Published in final edited form as:

J Am Soc Mass Spectrom. 2016 September ; 27(9): 1443–1453. doi:10.1007/s13361-016-1418-9.

Statistical examination of the a and $a+1$ fragment ions from 193 nm UVPD reveals local hydrogen bonding interactions

Lindsay J. Morrison, Jake A. Rosenberg, Jonathan P. Singleton, and Jennifer S. Brodbelt*
Department of Chemistry, University of Texas, Austin, TX 78712

Abstract

Dissociation of proteins and peptides by 193 nm ultraviolet photodissociation (UVPD) has gained momentum in proteomic studies due to the diversity of backbone fragments that are produced and subsequently unrivaled sequence coverage obtained by the approach. The pathways that form the basis for the production of particular ion types are not completely understood. In this study, a statistical approach is used to probe hydrogen atom elimination from $a+1$ radical ions, and different extents of elimination are found to vary as a function of the identity of the C-terminal residue of the a product ions and the presence or absence of hydrogen bonds to the cleaved residue.

Graphical Abstract

Introduction

Advances in high throughput bottom-up proteomics approaches [1, 2] have been driven in part by improvements to established ion activation methods and developments of new ones

[3, 4]. New activation techniques, such as electron-based [5–8] and photoactivation methods [9–11], have resulted in the production of different types and distributions of diagnostic fragment ions than ones generated by conventional collisional activation methods. Ideally, for sequencing applications, fragmentation of peptides would be restricted to one or more of the three bonds of the peptide backbone ($C_{\alpha} - C$, $C - N$, and $N - C_{\alpha}$), leaving sidechains and modifications intact, aside from those that successfully differentiate isobaric leucine and isoleucine (w-type ions). These are the dominant cleavages promoted by collisional, electron-mediated, and photon-based activation for most peptides, and *in silico*-based database search algorithms have proven successful for identification of thousands of peptides in numerous proteomics applications. As a consequence of extensive studies involving examination of large populations of peptides, CID fragmentation is well understood and can be explained by the mobile proton and the pathways in competition models [12, 13], the latter of which largely encompasses the mobile proton model. These studies have uncovered several preferential backbone cleavages, typically observed at proline and aspartic acid, which can be used in a predictive manner to facilitate peptide sequencing [14–17]. For example, upon collisional activation proline-containing peptides exhibit enhanced cleavage of the amide bond N-terminal to the proline residues, a process that is particularly prominent for peptides in higher charge states¹⁴. In contrast, preferential cleavage of the amide bond located C-terminal to acidic residues is exaggerated for peptides in low charge states [15, 16]. The statistical characterization of peptide fragmentation has been pursued in a number of other collisional activation studies [18–21]. The inroads in deciphering the fragmentation of peptides upon collisional activation has spurred significant interest in understanding the underpinnings of other activation processes, including both mechanistic aspects and the factors that contribute to variations in product ion abundances as a function of peptide length, sequence, and charge state.

UVPD is a relatively new activation approach that has rapidly developed in part because it yields unrivaled sequence coverage in large peptides and proteins [22, 23]. Activation by absorption of 193 nm ultraviolet photons promotes cleavage of all three backbone bonds, resulting in the formation of *a* and *x* (cleavage of the $C_{\alpha} - C$ bond), *c* and *z* (cleavage of the $N - C_{\alpha}$ bond), as well as *b* and *y* type ions. In addition to these six ion types, *a*+1, *x*+1, *x*+2, and *y*–2 ions are also prevalent due to various hydrogen atom migrations that can occur via radical-mediated pathways. Despite the promise of 193 nm UVPD, the mechanisms that govern the formation of these ions are not well understood. Consequently, algorithms and models that take advantage of 193 nm UVPD fragmentation are largely absent or under-developed.

The Reilly group has presented several elegant studies that examine UVPD mechanisms using fragmentation and modeling of small model peptides [24–30]. They have shown that activation of dipeptides by 157 nm UV radiation results in the elevation of an electron to a Rydberg orbital, which causes scission of the $C_{\alpha} - C$ bond by a Norrish type I mechanism to generate *a*+1 ion and *x*+1 radical ions depending on the location of an arginine residue [25–27]. Radical elimination via amide hydrogen or β -hydrogen elimination was shown to account for the *a* ions [25–27]. Here, we use lysN digestion to generate a ion series for a large number of peptides and use statistical analysis of the abundance of the *a* and *a*+1 ions to explore hydrogen elimination from *a*+1 ions.

Methods

Materials

Methanol, DTT, iodoacetamide, hemoglobin, beta-lactoglobulin, carbonic anhydrase, myoglobin, lysozyme, albumin, beta-casein, and kappa-casein were obtained from Sigma-Aldrich (St. Louis, MO). Guinea pig adrenocorticotrophic hormone (ACTH 1–39, sequence SYSMEHFRWGKPVGKKRRPVKVYANGAEEESAEAFPLEF) was obtained from Bachem (Torrence, CA). KGTDVLAWIRGCRL was obtained from ABI Scientific. Proteins lacking disulfide bonds were digested by incubation of LysN in 150 mM ammonium bicarbonate. The digestion was carried out using a 1:20 (m/m) ratio of enzyme to protein for two hours at 55 C. The resulting peptides were purified using Pierce (Grand Island, NY) C₁₈ spin columns, and the eluent was concentrated and re-suspended in 5% acetonitrile for LC analysis. An *E. coli* lysate and proteins having one or more disulfide bonds were reduced using dithiothreitol (DTT) and acetylated using iodoacetamide prior to LysN digestion. The peptides included in the tabulation of *a/a*+1 ions are summarized in Supplemental Table 1.

Mass spectrometry

All peptides were analyzed on a Thermo Scientific Orbitrap Fusion mass spectrometer (San Jose, CA) modified by addition of 193 nm excimer laser as described previously³¹. In brief, an optical window was added to the manifold of the vacuum chamber at the back of the dual linear ion trap. To enhance overlap of the laser beam with the ion cloud and thus improve UVPD efficiency, the laser beam was focused using two plano-convex optical lenses. The pulsing of the laser was timed to the normal CID activation period in the high pressure region of the linear ion trap and the laser was fired at 500 Hz, which equates to 2 ms per pulse. Detection in the Orbitrap mass analyzer is on a longer timescale, requiring 100 msec or longer including ion transfer times. For the extensive peptide data acquisition needed for the statistical analysis of hydrogen atom elimination in *a*+1 ions, liquid chromatography of LysN-digested peptides was performed using a Dionex nano-LC system coupled to the Orbitrap Fusion mass spectrometer. Peptides were separated on a 24 cm C18 column packed in-house and were eluted from the column using a 30 minute gradient during which the solvent composition was varied from 15% acetonitrile to 60% acetonitrile in water with 0.1 % formic acid. Ions were generated by ESI via application of 1800 V to a solvent-voltage tee and were transmitted to the linear ion trap for UVPD. UVPD was performed using 2 pulses with 1.5 mJ delivered per pulse. Precursor and product ion spectra were acquired using the Orbitrap mass analyzer; a resolving power of 30k was used for precursor ion spectra and a resolving power of 60k was used for UVPD product ion spectra.

