



RESEARCH ARTICLE

Tryptic y^{++} Fragment Ion Distributions Are Guided by Coulombic Repulsion

Karl K. Irikura,¹ John K. Merle,² Yamil Simón-Manso¹¹Chemical and Biochemical Reference Data Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA²Department of Chemistry, Winston-Salem State University, Winston-Salem, NC 27110, USA**Abstract**

Ideal tryptic peptides contain only a single basic residue, located at the C-terminus. Collisional fragmentation of their doubly- or triply-protonated ions generates doubly-charged y^{++} fragment ions with modest intensities. The size distribution of the y^{++} fragments, when averaged over many spectra, corresponds closely to the expectations from charge-directed backbone cleavage and a Coulomb-Boltzmann distribution of mobile protons. This observation should be helpful in developing mechanistic models for y^{++} formation.

Key words: Collision-induced dissociation, CID, Electrostatics, Fragmentation, Mobile proton model, Peptide ion, Proteomics, Quantum chemistry, Tandem mass spectrometry

Introduction

Tandem mass spectrometry of tryptic peptides is an important part of current proteomics studies [1]. Consequently, applied and basic research in this field now represents a substantial fraction of the publications in this journal and of the presentations at mass spectrometry conferences. A key step in applications is to infer the identity of a peptide ion from its fragment spectrum. Better results are expected from improved knowledge and understanding of fragmentation patterns. Although singly-charged y and b ions predominate among the collisional fragments of multiply charged peptide ions, doubly-charged y^{++} ions (Scheme 1) also appear. Their relative intensities, when averaged over many spectra, appear to follow a simple pattern, as described and rationalized herein.

Theory

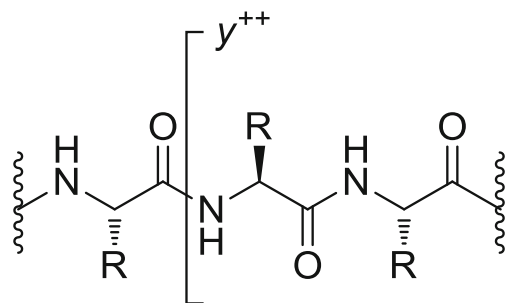
A minimal electrostatic model for a multiply-charged tryptic peptide is a one-dimensional, uniform lattice with a fixed positive charge at one end, corresponding to the protonated, basic residue at the C-terminus, and one or more mobile positive charges elsewhere. Each lattice point represents a backbone amide group upon which a mobile proton may reside. For any particular charge configuration (placement of the mobile protons), the energy is determined only by electrostatic repulsion. This model neglects some major effects, such as folding, side-chain chemistry, and charge delocalization. Nevertheless, as shown below, it represents some experimental observations surprisingly well.

To make the model concrete, let the number of lattice points (i.e., residues) be N . The total charge is given the symbol z , so there are $(z-1)$ mobile charges. The C-terminal residue, which carries the fixed charge, is numbered “1” for numerical convenience. This is illustrated by Figure 1 for a doubly-charged peptide. Numbering is from right to left because, by convention, peptide sequences are written with the N-terminus on the left. If two unit charges (+e, as appropriate for protons) are separated by a distance r , measured in nanometers (nm), then by Coulomb’s Law the repulsive potential energy between them is approximately

Electronic supplementary material The online version of this article (doi:10.1007/s13361-011-0293-7) contains supplementary material, which is available to authorized users.

Correspondence to: Karl Irikura; e-mail: karl.irikura@nist.gov

Received: 22 September 2010
Revised: 26 October 2011
Accepted: 29 October 2011
Published online: 20 December 2011



Scheme 1. Peptide bond cleavage results in a y^{++} ion when two charges remain on the C-side fragment

($138.94/\epsilon r$) kJ mol^{-1} , where ϵ is the dielectric constant. The lattice spacing is taken as 0.37 nm, which is reasonable for peptides. Then the contribution to the potential energy U in the model is

$$U(\Delta n) = \frac{375}{\epsilon \Delta n} \text{ kJ mol}^{-1}, \quad (1)$$

for each pair of unit charges separated by distance Δn , as measured in lattice units. The value of the dielectric constant (ϵ) to use in biomolecular simulations is a topic of consternation and debate [2]. We used ab initio electronic structure calculations, described below, to determine an appropriate value for the present application.

