J. Am. Soc. Mass Spectrom. (2011) 22:2222–2231
DOI: 10.1007/s13361-011-0241-6

RESEARCH ARTICLE

N-Protonated Isomers as Gateways to Peptide Ion Fragmentation

Fredrik Haeffner,[1] John K. Merle,[2] Karl K. Irikura[1]

[1]Computational Chemistry Group, Chemical and Biochemical Reference Data Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8320, USA
[2]Department of Chemistry, Winston-Salem State University, Winston-Salem, NC 27110, USA

Abstract

According to the popular "mobile proton model" for peptide ion fragmentation in tandem mass spectrometry, peptide bond cleavage is typically preceded by intramolecular proton transfer from basic sites to an amide nitrogen in the backbone. If the intrinsic barrier to dissociation is the same for all backbone sites, the fragmentation propensity at each amide bond should reflect the stability of the corresponding N-protonated isomer. This hypothesis was tested by using ab initio and force-field computations on several polyalanines and Leu-enkephalin. The results agree acceptably with experimental reports, supporting the hypothesis. It was found that backbone N-protonation is most favorable near the C-terminus. The preference for C-terminal N-protonation, which is stronger for longer polyalanines, may be understood in terms of the well known "helix macrodipole" in the corresponding helical conformations. The opposite stability trend is found for peptides constrained to be linear, which is initially surprising but turns out to be consistent with the reversed direction of the macrodipole in the linear conformation.

Key words: CID, Conformation, Fragmentation, Gas phase, Helix, Leu-enkephalin, Macrodipole, Mass spectrometry, Mobile proton model, Peptide, Polyalanine, Proton affinity, Quantum chemistry

Introduction

In the rapidly expanding field of proteomics, tandem mass spectrometry (MS/MS) has become an invaluable tool for protein identification. Developing methods to correctly assign the MS/MS spectra is pivotal, so understanding and predicting gas-phase fragmentation of peptide ions is of growing interest [1]. Fragmentation is thought to proceed by both charge-remote and charge-directed mechanisms. Charge-remote fragmentation is invoked to explain the characteristic cleavage near certain residues, such as aspartic acid and histidine, when the number of ionizing protons does not exceed the number of strongly basic residues [2–5]. The ionizing protons are held tightly by the basic side chains, and the proton necessary for directing the cleavage reaction is supplied by the Brønsted-Lowry acidic side-chain adjacent to the peptide bond.

Charge-directed fragmentation is rationalized using the mobile proton model [6, 7], in which the ionizing proton is transferred from a basic site (such as an arginine side chain or the N-terminus) to the point of fragmentation. This intramolecular proton transfer is endothermic. It occurs only upon activation, as during collision-induced dissociation (CID). The added energy is said to "mobilize" the proton, allowing it to move to less-stable sites along the peptide backbone and promote cleavage. Fragmentation spectra are usually rich in sequence information [8]. In tryptic peptides, which bear a C-terminal lysine or arginine residue, the process can be facilitated by formation of a "salt bridge" between the

Electronic supplementary material The online version of this article (doi:10.1007/s13361-011-0241-6) contains supplementary material, which is available to authorized users.

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

Received: 22 October 2010
Revised: 23 August 2011
Accepted: 24 August 2011
Published online: 24 September 2011

basic C-terminal side chain and a carboxylic acid moiety, either at the C-terminus or on a side chain [8–11].

When the number of ionizing protons exceeds the number of strongly basic residues, less energy is needed to mobilize a proton onto the backbone. For example, singly protonated Ala_5 can be dissociated more easily than singly protonated $ArgAla_4$ because it lacks the basic Arg side-chain, which sequesters the proton strongly [7]. Wysocki and coworkers noted that although the carbonyl oxygen of a peptide linkage is more basic than the nitrogen site, protonation at oxygen strengthens the peptide bond, while protonation at nitrogen weakens it, thus activating it for cleavage [7]. Thus, backbone cleavage is believed to occur only where a proton resides on an amide nitrogen atom.

Popular software for extracting sequence information from peptide MS/MS spectra, such as SEQUEST [12] and Mascot [13], assumes that peptides fragment with equal probability along the peptide backbone. This is a crude approximation, as demonstrated by Harrison and co-workers for polyalanines [14, 15]. For example, the b_4^+ ion (see Figure 1 for nomenclature) is observed as the predominant peptide fragment from Ala_5H^+. Paizs and Suhai investigated this system theoretically, computing both thermochemistry and kinetics for the competing fragmentation reactions [16]. They found that both basicities and activation barriers favor cleavage of the peptide bond near the C-terminus, supporting the experimental results [14, 15]. Proton migration from the N-terminus (N_1, the most basic site in the molecule) to the n^{th} peptide moiety leads either to the $b_{n\text{-}1}^+$ or the $y_{6\text{-}n}^+$ ion. No chemical explanation was presented for the favorable energetics of protonation near the C-terminus.

