International Journal of Mass Spectrometry 308 (2011) 89–97

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

A comprehensive investigation of proline fragmentation behavior in low-energy collision-induced dissociation peptide mass spectra

Nai-ping Dong, Liang-xiao Zhang, Yi-zeng Liang*

College of Chemistry and Engineering, Central South University, Changsha 410083, PR China

ARTICLE INFO

Article history:
Received 24 March 2011
Received in revised form 1 August 2011
Accepted 1 August 2011
Available online 10 August 2011

Keywords:
Proline effect
Protonated peptides
Proline fragmentation map
Specific cleavage effect
Statistical analysis

ABSTRACT

An investigation of more than 130 000 tandem mass spectra of proline-containing peptides extracted from Human and *E*coli peptide libraries in NIST Libraries (http://peptide.nist.gov/) is presented. In this study, fragmentation maps are drawn to show the fragmentation behavior of proline (Pro) in protonated peptides, taking the factors affecting the cleavage N-terminal to Pro as points of the map. In order to quantitatively characterize the fragmentation behavior of proline, probability of occurring selective cleavage at N-terminal side of Pro for each point is calculated. From the fragmentation maps, selective cleavage at N-terminal side of Pro is suppressed when all protons are sequestered by Args except at Asp-Pro and Glu-Pro. When protons are mobile, cleavage at N-terminal side of Pro is determined by pairwise cleavage Xxx-Pro and positions of proline in peptides. For triply and quadruply charged peptides, the Coulombic repulsion between protons affects the fragmentation of peptides and subsequently suppresses the cleavage at N-terminal side of Pro. Other fragmentation pathways influencing the fragmentation such as aspartic acid effect and y_{N-2}/b_2 pathway are also found. Investigation of multiple proline containing peptides shows similar influential factors and competition between multiple prolines.

1. Introduction

The development of mass spectrometry has greatly accelerated the study of proteins in recent two decades, making the characterization (i.e., identification, quantification) of thousands of proteins possible [1]. In this field, the most popular strategies using mass spectrometry for protein identification/quantification is bottom-up procedure. In this procedure, protein samples are first digested into peptides by proteases (e.g., enzyme trypsin), and then eluted by liquid chromatography to separate the resulted peptide mixtures. The fractions obtained at each retention time of chromatogram are ionized by soft ionization methods (such as ESI or MALDI) and detected by mass spectrometer to form MS^1 spectra, which mainly represent peptide molecular ion information. In order to obtain more information of peptide sequences, peptide ions are then selected and fragmented into ion pieces (tandem mass spectra). In this fragmentation process, low-energy collision induced dissociation (CID) method is often used. One of the advantages of low-energy CID spectra is that peptides are most frequently cleaved at amide bonds, producing sequence ion series named *y* (when protons are retained at C-terminal fragment), *b* (when protons are retained at N-terminal fragments) and *a* (CO loss from *b* ions) [2,3].

When tandem mass spectra are obtained, a database search strategy is used. In this strategy, protein database search tools search protein sequence databases and score every match between experimental spectrum and simulated spectrum of peptide in database to identify peptide sequences. For its simplicity and speed, this strategy becomes the most frequently used method to assign tandem mass spectra to peptide sequences [4]. However, one of the vital drawbacks of this database search procedure is that database search algorithms only consider *m*/*z* values leaving out of ion intensities, thus generate inaccurate predicted mass spectra for scoring the similarities to the experimental spectra. As a result, there exist a large number of false positives in search results. Furthermore, the errors would be enlarged when these ambiguous results are used to conjecture corresponding proteins [5]. Therefore, it is indispensable to understand the fragmentation pathways of peptides to confidently construct theoretical spectra to improve the accuracy of search algorithms.

To date, much knowledge has been gained in fragmentation of peptides and several pathways of the fragmentation are proposed. For low-energy fragmentation of peptides in gas phase, a large number of works have provided indubitable evidences that the mobile proton(s) in a protonated peptide play a critical role in dissociation of amide bonds. According to these discoveries, a model called 'mobile proton model' [6] was hypothesized. Mobile

* Corresponding author. Tel.: +86 731 88830824; fax: +86 731 88830831.
E-mail address: yizeng_liang@263.net (Y.-z. Liang).

1387-3806/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.ijms.2011.08.005

proton model classifies all protonated peptides into proton non-mobile and mobile categories according to the mobility of protons in peptides. The former is generally defined as number of Args exceeds or equals to number of protons that all protons are captured by side chains of Args and need significant energy to move to other sites of the peptide. Whereas when number of protons exceeds number of Args, protons could easily migrate to fragmentation sites from their initial locations (N-terminus, histidine or lysine) to induce the cleavage. Thus markedly different fragmentation behaviors are presented for peptides in these two categories. For its success in explaining why cleavage happens (such as histidine effect [7], aspartic acid effect [8]), mobile proton model becomes the most popular peptide fragmentation interpretation model. However, for predicting peptide fragment ion mass spectra, not only the information of why the cleavage happens but also how abundant the ions formed is required, that is, the quantitative characterization of fragment ion intensity should be provided. Since mobile proton model is a qualitative framework to predict a spectrum using appropriate rules [9], it fails in predicting quantitative intensity of a spectrum. Being aware of this limitation, a framework called 'pathways in competition' (PIC) was developed recently [10], in which several peptide fragmentation pathways carried out by a large number of previous works were summarized and considered as probability cases. It is reasonable to suppose the ion intensities in a spectrum are the probability of producing corresponding ions, which are related to the formation kinetics and stability of the ions. Thus, once the fragmentation pathways of a peptide and the relative stability of the resulting ions are fully understood, mass spectrum of the peptide can be accurately predicted. However, there needs much more works to consummate this idea since lots of obscurities still exist in characterization of peptide fragmentation behaviors.

Proline (Pro) effect is a well-known fragmentation in MS/MS spectra of peptides [11–14] or MS spectra of intact proteins [15], in which selective cleavage commonly occurs at the N-terminal side of Pro with mobile protons to form abundant *y* ions. Proline is the only proteinogenic amino acid whose side chain links to the α-amino group. Thus, proline is an imino acid and its pyrrolidine ring constructs the peptide backbone. This special molecular structure was firstly regarded as the cause of proline effect for pyrrolidine ring has high proton affinity [11,16]. However, Vaisar et al. found that proton affinity of proline was not the only reason for proline effect because substituting higher proton affinity residue for proline in pentapeptides could not produce same selective cleavage [17]. Recent studies showed that cleavage at N-terminal side of Pro was determined by proton mobility in peptides [18] and affected by identity of residue of preceding Xxx (Xxx denotes any residue) in Xxx-Pro [19], and proposed that this cleavage could be interpreted by PIC model [10]. Most recently, a systematical investigation of Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe and Trp) revealed that cleavage N-terminal to proline involved b_x–y_z pathway and the specific structure of proline could stabilize protonation at the nitrogen of Ala-Pro amide bond [20]. Besides the selective cleavage of N-terminal Pro in peptide, it seems that proline can also influence the neutral loss intensities [21] and the abundance of fragment ions [22]. Although lots of efforts have been made in trying to understand the proline effect and its effect on peptide fragmentation, it is still devoid of a convinced picture to show the fragmentation behaviors of proline. In addition, since proline effect always generates dominant base peak and low rest peaks in spectra that makes the mass spectra bad quality to be identified [23] and difficult to be sequenced [24], the understanding of whether proline-containing peptides could produce such spectra becomes essential. Additionally, the fragmentation behavior picture of proline in peptides would greatly assist the prediction of proline-containing peptides tandem mass spectra.