For each possible placement of mobile charge(s), or “charge configuration,” the total energy U is computed from equation 1 by summing over all pairs of charges. When $z=2$, as shown in Figure 1, there is one pair of charges (mobile-immobile), and when $z=3$ each there are three pairs of charges (mobile-mobile and two mobile-immobile). The Boltzmann probability for a charge configuration is proportional to $\exp(-U/RT)$, where R is the gas constant and T is an effective temperature. Each lattice site participates in one or more charge configurations (one configuration when $z=2$ and $N-2$ configurations when $z=3$). For each lattice site, the probabilities for all relevant charge configurations are summed. These sums are finally normalized together so their total is equal to the number of mobile charges. This yields a “mobile proton density” function, $\rho(n)$, along the lattice. This is simply the probability that a mobile proton will be located at lattice site n (disregarding correlations among mobile protons). At low temperature only the most stable charge configurations are populated, while at high temperature the charge distribution is more diffuse and delocalized. Assuming that fragmentation is charge-directed,

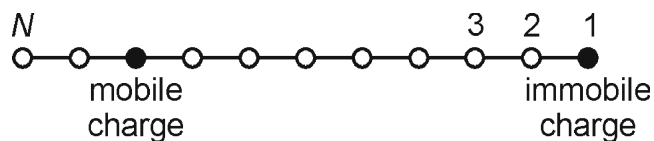


Figure 1. One-dimensional lattice model for a doubly protonated, ideal tryptic peptide

as in the popular “mobile proton” model [3], the probability $P(n)$ of fragmentation at a particular site is equal to the mobile proton density at that site, divided by the number of mobile protons. That is, $P(n) = \rho(n)/(z-1)$. Thus, the lattice model yields predictions about the length distribution of fragments. However, the model says nothing about the detailed chemistry or the disposition of charges between the fragments (e.g., $b^+ + y^+$, neutral $b + y^{++}$, $a^+ + x^+$, etc.).

Experimental Data and Computational Methods

Experimental data were taken from the NIST database of peptide MS/MS spectra [4]. For comparison with the lattice model, only ideal tryptic peptides with a C-terminal arginine, and no other basic residues (Arg, Lys, or His), were chosen. Spectra lacking y^{++} ions were discarded. “Site fragmentation propensities” were derived from each tandem mass spectrum. For backbone site n , the experimental fragmentation propensity $X(n)$ is defined as

$$X(n) = \frac{I(y_n^{++})}{\sum_{i=2}^{N-1} I(y_i^{++})} \quad (2)$$

where $I(y_n^{++})$ refers to the intensity of the y_n^{++} ion. For each number of residues (N) and total charge (z), the $X(n)$ distributions from all spectra were combined to produce average fragmentation propensities.

Ab initio calculations on neutral and protonated polyglycines (Gly $_n$, $n=5, 10, 15$) were done using HF/3-21G geometries and energies. Using more expensive geometries for neutral and cationic Gly $_5$, up to B3LYP/6-31+G(d,p), and more expensive single-point energies, up to B3LYP/6-311++G(d,p), gave very similar results so was deemed unnecessary. Peptide geometries were constrained to C_s symmetry to keep them linear, for comparison with the 1D lattice model. Thus, the optimized structures do not represent stationary points on their potential energy surfaces. The Coulombic effect at any backbone site was taken to be equal to the effect of the existing charge(s) on the energy for protonating that backbone site. For example, the energy to place a proton on N $_4$ in Gly $_5$ is taken as $E(\text{Gly}_5\text{H}_{\text{N}_4}^+) - E(\text{Gly}_5)$, where the subscript indicates that the proton has been placed on N $_4$. The corresponding protonation energy in the presence of the fixed, “spectator” charge is $E(\text{H}_{\text{N}_1}^+\text{Gly}_5\text{H}_{\text{N}_4}^+) - E(\text{H}_{\text{N}_1}^+\text{Gly}_5)$. The difference between these two protonation energies is ascribed to Coulombic repulsion in the dication.