Ample computational studies, particularly by Paizs and coworkers, have demonstrated successfully that the computational theory of (pseudo-thermal) peptide fragmentation is the same as for any other thermal reaction in organic chemistry [1]. Because of the larger size of peptides, computations are more difficult and expensive than for small organic compounds, but conventional thermochemical kinetics still applies [17]. Indeed, thermochemistry and kinetics have been used as the basis of a heavily parameterized method for predicting peptide ion fragmentation [18]. Thus, a predictive, ab initio theory for peptide ion fragmentation is available: quantum chemistry provides the parameters needed in detailed kinetics theories, which then yield values for the rate constants that dictate product branching fractions. Unfortunately, this theory is available only in a formal sense. Actual computations are far too slow to be included in a proteomics workflow. Shortcuts are needed.

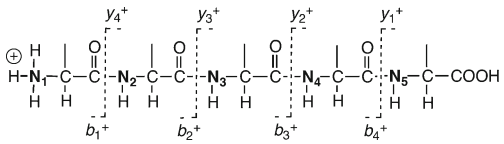


Figure 1. Singly protonated Ala_5H^+ and fragment-ion nomenclature. The ionizing proton occupies the most basic site (N-terminal amine). Upon mobilization it moves onto the peptide backbone, facilitating fragmentation

A similar challenge is found in other areas of chemistry. For example, comprehensive modeling of engineered combustion systems is desired so that computational design might increase the speed, and decrease the cost, of building new systems and accommodating new fuels. Progress is largely through improved shortcuts and approximations to more rigorous methods [19]. The most relevant for the present study are those involving the estimation of rate constants for elementary reactions [20, 21]. Linear free-energy relationships are often effective [22]. In particular, trends in rate constants are often the same as trends in thermochemistry.

In the present study, we make the simplifying assumption that all amide bonds break with equal ease, provided that they are first protonated on the nitrogen atom. This approximation avoids computations of transition structures, which are often difficult and time-consuming. In effect, we assume that the trend in cleavage rates along the backbone is the same as the trend in proton affinities (at nitrogen). This is similar to a hypothesis presented by Savitski et al. that cleavage propensity is proportional to the basicity of backbone oxygen atoms as inferred from a crystallographic database [23]. However, in our computations, the basicity depends upon sequence and conformation, and we attempt to predict fragmentation of individual peptides, not average behavior.

Attention here is limited to several non-tryptic peptides: singly protonated polyalanines ranging in size from Ala_3H^+ to $Ala_{11}H^+$, and protonated Leu-enkephalin (Tyr-Gly-Gly-Phe-Leu), one of the most studied peptides in mass spectrometry [24]. Relative proton affinities of backbone nitrogen sites are determined, and are discussed in light of previous experimental and theoretical reports [14–16]. We also suggest a structural reason for the trends in proton affinities.

Computational Details[1]

Locating low-energy minima of large and flexible molecules, such as peptides, is a challenging task and a topic of active research. To identify low-energy conformations of singly protonated peptides, we developed a procedure using a molecular mechanics-based Monte Carlo method with subsequent quantum-mechanical structure refinement and energy calculations of selected structures. Substantially similar procedures have been developed and used by others [8, 25–27]. Singly protonated structures for Ala_nH^+ (n=3–

[1]Certain commercial materials and equipment are identified in this paper in order to specify procedures completely. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

11) and protonated Leu-enkephalin were constructed by using an in-house Perl script. For each peptide, *n* constitutional isomers were generated by placing the proton on each nitrogen atom in turn. Each isomer was subjected to a conformational search using the Monte-Carlo Multiple Minima (MCMM) method as implemented in the MacroModel program [28, 29]. The OPLS2005 force-field [30] was used for these calculations.

To determine a reasonable number of MCMM steps, an initial round of calculations was carried out on Ala_5H^+ (protonated at N_3; see Figure 1 for labeling) with the number of MCMM steps at several values between 1000 and 500,000. No new low-energy minima were obtained when the number of steps exceeded 20,000. Based on this result, the number of steps was chosen to be 50,000. Conformers were discarded if they were at least 200 kJ/mol above the most stable conformer. All other parameters controlling the calculations were set to their default values. This procedure resulted in 600 to 12,000 conformers for each isomer of each chain length.

Since our model is based upon thermodynamics, it assumes that all degrees of freedom are equilibrated. In particular, the mobile proton samples all possible sites in proportion to their Boltzmann weights [31]. It is well known that the *cis-trans* isomerization barrier in amides is greatly reduced by N-protonation [32]. More quantitatively, for formamide we find from frozen-core MP2/6-311+G(d,p) calculations that the barrier is 68 kJ mol^{-1} when neutral, 172 kJ mol^{-1} when O-protonated, and 2 kJ mol^{-1} when N-protonated. Thus, during the conformational searches, amide C–N bonds and carboxylic hydroxyl groups were allowed to adopt both *cis*- and *trans*-conformations. The ramifications of this assumption are discussed in the following Section.