Here we investigated more than 130 000 low-energy CID spectra from NIST Peptide Mass Spectral Libraries (http://peptide.nist.gov/) to draw proline fragmentation picture. The advantages of adopting NIST libraries are: (i) all spectra are obtained by initially searching all collected spectra by four different search engines and then clustering multiple spectra assigned to same peptide to generate a representative spectrum for each peptide, hence these spectra are of high quality [25]; (ii) spectra in libraries have been fully annotated by *y*, *b*, *a* type ions, along with immonium ions, internal ions and neutral losses, which can easily and confidently be used to extract proline N- and C-terminal fragment ions for our investigation; (iii) the number of spectra is larger than previous statistical investigations [18,19] and would provide more sufficient information for drawing the outline of proline fragmentation behavior, especially for peptides at high charge states (more than +3) which are still rarely reported yet to date. In order to understand the factors which influence the fragmentation at N-terminal side of Pro as comprehensive as possible, we artificially drew fragmentation map as its points were possible factors and generated distribution of most abundant peak intensities produced by cleavage N-terminal to Pro for each point. Probabilities of occurrence of selective cleavage at that site were subsequently calculated according to the distributions, thus would provide prediction reference of selective cleavage at N-terminal side of Pro in proline-containing peptides.

2. Experimental

2.1. Data collection and processing

All data were collected from Human [26] and *E*coli [27] peptide mass spectra libraries in NIST Libraries downloaded from NIST website (http://peptide.nist.gov/). Spectra in both libraries are generated by collision-induced dissociation of tryptic peptides and recorded by ion trap (IT) spectrometer. All data processing and probability calculating for selective cleavage procedures (see Section 2.3) were performed on MatLab (version 7.1, http://www.mathworks.com), and fragmentation maps (see Section 2.2) were manually drawn by Microsoft Office Visio 2007 (http://office.microsoft.com/en-us/visio/).

We extracted all spectra assigned to peptides that contained proline in the libraries and then deleted all consensus spectra at same charge state for these spectra represented same peptides and might lead to ambiguous results. The rest spectra were made up of our investigation dataset, totally 133 169 'single' spectra with peptide charge state ranging from +1 to +4. Due to its few spectra number and uncommonly generated by low-energy CID, we did not include fivefold or more charged peptide mass spectra in our dataset. All peaks in a spectrum were normalized to base peak. In order to mention the characterization value more simply in following studies, we named the normalized peak intensity IIBR (*i*on peak *i*ntensity to *b*ase peak *r*atio). Since enhanced cleavage generally appears at N-terminal side of Pro and is always displayed as a single high peak in a mass spectrum, we used the maximum value of IIBRs of the ions produced by dissociation of proline N-terminal amide bond in our analysis, regardless of any ion types. For multiple proline-containing peptides, the maximum value of IIBRs of all proline N-terminal ions was used.

2.2. Construction of fragmentation maps

In order to investigate the factors impacting the cleavage at N-terminal side of Pro as comprehensive as possible, we collected a majority of knowledge that have influences in peptide fragmentation reported previously and from our experiences. After

comparing these factors, we manually constructed proline fragmentation maps as follows: Firstly, different charge states were considered separately for the ions and intensities formed at different charge states are significantly different, as will be discussed below. Thus, four fragmentation maps were constructed. Then we separated multiple prolines and single proline containing peptides for each charge state (denote as multi-Pro and mono-Pro in fragmentation maps). For single proline containing peptides, we split them up into two paths by *y* and *b* ions, that if the most abundant peaks generated by cleavage at N-terminal side of proline were *y* ion series, peptides were investigated along *y* ion path, otherwise peptides were investigated along *b* ion path. *a* type ions were excluded here because few *a* ions are more abundant than *y* or *b* ions in our dataset. For *y* and *b* ions respectively, proton mobility defined by correlation of number of Args and protons was considered. Highly charged peptides with number of Args equaled to number of protons minus 1 were also investigated separately for they have different fragmentation behaviors. Finally, factors such as positions of proline in peptides, the effect of other residues such as aspartic acid effect, glutamic acid effect, pairwise cleavage of Xxx-Pro, *etc.* were used to classify the suppressed and enhanced cleavage at N-terminal side of Pro, and only the factors that could do best classification were retained.

For multi-proline containing peptides, we firstly separated peptides with all prolines bonding together from interspersing among the peptide. Then similar to the construction of the fragmentation maps for single proline containing peptides, two paths were set for *y*, *b* ions. Finally, positions of proline in peptides and pairwise cleavage of Xxx-Pro were considered. It should be noted here that, all prolines in a peptide should be used for investigation of pairwise cleavage of Xxx-Pro. For example, since dissociation of Xxx-Pro amide bond is suppressed when Xxx is Gly or Pro, the suppression of cleavage at Xxx-Pro in multi-proline containing peptides should be strictly limited to all prolines N-terminally bonded to Gly or Pro to prevent other prolines in the peptides from producing enhanced cleavage.

Lastly, the fragmentation maps were drawn by fitting together all factors introduced above. The distribution of IIBR for each factor was also shown in fragmentation map under the point belonging to the factor.

2.3. Calculating probability for selective cleavage

Prior to calculate probability for occurring selective cleavage at N-terminal side of Pro, the selective cleavage should be defined to identify whether it happened to a peptide bond. Thus, the following two conditions were set. If the intensity of ions formed by cleavage N-terminal to Pro met one of the conditions, the cleavage was considered as selective.

Condition 1: IIBR = 1;
Condition 2: BCR $\geq$ 1/L and $R_{maxInt(N,Pro)} \leq 5$;

BCR stands for bond cleavage ratio, which is calculated by sum of all *y*, *b*, *a* type ions along with their neutral losses formed by cleavage N-terminal to Pro versus total ion current (TIC). This definition of BCR is little different from previous [19]. L is the length of peptide (i.e., number of amino acid residues). $R_{maxInt(N,Pro)}$ stands for the intensity rank of the most abundant ion formed by cleavage N-terminal to Pro in spectrum. We defined the selective cleavage according to these two conditions because of the following reasons: on the point of view of experience, an amide bond dissociation which is considered as selective should produce abundant ion peak in mass spectrum. Thus if the ion formed by this dissociation is the base peak, it could definitely be considered as selective. On the other hand, if the ion is not the most abundant in spectrum, it is difficult to identify the dissociation as selective for there exist a large number of spectra in which only a single abundant base peak is presented and intensities of all other peaks are very low, or many high abundant peaks are presented all together. When this situation occurs, BCR is used. Since dissociation of a peptide bond in peptide is considered as equal probability (can be easily expressed as 1/L, L is peptide length) [28], only the probability of occurring cleavage at one site higher than other sites in peptide is considered as selective. Accordingly, if the most abundant ion formed by cleavage N-terminal to Pro was not the base peak in spectrum, selective cleavage was identified when the ion was one the five highest peaks in spectrum and BCR of the cleavage was higher than 1/L.

Once the selective cleavage was defined, we calculated the probability of occurring such cleavage for each factor presented in fragmentation maps. For a frequency distribution, the probability can be simply estimated by dividing number of spectra appearing selective cleavage by total number of spectra satisfied the factor. However, if there are not enough data for this estimation, the result would be mistaken. Hence for such situation, bootstrap method [29] was used. Bootstrap is a resampling method useful in directly estimating parameters derived from unknown or too complicated distributions or in making statistical inference from insufficient sample size. The implementation of bootstrap is firstly constructs a new dataset called *bootstrap sample* from original dataset with replacement. That is, it randomly selects data (could be selected repeatedly) from original dataset with the number of data points same with the original one to form new dataset, which is named *bootstrap sample*. Take a dataset containing 6 data points [x_1, x_2, x_3, x_4, x_5, x_6] as an example, one can randomly select data points x_1, x_2, x_3, x_3, x_4, x_6 from the dataset to form a *bootstrap sample* [x_1, x_2, x_3, x_3, x_4, x_6] (data point x_3 is repeatedly selected) with data number of 6 which is same with the original one. Bootstrap generates a large number of such independent *bootstrap samples* (i.e., all *bootstrap samples* are not the same) and evaluates statistic of interest (e.g., mean, standard deviation of the sample, *etc.*) for each sample. Consequently, a distribution of the statistic could be drawn from the evaluation and used to infer conclusions for the original dataset (e.g., the confident interval of the mean). For our datasets, we generated 1000 *bootstrap samples* for each factor and calculated the number fraction of spectra occurring selective cleavage for each sample. The final probability value was simply estimated by averaging whole fraction values calculated from bootstrap samples.

