136 Da, respectively. The next abundant ion at m/z 775 would give a residue mass of 207 Da, which is too large for a single amino acid. Further, recalling that the precursor ion had an m/z of 784.5 allows the

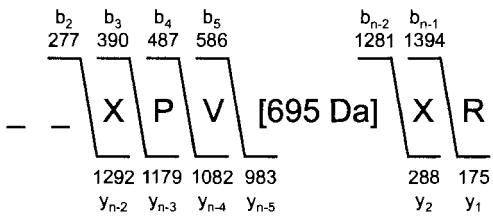

Figure 4.25. A summary of the on-going interpretation of the production ion spectrum in Figure 4.23. The value in brackets denotes the remaining mass difference between the assigned ions. The – indicates and unknown amino acid.

assignment of this ion as a doubly charged ion from the loss of water or ammonia from the precursor ion.

6. *Extend the b-ion series toward higher m/z.* Because of the difficulty extending the y-ion series, one might instead attempt to extend the b-ion series from the m/z 586 ion toward the C-terminus. The ions at m/z 640 and 658 correspond to a residue masses of 54 and 72 Da, respectively, which are not amino acid residue masses. The ion at m/z 683 gives a residue mass of 97 Da, a proline residue, although no corresponding y-ion is seen at m/z 885. As noted above, this pattern of abundance is often seen with fragmentation at proline residues, such that the proline assignment can be made. Regrettably, the b-ions series cannot be extended from the m/z 683 ion because no additional ions are seen in the spectrum that correspond to an amino acid residue mass. Figure 4.26 summarizes the interpretation to this point.

   In many instances the interpretation of a product ion spectrum might stop at this point. Because a complete interpretation has not been made, the molecular weight of the deduced peptide sequence cannot be calculated. Also, the amino acid content of the deduced sequence cannot by reconciled with the immonium ions because no immonium ions were observed. As a result, Steps 7 and 8 in the interpretation strategy are not relevant to this example at this point in the process.

9. *Attempt to identify all ions in the spectrum.* In this product ion spectrum, one should note the prominent ions at m/z 910 and m/z 658. The m/z 910 could

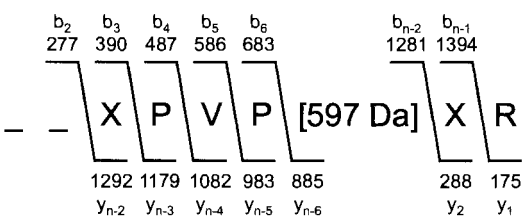

Figure 4.26. A summary of the on-going interpretation of the production ion spectrum in Figure 4.23. The value in brackets denotes the remaining mass difference between the assigned ions. The – indicates and unknown amino acid.

not be explained in the y-ions series and therefore must be a b-ion. The m/z 658 ion could not be explained in the b-ion series and therefore must be a y-ion. Further, one should note that these ions are a complementary b- and y-ion pair, because their combined mass equals 1568 Da. The recognition of this pair of ions gives another opportunity to continue the interpretation. As described above, the relative abundance of these two ions appears better than the surrounding ions in the respective ion series, a trait that is characteristic of fragmentation at a proline residue. This observation is supported by the b-ion at m/z 1007, and the y-ion seen at m/z 561. Figure 4.27 summarizes the interpretation to this point.

Other ions in the spectrum that remain unexplained are at m/z 326, m/z 474, m/z 522, and m/z 846. These ions have been excluded from the b- and y-ions series and do not appear to result from the loss of water or ammonia from nearby ions. Careful observation, however, should note that these ions differ by 64 Da from adjacent b-ions, the b_3-ion at m/z 390, the b_5-ion at m/z 586, and the b-ion at m/z 910, respectively. These neutral losses indicate that these b-ion contain an oxidized methionine and, based on the loss from the b_3-ion, this residue must be located in the first two N-terminal amino acids. Therefore, the first two amino acids must be either MoE– or EMo–. As stated above, no ion is seen in the spectrum that allows the assignment of the specific order of these amino acids, but this observation does add to the characterization of the sequence.

As seen in Figure 4.27, the interpretation to this point has produced a discontinuous peptide sequence with two defined sequence gaps. It is important to remember, however, that although the precise amino acid sequence in these gaps is not known at this time, the mass differences in the b- and y-ion series are known and provide significant limits on the possible amino acids contained in those gaps. The magnitudes of these mass differences are within the range of the two-amino acid-combinations tabulated in Table 4.2 as b_2-ions. From this table, the 227 Da difference would include the possible combinations of NX, QV, KV, and AR. No ions are observed in the spectrum that can differentiate among these possibilities, although one might discount the KV and AR possibilities because the additional fixed charged site they would create does not seem to be present in this peptide.

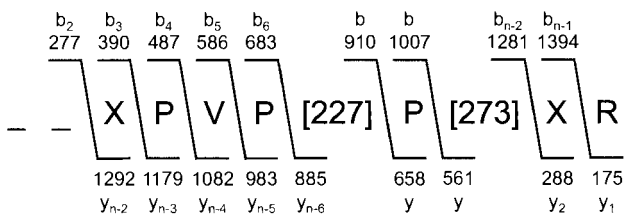

Figure 4.27. A summary of the on-going interpretation of the production ion spectrum in Figure 4.23. The value in brackets denotes the remaining mass difference between the assigned ions. The indicates and unknown amino acid.