Data Analysis

Peptide identifications were made manually from the single protein digests and using MaxQuant (v. 1.5) for the *E. coli* lysate, using a precursor tolerance of 10 ppm, a fragment tolerance 20 ppm, carbamidomethylation as a fixed modification, no variable modifications, and a false discovery rate of 0.01. The *E. coli* database was obtained from Uniprot. Determination of $\alpha_a ([a]/([a] + [a+1]))$ values for each peptide was undertaken by both manual and automated approaches. In the manual approach, theoretical isotope distributions for the fragment *a* ions were generated using Qualbrowser (Thermo Scientific, v. 2.2) and fit

to the experimental a and $a+1$ ion distributions using a solver function in Microsoft Excel. For the automated approach, mass spectra were converted to the mzML file format using MSconvert and a Python script based on the pymzML library was used to determine the abundances of all peaks belonging to a ion distributions [31]. Using the optimization module of SciPy [32], the experimental abundances were fit to average a and $a+1$ isotopic distributions generated with the fast Fourier transform method described by Palmblad, *et al* [33]. This provided an automated means to detect the peaks of the isotope distribution and correctly assign the abundances of the a and $a+1$ species. The algorithm can be found at: https://github.com/jarosenb/flask_ionratio_V2. The α_a from the two methods was typically within ± 0.05 , and often less. The automated approach was used for the bulk of the fragments and manual interpretation used as needed for cleavages featuring an unusual standard deviation between replicates, the cause of which was typically overlapping isotope distributions from other ions.

Results and Discussion

Distributions of the relative abundances of a_n and a_n+1 ions following fragmentation C-terminal to alanine, glycine, leucine, valine

From analysis of the UVPD mass spectrum of guinea pig adrenocorticotrophic hormone (ACTH 1–39, sequence SYSMEHFRWGKPVGKKRRPVKVYANGAEEESAEAFPLEF), the relative abundances of a_n and a_n+1 ions formed via cleavages of C-C $_{\alpha}$ backbone bonds were observed to vary with the identities of the C-terminal residue of the resulting a ions. Consequently, we applied a statistical approach to investigate the abundances of the a and $a+1$ ions following UVPD fragmentation of a large series of peptides. The Reilly group used hydrogen-deuterium exchange (HDX) reactions of peptides to demonstrate that a ion formation upon 157 nm UVPD occurred via elimination of either the backbone amide hydrogen or the β -hydrogen of the C-terminal amino acid of the original $a+1$ radical ion [25]. Because some residues, such as Gly and Pro, lack either an amide hydrogen or sidechain hydrogens, a dependence was observed in the rates of the two hydrogen atom elimination reactions as a function of amino acid identity [25]. In order to investigate the impact of hydrogen atom loss on the observed distribution of a and $a+1$ fragment ions upon 193 nm UVPD, the portion of the a ions relative to the summed a and $a+1$ abundances was tabulated as a function of the identity of the C-terminal amino acid of the a ion products from a series of 99 peptides obtained from an *E. coli* LysN digest; these peptides are listed in Supporting Information Table 1. Supporting Information Figure S1 shows the UVPD mass spectrum of peptide #1 from this dataset. Inclusion of each peptide into the dataset was contingent on several factors. First, the ability to quantify a and $a+1$ abundances requires good isotopic resolution, a factor dependent on initial ion signal. Consequently, only the most abundant peptides were typically used. In order to avoid peptides with highly mobile protons, peptides having an over-abundance of charge relative to basic sites were not included to avoid problems arising from CID-type fragmentation associated with the presence of one or more mobile protons. Peptides having only a single basic residue were excluded in order to favor longer peptides with multiple basic sites. This resulted in all peptides having a charge state greater than or equal to two. By the same logic, only peptides with more than 13 residues were considered, as characteristic UVPD patterns (e.g. dominant

production of a/x ions rather than conventional b/y ions) are only routinely observed for peptides of this size and larger [34]. Finally, peptides with a basic amino acid at the C-terminus were excluded from the data set due to concerns that N-terminal product ions (such as a ions) would be suppressed to some extent and thereby obfuscate the data.

To succinctly quantify the a and $a+1$ abundances, the relative abundance of the even electron a ion population was termed α_a and is given by: $[a]/([a] + [a+1])$. Because $a+2$ ions are thought to arise via a different pathway,³⁰ one independent of the $a+1 \rightarrow a$ hydrogen atom loss reaction, they were not included in the calculation of the α_a values. To account for the contribution of ^{13}C isotopes of the a ions to the $a + 1$ isotope distribution, the expected isotopic compositions of each fragment ion were calculated and fit by a three-component model to obtain the relative distributions of each of the a , $a+1$, and, if present, $a+2$ ions. It should be noted that the relative abundances of the a and $a+1$ fragments were not observed to change as a function of Orbitrap resolving power, suggesting the hydrogen elimination reaction was complete by the time of ion detection. To ensure that the contributions of the a and $a+1$ ions were properly deconvolved from the overlapping isotope distributions, the monoisotopic peak from a typical doubly protonated (2+) peptide was isolated and fragmented by UVPD. This removes the influence of overlapping ^{13}C isotopes and the relative abundance of the a_{11} and $a_{11}+1$ were subsequently compared to the distribution obtained from mathematical deconvolution of the overlapping isotopic envelopes; this comparison is shown in Supporting Information Figure S2. The percent difference in the calculated α_a from these two methods was determined to be less than 1%. The α_a values were plotted as histograms to visualize the distribution of α_a for a given amino acid. The a and $a+1$ ions compiled for each histogram terminate in one particular amino acid (i.e. a collection of a and $a+1$ ions of different sizes all terminating in Ala). In the top portion of Figure 1, the α_a histograms for Ala, Gly, Leu, and Val are shown as examples. The histograms of the α_a values for Gly and Leu are bimodal, and the Ala and Val histograms have asymmetrical Gaussian profiles, each suggesting that a shoulder component, approximately 0.2–0.3 below the main distribution, is present. This second population, present for each of these amino acids, is puzzling and it is appealing to consider it in regards to the HDX study published by the Reilly group[25]. Hydrogen atom elimination from the $a + 1$ ion was shown to involve the amide hydrogen or the β -hydrogen, and it is conceivable that different torsional conformations of a given amino acid could lead to the favoring of one pathway over the other [25]. However, because a bimodal distribution was observed for glycine, which lacks a sidechain and consequently cannot undergo β -hydrogen elimination, this explanation is inadequate.