According to fragmentation maps and the probabilities calculated, probability maps were manually drawn corresponding to fragmentation maps to quantitatively present proline fragmentation behavior.

3. Results and discussions

3.1. Data description and statistics

In our dataset, all peptides are charged from +1 to +4, with doubly and triply charged peptides being the major part (52.22%, 34.60% for +2 and +3 charge states respectively). This can be expected for it is well known that electrospray ionization method produces most amount +2 and +3 charged peptide precursor ions. Among the more than 130 thousand spectra, the lengths of peptides range from 5 to 79 amino acid residues with the average at 18 residues. The numbers of mono-proline and multi-proline containing peptides in our dataset are as follows: 6246(72.44%) and 2376(27.56%) spectra at charge state +1, with average peptide lengths at 11 and 13 residues; 42 911(61.70%) and 26 636(38.30%) spectra at charge state +2, with average peptide lengths at 14 and 16 residues; 22 229(48.25%) and 23 846(51.75%) spectra at charge state +3, with average peptide lengths at 18 and 22 residues;

3215(36.02%) and 6089(63.98%) spectra at charge state +4, with average peptide lengths at 23 and 27 residues. (Detailed description and statistics of our dataset are shown in Table S-1 in Supplemental Information.)

3.2. Fragmentation occurs at C-terminal side of proline

It is well known that cleavage at C-terminal side of proline is un-favored. In order to confirm this observation, we calculated the IIBR value generated by cleavage C-terminal to proline. The resulted distributions are shown in Fig. 1. The most number of IIBR values ranges from 0 to 0.2 indicating the prohibition of dissociation of proline C-terminal amide bond. It is interesting to find the enhanced cleavage at C-terminal side of Pro for few peptides in our dataset (IIBR equals to 1 in Fig. 1). The reason for this enhancement can be attributed to '$y_{N\text{-}2}$ rule' [30–32], that when proline locates at position 2 calculated from N-terminus (i.e., the second residue of the peptide), abundant $y_{N\text{-}2}$ ion could be produced at C-terminal side of Pro. The mechanism behind this selective cleavage may be via $y_{N\text{-}2}$/diketopiperazine pathway to form diketopiperazine structure of b_2 ion [10,32]. When peptides are highly charged, however, more other locations can produce high abundant ions at C-terminal side of proline (percent of IIBRs attributed to $y_{N\text{-}2}$ rule among all IIBRs larger than 0.8 is 90.51%, 70.61%, 45.61% and 16.67% for singly, doubly, triply and quadruply charged peptides respectively). This should be noted in prediction of peptide fragment ion spectra.

3.3. Statistics of ions generated by cleavage at N-terminal side of proline

Before investigating proline N-terminal fragmentation behavior, it is desirable to have an overview of the most abundant ions formed by cleavage N-terminal to Pro. As shown in Fig. 2, *y* and *b* type ions are the dominant ions with total frequency larger than 90%. Under some conditions, water or ammonia loss of *y* and *b* ions becomes the most abundant. It is worth noting here that cleavage N-terminal to proline can also produce high abundant *a* ions in few peptides, whereas no abundant neutral loss of such ions was observed. The reason why *a* ions become the dominant is not clear. It seems that it is formed randomly because only few peptides in our dataset are found to produce dominant *a* ions, and the average IIBR of the ions is much lower than *y* or *b* ions' (data not shown). For *y* and *b* ions, the probability of *y* ions being the dominant is far higher than *b* ions. This is consistent with the general observations that the proline-containing peptides often produce dominant *y* ions when selective cleavage occurred at N-terminal side of Pro. When peptides at high charge state, the multiply charged *y* and *b* ions occupy the highest probability (>50% for +3 charge state and >70% for +4 charge state, as shown in Fig. 2C and D). There are no significant differences of ion intensities formed by multiple proline and single proline containing peptides, though multiple proline containing peptides have a bit higher probability to lose neutral molecules.

Though the most abundant ions formed by cleavage at N-terminal side of proline are absolutely occupied by *y* and *b* ions, the average intensities have no such obvious evidence (see Fig. S-1 in supplemental information). The average IIBR of neutral losses of *y* and *b* ions are also abundant, especially when peptides are at higher charge states. This indicates that the neutral losses of *y* and *b* ions also have high probability to become the dominant ions when selective cleavage occurred at N-terminal proline. Comparing all ions among four charge states, the highest average IIBR at +2 and +3 charge states is y^{2+} and y^{3+} ions (nearly 0.9 and 0.8 for multiple proline and single proline containing peptides) respectively, and is higher than other ions at other charge states, indicating that the selective cleavage could most probably be happened at +2 and +3 charge states. This obviously selective cleavage in peptides charged +2 and +3 charge states could be attributed to more free mobility of the protons than singly charged and less effected by Coulombic repulsion than quadruply charged. It is worth noting here that, for +4 charge state, the relatively low average IIBRs of *y* and *b* ions reveal the influence of Coulombic repulsion [13,33] and other competitive pathways to the selective cleavage at N-terminal side of proline. When more than one proline is contained in a peptide, the probability of selective cleavage significantly increases, even two times more than single proline containing peptides at high charge states. This may due to the high proton affinity of proline that multiple prolines spreading in peptide could increase the probability of sequestering protons to induce the dissociation of proline N-terminal peptide bond [34].

In addition to statistical analysis of the most abundant ions formed, we also investigated the distribution of *y* and *b* ions according to the position of proline in peptide (Fig. S-2 in supplemental information). From the investigation, the trends for formation of *y* and *b* ions are clear: *y* ions are abundant when proline locates near the N-terminus of peptides, and *b* ions are shown up when proline locates towards C-terminus, as obviously and simply shown at charge state +1 (Fig. S-2A in supplemental information). This indicates that the fragmentation of proline in peptide tends to form large ions. However, it becomes complex at higher charge states, that *y* ions are more abundant than *b* ions even when proline locates near C-terminus, especially at +2 and +4 charge states. This reduction of *b* ions may indicate the fragility of large *b* ions. For quadruply charged peptides, the nearly uniform distribution of *b* and *y* ions can be due to Coulombic repulsion. We also found that when proline stays at other sites in peptide except position 2, one of the strongly competitive pathways to N-terminal cleavage of proline is the formation of $y_{N\text{-}2}$ ions (data not shown). This phenomenon has been discussed recently by Harrison [30] and is named '$y_{N\text{-}2}$ rule' [31,32] which is proposed via b_2–$y_{N\text{-}2}$ pathway to form traditional oxazolone structure b_2 ion and $y_{N\text{-}2}$ ion [35,36].