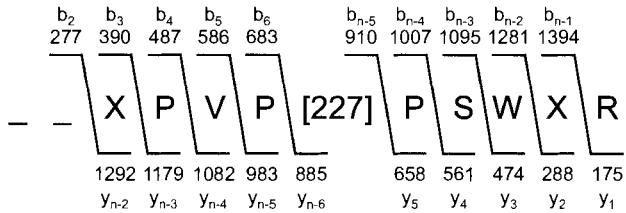

Figure 4.28. A summary of the on-going interpretation of the production ion spectrum in Figure 4.23. The value in brackets denotes the remaining mass difference between the assigned ions. The indicates and unknown amino acid.

The possible two-amino acid-combinations for the 273 Da gap include C[a]V and WS. In this case, the m/z 474 ion is consistent with a y-ion that gives a residue mass of 186 Da, assigning the –WS– sequence and giving the interpretation shown in Figure 4.28. As described above, inclusion of the tryptophan residue in the interpretation raises the possibility of two amino acids at this position. Although no ions are observed that are consistent with two amino acids at this position, the nature of this spectrum makes this lack of data a less conclusive result than that seen in the interpretation problem number two.

The complete amino acid sequence of this peptide is shown in Figure 4.29. One sees that every aspect of the sequence is consistent with the spectrum in Figure 4.23. Unfortunately, the absence of a couple of key ions has prevented the complete interpretation. This loss of information reflects the preferential fragmentation at the N-terminal side of proline residues that produces an overly abundant ion that detracts from the abundance of the other, needed ions.

Sample Interpretation Problem Number Four The product ion spectrum for the next interpretation problem is shown in Figure 4.30. This product ion spectrum is presented as an example of another common aspect of the fragmentation of protonated peptide ions, the poor fragmentation often seen at glycine residues. In many ways, the lack of fragmentation at glycine residues is a more difficult problem because, in contrast to the facile fragmentation at proline residues described above,

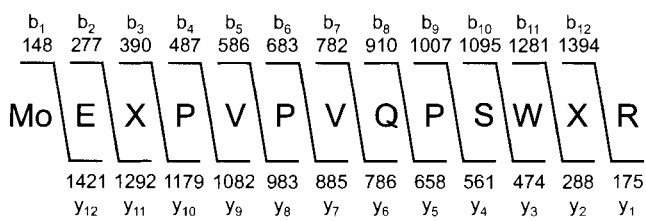

Figure 4.29. A summary of the complete interpretation of the product ion spectrum shown in Figure 4.23. Note that because the length of the peptide is now known, the b- and y-ion notation has changed to reflect the position in the peptide sequence.

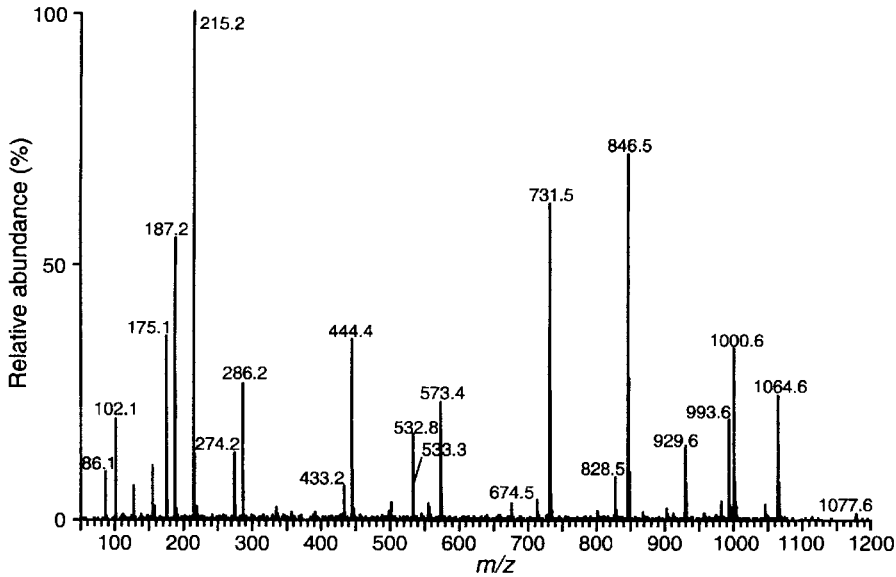

Figure 4.30. Interpretation problem number four. The product ion spectrum was acquired using a quadrupole-time-of-flight mass spectrometer by collisionally induced dissociation of a doubly charged ion, m/z 639.8. The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 2D electrophoresis gel. All of the ions in the spectrum are normalized to the most abundant ion in the spectrum. The ion at nominal mass 533 is labeled to indicate that the m/z spacing seen in the isotope cluster of the ion is 0.5 Th. No additional ions were observed outside of the displayed m/z range.

there is rarely an accompanying high-abundance ion to presage the anomaly. More often one will suspect a glycine residue when the on-going interpretation reaches an apparent gap. The product ion spectrum shown in the figure was obtained by fragmentation of a doubly charged ion with an m/z of 639.8. The measured molecular weight of the peptide is 1278.6 Da, $(M + H)^+$. The experiment was carried out with a quadrupole-time-of-flight mass spectrometer system with an m/z resolution of $\sim$7000.