OPLS2005 is a well established force field for generating geometries and energies of peptides and proteins. However, it was not parameterized for gas-phase ions, reducing our confidence in its predictions. Therefore, further structural and energetic refinement was done using ab initio methods. For each constitutional isomer, the large set of conformations from the MCMM/OPLS2005 calculations was pruned in stages. Each set was sorted by OPLS2005 energy. All structures below 41.8 kJ/mol, plus 100 structures uniformly distributed between 41.8 and 200 kJ/mol, were subjected to single-point energy evaluations using the Hartree-Fock method and the 3-21G basis set (HF/3-21G). We found the HF/3-21G energies to be well correlated with the OPLS2005 energies (0.98>linear correlation coefficients>0.86; see Supporting Information), increasing our confidence in the OPLS2005 energies.

The pruned set was sorted by HF/3-21G energy and pruned further to include only conformers with energies below 50 kJ/mol, discarding conformations that were structurally similar. Two conformations were considered similar if none of their independent dihedral angles (including heavy atoms and polar hydrogens) differed by more than 50°. The geometries of the selected conformers were fully optimized at the HF/3-21G level. To increase confidence in the energies (e.g., by considering electron correlation, including van der Waals interactions), each resulting structure was finally subjected to a single-point RI-LMP2/cc-pVDZ energy calculation (localized-orbital, second-order perturbation theory with the resolution-of-the-identity approximation). Relative proton affinities were computed as the difference in RI-LMP2/cc-pVDZ energies. No thermal or zero-point corrections were made in estimating the relative proton affinities, to avoid the expensive computation of vibrational frequencies. Since we are comparing similar isomers, neglecting such effects is a reasonable approximation [33, 34]. For the same reason, we neglect entropic effects and discuss proton affinities (enthalpies) instead of gas-phase basicities (Gibbs energies). The Gaussian software package [35] was used for the Hartree-Fock calculations and the QChem package [36, 37] for the RI-LMP2 calculations.

As a test of the conformational predictions obtained using OPLS2005, a systematic (but not exhaustive) search was done for Ala_3H^+ (protonated on N_2) on the HF/3-21G potential energy surface. There were 5,832 initial structures generated on a coarse grid in (independent) dihedral angle space (120° or 180° increments, as appropriate). Each initial structure was energy-minimized while constrained to its grid point. A subsequent, full optimization was done without any constraints. Among the resulting structures, two were considered identical if their electronic energies, nuclear repulsion energies, geometric-mean rotational constants, and dipole moments were the same within certain tolerances (10^{-5} hartree, 0.15 hartree, 0.001 GHz, and 0.03 Debye, respectively). Finally, structures were discarded if any bond had stretched more than 50%, indicating isomerization. There were 1,065 distinct conformers found. The most stable conformer was the same as that found using the force-field-based protocol described above, sustaining our confidence in the OPLS2005-based protocol.

Results and Discussion

The purpose of this study is to test the hypothesis that the fragment ion distribution reflects the relative stabilities of the corresponding N-protonated isomers. The hypothesis is illustrated by Figure 2. Collisional or radiative activation raises the effective temperature of the ion ensemble, creating small populations of high-energy isomers, including some that are protonated on backbone N atoms. The N-protonated isomers have weakened amide bonds and are poised for fragmentation. As stated above, the energetic barrier to dissociation, denoted E^* in Figure 2, is presumed equal for all such isomers. This is the simplifying approximation that permits us to avoid explicit computation of transition structures and barrier heights. Thus, the rate of dissociation at each peptide bond is predicted from the stability of the corresponding N-protonated isomer. There are probably intermediates involved in proton migration, such as O-protonated isomers, but they are not of concern here; final

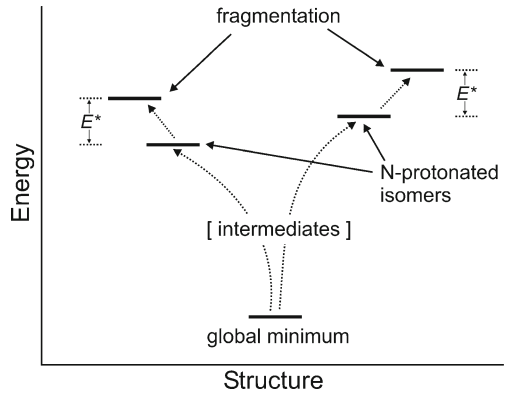


Figure 2. The fragmentation hypothesis examined in this study derives from the well-known mobile proton model, with additional, simplifying assumptions

dissociation is presumed to be the rate-limiting step that dictates the product branching fractions.