Upon assessment of the α_a values for each amino acid and for specific peptide sequences, we observed that a reduction of the α_a value commonly occurred when amino acids capable of hydrogen bonding were found in positions adjacent to the C-C $_{\alpha}$ cleavage site along the peptide backbone. To explore this observation, the α_a values were binned according to whether the amino acids flanking the terminal residue of the targeted a ion could engage in hydrogen bonding (Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, and Thr) or not (Ala, Phe, Gly, Ile, Leu, Met, Pro, Val, Cys, Tyr and Trp). Note that Tyr and Trp have hydrogen bonding sidechains, but cannot easily hydrogen bond to the adjacent residue due to steric constraints. The segregated histograms for Leu, Val, Gly, and Ala based on hydrogen-bonding capability

of the flanking residues are shown in the middle and bottom panels of Figure 1. Analogous histograms for the remaining 16 amino acids are shown in Supporting Information Figure S3. Comparison of each pair of segregated histograms suggests that the bimodal or asymmetric character of the original α_a value histograms (Figure 1) can be largely accounted for by considering the influence of adjacent hydrogen bonding residues. For example, the lower α_a distribution is virtually eliminated in ULeuZ and UValZ sequences by controlling for aliphatic residues in the flanking positions (where U and Z are the flanking residues) and reduced by roughly 50 and 75% in UGlyZ and UAlaZ sequences, respectively. This suggests that hydrogen bonding, either to the amide nitrogen or amide oxygen of the amino acid at which C-C α cleavage occurs, results in a suppression of the hydrogen loss reaction that converts $a+1$ to a ions. On average, this effect reduces the resulting α_a value by 0.2–0.3.

It is interesting to note that many of the cleavages corresponding to the lower α_a mode in Figure 11 were the consequence of AlaAla cleavages (i.e. the flanking amino acid was Ala, a non-hydrogen bonding residue) or AlaPro or ProAla cleavages. It is possible that hydrogen bonding from non-adjacent residues or the peptide backbone accounts for the reduction in the α_a in these instances. Polyalanine has a high propensity to form helical secondary structures in the gas phase[35]; consequently, it is possible that small gas-phase helices form in the aliphatic regions having high Ala content. Similarly, four of the five instances of UGlyZ cleavages in which both U and Z were aliphatic and the observed α_a was less than 0.1 had PheGlyAla or AlaGlyPhe motifs. Although the peptides in the present study were sprayed from denaturing conditions, there is substantial evidence that α -helices are unusually stable in the gas phase, and the enhanced hydrogen bonding associated with the helical structure may result in the observed reduction in α_a values[36–39].

To further explore the notion that hydrogen bonding resulting from gas-phase helices and other structures accounts for many of the reduced α_a instances in the lower α_a distributions of Figure 1i and 11, the difference in the α_a value for a given cleavage C-terminal to an amino acid X with flanking residues U and Z and the larger mean α_a for residue X was termed α_a . In Figure 2, histograms are plotted for UXZ cleavages (for X = Ala, Asp, Phe, Leu, Asn, Gln, Val, and Trp) for a) all cleavages in which U and Z are a) aliphatic, b) all cleavages in which either or both U and Z are Pro, c) cleavages in which two or more residues of the UXZ sequence are Ala, and d) all cleavages in which U and Z are aliphatic (non-hydrogen bonding), excluding the instances shown in b) and c). The distribution in Figure 2a is observed to strongly tail to the left, indicating the presence of a negative α_a population that is consistent with the populations observed at lower α_a values in Figure 1. In Figure 2b, the number of adjacent Pro cleavages is relatively low, but two or three distributions are evident, centered at approximately $\alpha_a = 0.0$, -0.1 , and -0.3 . Significantly, the relative abundance of the more negative α_a distributions seems to be much higher than in Figure 2a, suggesting that nearly half of all cleavages involving an adjacent Pro feature a reduction in α_a associated with hydrogen bonding of the cleaving residue. Unlike most amino acids, the Pro amide bond frequently occupies a *cis* isomerization state [40, 41], which can result in turns in the backbone of peptides and proteins that may facilitate hydrogen bonding from non-adjacent residues. This concept may also explain why glycine only exhibits an approximately 50% reduction in the lower α_a distribution by controlling for adjacent hydrogen bonding residues, shown in Figure 1i, as glycine has the largest

conformational freedom of the 20 natural amino acids and readily engages in turn motifs. Such structures may be present in the denatured peptides studied herein and contribute to non-adjacent residue hydrogen bonding. In Figure 2c, only ones with two or more Ala residues in the UXZ sequence were considered, and as in Figure 2b, two to three distributions seem to contribute to the composite α_a histogram. Interestingly, a well-defined distribution with a mean of $\alpha_a = -0.4$ is present, suggesting that in some cases, reduction in α_a can be greater than 0.3. For these cases, the reduction in α_a may be caused by gas-phase alpha helices, as the non-capping residues in helices are typically both NH and O hydrogen bonded. In Figure 2d, exclusion of the cleavage motifs shown in Figure 2b and 2c is shown to somewhat restore a Gaussian shape to the distribution, although mild tailing to the left is still observed.

Hydrogen atom elimination from a+1 ions

If the stabilization of the amide hydrogen via hydrogen-bonding from side-chains of flanking residues explains the suppression of α_a values for certain peptide sequences, then the α_a values evolving from cleavages without the hydrogen bonding interactions can also be dissected to illuminate the source(s) of hydrogen elimination in the formation of a ions from $a+1$ ions. Presumably, the higher α_a population accounts for these C-C $_{\alpha}$ bond cleavages that are not modulated by hydrogen bonding. The mean α_a value of each residue was therefore evaluated. Because of the overlap in the α_a values (for those a ions with and without flanking hydrogen-bonding residues), Gaussian mixed modeling (GMM) using the Expectation Maximum (EM) algorithm with equal and unequal variances was used to fit Gaussian distributions to the data in order to obtain accurate averages and variances for the higher and lower α_a populations for each amino acid. The Bayesian Information Criterion (BIC) was used as the selection criterion for the fits. The histograms of the data and GMM fits for Ala, Gly, Leu, and Val are shown in Supporting Information Figure S4. Histograms and fits for the 16 remaining amino acids are shown in Supporting Information Figure S5. These means and standard deviations of the Gaussian fits are listed in Table 1.