Results and Discussion

Ab Initio Model Refinement

To mimic linear tryptic peptides, the “fixed” proton on polyglycines was placed on the basic N-terminus (which is, therefore, numbered “1” for the ab initio calculations), in the

plane of symmetry. The effect of this fixed, spectator charge on the protonation energy of each backbone site was interpreted as a purely electrostatic effect in the resulting doubly-charged peptide. Numerical results are compiled in Table 1 for Gly₅, Gly₁₀, and Gly₁₅. The repulsion energies are insensitive to the total chain length and are nearly equal for protonation at N_{*m*} (the nitrogen atom of the *m*th residue) and at O_{*m-1*}. This equality makes sense because those two atoms are within the same amide moiety, approximately the same distance from the fixed charge.

In a computational study of Coulombic repulsion within doubly protonated α,ω -diamines and doubly-deprotonated α,ω -diols, Gronert found that the effective dielectric constant was approximately 0.85 [5]. This seemingly unphysical value was attributed to charge delocalization, which brings the charge centers closer than would be expected from a simple Lewis structure. Fitting the data of Table 1 to equation 1 yields values $\epsilon \approx 0.86$ and 0.88 for protonation on backbone N and O atoms, respectively. Although this is the same as Gronert's result, it is more surprising because in these molecules, unlike the α,ω -diamines, charge delocalization could move the charges farther apart instead of bringing them closer. We adopt $\epsilon = 0.86$ in the Coulombic lattice model.

Averaged Intensities of y^{++} Ions

As an example, for $N=14$ (a tetradecapeptide) and $z=2$ (a doubly-charged peptide), the symbols in Figure 2 show the experimental fragmentation propensity, $X(n)$, averaged over 816 ideal tryptic peptides. The solid curve shows the mobile proton density predicted by the 1D lattice model. As expected, the mobile proton is localized near the N-terminus because it is repelled by the fixed positive charge at the C-terminus. As discussed above, fragmentation is assumed to be charge-directed, so the mobile proton density at a lattice site corresponds to the fragmentation propensity at that site.

Table 1. Coulombic Repulsion Energies (kJ mol⁻¹) Derived from HF/3-21G Calculations on Doubly-Protonated, Linear Polyglycines, Gly_{*n*}

Residue ^a	NH ⁺			OH ⁺		
	G ₅	G ₁₀	G ₁₅	G ₅	G ₁₀	G ₁₅
1				402	401	401
2	415	414	414	237	235	235
3	241	241	241	160	159	159
4	162	161	161	120	117	117
5	120	118	118	98	91	91
6		92	92		74	74
7		75	74		63	62
8		63	63		54	54
9		54	54		48	47
10		48	47		43	42
11			42			38
12			38			35
13			35			32
14			32			29
15			29			28

^a The fixed charge is on N₁, the terminal amine group.

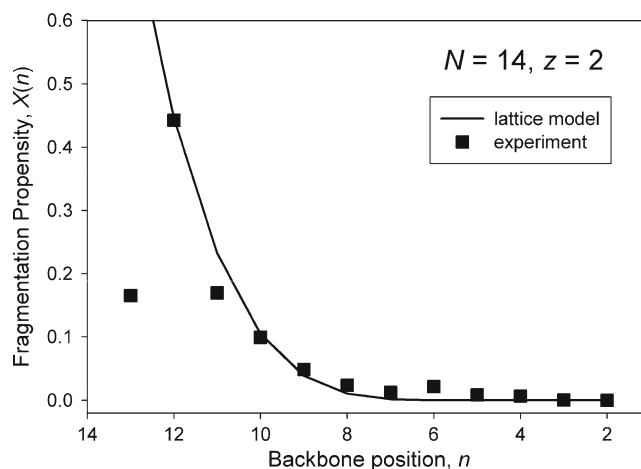


Figure 2. Comparison of lattice model with experimental y^{++} ion intensities from 816 doubly-charged tetradecapeptides