3.4. Proline fragmentation map

3.4.1. *y* ions generated by cleavage N-terminal to Pro in single proline containing peptides

Abundant *y* ions produced by selective cleavage at N-terminal side of Pro have been generally observed [11,12,14–17] and regarded as the phenomenon of proline effect [17]. For all ions formed by cleavage N-terminal to Pro in our dataset, number of spectra with *y* ions being the most abundant is much more than the number of spectra with *b* ions (even nearly 6 times more, as is shown in Fig. 3A and B). This is consistent with the generally observations. In this analysis, since cleavage N-terminal to Pro has been proved to be determined by proton mobility in peptides, investigation on the relation between number of Arg and protons was made firstly. When protons are all sequestered by Args (point No.R $\geq$ +2 in Fig. 3A), which is generally defined as proton non-mobile model, the selective cleavage at N-terminal side of Pro is suppressed. However, the significant enhancement is found for Asp-Pro and Glu-Pro (points D-P and E-P in Fig. 3A), indicating the positive influence of aspartic acid effect on selective cleavage at N-terminal side of proline. In fact, the probability of occurrence of selective cleavage at Asp-Pro and Glu-Pro is generally more than 0.7 among four charge states. Besides, enhanced cleavage is also found at Arg-Pro when peptides are doubly charged (point R-P_IN along path *y* ions from No.R $\geq$ 2 to X-P in Fig. 3A). The mechanism behind the enhanced cleavage may be via nucleophilic attack of εN of Arg side chain on the carbon of amide bond to form 6-member ring lactam structure [37]. When protons are non-mobile for *y* ions, it seems that the charge-directed cleavage at N-terminal side of Pro competes with charge-remote cleavage of Asp and Glu (probability 0.7592 versus 0.3956 for points D/E_IN and (D & E)_EX respectively in Fig. S-

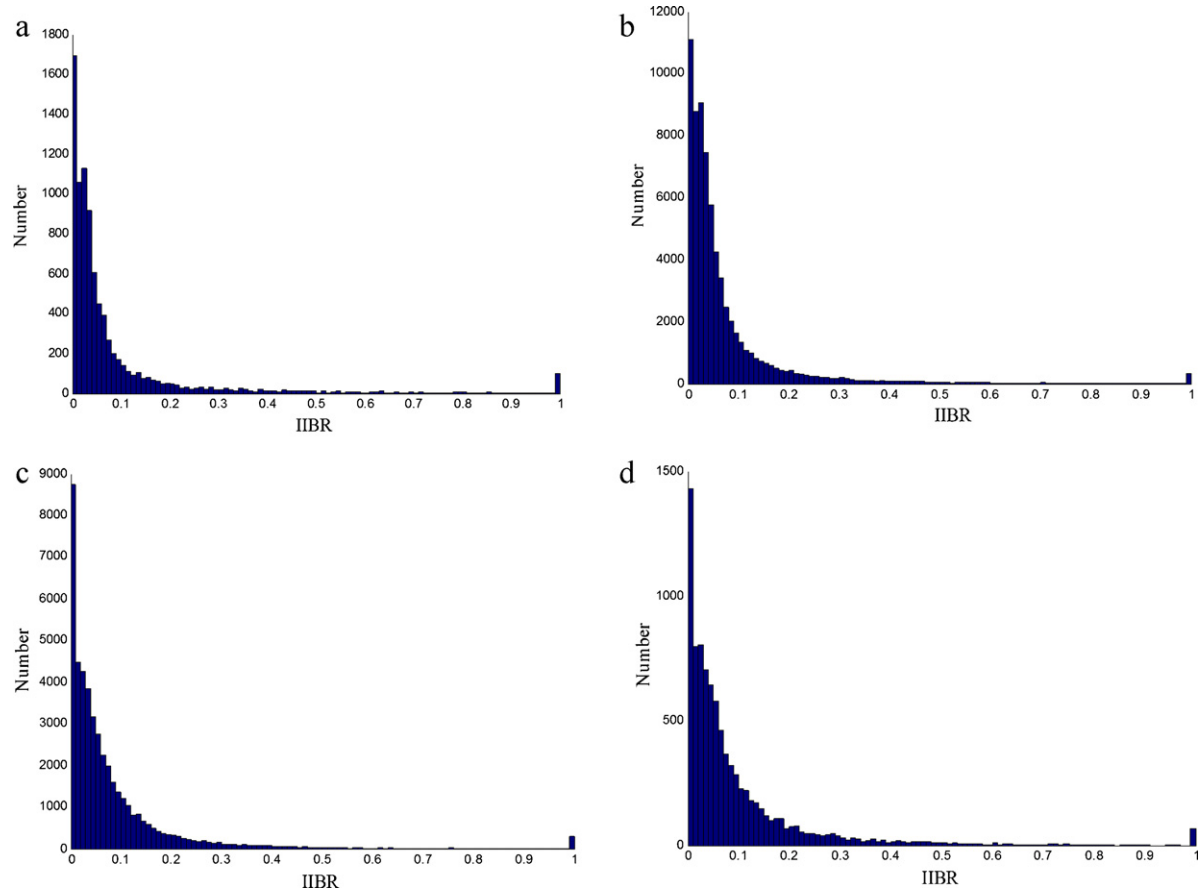

Fig. 1. Distributions of IIBR value of C-terminal proline for +1(A), +2(B), +3(C) and +4(D) charge states.

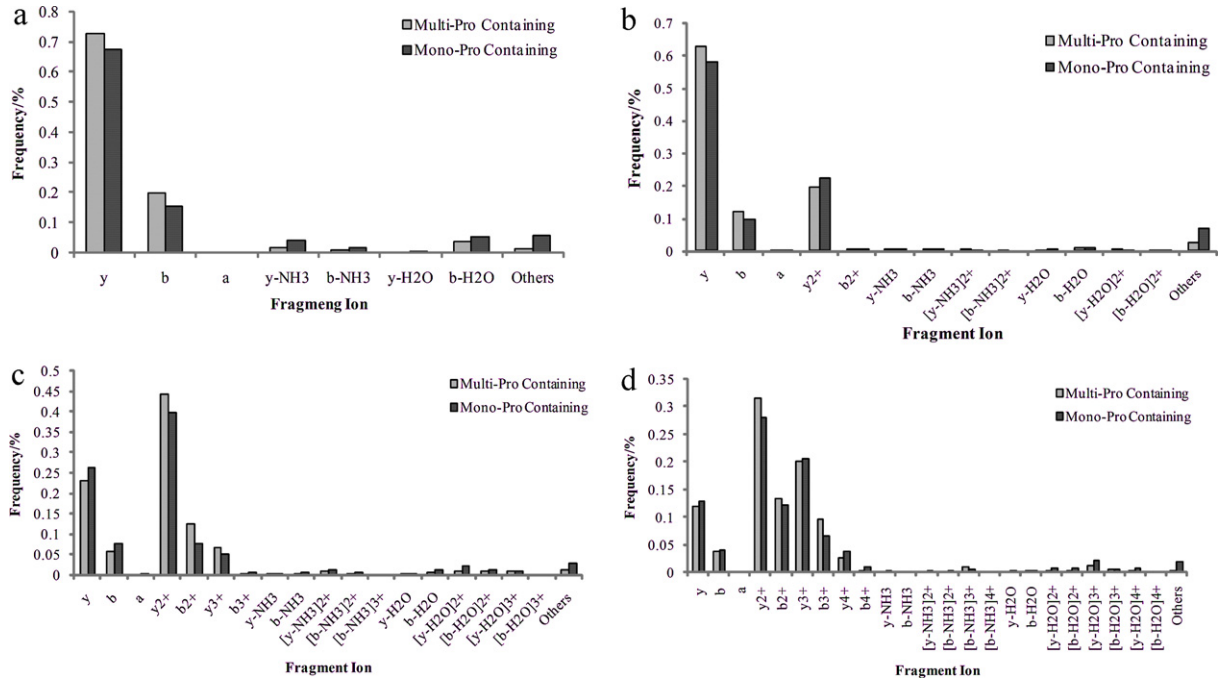

Fig. 2. Frequency of ions formed at N-terminal side of proline at +1(A), +2(B), +3(C) and +4(D) charge states. Multi-Pro Containing and Mono-Pro Containing in legend stand for ions formed in multi-proline and mono-proline containing peptides. The ions listed for analysis are *y*, *b*, *a* and their water and ammonia losses. When charge state of peptides range from +2 to +4, the possible multi-charged *y*, *b*, *a* and their neutral losses are also analyzed.