The majority of the amino acids, shown in Table 1, were best fit to bimodal distributions, however, there were several notable exceptions. Asp and Asn were best fit to unimodal distributions with an average α_a of 0.49 and 0.49; however, separation of the histograms based on flanking hydrogen-bonding residues, shown in Supporting Information Figure S3, indicates the presence of lower α_a distributions for both amino acids, centered approximately 0.2 below the higher α_a distribution. Restriction of the algorithm to bimodal and higher order distributions resulted in the data being best fit to a bimodal distribution. As the bimodal distribution best reflects the histograms separated based on flanking hydrogen bonding residues, Table 1 shows averages and standard deviations for the bimodal fit of Asp and Asn. A similar situation arose for Met, in which a unimodal fit was slightly more favorable than bimodal, and is likely a consequence of too few instances of Met cleavages, which is known to be a limitation for determination of bimodality by GMM, as inspection of the distribution suggests it is bimodal. Consequently, the Met data was also restricted to bimodal and higher order distributions, and Table 1 shows two means and standard deviations. His, Trp, and Pro were also best fit to unimodal distributions. Separation of the His and Trp data based on flanking residues also indicates unimodal distributions, although

the number of instances of cleavage C-terminal to both residues may be too low to clearly define two distributions, if present. In contrast, Pro features more complexity, and it is possible that as many as four distributions are present. Again, however, the statistics are too limited to make any strong conclusions. It bears noting that Pro commonly features *cis* as well as *trans* isomerization of the amide bond, and the conformational consequences of these two isomerization states may contribute to the broad, ill-defined distributions. Iso and Thr were both best fit to trimodal distributions.

Based on hydrogen-deuterium exchange of several model peptides, the Reilly group proposed that hydrogen elimination may involve both the amide and β -hydrogens of the terminal amino acid of $a+1$ ion fragments [25]. A similar HDX experiment was performed on singly protonated YFMRF, in which UVPD was performed following monoisotopic selection of the normal and fully deuterated peptide and the resulting losses from the a_4+1 ion monitored. These spectra are shown in Supporting Information Figure S6 and show the loss of both deuterium and hydrogen from the deuterated a_4+1 , suggesting both amide hydrogen and β -hydrogen elimination are active pathways in 193 nm UVPD. As adapted from Reilly's work [25], Scheme 1 shows possible pathways for formation of a ions illustrated for a representative valine-containing peptide via elimination of a) the β -hydrogen and b) the amide hydrogen. A third route is shown in Scheme 1c in which a hydrogen atom is eliminated from the terminal γ -C of the Val side-chain. This pathway is inherently more complex and requires a simultaneous hydrogen transfer of the β -hydrogen to the alpha carbon. In comparison of the mean α_a value resulting from cleavages C-terminal to Ala, Gly, Val, it is useful to consider the factors that contribute to and consequences of the different elimination pathways shown in Scheme 1. Elimination of an amide hydrogen is expected to be independent of the sidechain identity. In contrast, elimination of a β -hydrogen or terminal hydrogen (γ in the case of Val) from a given amino acid side-chain is expected to be sidechain dependent and vary as a function of the orientation of the individual hydrogen(s) and the number of hydrogen atoms available for elimination. Hence, sidechains with a larger number of β -hydrogens would be expected to exhibit a higher probability for elimination of a β -hydrogen and sidechains with more branching and consequently a larger number of terminal hydrogen atoms would be expected to exhibit a higher probability for terminal hydrogen elimination.

In order to illuminate the sources of hydrogen elimination in $a+1$ ions in 193 nm UVPD, the α_a values of Gly, Val, Leu, and Ala were examined. These four residues are aliphatic but feature differing numbers of β -hydrogens: 0, 1, 2, and 3, respectively. From inspection of Table 1 and Figure 1g, the average of the higher α_a population of Gly is clearly both non-zero and significantly lower than the higher α_a population of Val, Leu, or Ala. As Gly lacks a sidechain, the non-zero α_a suggests that amide hydrogen elimination is an active pathway in 193 nm UVPD, as in 157 nm UVPD. This pathway is likely to occur for all $a+1$ ions, regardless of the identity of the terminal residue. Interestingly, the Reilly group has shown a and $a+1$ ion distributions for cleavages C-terminal to Gly [25,42]; notably the α_a for the a_5 ion from RGGQGG was observed to be approximately 0.3, which is in excellent agreement with our results. Other cleavages, such as C-terminal to Phe, Glu, and others also seemed to be relatively consistent with the results from 193 nm UVPD [25]. The difference between the average upper α_a population of glycine and the average upper α_a of Val, Leu, and Ala is

considerable at 0.3–0.4, and indicates that amide hydrogen elimination is not the only route for conversion of $a+1$ ions to a ions. The average of the higher α_a populations for the Ala, Leu, and Val histograms are much more similar, 0.63, 0.70, and 0.75, respectively. An ANOVA analysis was carried out using the means, variances, and population weights from the GMM fits. The determined F-value of 28.85 is significantly greater than the 99% confidence critical value of 4.68, suggesting that the three populations are statistically different. The α_a for Val, Leu, and Ala decreases as the number of β -hydrogens available for elimination increases. For example, Ala, which features the largest number of β -hydrogens, has the smallest α_a , while Val features the largest α_a . This is in direct contradiction of what would be expected for β -hydrogen elimination and suggests one or more additional hydrogen elimination pathways may be present.

Akin to the observation from the Gly dataset that amide hydrogen elimination must be active in 193 nm UVPD, examination of dataset of Ala, which has a sidechain featuring exclusively β -hydrogens, is instructive. The mean of the higher α_a distribution of the Gly data is 0.34, whereas the mean of the higher α_a distribution of the Ala histogram is 0.63. It is clear from this difference that β -hydrogen elimination must occur and accounts for roughly half of the observed hydrogen elimination in Ala-terminated a ions. However, the previous comparison between the higher α_a distributions of Ala, Leu, and Val suggests a third hydrogen transfer pathway also contributes to formation of a ions from $a+1$ ions. Both Val and Leu have branched sidechains with six terminal hydrogens, respectively. Hence, the number of terminal hydrogens cannot explain the difference in the means of the α_a values. Hydrogen atom elimination by the pathway shown in Scheme 1c requires a 5-membered ring transition state for Val-terminated fragments and a 6-membered ring transition state for Leu-terminated fragments. Given that 5-membered rings feature faster formation kinetics, it is possible that the higher mean α_a found for Val is a consequence of the kinetics of a pathway such as the one shown in Scheme 1c. Note that modeling is needed to confirm the details of this pathway. Interestingly, although this terminal hydrogen atom elimination pathway is relatively minor, contributing only about 15% to the total hydrogen atom elimination from Val sidechains, it has an important consequence. From inspection of Table 1, it is evident that relatively unique α_a values are associated with each of the different amino acids. This is illustrated more clearly in bar graph format in Supporting Information Figure S7 (which shows each α_a value for each amino acid) This semi-uniqueness of the α_a may provide an avenue to pursue use of α_a values as a validation strategy for *de novo* sequencing of 193 nm UVPD spectra. As there are currently no available *de novo* sequencing algorithms that are designed to handle the complexity of UVPD spectra, this may offer an opportunity to exploit the rich information that can be obtained from 193 nm UVPD while simultaneously taking advantage of the database independent properties of a *de novo* approach.