The model and experimental trends are similar, with the exception of y_{13}^{++} . For y_{13}^{++} ions, the average experimental fragmentation probability is much less than for the y_{12}^{++} ions. In contrast, the model predicts that y_{13}^{++} should have the greatest intensity (0.77 on the scale of Figure 2). This disagreement is the biggest discrepancy between the model and the experimental data. A y_{13}^{++} ion corresponds to protonation at lattice position $n=13$ and loss of the N-terminal residue as a neutral b_1 fragment or its mass-equivalent. We attribute the low y_{13}^{++} intensities to the relative instability of the companion b_1 fragments, which are too small to form stable oxazolone or diketopiperazine structures [6]. Acknowledging this discrepancy, the model was fitted to the experimental points excluding y_{13}^{++} and using temperature as the fitting parameter; this is the solid curve shown in Figure 2. The resulting effective temperature is $T=732$ K. In experimental studies, effective temperatures in the range 500 to 1100 K are typical [7].

In the triply-charged case ($N=14$, $z=3$), for which the tandem spectra are often harder to interpret, Figure 3 compares the Coulombic model with averaged experimental data for 143 peptides. As before, the experimental abundance of y_{13}^{++} is low (the model predicts a relative abundance of 0.32 on the scale of Figure 3); here it corresponds to loss of a b_1^+ ion. In this case, the best fit to the experimental data (shown in Figure 3), again excluding y_{13}^{++} from the fit, is for $T=4344$ K. This unphysical value presumably results from effects that are missing from the simple lattice model and that tend to diffuse the mobile proton distribution, such as peptide folding, charge-remote fragmentation, and variability in the kinetic barrier for dissociation.

Another possible explanation of the weaker agreement for $z=3$ could be a different fragmentation mechanism. In particular, a triply charged peptide could fragment directly to (b^+, y^{++}) or could initially form (b^{++}, y^+) , followed by a proton transfer to produce y^{++} . Similarly, initial fragmentation could form y^{+3} , followed by secondary fragmentation of this shorter ion to give y^{++} . However, in the experimental spectra we did not observe significant b^{++} or y^{+3} ions,

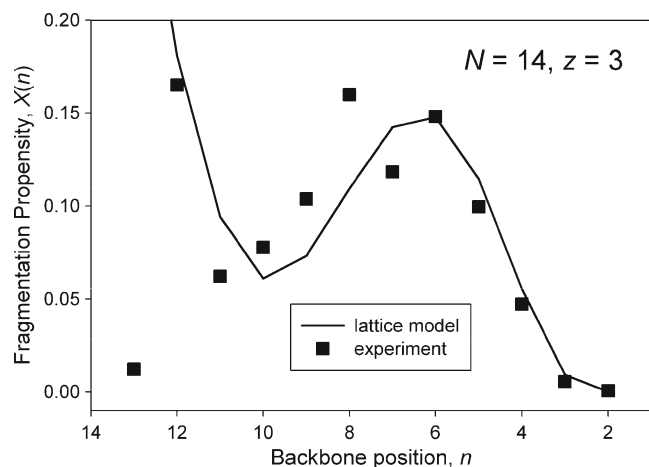


Figure 3. Comparison of lattice model with experimental y^{++} ion intensities from 143 triply-charged tetradecapeptides

suggesting that these alternative mechanisms are not important.