For each peptide, Ala_nH^+ (n=3–11), the energy was computed for the most stable conformer of each constitutional (N-protonated) isomer using the RI-LMP2/cc-pVDZ//HF/3-21G computational model. Energies are shown in Table 1, relative to the isomer protonated at the terminal amine. There is a clear trend: the most favorable site of amide nitrogen protonation is that neighboring the C-terminus. The one exception is Ala_7H^+, for which protonation at N_5 is more favorable than at N_7. There is no obvious reason for n=7 to be exceptional, and we made extra, unsuccessful efforts to find an N_7-protonated isomer of greater stability. We believe this exception does not indicate a real difference, but instead reveals a weakness in our computational protocol.

Except for the N-terminus, all the backbone nitrogen sites are the same functional group and would be expected to have equal proton affinities. Thus, the surprisingly large differences in proton affinity must arise from conformational effects. All HF/3-21G-optimized structures were inspected visually. Interestingly, most structures with the C-terminal amide nitrogen protonated adopt an α-helix (a right-handed helix, where the N–H group of one amino acid forms a hydrogen bond with the C=O group of the amino acid positioned four residues nearer the N-terminus, so-called i+4 → i hydrogen bonding). In contrast, all isomers protonated at the N-terminus (that is, the amino group) favor globular conformations, with the protonated amine solvated by carbonyl oxygens. Thus, a conformational change is required to mobilize the proton from the N-terminus to the C-terminal amide nitrogen. It is known that globule-helix conformational transitions of peptides are rapid [38, 39]. Seemingly at odds with our conclusion, experimental and molecular dynamics (MD) studies have suggested that no helical structures should be expected from protonated $Ala_{15}H^+$ [40]. However, those studies were restricted to the predominant, low-energy isomers. The experimental study was carried out at room temperature (undetectable populations of high-energy isomers) and the MD calculations only considered N-terminal protonation (protons are immobile in a classical MD simulation).

Helix formation is not surprising in these systems. It is known that alanine polypeptides form unusually stable α-helices in the gas phase when they contain a positively charged lysine at the C-terminus [41–44]. Most dramatically, Jarrold and co-workers discovered that $Ala_{15}LysH^+$ and $Ala_{20}LysH^+$ maintain helical conformations up to the fragmentation temperature of 725 K [44]. Quantum chemical computations by Dannenberg and coworkers showed that a positive charge placed at the C-terminus stabilizes polyalanine helices [45]. The stability of the α-helix and its interactions with ions has been rationalized in terms of charge-dipole interactions and polarization, sometimes termed collectively the "macrodipole" [46–50]. Thus, protonation near the C-terminus and helix formation are mutually stabilizing. However, the influence of peptide conformation on the fragmentation pattern has been discussed in the literature only to a minor extent [51, 52].

It should be emphasized that most of the experimental peptide population has the proton located at the most stable position, the N-terminus, at all internal energies (i.e., at all effective temperatures), even at the threshold of dissociation (barring entropic factors). Other low-energy forms, such as many O-protonated isomers, will also be substantially populated. The higher-energy, N-protonated isomers represent only a small fraction of the ion population. For

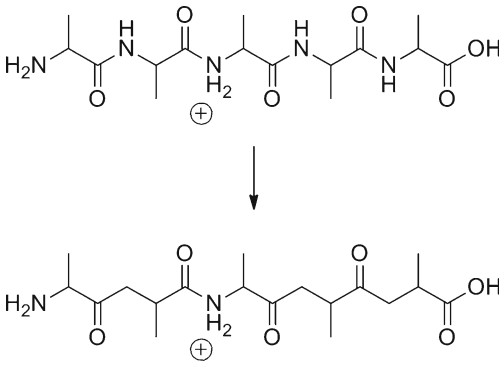
Table 1. Relative RI-LMP2/cc-pVDZ//HF/3-21G Energies (in kJ/mol) of Backbone N-protonated Isomers of Ala_nH^+ (n=3–11) and of Leu-enkephalin; N_1 is the Terminal Amino Group. The Most Favorable Peptide Site is in Boldface

	n=11	10	9	8	7	6	5	4	3	Leu-enk
N_1	0	0	0	0	0	0	0	0	0	0
N_2	114	97	85	90	100	82	98	100	74	87
N_3	125	106	103	97	108	92	105	95	**71**	87
N_4	126	89	112	110	107	83	107	**69**		90
N_5	133	87	105	98	**75**	75	**61**			**64**
N_6	108	98	108	82	98	**63**				
N_7	103	90	83	92	87					
N_8	98	58	84	**70**						
N_9	91	59	**55**							
N_{10}	73	**28**								
N_{11}	**32**									

example, using relative energies from Table 1, at an effective temperature of 800 K the Boltzmann weight for the N_5-protonated isomer of Ala_5H^+ is only 10^{-5}. These rare isomers are unlikely to be detected in experiments such as ion mobility spectrometry. Nevertheless, they are of primary importance in fragmentation because they are the gateway isomers for backbone cleavage, according to the mobile proton model.