To verify that a third elimination pathway contributes to a ion formation from radical elimination of $a+1$ ions in 193 nm UVPD, an ANOVA analysis was performed on the higher α_a population of Asp, Asn, and Leu, all of which feature identical branching structures but different terminal functional groups. The mean α_a of the higher Asn and Asp populations are 0.60 and 0.59, respectively, and were not found to be significantly different from each other based on a student's *t*-test. However, an ANOVA analysis of the three populations (Leu

(0.70), Asn (0.60), and Asp (0.59)) returned an F-value of 19.35, which is greater than the 99% confidence critical value of 4.70, suggesting that the average upper α_a of Leu is significantly different from the average higher α_a of Asp and Asn. Asp and Asn have similar terminal functional groups (amide vs. a carboxylic acid) in which the one to two terminal hydrogen atoms present cannot undergo transfer by the pathway shown in Scheme 1c because Asp and Asn lack a γ -hydrogen. In contrast, the terminal methyl groups of the Leu sidechain feature six hydrogens, any of which can be transferred by a Scheme 1c-type pathway because of the available γ -hydrogen. The increased average α_a of Leu relative to Asp and Asn can thus be accounted for by consideration of a third hydrogen elimination pathway from the terminal end of aliphatic residues, and supports a route such as the one shown in Scheme 1c. Note that additional modeling is needed to fully illuminate the source of the second hydrogen atom transfer. It is interesting to note that the mean α_a values of the Asn and Asp datasets (higher α_a populations) are less than that of the Ala population by approximately 0.05, suggesting that the third β -hydrogen atom of the Ala sidechain contributes approximately 0.05 to the observed hydrogen elimination.

It is interesting to consider the α_a values in Table 1, particularly the higher α_a distributions associated with the cleaving residue not being hydrogen bonded, in the context of the stability of the a and $a+1$ radical. In particular, the α_a values for Tyr and Trp are higher than almost every other residue (0.86 and 0.84, respectively). Both of these residues are capable of amide and β -hydrogen elimination but cannot eliminate a terminal hydrogen due to the aromaticity of the sidechain. Loss of a β -hydrogen, however, results in the formation of an extended conjugation that likely stabilizes the a ion. This stabilization possibly drives the production of the a ion in fragments terminated by these two residues. In contrast, phenylalanine, which is also aromatic, features an α_a value that is significantly lower (0.63). One possible explanation for this is that the phenylalanine radical is more stable and that the stability of the radical results in less favorable formation of the a ion product compared to Tyr and Trp.

Shifts in α_a as a function of amide hydrogen bonding

In Figure 1, the abundance of the lower mode of the bimodal α_a distributions from Ala, Val, and Leu was shown to be significantly reduced by considering only cleavages in which the adjacent residues were aliphatic (non-hydrogen bonding). We therefore postulated that hydrogen bonding to either the amide oxygen or amide nitrogen adjacent to the $C_\alpha - C$ cleavage site contributes to this reduction. From statistical analysis of the average higher α_a values of Ala, Leu, Val, Gly, Asn, and Asp, hydrogen atom elimination is proposed to occur from the amide hydrogen, the β -hydrogen, and the terminal hydrogens of some amino acid sidechains. It is reasonable that hydrogen bonding to the amide nitrogen could prevent amide hydrogen elimination to some extent, shifting the average α_a value. However, it is also possible that hydrogen bonding to the amide oxygen has an effect on the observed α_a value. To explore whether amide nitrogen or amide oxygen (or both) hydrogen bonding accounts for shifts in α_a values, $\Delta\alpha_a$, the difference between the observed α_a value for a given cleavage C-terminal to X in a sequence UXZ and the average α_a value associated with the amino acid X, was calculated for all cleavages in which U and Z were Ser and Thr and X was Ala, Asp, Phe, Leu, Asn, Gln, and Val. These cleavages were selected because they

featured bimodal α_a distributions with relatively equal modes. Ser and Thr were chosen as the adjacent residues because they have short sidechains that restrict the hydrogen bonding interactions in which they can engage. When Ser precedes X, for example, it naturally forms a stable hydrogen bond to the amide nitrogen. Hydrogen bonding from a preceding Ser is also favorable to the amide oxygen of residue X. In contrast, hydrogen bonding to the amide oxygen from a succeeding Ser is possible via the formation of a seven-membered ring, however, hydrogen bonding to the amide nitrogen is not generally possible due to the orientation of the peptide backbone which places the amide hydrogen 180° from the Ser OH group. Thus, preceding Ser/Thr residues can hydrogen bond to the amide nitrogen or amide oxygen whereas succeeding Ser/Thr residues can only hydrogen bond to the amide oxygen.

In Figure 3, histograms are shown for UXZ cleavages in which a Ser or Thr precedes the cleavage site (a) and succeeds the cleavage site (b). Note that the other adjacent residue (U or Z) was only permitted to be aliphatic. Stick structures depicting the different hydrogen bonding interactions of a preceding and succeeding Ser are shown in Figure 3e–h. Similar ball and stick depictions of these hydrogen bonding interactions are shown in Supporting Information Figure S8. In Figure 3c, histograms are shown for cleavages with hydrogen bonding residues in both the U and Z position in order to determine if hydrogen bonding at both the amide NH and O has a compounding effect on α_a . Because our intention was to control for cleavages involving two hydrogen bonds, Gln, Asn, Glu, and Asp, were included as adjacent residues (occupying the U and Z positions) in addition to Ser and Thr, which also provided a larger pool of cleavages and enhanced the statistical confidence of the comparison. Note that these residues were excluded from the histograms in Figure 3a and 3b because the amide and acid side-chains can conceivably hydrogen bond to the amide NH and O sites simultaneously. Finally, in Figure 3d, the histogram from Figure 2d is reprinted and shown as a control wherein only aliphatic residues were permitted in the positions adjacent to the cleavage site. Figure 3a and 3b feature clear bimodal distributions. Despite the relatively few instances, both datasets were well fit to bimodal distributions by GMM and means of $\alpha_a = -0.23$ and 0.03 were found for cleavages with a preceding Ser/Thr and means of $\alpha_a = -0.32$ and -0.01 were found for cleavages with a succeeding Ser/Thr. A t-test was performed on the higher α_a (-0.01 and 0.03) and lower α_a (-0.23 and -0.32) populations. While the higher α_a populations were not found to be significantly different, the means of the lower α_a populations were found to be statistically different at the 99% confidence level, suggesting that controlling for a preceding vs. succeeding Ser/Thr has a significant influence on the extent to which α_a is reduced for a given cleavage. Because a succeeding Ser/Thr can *only* hydrogen bond to the amide oxygen of the cleaving residue, it is clear that amide oxygen hydrogen bonding is at least partially responsible for the reduction in α_a . It is also noteworthy that the α_a population centered at -0.32 is significantly more abundant than the small population observed in the tailing distribution in Figure 3d.