1. *Inspect the low-mass region for immonium ions.* Because of the use of the quadrupole-time-of-flight instrument, the entire mass range was acquired. Immonium ions are detected at m/z 86, m/z 102, and m/z 129, which are consistent with leucine or isoleucine, glutamate, and arginine being present in the peptide.
2. *Inspect the low-mass region for the* b_2*-ion.* A b_2-ion/a_2-ion pair is clearly seen at m/z 215 and m/z 187, respectively. Recognition of these ions allows calculation of the y_{n-2}-ion at m/z 1065. Inspection of the spectrum finds this ion in the spectrum. Based on the tabulated values in Table 4.2, the possible combinations for the first two amino acids are XT and VD.
3. *Inspect the low-mass region for the* y_1*-ion.* The y_1-ion is clearly recognizable at m/z 175, indicating that the C-terminal amino acid is an arginine.

Calculation of the expected b_{n-1}-ion and inspection of the spectrum, however, cannot locate this b-ion in the spectrum. The lack of the high b-ions is not uncommon, but this lack of a confirming ion should lead one to check the spectrum, unsuccessfully in this case, for any indication of a lysine residue at the C-terminus. As a result, the arginine assignment is made. This beginning of the interpretation process is summarized in Figure 4.31. As stated in the first interpretation example, this stage in the interpretation is an important one because these first steps have established the N- and C-terminal orientation of the peptide and the beginning of the two-ion series.

4. *Inspect the high-mass region to identify the* y_{n-1}*-ion.* An inspection of the high-mass region of the spectrum for a possible y_{n-1}-ion does not readily identify any ions in this region other than the previously noted y_{n-2}-ion. However, because the b_2-ion has given a relatively limited number of possibilities for the first two amino acids, a methodical search is possible that focuses on four possibilities: m/z 1178, m/z 1166, m/z 1180, and m/z 1164, for T, X, V, and D, respectively, as the N-terminal amino acid. A low but apparently significant abundance ion is seen at m/z 1178, consistent with the N-terminal amino acid being a threonine residue. This observation also allows the assignment of the second amino acid as a leucine or isoleucine residue (designated with an X). As described for the previous interpretation, the abundance of this ion does not appear much greater than the other low-abundance ions in this region that have no apparent significance. The fact that the b_2-ion gives a limited number of possibilities, however, gives this ion more significance and allows its use in the interpretation.
5. *Extend the y-ion series toward lower m/z.* The extension of the y-ion series from the ion at m/z 1065 has a number of high-abundance ions to consider. The ion at m/z 1001 gives a residue mass of 64 Da, which is not an amino acid. This loss, however, is produced by an oxidized methionine losing $HSOCH_3$, alerting one to the presence of this amino acid in the peptide sequence. The next prominent ion at m/z 994 is consistent with a residue mass of 71 Da for an alanine. The corresponding b_3-ion is calculated as m/z 286, which is observed in the spectrum. The next abundant ion in the high-mass region is m/z 930. This ion represents a loss of 64 Da from the m/z 994 ion, further evidence of the presence of the oxidized methionine. The next abundant ion in this region is the m/z 847 ion that gives a residue mass of

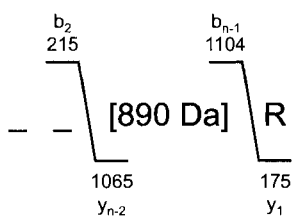

Figure 4.31. A summary of the on-going interpretation of the production ion spectrum in Figure 4.30. The value in brackets denotes the remaining mass difference between the assigned ions. The – indicates and unknown amino acid.

147 Da. This residue mass indicates either a phenylalanine or oxidized methionine residue at this position. The preceding losses of 64 Da have indicated the presence of an oxidized methionine in the peptide. Also, since no immonium ion at m/z 120 was seen, no phenylalanine is expected. Finally, no additional losses of 64 Da are noted for the remainder of the spectrum. These observations allow a confident assignment of oxidized methionine as the next residue. The corresponding b_4-ion is calculated as m/z 433, which is observed in the spectrum. Observation of the ion at m/z 731 allows the assignment of the next amino acid as an aspartate residue (a residue mass of 115 Da). The calculated b_5-ion is calculated as m/z 548 but is not observed in the spectrum. The lack of a b_5-ion is not surprising, especially considering the fading abundance of the b-ion series from the b_2- to b_3- to b_4-ions. Further extension of the y-ion series now becomes difficult. If the ion at m/z 573 is considered as the next ion in the y-ion series, then the calculated residue mass would be 158 Da, which is not an amino acid. The interpretation to this point is summarized in Figure 4.32.