To explore the effect of conformation upon proton affinity, we investigated linear, extended conformations of $Ala_{11}H^+$ isomers. These structures were generated by constraining the peptide backbone to remain planar during the geometry optimizations. Surprisingly, the trend in backbone proton affinities is the opposite of that described above. The energy increases as the proton is moved from the amide nitrogen near the N-terminus (N_2) towards the C-terminus (N_{11}) (see Figure 3, "linear"). The relative energies of fully optimized structures, from Table 1, are included in Figure 3 ("helix") for comparison. The amide C(O)–N dipoles are aligned in the extended conformation. This is similar to the helical situation but with the dipoles pointed in the opposite direction. Protonation is favored toward the N-terminus of the molecule. Fixed-geometry calculations were done with the neutral, spectator amide NH groups mutated into CH_2 groups (see Scheme 1), thus eliminating the C-N dipoles but retaining the perpendicular C=O dipoles. This change erased the stability trend (Figure 3, "linear polyketone"), confirming the effect of the macrodipole in the linear polypeptide.

The electrostatic models are summarized pictorially in Figure 4 (dipolar arrows point from positive to negative). As a numerical test, we computed the dipole moment of neutral Gly_{11} in both helical and extended conformations, at the MP2/6-31+G(d)//HF/3-21G level. For the helix, the moment is 47 D (1 D $\approx 3.336\times10^{-30}$ C m), with the negative end toward the C-terminus, while for the extended conformation the moment is 25 D, with the negative end toward the N-terminus. For comparison, the extended polyketone has a moment of only 1.2 D. Conformational stability depends upon the interaction between amide dipoles and the monopole of the proton, and upon hydrogen bonding. Dipole cooperativity (polarizability) is important in α-helices [53–56] and may also play a role in stabilizing other conformations as has been suggested, for example, in β-amyloid peptide aggregation [57].



Scheme 1. Removing the effects of spectator amide C–N dipoles

To proceed beyond the artificial polyalanine systems, we investigated protonated Leu-enkephalin, (Tyr-Gly-Gly-Phe-Leu)H^+. Our calculations (included in Table 1) indicate that protonation at N_5 is favored by 23 to 26 kJ/mol relative to the other amide nitrogens. The helical structures of protonated polyalanines, as discussed above, should not be surprising because alanine is a known alpha-helix former [58]. Leu-enkephalin does not contain alanine, but its N_5-protonated form shows a helical structure (Figure 5). (For

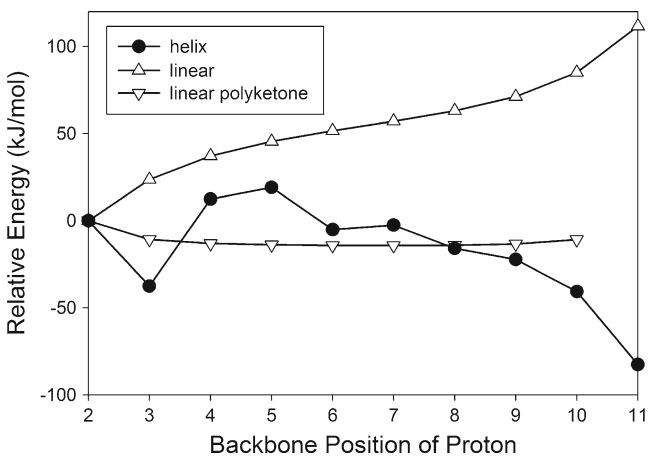


Figure 3. Relative RI-LMP2/cc-pVDZ//HF/3-21G energies (kJ/mol) of N-protonated isomers of optimized (mostly helical) and extended $Ala_{11}H^+$ and the linear polyketone analogue

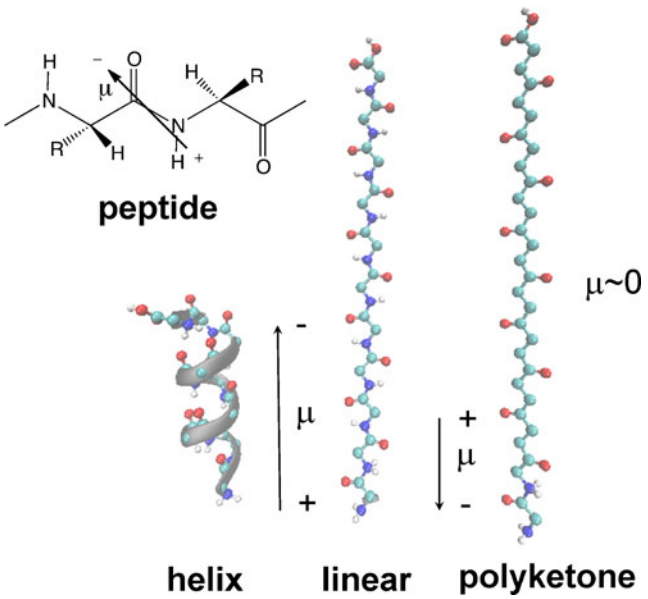


Figure 4. Macrodipole effects in helical peptides, extended peptides, and extended polyketone amides. The alignment of dipoles from individual functional groups determines the overall direction and strength of the electrostatic field