Examination of the distribution in Figure 3c is useful to further elucidate the effects of hydrogen bonding on α_a . It is clear from Figure 3b that amide oxygen hydrogen bonding results in a reduction of α_a by approximately 0.3, however the lower distribution in Figure 3a features a mean of only $\alpha_a = -0.2$. Because a preceding Ser/Thr can hydrogen bond to either the amide NH or amide O, it is possible that this distribution represents an average

α_a , resulting from a -0.3 shift for all sequences with amide oxygen hydrogen bonding and some other shift associated with amide nitrogen hydrogen bonding. This assertion is supported by histogram shown in Figure 3c, which features a broad distribution spanning from $\alpha_a = -0.5$ to 0.1 . Inspection of this distribution and comparison to the widths of the distributions in Figure 3a and 3b suggests at least two or three distributions are present. GMM was unfortunately unsuccessful at fitting the data (it suggested a unimodal distribution), which is likely a consequence of too few instances and too much overlap of the apparent distributions. For this reason, Gaussian envelopes were manually fit to the distribution. Supporting Information Figure 9, 10, and 11 show histograms for four, three, and two component fits. The sum of the variance between the fit and the experimental ($\Sigma\chi^2$) was used to parameterize the fits and the $\Sigma\chi^2$ is reported for each plot. The four and three component fits by far outscored the two component fit ($\Sigma\chi^2$ of 27.0, 29.4, and 39.24, respectively), but were relatively similar to one another. However, because the four component model best fit the data, the means and variances associated with that fit were used. The means associated with the four component model are also the most consistent with the means found for the histogram data in Figure 3a and 3b. Using the four component model, means were found at $\alpha_a = 0.025, -0.095, -0.31$, and -0.45 . The populations at ≈ 0.0 and -0.3 are in agreement with the populations means observed for cleavages with a succeeding Ser/Thr. It is possible that the remaining two means, at $\alpha_a \approx -0.1$ and -0.45 , are the result of amide NH hydrogen bonding and combined amide NH and O hydrogen bonding, respectively. From Figure 3b, amide oxygen hydrogen bonding is predicted to result in a -0.32 shift in α_a ; if amide NH hydrogen bonding results in a reduction of α_a by ≈ 0.1 , the combined effect of amide NH and O hydrogen bonding would be expected to be approximately -0.42 , which is consistent with the -0.45 population only observed for cleavages with two adjacent hydrogen bonding residues.

Examination of Hydrogen Bonding Effect

To verify in a single peptide that hydrogen bonding can result in a reduction in the observed α_a value for a given cleavage, a model peptide from the database KGTDVQAWIRGCRRL was examined in further detail (Figure 4). The a_5 and a_7 ions were observed to feature α_a values consistent with the lower modes of the Val and Ala histograms, 0.36 and 0.19 , respectively. We postulated that Gln6 could be hydrogen bonding to the amide oxygen of either Val5 or Ala7, causing the apparent reduction in α_a . Hence, a modified peptide, KGTDVLAWIRGCRRL (Q6L) was synthesized and the α_a values determined. The α_a values for each cleavage of WT KGTDVQAWIRGCRRL and the Q6L variant are shown in Figure 4. Consistent with expectations, substitution of Leu for Gln6 resulted in an increase (0.15) in the α_a of Val5, suggesting that Gln6 is engaged in a hydrogen bond with the amide bond of this residue. This is in direct contrast with the behavior of the C-terminal residues of the two peptides, which exhibit near identical α_a values. Interestingly, Thr3 and Asp4 feature somewhat different α_a values in the wildtype and Q6L peptides; it is possible that substitution of Leu resulted in structural changes at the N-terminus of the peptide. The α_a values corresponding to the a_6 product ions were also observed to be different for both peptides, which is not unexpected given that the a_6/a_6+1 ions are terminated by different residues (Q versus L). These generally low α_a values for the a_6 ions support that the sixth residue is hydrogen bonded in both peptides.

Conclusions

The extent of atom elimination from $a+1$ ions to generate a ions, presented as α_a distributions, was statistically evaluated from 99 peptides from proteolytic digestion using LysN. Examination of the α_a distributions suggests that a bimodality is present in the data, with one population representing more extensive hydrogen atom elimination and a second representing suppressed hydrogen elimination. Restriction of the amino acids adjacent to a given cleavage site to those with aliphatic side chains largely reduces and in some cases virtually eliminates the lower α_a population associated with less extensive hydrogen atom elimination from the $a+1$ radical ion. Restriction of adjacent residues to Ser/Thr further suggests that both amide oxygen and amide nitrogen hydrogen bonding account for reduction observed in α_a values, with amide oxygen hydrogen bonding accounting for a shift of -0.3 in the α_a value and amide NH hydrogen bonding accounting for a shift of -0.1 . It may be possible to use these shifts to probe the backbone hydrogen bonding networks in native proteins. Further assessment of the α_a distributions also suggests a third hydrogen elimination pathway is active that results in mean α_a values that are unique to the identity of the C-terminal amino acid of each a ion product. Moreover, the semi-unique mean α_a values associated with the C-terminal amino acid of each a product ion may provide a means to assist *de novo* sequencing using 193 nm UVPD, an appealing approach that is inherently challenging due to the complex nature of the MS/MS spectra. Efforts are thus being undertaken to make use of the α_a values to validate *de novo* methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding from the NSF (CHE-1402753), the Welch Foundation (F-1155) and NIH 1K12GM102745 (fellowship to LM) is acknowledged.