6. *Extend the b-ion series toward higher m/z.* Normally, it would be recommended that one begin to consider the b-ion series at this point and attempt to extend the b-ion series toward higher m/z values. This approach is not productive in this case because of the lack of b-ions that has been encountered.
7. *Calculate the mass of the peptide.* Because the interpretation cannot be completed, this step is irrelevant.
8. *Reconcile the amino acid content with the spectrum data.* The amino acid sequence that has been deduced to this point contains a leucine or isoleucine and an arginine. The glutamate that was also seen in the immonium ion data, however, has not yet been assigned.
9. *Attempt to identify all ions in the spectrum.* The significant abundance of the ions at m/z 274, m/z 444, and m/z 573 cannot be overlooked. (One will note that in this spectrum, the ion at m/z 532 can be overlooked because it can be seen as doubly charged due to the resolution of isotope cluster by the time-of-flight mass analysis.) Based on the calculated b-ions, the ions at m/z 274 and m/z 444 would definitely appear to be y-ions. Further, the ion at m/z 573 is very likely a y-ion based on the fact that the calculated difference from the last b-ion, m/z 433, is 140 Da; a value that cannot be explained by any amino acid or two-amino acid-combinations. The m/z 274 ion can be assigned as the y_2-

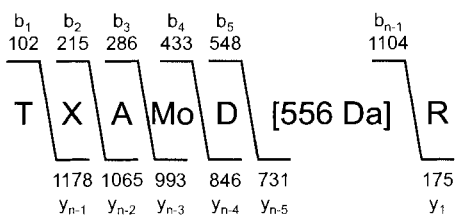

Figure 4.32. A summary of the on-going interpretation of the production ion spectrum in Figure 4.30. The value in brackets denotes the remaining mass difference between the assigned ions.

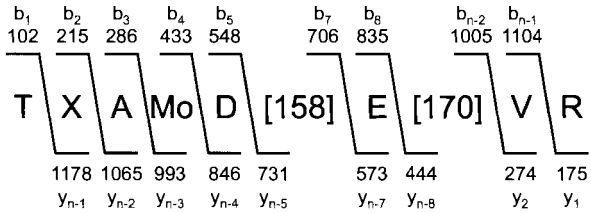

Figure 4.33. A summary of the on-going interpretation of the production ion spectrum in Figure 4.30. The value in brackets denotes the remaining mass difference between the assigned ions.

ion making the C-terminal amino acids -VR. Also, the difference between the m/z 573 and the m/z 444 ions is 129 Da, the residue mass of a glutamate. The corresponding b-ions are calculated for each of these assignments but none of the b-ions are observed in the spectrum. These observations can be incorporated into the interpretation as summarized in Figure 4.33.

As noted in Figure 4.33, two gaps are present in the interpretation. The magnitude of these gaps is not consistent with any single amino acid residue mass but is consistent with a number of two-amino acid-combinations. Table 4.2 can be used at this point to compile a list of possible amino acid combinations to bridge these gaps, remembering that the values contained in the table are for b_2-ions and will be 1 Da greater that the internal two-amino acid-combinations needed here. For the 158 Da gap, the possible combinations are –SA– and –GT–. For the 170 Da gap, the possible combinations are –VA– and –XG–. At this time, the knowledge that fragmentation at glycine residues is often quite poor can be used in the interpretation since each gap can be explained by a glycine-containing pair of amino acids. Indeed, for the 158 Da sequence gap, a low- abundance ion is observed at m/z 674 that is consistent with the –GT– assignment. This extension of the interpretation is shown in Figure 4.34. This assignment is significant because it extends the largest portion of interpreted sequence from five consecutive amino acids, a portion of sequence that is generally difficult to use in database search queries, to an eight amino acid sequence, which works well in such queries. A similar inspection of the spectrum for ions that would distinguish between the –VA– and –XG– pairs in the 170 Da sequence gap cannot identify any informative ions. Again, one suspects that the lack of data is a reflection of the poor

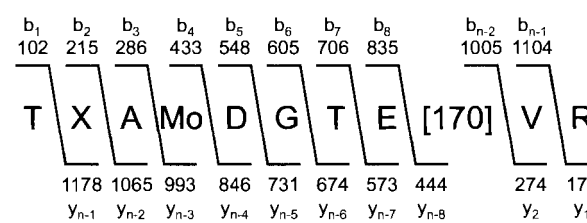

Figure 4.34. A summary of the on-going interpretation of the production ion spectrum in Figure 4.30. The value in brackets denotes the remaining mass difference between the assigned ions.

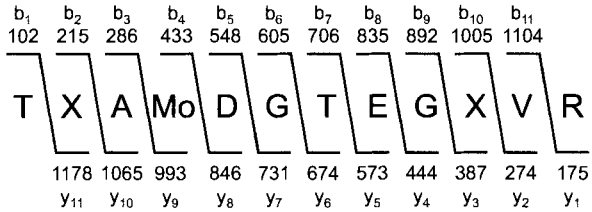

Figure 4.35. A summary of the complete interpretation of the product ion spectrum shown in Figure 4.30. Note that because the length of the peptide is now known, the b- and y-ion notation has changed to reflect the position in the peptide sequence.

fragmentation around a glycine residue but the total lack of data makes it difficult to make that assignment in this case. As it turns out, the proper assignment is the –GX– assignment. The complete interpretation of this spectrum is summarized in Figure 4.35.