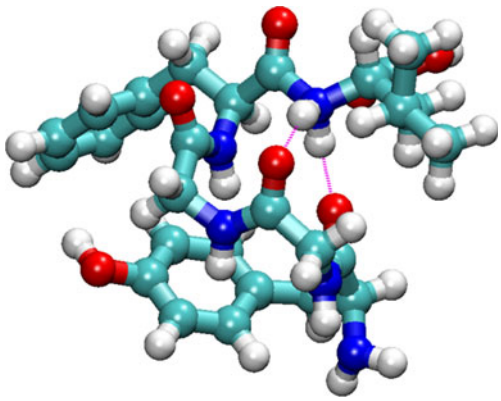


Figure 5. The most stable amide-N-protonated isomer of Leu-enkephalin-H^+

Leu-enkephalin, a large set of 100 low-energy N_5-protonated conformers was considered, instead of only 10; the conformer with the lowest energy was absent from the smaller set of 10.) The isomers protonated at N_2, N_3, and N_4 display cation-π interactions between the protonated amide and the aromatic ring of the Tyr residue [22], but these interactions evidently are not strong enough to compete with protonation at N_5.

As described above, in our computations we assumed that the mobile proton will catalyze *trans-cis* isomerization of peptide bonds, so we included both possibilities in our Monte Carlo conformational searches. However, we also considered a modified model in which all peptide bonds are restricted to the *trans* conformation. The results (see Supporting Information) differ only slightly from those of our more liberal model. In particular, the most stable conformations of the most stable isomers are all *trans*. Thus, our results reveal nothing about the facility of *trans-cis* isomerization in peptide ions.

Comparison with Experiments

Before comparing our model with experimental observations, it is necessary to point out three difficulties in making the comparison. (1) Our computations are approximate. Even if the underlying model is excellent, our numerical predictions may not be. (2) The theoretical model presented here is intended to identify the sites of fragmentation, but is mute about the ultimate location of the ionizing proton. That is, we say nothing about whether cleavage will result in a *b* ion, a *y* ion, or some other ion. The fragment with the higher proton affinity will retain the proton [1, 59], analogously to "Stevenson's Rule" in electron-ionization mass spectrometry [60], but fragment proton affinities are not part of the model presented here. (3) Experimental spectra include not only the primary product ions that result from fragmentation of the parent peptide ion, but also secondary products that result from subsequent dissociation of fragment ions. For example, *b* ions fragment sequentially to smaller *b* ions in the sequence $b_n^+ \rightarrow b_{n-1}^+ \rightarrow b_{n-2}^+ \rightarrow \ldots$ [14]. Thus, spectra depend upon the amount of energy deposited in the parent ion and the timescale of the experiment, with higher energy (and, therefore, certain instrumentation [61]) favoring smaller fragment ions. To identify the initial fragment ions, double resonance [62] is necessary to establish ion parentage. For example, a CID spectrum of Leu-enkephalin suggested that only half the primary product ions were b_4, but the correct fraction is 93%, as determined by double-resonance experiments [63]. A reasonable alternative is to measure spectra over a range of excitation energies, that is, to measure the breakdown curve. The branching fractions in the low-energy limit are likely to correspond to primary fragmentation. Unfortunately, data are seldom available from either of these experimental techniques. The best experiment for comparison with our model is threshold CID, which yields the zero-temperature energy threshold for each primary product ion (and sometimes also for secondary ions) [64–66]. However, data analysis is complex for such experiments; tripeptides are the largest peptides studied so far.

Our model, as presented above, predicts that the intensities of primary fragment ions will correspond to the stabilities of the corresponding N-protonated isomers. For equilibrium abundances or kinetic branching fractions, Boltzmann weights are more appropriate than raw energies. These are listed in Table 2, as derived from the energies in Table 1 assuming an effective temperature of 800 K. For comparison, available experimental data are summarized in Table 3. In Table 3 the fractions indicate the prevalence of fragmentation observed at each peptide bond, irrespective of

Table 2. Normalized Boltzmann Weights (T_{eff}=800 K) for Competing, Amide N-protonated Isomers of Ala_nH^+ (n=3–11) and of Leu-enkephalin; Energies are from Table 1