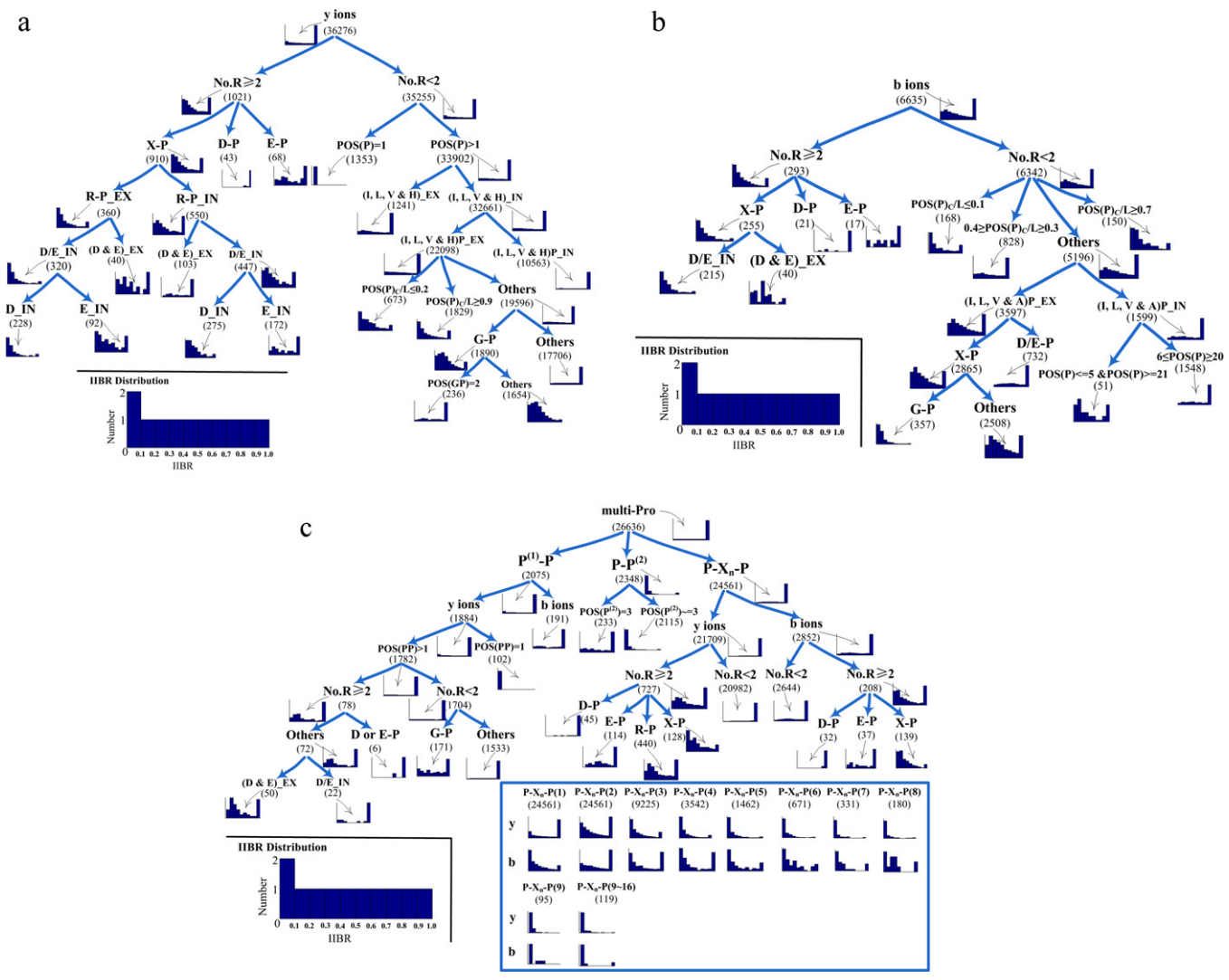

Fig. 3. Fragmentation map for cleavage N-terminal to Pro in doubly charged peptides. (A) When ions with maximum IIBR formed by cleavage N-terminal to Pro are *y* ions, (B) ions with maximum IIBR are *b* type ions, and (C) fragmentation behaviors of multi-prolines. Meanings of symbols in the figure are: No.R is number of Args in peptides; D-P, E-P and X-P represent Asp-Pro, Glu-Pro and any other residues except Asp and Glu N-terminally bonded to Pro exist; G-P represents Gly-Pro exists; D_IN and E_IN represent residue Asp and Glu exist; D/E_IN represents Asp or Glu exists; (D & E)_EX represents Asp and Glu are excluded; R-P_IN and R-P_EX represent Arg-Pro is included or excluded; (I, L, V & H)_IN represents any of residues Ile, Leu, Val and His is included and (I, L, V & H)_EX represents all of the residues Ile, Leu, Val and His are excluded; (I, L, V & H)P_IN represents any of Ile-Pro, Leu-Pro, Val-Pro and His-Pro exists and (I, L, V & H)P_IN represents all the Ile-Pro, Leu-Pro, Val-Pro and His-Pro are excluded; POS(P) and POS(GP) are positions of proline and Gly-Pro in peptides calculated from N-terminus and $POS(P)_C/L$ is relative position calculated by dividing position of proline computed from C-terminus by peptide length. For multiple prolines containing peptides, P-P or P-X_n-P means all prolines bonded together or spreading in peptides, and the number in superscript represents relative position of proline to other prolines in peptides. The box inserted is the distributions of *y* and *b* ions for different relative positions of prolines in P-X_n-P containing peptides (this relative position is defined by position relative to the proline nearest N-terminus of peptide and the order of the position is included in bracket at the end of symbol P-X_n-P). Number of spectra satisfied the condition is also included in bracket.

4B in supplemental information), and the effect of Asp is more stronger than Glu (probabilities for occurring selective cleavage at N-terminal side of Pro are 0.2616 and 0.6106 respectively). This conclusion conflicts from previous work that Aspartic acid effect did not compete with proline effect [7]. The reason of this confliction may be due to the few number of spectra investigated or even the incorrect assignment of the spectra in our dataset or previous work. The same conclusion could not be drawn for triply and quadruply charged peptides for there is not enough data for investigation.

When protons are mobile (No.Arg ≤ +2 in Fig. 3A), cleavage of peptides is dominated by selective cleavage occurs at N-terminal side of Pro (probability is generally more than 0.5 for all four charge states, Fig. S-4 section I in supplemental information). It has been reported that for pairwise cleavage of Xxx-Pro (Xxx denotes any residue), cleavage enhanced when Xxx is Ile, Leu or Val and suppressed when Xxx is Gly or Pro [19]. Our investigation obtains same trends and the probability for occurring selective cleavage could be up to more than 0.8 at Ile, Leu, Val-Pro (point (I, L & V)P_IN in Fig. S-4B section I) and down to less than 0.01 at Gly-Pro (point G-P in Fig. S-4A section I). It should be noted that the pairwise fragmentation of G-P is enhanced when G-P locates at position 2 calculated from N-terminus, which can be attributed to 'y_{N-2} rule'. Surprisingly, for triply and quadruply charged peptide, cleavage is most suppressed at Arg-Pro, Lys-Pro and His-Pro among all 20 pairwise Xxx-Pro. This may be due to the Coulombic repulsion between the protons located at these basic residue side chains and the freely mobile protons that migrate to Pro N-terminal site, thus prohibiting the induction of cleavage. In addition to residues in pairwise Xxx-Pro, the position of proline in peptides also greatly affects the fragmentation of proline. As point $POS(P)_C/L$ shows in

Fig. 3A, when $POS(P)_C/L$ value ranges from 0.2 to 0.9, probability of selective cleavage at N-terminal side of Pro is significantly higher than other sites (0.7359 versus 0.1708 and 0.2033). Similar results are found for the other three charged peptides, that proline effect is suppressed when proline locates near two termini of peptides, especially for highly charged peptides.