A Summary of Interpretation Problems Three and Four These two interpretation problems were chosen to illustrate situations in which high-quality product ion spectra fail to produce a complete interpretation due to a lack of a complete set of b- or y-ions. The lack of information in these spectra was due to either proline or glycine residues in the peptide sequences. These residues tend to affect the appearance of product ion spectra as shown in the examples. In other spectra, information will be lacking even without such explanations. It is also possible that the informative ions are masked by uninformative ions generated through rearrangement reactions. In any case, the result of an incomplete set of informative ions is an inability to deduce the entire sequence of the peptide. It is important to emphasize in these situations that there is often still a great deal of sequence-specific information in the product ion spectra and the spectra may still yield valid, confident matches when used in database search queries as described in Chapter 8. Alternatively, as discussed in Chapter 9, additional experiments can be performed that are designed to produce more interpretable sequence information and, ideally, effect a complete interpretation of the spectrum.

4.3.7. Interpretation of Product Ion Spectra from Triply Charged Ions

All of the previous examples have involved product ion spectra of doubly charged peptide ions. Although doubly charged ions are more common in the tryptic digests of interest in this discussion, several triply charged peptides will be present in every digest. The important difference in the structure of triply charged peptide ions is the presence of an internal lysine, arginine, or histidine residue in the peptide that gives a highly basic site for attaching a fixed proton during the ionization process. This fixed internal charge site, in turns, affects the migration of the mobile proton along the length of the peptide. A simple model of this effect is that the mobile proton, initially associated with the N-terminal amine moiety, cannot migrate close to or past the fixed internal protonation site. One result of this limit on the movement of the mobile

proton is a corresponding limit on amide bonds that are broken in the charge site-initiated fragmentation reactions. The second result is that many of the products of these reactions will be doubly charged. Therefore, the question of the charge state of a product ion is particularly important in these spectra.

Sample Interpretation Problem Number Five The product ion spectrum for the next interpretation problem is shown in Figure 4.36.A. This spectrum was acquired from a triply charged ion with an m/z of 802.8 using an ion trap mass spectrometer. The calculated average molecular weight of the peptide is 2406.4 Da, $(M + H)^+$. The monoisotopic molecular weight can also be calculated, based on a doubly charged ion detected in the mass spectrum, as 2405.6 Da, $(M + H)^+$. This example illustrates the fact that sequence information can be present in the product ion spectra of triply charged ions although that information is rarely complete and is often in the form of a series of doubly charged product ions.

In this spectrum, no ions with m/z below ~600 are observed. As a result, no immonium ions, b_2-ion, or y_1-ion are observed and steps 1, 2, and 3 in the interpretation strategy are irrelevant. Inspection of this spectrum should note a series of product ions that are reminiscent of a series of informative product ions. The differences between these ions, however, are apparently not consistent with residue masses. For example, beginning with the m/z 1147 ion, the ion at m/z 1089 would indicate a residue mass of 57 for a glycine, but the ion at m/z 1016 would then indicate a non-existent residue mass of 74 Da and the ion at m/z 951 would indicate a non-existent residue mass of 138 Da. Similar problems would be encountered throughout this series of ions for any attempt to derive sequence information assuming singly charged ions. In short, these problems should lead one to consider the possibility that these ions are doubly charged. Exactly which ions are doubly charged would be apparent if the experiment were carried out using a quadrupole-time-of-flight instrument because of the resolution of the time-of-flight mass analyzer. However, as shown in this example, the deduction of the +2 charge state is possible in many instances without the high resolution.

Figure 4.36.B shows the critical region of the spectrum expanded for clarity. For each of the more abundant ions in this region of the spectrum, the m/z has been calculated as a singly charged ion and added to the spectrum as an underlined value. The first ion in this series, a calculated singly charged m/z of 2292.4, is 113.2 Da less than the measured monoisotopic mass of 2405.6 Da. This difference corresponds to a leucine or isoleucine as the first amino acid. This amino acid is assumed to be the N-terminal amino acid because the peptide was formed in a tryptic digestion. The use of the next series of calculated singly charged m/z (the underlined values) allows the calculation of a series of residue masses as 114.8 Da, 147.2 Da, 129.0 Da, 115.0 Da, 99.4 Da, 112.6 Da, 71.4 Da, and 128.8 Da, respectively. These residue masses indicate, in order, D, F (or Mo), E, D, V, X, A, and E, to give an N-terminal amino acid sequence of XDFEDVXAE. One should appreciate that the calculated residue mass are not as precise when using doubly charged ions as the values calculated with singly charged ions. This deterioration of the precision is due