In Figures 2 and 3, both experimental fragmentation probabilities, $X(n)$, and mobile charge densities from the model, $P(n)$, are plotted as functions of cleavage position, n . A more concise comparison is to plot $X(n)$ and $P(n)$ against each other directly. Since we have already established that the largest ions, y_{N-1}^{++} , do not conform to the lattice model, they are omitted from this comparison, i.e., $P(n)$ is renormalized as in Figures 2 and 3 so that $\sum_{n=2}^{N-2} P(n) = \sum_{n=2}^{N-2} X(n)$. The lattice model was fitted to averaged experimental data for each (N, z)

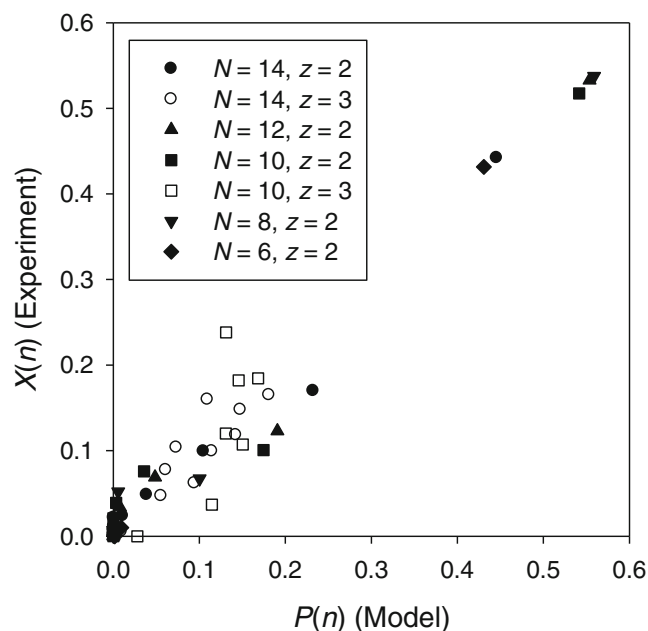


Figure 4. Comparison of fragmentation propensity (from aggregated observations) with normalized charge density as computed using the model parameters in Table 2. The largest ion, y_{N-1}^{++} , is excluded from the analysis

Table 2. Parameters for Data Shown in Figure 4: Peptide Length, Charge, Number of Peptides, Fitted Effective Temperature

N	z	No. of spectra	T/K
14	2	816	732
14	3	143	4344
12	2	842	684
10	2	502	1104
10	3	8	16210
8	2	863	1526
6	2	72	2323

combination; the relevant parameters are summarized in Table 2. The resulting comparison is shown in Figure 4. The overall agreement appears reasonable, but is worst for the series with the highest charge density, $N=10$ and $z=3$, for which only eight experimental spectra are available. For predictive purposes, the fitted temperatures in Table 2 can be reproduced fairly well by the empirical expression

$$T \approx cN^{-s}, \quad (3)$$

where $c=39400$ K and $s=1.57$ when $z=2$ and $c=1.33 \times 10^8$ K and $s=3.91$ when $z=3$.