When peptides are singly charged, N-terminus glutamic acid (Glu) is observed to significantly prohibit selective cleavage at N-terminal side of Pro (points POS(E) = 1 versus POS(E)∼ = 1 in Fig. S-3A section I in supplemental information). It has been reported that peptide with glutamic acid locating at its N-terminus has a strong effect to lose water or ammonia molecules to form pyroglutamic acid derivatives [38], which prohibits cleavage at other sites and generates high abundant peak for water loss of precursor ion [39]. Therefore, cleavage N-terminal to Pro is suppressed. However, the impact of N-terminus Glu is greatly reduced when charge state increased. This may be ascribed to the induction of cleavage at other sites by other freely mobile protons and hence cleavage at N-terminal side of Pro becomes selective. Additionally, for highly charged peptides, when number of Arg is equal to number of protons minus 1, probability of occurrence of selective cleavage N-terminal to Pro is the lowest comparing to conditions of proton non-mobile and more than 2 protons are freely mobile. Nevertheless, it is enhanced when position of Arg is more than two residues away from position of Pro or disappeared at N-terminal part of Pro in peptides (point $POS(P\text{ to }R)_{REL} > 2$ along path No.R = 2 in Fig. S-3C section I in supplemental information). As mentioned above, though cleavage at N-terminal side of Pro requires location of proton at its site, highly charged peptides are strongly influenced by Coulombic repulsion of multiple protons. Consequently, when Arg positions at or near N-terminal side of Pro, repulsion between protons sequestered by Arg and N-terminal side of peptides weakens the inducement of dissociation of the peptide bond. This should be noticed for spectral prediction of highly charged peptides.

3.4.2. *b* ions generated by cleavage N-terminal to Pro in single proline containing peptide

For *b* ions formed by cleaving at N-terminal side of Pro, same set of factors affecting the selective cleavage were investigated as *y* ions and similar factors were found (Fig. 3B). When protons are sequestered by side chains of Arg, strongly selective cleavage occurs at Asp-Pro and Glu-Pro attributed to aspartic acid effect, and the relative lower extent of selective cleavage at Glu-Pro than Asp-Pro in non-proton mobile model confirms the depressed effect of glutamic acid effect [40] (points D-P and E-P along the path No.R ≥ 2 in Fig. 3B). Besides the selective cleavage at Asp-Pro and Glu-Pro, the existence of His in singly charged peptides could also show up the intensity of *b* ions when protons are non-mobile, increasing the probability of occurrence of selective cleavage at N-terminal side of Pro (point H_IN along the path R_IN in Fig. S-3A section II in supplemental information). It has been documented that selective cleavage could be occurred at C-terminal side of His when protons are mobile [41] and position of His in peptides could affect the cyclization/reopening of *b* ions [42]. However, no evidence in our dataset shows the influence of position of His in peptides on the enhancement of cleavage at N-terminal side of Pro (data not shown). Thus, the increased probability of the cleavage in singly charged peptides could not definitely be attributed to the His effect.

For pairwise cleavage, when protons are freely mobile through peptide (path No.R < charge state in Fig. 3B and S-3A, S-3C and S-3D section II in supplemental information), selective cleavage at N-terminal side of Pro is suppressed at Gly-Pro and enhanced at Ala, Ile, Leu, Val-Pro amide bond (points G-P and (I, V, L & A)P), indicating the general suppression and enhancement of cleavage at Gly-Pro and Ile-Pro, Leu-Pro, Val-Pro. The dissociation of Asp-Pro and Glu-Pro amide bond is also enhanced for doubly and triply charged peptides (D/E-P along path No.R < 2 in Fig. 3B and Fig. S-3C section II in supplemental information). It has been reported that when Lys stayed at the C-terminus of peptide, cleavage at Asp-Xxx was selected and subsequently suppressed cleavage at Xxx-Pro, whereas Xxx-Pro would be selected when Lys stayed at the middles site of peptide [18]. Our data do not show same trends but enhanced cleavage also was generally observed when more than two basic residues existed in peptides with Lys stayed at C-terminus (data not shown).

As discussed in Section 3.4.1, selective cleavage most probably occurs when proline positions at middle sites of peptides and is suppressed near two terminuses. For *b* type ions, similar trends are found when peptides are doubly or higher charged. However in triply charged peptides, the trends are a bit changed for the dissociations of Ile-Pro, Leu-Pro, Val-Pro and Asp-Pro, Glu-Pro amide bonds, that the cleavage is suppressed when proline locates near N-terminus of peptide (point $POS(P)_C/L > 0.6$ along path No.R < 2 in Fig. S-3C section II in supplemental information). It is not clear why it happens for these amide bonds but not for other Xxx-Pro amide bonds in triply charged peptides. For quadruply charged peptides, selective cleavage at N-terminal side of Pro is not as strong as other charged peptides, even for dissociation of Ile-Pro, Leu-Pro and Val-Pro amide bonds (probability is less than 0.6 for point (I, L & V)P in Fig. S-4D section II in supplemental information) and nearly does not occur at His, Lys and Arg-Pro (probability 0.0582, point (H, K & R)P). This indicates the dominant effect of Coulombic repulsion on peptide fragmentation when peptides are highly charged. In addition to influences of pairwise cleavage of Xxx-Pro and position of Pro in peptides, cleavage at N-terminal side of Pro is weakened when Glu positioned at N-terminus of peptides when peptides are singly charged, as has been demonstrated for *y* type ions.

3.4.3. Fragmentation of proline in multiple prolines containing peptides

Cleavage at N-terminal side of Pro in multiple prolines containing peptides is much more complex, as shown in Fig. 3C and Fig. S-3A, S-3C and S-3D section III in supplemental information. When prolines in peptides are bonded together, the 'active' N-terminal amide bond only belongs to the Pro nearest to the N-terminus of peptide for other Pros are bonded to C-terminal side of previous Pros at which cleavage would is prohibited (point $P\text{-}P^{(2)}$) [19]. As a result, these multiple prolines could be considered as one 'active' Pro and may present similar fragmentation behavior with single proline containing peptides. Though there has not enough data for comprehensive investigation of multiple prolines containing peptides to draw such conclusion, similar factors affect selective cleavage at N-terminal side of 'active' Pro are still found, as shown along path $P\text{-}P^{(1)}$ in Fig. 3C and Fig. S-3A, S-3C and S-3D section III. When protons are mobile, cleavage at N-terminal side of 'active' Pro is probably selective except Gly-Pro (point G-PP in Fig. 3C). Contrarily, cleavage is suppressed when protons are non-mobile and only enhanced at Asp-Pro and Glu-Pro. However, dissociation of Gly-Pro and Pro-Pro amide bonds could be enhanced when they positioned at 2 calculated from N-terminus, which is due to the y_{N-2} rule. When prolines disperse in peptides, the probability of occurrence of selective cleavage at N-terminal side of Pro is notably increased, on the verge of 0.9 (point P-Xn-P in Fig. S-4B, section III in Supplemental Information). This is reasonable because all prolines in peptides are 'active' sites for occurring selective cleavage, hence increases the probability of dissociation of Pro N-terminal amide bond. As expected, when protons are non-mobile or all Xxx-Pro amide bonds in peptide are fixed by non-active cleavage pairwise such as

Gly-Pro, Arg-Pro, Lys-Pro, *etc.*, cleavage at N-terminal sides of all Pros are prohibited (points (G, R, K & H)$_{co}$P and (G, R & K)$_{co}$P in Fig. S-3B and S-3C in supplemental information respectively). In order to identify whether there exists competitive cleavage among all prolines, distribution of IIBR for different relative positions of multiple prolines in peptides is presented in each fragmentation map as an insert box. As is shown, the probability of occurring selective cleavage at N-terminal side of each proline is low, indicating the competitive cleavage between prolines in peptides. The further decreased intensities of *b* and *y* ions when prolines locate towards C-terminus of peptides imply that the selective cleavage most probably appears at N-terminal side of Pro near the N-terminus.