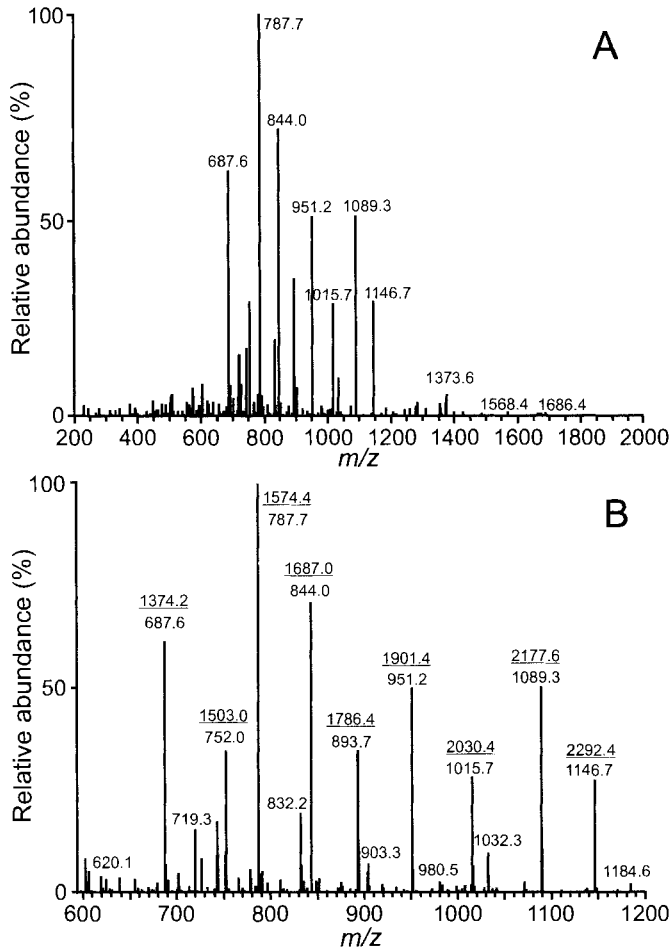

Figure 4.36. Interpretation problem number five. The product ion spectrum was acquired by using an ion trap mass spectrometer by collisionally induced dissociation of a triply charged ion, m/z 802.8. The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 1D electrophoresis gel. All of the ions in the spectrum are normalized to the most abundant ion in the spectrum. (A) The complete product ion spectrum. No additional ions were observed outside of the m/z range shown in this spectrum. (B) A selected portion of the spectrum shown in Part A of the figure to show details of the ions that are seen. Note that many of the ions are labeled with both the measured m/z and the calculated m/z for a corresponding singly charged product ion.

to the additional error created by the conversion of measured doubly charged m/z to calculated singly charged m/z. Calculation of the corresponding b-ion series for these assignments allows the identification of the m/z 620, m/z 719, m/z 832, m/z 903, and m/z 1032 ions as b-ions. This interpretation of the spectrum is summarized in Figure 4.37. As described in Chapter 8 of this volume, this portion of sequence information is sufficient to allow matching to the database sequence IDFEDVIAE-PEGTHSFDGIWK from the protein caveolin.

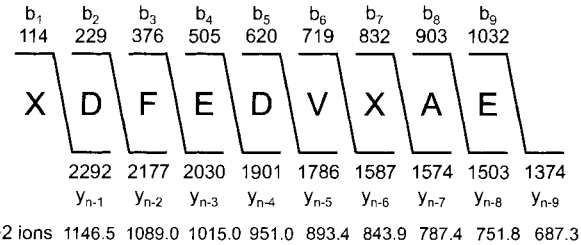

Figure 4.37. A summary of the interpretation of the product ion spectrum shown in Figure 4.36. This figure shows as much sequence information as could be deduced from this spectrum. The entire sequence of this peptide is IDFEDVIAEPEGTHSFDGIWK. The calculated *m/z* for the doubly charged y-ions have been added below the y-ion series.

It is interesting to note the position of the histidine residue, and therefore the internal fixed protonation site, in this peptide sequence. The histidine location, 14 residues removed from the N-terminus, allows the mobile proton located at the N-terminus to migrate over the 10 amide bonds at which fragmentation was observed without the electrostatic repulsion that the protonated histidine would create. This migration produces the fragmentation observed in the spectrum. Also, the additional charge located on the histidine residue means that the fragment ions that are formed are doubly charged.

In contrast, the product ion spectrum shown in Figure 4.38 is from the peptide GHYTEGAELVDSVLDVVR from β-tubulin. For this product ion spectrum, a triply

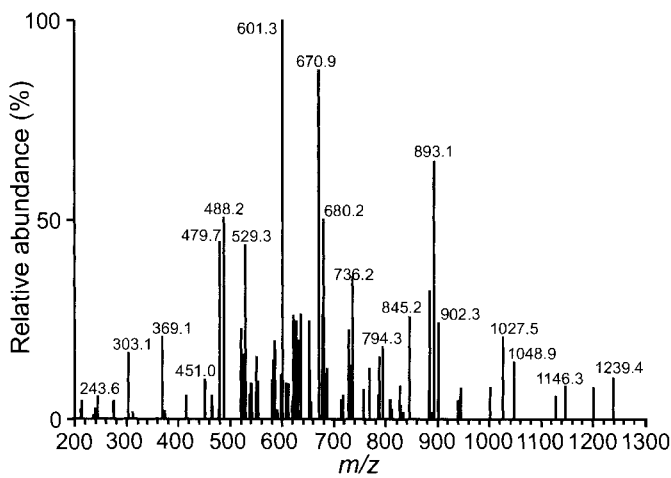

Figure 4.38. The collisionally induced dissociation spectrum of a triply charged peptide ion. The product ion spectrum was acquired using an ion trap mass spectrometer by collisionally activated dissociation of a triply charged ion, *m/z* 654.1. The peptide was produced by an in-gel tryptic digestion of a Coomassie blue-stained protein band in a 2D electrophoresis gel. All of the ions in the spectrum are normalized the most abundant ion in the spectrum. No additional ions were observed outside of the *m/z* range shown in this spectrum. The sequence of the peptide is GHYTEGAELVDSVLDVVR. The uninformative nature of the spectrum is due to the deleterious effect of the histidine protonation site.