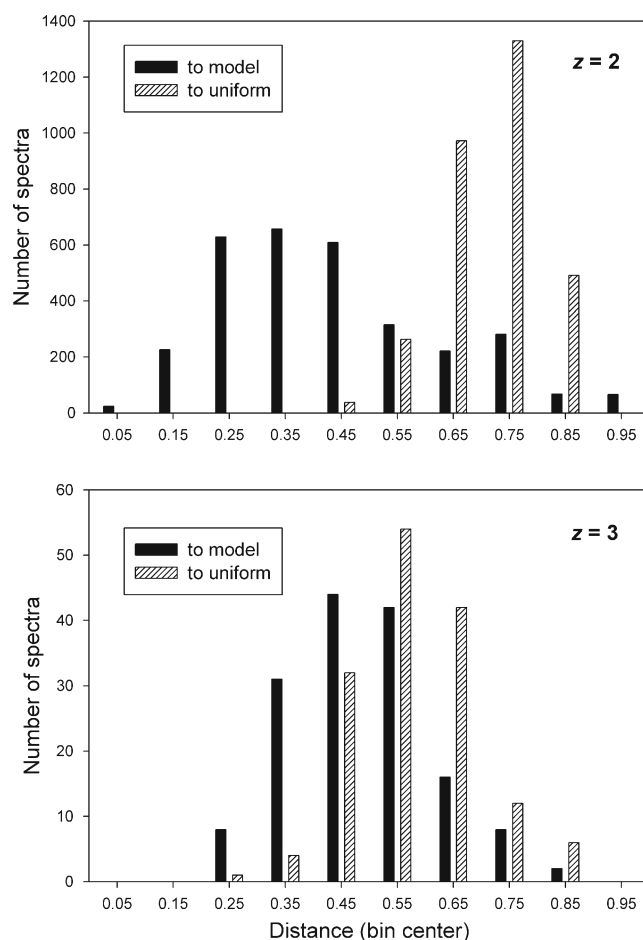


Figure 5. Histograms of distances, D , of individual y^{++} spectra from the fitted lattice model (black bars) and from a uniform distribution (gray bars) for charge states $z=2$ (top) and $z=3$ (bottom)

Variability in y^{++} Intensities

The simple Coulombic model provides a semiquantitative description of relative y^{++} intensities when averaged over many spectra. However, individual spectra may show wide deviations from the average (see Supporting Information). To provide quantitative comparisons, a means is needed to measure the difference between spectra, or between a spectrum and the predictions of the model. Here we use the Jensen-Shannon divergence metric, D [8]. For discrete probability distributions P and Q over the same set of outcomes, this metric is defined by

$$D^2 = \sum_i \left(p_i \log \frac{2p_i}{p_i + q_i} + q_i \log \frac{2q_i}{p_i + q_i} \right). \quad (4)$$

We have chosen the base of the logarithm to be 4, so that the range of D is [0, 1].

As two points of reference, we choose (1) the probability distribution from the fitted lattice model and (2) a uniform distribution. For each experimental spectrum, we compute its distances, D , to the lattice model (parameters from Table 2) and to the uniform distribution. As above, the y_{N-1}^{++} ions are excluded. The distribution of distances depends more strongly on the charge state than on the length of the parent peptide. Figure 5 shows histograms of the distance distributions both for doubly charged and for triply charged peptide ions. For $z=2$, the experimental spectra are dramatically closer to the lattice model than to the uniform distribution. This indicates that the Coulombic model is substantially better than the uniform distribution. For $z=3$ the difference is smaller, but still clear. For both charge states, there is a fairly wide distribution of distances. This distribution may be attributed to chemical (i.e., sequence-dependent) variations in fragmentation propensities.

Prior Work

Neta and Stein recently published a detailed analysis of y^{++}/y^+ ratios [9]. Because their data set was so large, they were able to analyze differences by residue pairs, as well as other chemical effects. However, they did not analyze y^{++} distributions in isolation, as we have done here. We are unaware of any published analyses of y^{++} ion distributions, although several earlier models for peptide ion fragmentation have included Coulombic effects. These models were generally motivated by the observation, now well known, that heavily protonated ions dissociate more easily than ions with smaller charge [10, 11]. Rockwood et al. considered a simple one-dimensional, Coulombic model in which the charges are fixed in uniformly spaced positions [12]. They predicted that fragmentation, in general, should be favored near the ends of the peptide for low charge states and near the middle for high charge states. Zhang developed a detailed, kinetics-based model of peptide fragmentation that accommodates multiple charges, but Coulombic effects were not specifically of interest [13]. Relevant results

have also been obtained in non-peptide systems. In particular, Cerda et al. found that multiply protonated polyethylene glycols fragment preferentially at approximately uniformly spaced intervals along the oligomer, indicating that it is not the Coulombic repulsion energy, but the Coulomb-directed location of the protons that determines the cleavage sites [14]. This is similar to the present model for peptide ions. In small amine dendrimers, de Maaijer-Gielbert et al. found that increased protonation does not facilitate collisional fragmentation [15]. This striking contrast with peptide ion fragmentation was ascribed to a reaction mechanism that does not require proton mobility.

Conclusions

A simple Coulombic model can rationalize average intensities of y^{++} ions in the MS/MS spectra of ideal tryptic peptides carrying two or three charges, except for the largest such ion, y_{N-1}^{++} , which is much less abundant than expected from the model. The best fit between averaged data and the model is obtained for effective temperatures that are sometimes unphysically high, presumably because of neglected effects that spread out the charge distribution.