References

1. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews*. 2013; 113:2343–2394. [PubMed: 23438204]
2. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD. Mass spectrometry-based proteomics: existing capabilities and future directions. *Chem Soc Rev*. 2012; 41:3912–3928. [PubMed: 22498958]
3. Jones AW, Cooper HJ. Dissociation techniques in mass spectrometry-based proteomics. *Analyst*. 2011; 136:3419–3429. [PubMed: 21698312]
4. Brodbelt JS. Ion Activation Methods for Peptides and Proteins. *Analytical Chemistry*. 2016; 88:30–51. [PubMed: 26630359]
5. Cooper HJ, Hakansson K, Marshall AG. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom Rev*. 2005; 24:201–222. [PubMed: 15389856]
6. Zhurov KO, Fornelli L, Wodrich MD, Laskay UA, Tsybin YO. Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem Soc Rev*. 2013; 42:5014–5030. [PubMed: 23450212]
7. Sarbu M, Ghiulai R, Zamfir A. Recent developments and applications of electron transfer dissociation mass spectrometry in proteomics. *Amino Acids*. 2014; 46:1625–1634. [PubMed: 24687149]

8. Kim MS, Pandey A. Electron transfer dissociation mass spectrometry in proteomics. *Proteomics*. 2012; 12:530–542. [PubMed: 22246976]
9. Brodbelt JS. Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem Soc Rev*. 2014; 43:2757–2783. [PubMed: 24481009]
10. Reilly JP. Ultraviolet photofragmentation of biomolecular ions. *Mass Spectrom Rev*. 2009; 28:425–447. [PubMed: 19241462]
11. Ly T, Julian RR. Ultraviolet Photodissociation: Developments towards Applications for Mass-Spectrometry-Based Proteomics. *Angew Chem, Int Ed*. 2009; 48:7130–7137.
12. Wysocki VH, Tsaprailis G, Smith LL, Brei LA. Mobile and localized protons: a framework for understanding peptide dissociation. *J Mass Spectrom*. 2000; 35:1399–1406. [PubMed: 11180630]
13. Paizs B, Suhai S. Fragmentation Pathways of Protonated Peptides. *Mass Spectrometry Reviews*. 2005:508–548. [PubMed: 15389847]
14. Brei LA, Tabb DL, Yates JR 3rd, Wysocki VH. Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra. *Anal Chem*. 2003; 75:1963–1971. [PubMed: 12720328]
15. Gu C, Tsaprailis G, Brei L, Wysocki VH. Selective Gas-Phase Cleavage at the Peptide Bond C-Terminal to Aspartic Acid in Fixed-Charge Derivatives of Asp-Containing Peptides. *Analytical Chemistry*. 2000:5804–5813. [PubMed: 11128940]
16. Tsaprailis G, Somogyi Á, Nikolaev EN, Wysocki VH. Refining the model for selective cleavage at acidic residues in arginine-containing protonated peptides. *International Journal of Mass Spectrometry*. 2000; 195–196:467–479.
17. Li W, Ji L, Goya J, Tan G, Wysocki VH. SQID: An Intensity-Incorporated Protein Identification Algorithm for Tandem Mass Spectrometry. *Journal of Proteome Research*. 2011; 10:1593–1602. [PubMed: 21204564]
18. Tabb DL, Smith LL, Brei LA, Wysocki VH, Lin D, Yates JR 3rd. Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides. *Anal Chem*. 2003; 75:1155–1163. [PubMed: 12641236]
19. Huang Y, Triscari JM, Tseng GC, Pasa-Tolic L, Lipton MS, Smith RD, et al. Statistical Characterization of the Charge State and Residue Dependence of Low-Energy CID Peptide Dissociation Patterns. *Anal Chem*. 2005; 77:5800–5813. [PubMed: 16159109]
20. Michalski A, Neuhauser N, Cox J, Mann M. A systematic investigation into the nature of tryptic HCD spectra. *J Proteome Res*. 2012; 11:5479–5491. [PubMed: 22998608]
21. Shao C, Zhang Y, Sun W. Statistical characterization of HCD fragmentation patterns of tryptic peptides on an LTQ Orbitrap Velos mass spectrometer. *Journal of proteomics*. 2014; 109:26–37. [PubMed: 24981973]
22. Shaw JB, Li W, Holden DD, Zhang Y, Griep-Raming J, Fellers RT, et al. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *Journal of the American Chemical Society*. 2013; 135:12646–12651. [PubMed: 23697802]
23. Madsen JA, Boutz DR, Brodbelt JS. Ultrafast Ultraviolet Photodissociation at 193 nm and its Applicability to Proteomic Workflows. *Journal of proteome research*. 2010; 9:4205–4214. [PubMed: 20578723]
24. Cui W, Thompson MS, Reilly JP. Pathways of Peptide Ion Fragmentation Induced by Vacuum Ultraviolet Light. *Journal of the American Society for Mass Spectrometry*. 2005; 16:1384–1398. [PubMed: 15979330]
25. Zhang L, Cui W, Thompson MS, Reilly JP. Structures of α -Type Ions Formed in the 157 nm Photodissociation of Singly-Charged Peptide Ions. *Journal of the American Society for Mass Spectrometry*. 2006; 17:1315–1321. [PubMed: 16857381]
26. Thompson MS, Cui W, Reilly JP. Factors That Impact the Vacuum Ultraviolet Photofragmentation of Peptide Ions. *Journal of the American Society for Mass Spectrometry*. 2007; 18:1439–1452. [PubMed: 17543535]
27. Parthasarathi R, He Y, Reilly JP, Raghavachari K. New Insights into the Vacuum UV Photodissociation of Peptides. *Journal of the American Chemical Society*. 2010; 132:1606–1610. [PubMed: 20078132]