	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}
Ala_3	0.4	0.6								
Ala_4	0.01	0.02	0.97							
Ala_5	0.004	0.001	0.001	0.99						
Ala_6	0.04	0.01	0.04	0.1	0.8					
Ala_7	0.02	0.006	0.006	0.8	0.03	0.1				
Ala_8	0.04	0.01	0.002	0.01	0.1	0.03	0.8			
Ala_9	0.01	8×10^{-4}	2×10^{-4}	6×10^{-4}	4×10^{-4}	0.01	0.01	0.96		
Ala_{10}	3×10^{-5}	8×10^{-6}	1×10^{-4}	1×10^{-4}	3×10^{-5}	9×10^{-5}	0.01	0.01	0.98	
Ala_{11}	5×10^{-6}	8×10^{-7}	7×10^{-7}	3×10^{-7}	1×10^{-5}	2×10^{-5}	5×10^{-5}	1×10^{-4}	0.002	1.0
Leu-enk	0.03	0.03	0.02	0.93						

the identity of the resulting ion. For example, for Ala_5H^+ the value for N_2 includes b_2^+, y_3^+, and a_2^+ together. In the case of Leu-enkephalin, some ions are internal, resulting from the cleavage of two peptide bonds. For such ions, half the intensity is assigned to each peptide site. Unassigned ions are ignored. CID breakdown data are summarized in Table 3 by a value at low energy, with an inequality symbol to indicate the slope of the curve. For example, a value listed as "<15" means that the cleavage site accounts for 15% of ions observed, but with a positive slope, so that a smaller percentage is expected in the low-energy limit. When the slope is close to zero, no trend is specified in the table. No experimental data are available for Ala_9H^+, $Ala_{10}H^+$, or $Ala_{11}H^+$.

Comparing Tables 2 and 3, we consider each peptide in turn. For Ala_3H^+, the computations predict nearly equal fragmentation at N_2 and N_3, while the observations are that N_3 predominates heavily. For Ala_4H^+, the computations predict N_4 to dominate, while experiments show $N_3 > N_4 >> N_2$. For Ala_5H^+, the computations predict N_5 to dominate, in agreement with four of the five experimental reports. For Ala_6H^+, the computations predict N_6 to dominate; experi-

Table 3. Experimental Branching among Backbone Cleavage Sites

	N_2	N_3	N_4	N_5	N_6	N_7	N_8	Notes
Ala_3	0.0	1.0						a
	0.26	0.74						b
	0.0	1.0						c
Ala_4	0.0	<0.67	>0.33					a
	0.0	<0.60	>0.40					d
	0.0	0.66	0.34					e
	0.0	<0.58	>0.42					f
	0.04	<0.56	>0.39					g
	0.11	0.56	0.33					h
	0.0	0.6	0.4					c
Ala_5	0.0	0.32	0.41	0.27				i
	0.0	<0.15	<0.18	>0.67				j
	0.0	0.17	0.19	0.64				e
	0.0	<0.05	<0.17	>0.77				f
	0.0	<0.02	<0.03	>0.95				c
Ala_6	0.0	0.09	0.21	0.35	0.35			k
	0.0	0.0	0.0	<0.18	>0.82			l
	0.0	<0.02	<0.07	>0.29	>0.63			g
Ala_7	0.0	0.0	0.03	0.08	0.30	0.58		m
Ala_8	0.0	0.09	0.12	<0.04	<0.16	>0.28	0.31	g
Leu-enk	0.03	0.10	0.18	0.70				e
	0.0	0.11	<0.11	>0.78				f
	0.06	0.13	0.24	0.57				n
	0.0	<0.09	<0.19	>0.73				o
	0.0	0.0	<0.06	>0.94				p
	0.03	0.11	0.22	0.64				q
	0.0	0.0	0.02	0.98				r
	0.0	0.0	0.05	0.95				s
	0.05	0.04	0.05	0.86				t
	0.27	0.09	0.33	0.31				u
	0.0	0.0	0.03	0.97				v

[a]CID at 6 eV [14]
[b]CID at 5.8 keV [67]
[c]SORI-CID at low energy [68]
[d]In-source at 21 V [14]
[e]Metastable decay [15]
[f]CID at 5 eV [15]
[g]CID at 10 eV [69]
[h]CID at 5.2 keV [67]
[i]CID at 20 eV [14]
[j]In-source at 24 V [14]
[k]CID at 20 eV [14]
[l]In-source at 27 V [14]
[m]SORI-CID [70]
[n]CID at 15 eV from Figure 1 of ref. [24]
[o]CID at 9 eV from Figure 5 of ref. [24]
[p]CID at 8 eV from Figure 7a of ref. [24]
[q]CID in QQQ instrument, 50% survival, from Figure 8b of ref. [24]
[r]CID in QIT instrument, 50% survival, from Figure 8c of ref. [24]
[s]SORI-CID, 50% survival, from Figure 8e of ref. [24]
[t]IRMPD, 50% survival, from Figure 8f of ref. [24]
[u]SID at 30 eV, from Figure 8g of ref. [24]
[v]Primary fragmentation from double resonance [63]

ments do show N_6 predominant, but not as strongly. For Ala_7H^+, the computational results (which are suspect, as noted above) predict N_5 to dominate, but the experiment shows most fragmentation at N_7. For Ala_8H^+, the computations predict N_8 to dominate; the experiment shows $N_8 \approx N_7$. For Leu-enkephalin, the computations predict N_5 to predominate, in agreement with the experiments. Overall, the computational predictions agree with experiments well enough to support the underlying model.