3.5. Interpretation of proline effect

Proline effect is still an unsolved problem in peptide fragmentation research. Paizs and Suhai [10] concluded that the specific activity of proline is exerted via affecting otherwise non-specific fragmentation pathways and can be fully explained by b_x–y_z pathway. In other words, the selective cleavage occurs at N-terminal side of Pro requires proton locates at its site and would be suppressed when other specific fragmentation pathways appear. The suppression of proline effect when the protons in peptides are non-mobile or other specific effects (e.g., aspartic acid effect, N-terminus glutamic acid effect, *etc.)* are happened confirms this explanation. Additionally, in order to rationalize the general suppression of cleavage at Gly-Pro and Pro-Pro and enhancement of cleavage at Ile-Pro, Leu-Pro, Val-Pro, and under certain condition Ala-Pro, Breci et al. [19] resorted to different conformations of Xxx-Pro. It has been reported that proline in Xxx-Glu-Gly-Pro may stabilize the reacting conformation for facile cleavage of Glu-Gly in proline-rich peptides to enhance the cleavage [43,44]. Furthermore, proline under different conditions can adopt different conformations [45]. Therefore, since there has no specifically chemical active side chains in proline, it is reasonable to assume that the proline's cyclic side chain has some special effect on cleavage transition structures that facile the cleavage N-terminal to proline, i.e., lowering the relative energy of the structures, and the cleavage at N-terminal side of proline follows the traditional b_x–y_z pathways. This hypothesis has been proved in most recently published work that proline side chain in protonated peptides can stabilize the protonation at N-terminal side of Pro [20]. Thus the suppression of proline effect when proline positions near two termini of peptides can be due to the different conformation of proline in cleavage transition structure from proline positions at middle sites of peptides. And the reason for less competitive than chemically driven effect (e.g., aspartic acid effect, N-terminus glutamic acid effect, histidine effect, *etc.)* can be rationalized by the lower energy for occurring latter reaction. However, theoretical studies should be carried out to prove these assumptions.

4. Conclusions

Enhanced cleavage at N-terminal side of proline termed proline effect is commonly observed in proline-containing peptide fragment ion mass spectra. Statistical analysis of more than 130, 000 high quality spectra produced by singly, doubly, triply and quadruply charged peptides definitely shows the strong N-bias of proline [46]. Besides, abundant *y* ions being the major part of fragment ions produced by cleavage at N-terminal side of Pro confirms the common observations of proline effect. Comparing the factors that enhance and suppress the selective cleavage, the trends become apparent:

(1) Cleavage at N-terminal side of Pro is prohibited when protons are non-mobile, whereas significantly selective cleavage occurs at Asp-Pro and Glu-Pro due to aspartic acid effect;
(2) Cleavage at N-terminal side of Pro is clearly selective when protons are mobile. Under this condition, the cleavage is determined by pairwise Xxx-Pro and position of proline in peptides. For the former, cleavage is selective when preceding Xxx is Ile, Leu or Val and suppressed when Xxx is Gly or Pro. And for the latter, selective cleavage most probably occurs at middle site of peptide. It should be noticed that when other specific fragmentation pathways appear (such as $y_{N\text{-}2}$ rule, N-terminus Glu effect, *etc.*), proline effect is suppressed, indicating the influence of cleavage competition on the fragmentation of proline;
(3) When peptides are highly charged, cleavage at N-terminal side of Pro is depressed comparing to lower charged peptides and more strongly affected by position of proline in peptides. The rarely cleavage at Xxx-Pro when preceding Xxx is Arg, Lys or His and effect of position of proline on proline effect indicate the influence of Coulombic repulsion in highly charged peptides, which could also inhibit other specific fragmentation pathways;
(4) For multiple proline containing peptides, selective cleavage occurs at N-terminal side of Pro is stronger than single proline containing peptides, especially when prolines disperse in peptides. Applying the factors affecting proline effect drawn from single proline containing peptides to multiple proline containing ones, similar results are found when prolines are bonded together. However, when prolines disperse in peptides, selective cleavage is still observed even if cleavage at N-terminal side of several prolines is inhibited because the remaining prolines could still produce selective cleavage. Additionally, competitive cleavage is found between prolines in same peptide and the highest probability of occurring selective cleavage belongs to proline nearest to N-terminus of peptide;
(5) '$y_{N\text{-}2}$ rule' is generally found in our dataset and increases the probability of occurring selective cleavage when proline positions at 3 calculated from N-terminus, even for Gly-Pro and Pro-Pro amide bonds.

As proline effect often leads to single most abundant peak that makes other peaks too low to distinguish from signal to background noise in peptide fragment ion spectra, it is not clear whether this effect happens in all proline-containing peptides. Hence the theoretical spectra predicted for proline-containing peptides are still susceptible. This statistical analysis is comprehensive and gives an insight into proline fragmentation behavior in protonated peptides and can be integrated into spectral prediction framework. Furthermore, the probability of occurring selective cleavage at N-terminal side of Pro calculated for each fragmentation behavior provides a quantitative characterization of the cleavage which would be helpful in prediction of mass spectra and potentially be a reference data in *de novo* sequencing of proline containing peptides.

Acknowledgement

This work was financially supported by the National Nature Foundation Committee of P.R. China (Grants Nos. 20875104 and 21075138).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijms.2011.08.005.