28. He Y, Webber N, Reilly JP. 157 nm photodissociation of a complete set of dipeptide ions containing C-terminal arginine. *J Am Soc Mass Spectrom.* 2013; 24:675–683. [PubMed: 23378257]
29. Webber N, He Y, Reilly JP. 157 nm photodissociation of dipeptide ions containing N-terminal arginine. *J Am Soc Mass Spectrom.* 2014; 25:196–203. [PubMed: 24310819]
30. Kim TY, Valentine SJ, Clemmer DE, Reilly JP. Gas-phase conformation-specific photofragmentation of proline-containing peptide ions. *J Am Soc Mass Spectrom.* 2010; 21:1455–1465. [PubMed: 20483641]
31. Klein DR, Holden DD, Brodbelt JS. Shotgun Analysis of Rough-Type Lipopolysaccharides Using Ultraviolet Photodissociation Mass Spectrometry. *Anal Chem.* 2016; 88:1044–1051. [PubMed: 26616388]
32. Bald T, Barth J, Niehues A, Specht M, Hippler M, Fufezan C. pymzML—Python module for high-throughput bioinformatics on mass spectrometry data. *Bioinformatics.* 2012; 28:1052–1053. [PubMed: 22302572]
33. Palmblad M, Buijs J, Håkansson P. Automatic analysis of hydrogen/deuterium exchange mass spectra of peptides and proteins using calculations of isotopic distributions. *Journal of the American Society for Mass Spectrometry.* 2001; 12:1153–1162. [PubMed: 11720389]
34. Robotham SA, Brodbelt JS. Comparison of Ultraviolet Photodissociation and Collision Induced Dissociation of Adrenocorticotrophic Hormone Peptides. *J Am Soc Mass Spectrom.* 2015; 26:1570–1579. [PubMed: 26122515]
35. Kohtani M, Jones TC, Schneider JE, Jarrold MF. Extreme Stability of an Unsolvated α -Helix. *J Am Chem Soc.* 2004; 126:7420–7421. [PubMed: 15198571]
36. Hudgins RR, Ratner MA, Jarrold MF. Design of Helices That Are Stable in Vacuo. *Journal of the American Chemical Society.* 1998; 120:12974–12975.
37. Hudgins RR, Mao Y, Ratner MA, Jarrold MF. Conformations of GlynH⁺ and AlanH⁺ Peptides in the Gas Phase. *Biophysical Journal.* 1999; 76:1591–1597. [PubMed: 10049339]
38. Kohtani M, Schneider JE, Jones TC, Jarrold MF. The Mobile Proton in Polyalanine Peptides. *J Am Chem Soc.* 2004; 126:16981–16987. [PubMed: 15612736]
39. Morrison LJ, Wysocki VH. Gas-phase helical peptides mimic solution-phase behavior. *J Am Chem Soc.* 2014; 136:14173–14183. [PubMed: 25203898]
40. Grathwohl C, Wüthrich K. Nmr studies of the rates of proline cis–trans isomerization in oligopeptides. *Biopolymers.* 1981; 20:2623–2633.
41. Maia HL, Orrell KG, Rydon HN. Evaluation of the thermodynamic parameters for the interconversion of the conformational isomers of two N-acylprolines by nuclear magnetic resonance spectroscopy. *J Chem Soc D.* 1971:1209–1210.
42. Zhang L, Reilly JP. Radical-Driven Dissociation of Odd-Electron Peptide Radical Ions Produced in 157 nm Photodissociation. *J Am Soc Mass Spectrom.* 2009; 20:1378–1390. [PubMed: 19477139]

Figure 1.

Histograms of the α_a values for UXZ cleavages in which X is Leu in a), b) and c), Val in d), e) and f), Gly in g), h), and i), and Ala in j), k), and l). α_a is defined as the portion of a ions: $[a]/([a] + [a+1])$. In a), d), g), and j) all UXZ cleavages are included, in b), e), h) and k) either or both U and Z are capable of hydrogen bonding, and in c), f), i), and l) U and Z are both aliphatic (and thus not capable of hydrogen bonding). All $a/a+1$ ions in a given histogram terminate in residue X. U and Z are the residues that flank the X residue. Duplicate measurements of α_a were averaged using a single charge state of the precursor peptide.

Figure 2.

Histograms of α_a for cleavages C-terminal to X in UXZ sequences in which a) U and Z are both aliphatic, b) U and/or Z is a Pro, c) the three amino acid sequence is comprised of two or more Ala residues, and d) U and Z are all aliphatic residues excluding Pro and instances of two or more Ala residues.

Figure 3.

Histograms of UXZ cleavages wherein U is Ser/Thr in a) and Z is Ser/Thr in b). In c), both U and Z are hydrogen bonding, including Glu, Asp, Gln, Asn, Thr, and Ser. In d), only cleavages in which both X and Z are aliphatic are included in the histogram data and instances of an adjacent Pro and sequences with 2 or more Ala residues were excluded. Stick structures depicting the hydrogen bonding interactions are shown for a preceding (e and f) and succeeding Ser (g and h) in e – h.

Figure 4. α_a values are shown for each backbone cleavage (generating a ions) for KGTDVQAWIRGCRL, shown in blue, and the Q6L variant, shown in orange.

Scheme 1.

Possible pathways for hydrogen elimination from an $a+1$ ion terminated with a Val residue. In a), the β -hydrogen is eliminated, in b) the amide hydrogen is eliminated, and in c) the γ -hydrogen is eliminated. The first two pathways are adapted from reference 25 (Reilly).

Table 1

Average and standard deviations of the α_a values of each distribution of each amino acid determined from GMM fits to the α_a distributions (using all cleavages) of each amino acid. The number of instances included in each population is listed in parentheses after the standard deviation. The amino acids are listed in alphabetical order based on their one letter code.

Amino Acid	Average α_a	Std Dev. (N)	Amino Acid	Average α_a	Std. Dev. (N)
Ala ₁	0.30	0.13 (78)	Met ₁	0.39	0.12 (19)
Ala ₂	0.63	0.13 (114)	Met ₂	0.79	0.12 (6)
Cys ₁	0.98	0.06 (8)	Asn ₁	0.33	0.11 (28)
Cys ₂	0.72	0.06 (21)	Asn ₂	0.60	0.11 (41)
Asp ₁	0.33	0.14 (52)	Pro	0.40	0.23 (89)
Asp ₂	0.59	0.14 (87)	Gln ₁	0.25	0.12 (33)
Glu ₁	0.41	0.13 (80)	Gln ₂	0.57	0.12 (48)
Glu ₂	0.81	0.13 (72)	Arg ₁	0.37	0.13 (28)
Phe ₁	0.33	0.10 (44)	Arg ₂	0.76	0.13 (31)
Phe ₂	0.63	0.10 (49)	Ser ₁	0.28	0.16 (16)
Gly ₁	0.05	0.04 (40)	Ser ₂	0.75	0.16 (54)
Gly ₂	0.34	0.13 (58)	Thr ₁	0.27	0.08 (6)
His	0.66	0.14 (40)	Thr ₂	0.76	0.16 (64)
Ile ₁	0.25	0.05 (21)	Thr ₃	0.99	0.02 (28)
Ile ₂	0.47	0.04 (63)	Val ₁	0.46	0.11 (56)
Ile ₃	0.72	0.11 (16)	Val ₂	0.75	0.11 (87)
Lys ₁	0.09	0.06 (30)	Trp	0.84	0.13 (23)
Lys ₂	0.33	0.14 (26)	Tyr ₁	0.55	0.11 (35)
Leu ₁	0.41	0.11 (93)	Tyr ₂	0.86	0.11 (29)
Leu ₂	0.70	0.11 (90)			