charged ion with m/z 654.1 was fragmented in an experiment carried out using an ion trap mass spectrometer system. Two aspects of this spectrum are noteworthy. First, although difficult to see in this presentation, the abundance of ions in the spectrum is significantly lower than expected based on the abundance of the precursor ions. Second, the ions that are formed provide little information that can be used to deduce the sequence. These characteristics reflect the effect of the fixed protonation at the histidine residue on the migration of the mobile proton at the N-terminus, preventing its movement down the length of the peptide. Without this movement, not only is little fragmentation observed (4.28). hence the low abundance of the product ion spectrum, but the fragmentation that is observed is not due to the predictable and informative reactions summarized in Section 4.2 of this chapter and discussed in the previous interpretations.

4.4. SUMMARY

The primary purpose of this chapter has been to describe a methodical approach to the interpretation of the product ion spectra of the peptides formed in the tryptic digestion of a protein. A nine-step strategy has been described that is designed to facilitate the recognition of the different b- and y-ions in a product ion spectrum and to calculate the residue mass of the amino acid at each position from those ions. This method is a cyclic process that actively verifies each assignment through complementary observations in the two-fragment ion series. The method is aided by the preferential fragmentation of the amide bonds in peptides, the fact that there are only 20 different peptides and the fact that the residue masses of the 20 peptides vary over a wide range.

The peptides produced by a tryptic digestion have structures that are ideally suited to peptide sequencing by electrospray ionization-tandem mass spectrometry. These peptides typically contain only two strongly basic groups in each peptide, the N-terminus and the side chain of the C-terminal lysine or arginine at which proteolysis occurs. This type of peptide structure forms abundant doubly charged ions that fragment in a systematic and predictable manner following collisional activation. As a result, the product ion spectra that are recorded are readily interpretable to deduce the amino acid sequence of the peptide. This situation does not imply that the product ion spectrum of every doubly charged peptide ion is completely interpretable. In fact, most product ion spectra are not completely interpretable. In the analysis of any given digest, however, several product ion spectra will be completely interpretable and these spectra will form the basis of a confident, multifaceted identification of the source protein. The majority of remaining product ion spectra will be sufficiently informative that they can be retrospectively reconciled with the protein sequence that is identified, further strengthening that identification.

4.5. REFERENCES

4.1. Biemann, K. Appendix 5. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol.* 193:886–887, 1990.

4.2. McCormack, A.L.; Jones, J.L.; Wysocki, V.H. Surface-induced dissociation of multiply charged protonated peptides. *J. Am. Soc. Mass Spectrom.* 3:859–862, 1992.

4.3. McCormack, A.L.; Somogyi, A.; Dongre, A.R.; Wysocki, V.H. Fragmentation of protonated peptides: Surface-induced dissociation in conjunction with a quantum mechanical approach. *Anal. Chem.* 65:2859–2872, 1993.

4.4. Jones, J.L.; Dongre, A.R.; Somogyi, A.; Wysocki, V.H. Sequence dependence of peptide fragmentation efficiency curves determined by electrospray ionization/surface-induced dissociation mass spectrometry. *J. Am. Chem. Soc.* 116:8368–8369, 1994.

4.5. Dongre, A.R.; Somogyi, A.; Wysocki, V.H. Surface-induced dissociation: An effective tool to probe structure, energetics, and fragmentation mechanism of protonated peptides. *J. Mass Spectrom.* 31:339–350, 1996.

4.6. Dongre, A.R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. *J. Am. Chem. Soc.* 118:8365–8374, 1996.

4.7. Cox, K.A.; Gaskell, S.J.; Morris, M.; Whiting, A. Role of the site of protonation in the low-energy decompositions of gas-phase peptide ions. *J. Am. Soc. Mass Spectrom.* 7:522–531, 1996.

4.8. Yalcin, T.; Khouw, C.; Csizmadia, I.G.; Peterson, M.R.; Harrison, A.G. Why are b- ions stable species in peptide spectra? *J. Am. Soc. Mass Spectrom.* 6:1165–1174, 1995.

4.9. Yalcin, T.; Khouw, C.; Csizmadia, I.G.;Peterson, M.R.; Harrison, A.G. The structure and fragmentation of $b_n (n \geq 3)$ ions in peptide spectra. *J. Am. Soc. Mass Spectrom.* 7:233–242, 1996.