The distributions of b^+ and y^+ ions do not appear similar to those of y^{++} ions and do not follow the Coulombic model presented here. This suggests that there are different reaction mechanisms for the competing y^{++} and (b^+ , y^+) pathways.

Acknowledgments

This work grew from the suggestion by Dr. S. E. Stein that we investigate the fragmentation of triply charged peptide ions. Some of the computations were done on the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>). The authors thank the referees for several helpful suggestions.

References

1. Eidhammer, I., Flikka, K., Martens, L., Mikalsen, S.-O.: Computational Methods for Mass Spectrometry Proteomics. John Wiley and Sons, Chichester (2007)
2. Mellor, B.L., Cortes, E.C., Busath, D.D., Mazzeo, B.A.: Method for estimating the internal permittivity of proteins using dielectric spectroscopy. *J. Phys. Chem. B* **115**, 2205–2213 (2011)
3. Dongré, A.R., Jones, J.L., Somogyi, A., Wysocki, V.H.: Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. *J. Am. Chem. Soc.* **118**, 8365–8374 (1996)
4. Stein, S. E., Rudnick, P. A., Eds. *NIST Libraries of Peptide Tandem Mass Spectra, H. Sapiens, Ion Trap*, built Feb. 4, 2009; National Institute of Standards and Technology, www.peptide.nist.gov, accessed June 17, 2010.
5. Gronert, S.: Coulomb repulsion in multiply charged ions: A computational study of the effective dielectric constants of organic spacer groups. *Int. J. Mass Spectrom.* **187**, 351–357 (1999)
6. Harrison, A.G.: To B or not to B: The ongoing saga of peptide B ions. *Mass Spectrom. Rev.* **28**, 640–654 (2009)
7. Meot-Ner, M., Somogyi, A.: A thermal extrapolation method for the effective temperatures and internal energies of activated ions. *Int. J. Mass Spectrom.* **267**, 346–356 (2007)

8. Endres, D.M., Schindelin, J.E.: A new metric for probability distributions. *IEEE Trans. Inf. Theory* **49**, 1858–1860 (2003)
9. Neta, P., Stein, S.E.: Charge states of γ ions in the collision-induced dissociation of doubly charged tryptic peptide ions. *J. Am. Soc. Mass Spectrom.* **22**, 898–905 (2011)
10. Busman, M., Rockwood, A.L., Smith, R.D.: Activation-energies for gas-phase dissociations of multiply charged ions from electrospray ionization mass-spectrometry. *J. Phys. Chem.* **96**, 2397–2400 (1992)
11. Ishikawa, K., Nishimura, T., Koga, Y., Niwa, Y.: Role of coulomb energy in promoting collisionally activated dissociation of multiply-charged peptides formed by electrospray-ionization. *Rapid Commun. Mass Spectrom.* **8**, 933–938 (1994)
12. Rockwood, A.L., Busman, M., Smith, R.D.: Coulombic effects in the dissociation of large highly charged ions. *Int. J. Mass Spectrom.* **111**, 103–129 (1991)
13. Zhang, Z.: Prediction of low-energy collision-induced dissociation spectra of peptides with three or more charges. *Anal. Chem.* **77**, 6364–6373 (2005)
14. Cerda, B.A., Breuker, K., Horn, D.M., McLafferty, F.W.: Charge/radical site initiation versus coulombic repulsion for cleavage of multiply charged ions. Charge Solvation in Poly(Alkene Glycol) Ions. *J. Am. Soc. Mass Spectrom.* **12**, 565–570 (2001)
15. de Maaier-Gielbert, J., Gu, C.G., Somogyi, A., Wysocki, V.H., Kistemaker, P.G., Weeding, T.L.: Surface-induced dissociation of singly and multiply protonated polypropylenamine dendrimers. *J. Am. Soc. Mass Spectrom.* **10**, 414–422 (1999)