References

[1] J. Cox, M. Mann, Is proteomics the new genomics? Cell 130 (2007) 395–398.
[2] P. Roepstoff, J. Fohnman, Proposal for a common nomenclature for sequence ions in mass spectra of peptides, Biol. Mass Spectrom. 11 (1984) 601.
[3] K. Bieman, Contributions of mass spectrometry to peptide and protein structure, Biomed. Environ. Mass Spectrom. 16 (1988) 99–111.
[4] A.I. Nesvizhskii, O. Vitek, R. Aebersold, Analysis and validation of proteomic data generated by tandem mass spectrometry, Nat. Biotechnol. 4 (2007) 787–797.
[5] H. Steen, M. Mann, The ABC's (and XYZ's) of peptide sequencing, Nat. Rev. Mol. Cell Biol. 5 (2004) 699–711.
[6] J.L. Jones, A.R. Dongré, Á. Somogyi, V.H. Wysocki, Sequence dependence of peptide fragmentation efficiency curves determined by electrospray ionization/surface-induced dissociation mass spectrometry, J. Am. Chem. Soc. 116 (1994) 8368–8369.
[7] Y. Huang, V.H. Wysocki, D.L. Tabb, J.R. Yates III, The influence of histidine on cleavage C-terminal to acidic residues in doubly protonated tryptic peptides, Int. J. Mass Spectrom. 219 (2002) 233–244.
[8] C. Gu, G. Tsaprailis, L.A. Breci, V.H. Wysocki, Selective gas-phase cleavage at the peptide bond C-terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides, Anal. Chem. 72 (2000) 5804–5813.
[9] V.H. Wysocki, G. Cheng, Q. Zhang, K.A. Hermann, R.L. Beardsley, A.E. Hilderbrand, Peptide fragmentation overview, in: J. Laskin, C. Lifshitz (Eds.), Principles of Mass Spectrometry Applied to Biomolecules, John Wiley and Sons, Hoboken, NJ, 2006, p. 279.
[10] B. Paizs, S. Suhai, Fragmentation pathways of protonated peptides, Mass Spectrom. Rev. 24 (2005) 508–548.
[11] D.F. Hunt, J.R. Yates III, J. Shabanowitz, S. Winston, C.R. Hauer, Protein sequencing by tandem mass spectrometry, Proc. Natl. Acad. Sci. U. S. A. 82 (1986) 6233–6237.
[12] D.J. Harvan, J.R. Hass, W.E. Wilson, C. Hamm, R.K. Boyd, H. Yajima, D.G. Klapper, Fast atom bombardment/tandem mass spectrometry of physalaemin-like peptides, Biomed. Environ. Mass Spectrom. 14 (1987) 281–287.
[13] X.J. Tang, P. Thibault, R.K. Boyd, Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation, Anal. Chem. 65 (1993) 2824–2834.
[14] Y. Wang, J. Johansson, W.J. Griffiths, Characterisation of variant forms of prophenin: mechanistic aspects of the fragmentation of proline-rich peptides, Rapid Commun. Mass Spectrom. 14 (2000) 2182–2202.
[15] J.A. Loo, C.G. Edmonds, R.D. Smith, Tandem mass-spectrometry of very large molecules 2. Dissociation of multiply charged proline-containing proteins from electrospray ionization, Anal. Chem. 65 (1993) 425–438.
[16] B.L. Schwartz, M.M. Bursey, Some proline substituent effects in the tandem mass spectrum of protonated pentaalanine, Biol. Mass Spectrom. 21 (1992) 92–96.
[17] T. Vaisar, J. Urban, Probing the proline effect in CID of protonated peptides, J. Mass Spectrom. 31 (1996) 1185–1187.
[18] Y. Huang, J.M. Triscari, G.C. Tseng, L. Pasa-Tolic, M.S. Lipton, R.D. Smith, V.H. Wysocki, Statistical characterization of the charge state and residue dependence of low-energy CID peptide dissociation patterns, Anal. Chem. 77 (2005) 5800–5813.
[19] L.A. Breci, D.L. Tabb, J.R. Yates, V.H. Wysocki, Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra, Anal. Chem. 75 (2003) 1963–1971.
[20] C. Bleiholder, S. Suhai, A.G. Harrison, B. Paizs, Towards understanding the tandem mass spectra of protonated oligopeptides. The proline effect in collision-induced dissociation of protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp), J. Am. Soc. Mass Spectrom. 20 (2) (2011) 1032–1039.
[21] D.B. Martin, J.K. Eng, A.I. Nesvizhskii, A. Gemmill, R. Aebersold, Investigation of neutral loss during collision-induced dissociation of peptide ions, Anal. Chem. 77 (2005) 4870–4882.
[22] S.J. Barton, J.C. Whittaker, Review of factors that influence the abundance of ions produced in a tandem mass spectrometer and statistical methods for discovering these factors, Mass Spectrom. Rev. 28 (2009) 177–187.
[23] R.J. Simpson, L.M. Connolly, J.S. Eddes, J.J. Pereira, R.L. Moritz, G.E. Reid, Proteomic analysis of the human colon carcinoma cell line (LIM 1215): development of a membrane protein database, Electrophoresis 21 (2000) 1707–1732.
[24] N. Leymarie, E.A. Berg, M.E. McComb, P.B. O'Connor, J. Grogan, F.G. Oppenheim, C.E. Costello, Tandem mass spectrometry for structural characterization of proline-rich proteins: application to salivary PRP-3, Anal. Chem. 74 (2002) 4124–4132.
[25] NIST Libraries of Peptide Ion Fragmentation Spectra, http://peptide.nist.gov/docs/NIST_PepLib_08.pdf.
[26] S.E. Stein, P.A. Rudnick (Eds.), NIST Peptide Tandem Mass Spectral Libraries. Human Peptide Mass Spectral Reference Data, H. sapiens, ion trap, Official Build Date: Feb. 4, 2009, National Institute of Standards and Technology, Gaithersburg, MD, 2009, Downloaded from http://peptide.nist.gov on June 23, 2009.
[27] S.E. Stein, P.A. Rudnick, NIST Peptide Tandem Mass Spectral Libraries. Ecoli Peptide Mass Spectral Reference Data, E. coli, ion trap, Official Build Date: Mar. 21, 2009, National Institute of Standards and Technology, Gaithersburg, MD, 2009, Downloaded from http://peptide.nist.gov on June 23, 2009.
[28] B. Paizs, S. Suhai, Towards understanding some ion intensity relationships for the tandem mass spectra of protonated peptides, Rapid Commun. Mass Spectrom. 16 (2002) 1699–1702.
[29] E. Bradly, R.J. Tibshirani, An Introduction to the Bootstrap, first ed., Chapman and Hall, London, 1994.
[30] G.A. Harrison, Effect of proline position on symmetric versus asymmetric fragmentation of doubly-protonated tryptic-type peptides, Int. J. Mass Spectrom. (2010), doi:10.1016/j.ijms.2010.10.015.
[31] G.A. Unnithan, M.J. Myer, C.J. Veale, A.S. Danell, MS/MS of protonated polyproline peptides: the influence of N-terminal protonation on dissociation, J. Am. Soc. Mass Spectrom. 18 (2007) 2198–2203.
[32] M.M. Savitski, M. Fälth, Y.M. Fung, C.M. Adams, R.A. Zubarev, Bifurcating fragmentation behavior of gas-phase tryptic peptide dications in collisional activation, J. Am. Soc. Mass Spectrom. 19 (2008) 1755–1763.
[33] R.D. Smith, J.A. Loo, R.R.O. Loo, M. Busman, H.R. Udseth, Principles practice of electrospray ionization-mass spectrometry for large polypeptides and proteins, Mass Spectrom. Rev. 10 (1991) 359–451.
[34] R.N. Grewal, H. El Aribi, A.G. Harrison, K.W.M. Siu, A.C. Hopkinson, Fragmentation of protonated tripeptides: the proline effect revisited, J. Phys. Chem. B 108 (2004) 4899–4908.
[35] E.A. Kapp, F. Schütz, G.E. Reid, J.S. Eddes, R.L. Moritz, R.A. O'Hair, T.P. Speed, R.J. Simpson, Mining a tandem mass spectrometry database to determine the trends and global factors influencing peptide fragmentation, Anal. Chem. 75 (2003) 6251–6264.
[36] B.J. Bythell, S. Suhai, A. Somogyi, B. Paizs, Proton-driven amide bond-cleavage pathways of gas-phase peptide ions lacking mobile protons, J. Am. Chem. Soc. 131 (2009) 14057–14065.
[37] B.J. Bythell, I.P. Csonka, S. Suhai, D.F. Barofsky, B. Paizs, Gas-phase structure and fragmentation pathways of singly protonated peptides with N-terminal arginine, J. Phys. Chem. B 114 (2010) 15092–15105.
[38] A.G. Harrison, Fragmentation reactions of protonated peptides containing glutamine or glutamic acid, J. Mass Spectrom. 38 (2003) 174–187.
[39] B. Godugu, P. Neta, Y. Simón-Manso, S.E. Stein, Effect of N-terminal glutamic acid and glutamine on fragmentation of peptide ions, J. Am. Soc. Mass Spectrom. 21 (2010) 1169–1176.
[40] G. Tsaprailis, A. Somogyi, E.N. Nikolaev, V.H. Wysocki, Refining the model for selective cleavage at acidic residues in arginine-containing protonated peptides, Int. J. Mass Spectrom. 196 (2000) 467–479.
[41] G. Tsaprailis, H. Nair, W. Zhong, K. Kuppannan, J.H. Futrell, V.H. Wysocki, A mechanistic investigation of the enhanced cleavage at histidine in the gas-phase dissociation of protonated peptides, Anal. Chem. 76 (2004) 2083–2094.
[42] B.J. Bythell, M. Knapp-Mohammady, B. Paizs, A.G. Harrison, Effect of the His residue on the cyclization of b ions, J. Am. Soc. Mass Spectrom. 21 (2010) 1352–1363.
[43] W.J. Griggiths, A.P. Jonsson, Gas phase conformation can have an influence on peptide fragmentation, Eur. J. Mass Spectrom. 7 (2001) 88–99.
[44] A.P. Jonsson, T. Bergman, H. Jörnvall, W.J. Griffiths, P. Bratt, N. Strömberg, Gln-Gly cleavage: correlation between collision-induced dissociation and biological degradation, J. Am. Soc. Mass Spectrom. 12 (2001) 337–342.
[45] A.E. Counterman, D.E. Clemmer, Anhydrous polyproline helices and globules, J. Phys. Chem. B 108 (2004) 4885–4898.
[46] D.L. Tabb, L.L. Smith, L.A. Breci, V.H. Wysocki, D. Lin, J.R. Yates III, Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides, Anal. Chem. 75 (2003) 1155–1163.