4.10. Tang, X-J.; Boyd, R.K. An investigation of fragmentation mechanisms of doubly charged protonated tryptic peptides. *Rapid Comm. Mass Spectrom.* 6:651–657, 1992.

4.11. Covey, T.R.; Huang, E.C.; Henion, J.D. Structural characterization of protein tryptic peptides via liquid chromatography/mass spectrometry and collision-induced dissociation of their doubly charged molecular ions. *Anal. Chem.* 63(13):1193–200, 1991.

4.12. Cardenas, M.S.; van der Heeft, E.; de Jong A.P. On-line derivatization of peptides for improved sequence analysis by micro-column liquid chromatography coupled with electrospray ionization-tandem mass spectrometry. *Rapid Comm. Mass Spectrom.* 11:1271–1278, 1997.

4.13. Burlet, O.; Orkisqewski, R.S.; Ballard, K.D.; Gaskell, S.J. Charge promotion of low-energy fragmentations of peptide ions. *Rapid Comm. Mass Spectrom.* 6:658–662, 1992.

4.14. Summerfield, S.G.; Bolgar, M.S.; Gaskell, S.J. Promotion and stabilization of b_1 ions in peptide phenythiocarbamoyl derivatives: Analogies with condensed-phase chemistry. *J. Mass Spectrom.* 32:225–231, 1997.

4.15. Vachet, R.W.; Ray, K.L.; Glish, G.L. Origin of product ions in the MS/MS spectra of peptides in a quadrupole ion trap. *J. Am. Soc. Mass Spectrom.* 9:341–344, 1998.

4.16. Brown, R.S.; Lennon, J.J. Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. *Anal. Chem.* 67:3990–3999, 1995.

4.17. Brown, R.S.; Carr, B.L.; Lennon, J.J. Factors that influence the observed fast fragmentation of peptides in matrix-assisted laser desorption. *J. Am. Soc. Mass Spectrom.* 7:225–232, 1996.

4.18. Spengler, B.; Kirsch, D.; Kaufmann, R.; Jaeger, E. Peptide sequencing by matrix-assisted laser-desorption mass spectrometry. *Rapid Comm. Mass Spectrom.* 6:105–108, 1992.

4.19. Chaurand, P.; Luetzenkirchen, F.; Spengler, B. Peptide and protein identification by matrix-assisted laser desorption ionization (MALDI) and MALDI-post-source decay time-of-flight mass spectrometry. *J. Am. Soc. Mass Spectrom.* 10:91–103, 1999.

4.20. Domingues, M.R.M.; Marques, M.G.O.S.; Vale, C.A.M.; Neves, M.G.; Cavaleiro, J.A.S.; Ferrer-Correia, A.J.; Nemirovskiy, O.V.; Gross M.L. Do charge-remote fragmentations occur under matrix-assisted laser desorption ionization post-source decompositions and matrix-assisted laser desorption ionization collisionally activated decompositions? *J. Am. Soc. Mass Spectrom.* 10:217–223, 1999.

4.21. Liao, P-C.; Huang, Z-H.; Allison, J. Charge remote fragmentation of peptides following attachment of a fixed positive charge: A matrix-assisted laser desorption/ionization postsource decay study. *J. Am. Soc. Mass Spectrom.* 8:501–509, 1996.

4.22. Kaufmann, R.; Spengler, B.; Lutzenkirchen, F. Mass spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass spectrometer using matrix-assisted laser desorption ionization. *Rapid Comm. Mass Spectrom.* 7:902–910, 1993.

4.23. Kaufmann, R.; Chaurand, P.; Kirsch, D.; Spengler, B. Post-source decay and delayed extraction in matrix-assisted laser desorption/ionization-reflectron time-of-flight mass spectrometry. Are there trade-offs? *Rapid Comm. Mass Spectrom.* 10:1199–1208, 1996.

4.24. Griffin, P.R.; MacCoss, M.J.; Eng, J.K.; Blevins, R.A.; Aaronson, J.S.; Yates, J.R. III. Direct database searching with MALDI-PSD spectra of peptides. *Rapid Comm. Mass Spectrom.* 9:1546–1551, 1995.

4.25. A regularly presented short course on protein sequencing by tandem mass spectrometry, presented by Professor Donald F. Hunt and his research group at the University of Virginia Department of Chemistry.

4.26. Mandal, A.; Naaby-Hansen, S.; Wolkowicz, M.J.; Klotz, K.; Shetty, J.; Retief, J.D.; Coonrod, S.A.; Kinter, M.; Sherman, N.; Cesar, F.; Flickinger, C.J.; Herr, J.C. FSP95, a testis-specific 95-kilodalton fibrous sheath antigen that undergoes tyrosine phosphorylation in capacitated human spermatozoa. *Biol. Reproduct.* 61:1184–1197, 1999.

4.27. Mehta, A.; Kinter, M.T.; Sherman, N.E.; Driscoll, D.M. Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Molec. Cell. Biol.* 20:1846–1854, 2000.

4.28. Summerfield, S.G.; Gaskell, S.J. Fragmentation efficiencies of peptide ions following low-energy collisional activation. Int. *J. Mass Spectrom. Ion Processes* 165/166:509–521, 1997.