For all the peptides studied here, which contain only nonpolar or weakly polar residues, the most favorable site for backbone N-protonation is that closest to the C-terminus. This isomer adopts a helical conformation and is stabilized by the corresponding macrodipole. It is tempting to speculate that this is a general phenomenon. Additional work is needed to explore the prevalence of this conformational, tautomeric pattern.

In proteomics, it is tryptic peptides that are most interesting, that is, peptides bearing either arginine or lysine at the C-terminus. However, the present study includes no tryptic peptides. Further work is necessary to determine whether the fragmentation of tryptic peptides is correlated with the energetics of N-protonated isomers, as it is for the non-tryptic peptides studied here.

Conclusions

The mobile proton model suggests that backbone fragmentation is preceded immediately by the migration of a proton to an amide nitrogen position. This suggests, in turn, that the thermodynamic stability of competing N-protonated isomers will dictate the sites of backbone cleavage. The evidence for polyalanines and Leu-enkephalin supports this hypothesis. For these systems, a helical macrodipole favors protonation adjacent to the C-terminus and predominance of b_{n-1}^+ ions.

Acknowledgments

Most of the computations were carried out on the Biowulf computer cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov). The authors thank Professor Peter Armentrout for a preprint of reference [66] and are especially grateful to the anonymous reviewers for helpful comments and suggestions.

References

1. Paizs, B., Suhai, S.: Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **24**, 508–548 (2005)
2. Bakhtiar, R., Wu, Q., Hofstadler, S.A., Smith, R.D.: Charge state specific facile gas-phase cleavage of Asp75-Met76 peptide bond in the α-Chain of human apohemoglobin probed by eletrospray ionization mass spectrometry. *Biol. Mass Spectrom.* **65**, 707–710 (1994)
3. Farrugia, J.M., Taverner, T., O'Hair, R.A.J.: Side-chain involvement in the fragmentation reactions of the protonated methyl esters of histidine and its peptides. *Int. J. Mass. Spectrom.* **209**, 99–112 (2001)
4. Wysocki, V.H., Tsaprailis, G., Smith, L.L., Breci, L.A.: Mobile and localized protons: a framework for understanding peptide dissociation. *J. Mass Spectrom.* **35**, 1399–1406 (2000)
5. Yu, W., Vath, J.E., Huberty, M.C., Martin, S.A.: Identification of the facile gas-phase cleavage of the Asp-Pro and Asp-Xxx peptide bonds in matrix assisted laser desorption time-of-flight mass spectrometry. *Anal. Chem.* **65**, 3015–3023 (1993)
6. Boyd, R., Somogyi, A.: The mobile proton hypothesis in fragmentation of protonated peptides: a perspective. *J. Am. Soc. Mass Spectrom.* **21**, 1275–1278 (2010)
7. 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)
8. Bythell, B.J., Suhai, S., Somogyi, A., Paizs, B.: Proton-driven amide bond-cleavage pathways of gas-phase peptide ions lacking mobile protons. *J. Am. Chem. Soc.* **131**, 14057–14065 (2009)
9. Calvo, F., Dugourd, P.: Theoretical evidence for temperature-induced proton mobility in isolated lysine-rich polyalanines. *J. Phys. Chem. A* **112**, 4679–4687 (2008)
10. Farrugia, J.M., O'Hair, R.A.J.: Involvement of salt bridges in a novel gas phase rearrangement of protonated arginine-containing dipeptides which precedes fragmentation. *Int. J. Mass. Spectrom.* **222**, 229–242 (2003)
11. Paizs, B., Csonka, I.P., Lendvay, G., Suhai, S.: Proton mobility in protonated glycylglycine and N-Formylglycylglycinamide: a combined quantum chemical and RKKM study. *Rapid Commun. Mass Spectrom.* **15**, 637–650 (2001)
12. Eng, J.K., McCormack, A.L., Yates III, J.R.: An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994)
13. Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S.: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999)
14. Harrison, A.G., Young, A.B.: Fragmentation of protonated oligoalanines: amide bond cleavage and beyond. *J. Am. Soc. Mass Spectrom.* **15**, 1810–1819 (2004)
15. Yalcin, T., Csizmadia, I.G., Peterson, M.B., Harrison, A.G.: The structure and fragmentation of B_n ($n \geq 3$) ions in peptide spectra. *J. Am. Soc. Mass Spectrom.* **7**, 233–242 (1996)
16. Paizs, B., Suhai, S.: Towards understanding the tandem mass spectra of protonated oligopeptides. 1: mechanism of amide bond cleavage. *J. Am. Soc. Mass Spectrom.* **15**, 103–113 (2004)
17. Walsh, R.: Thermochemical kinetics: does it still give insights? *Chem. Soc. Rev.* **37**, 686–698 (2008)
18. Zhang, Z.Q.: Prediction of low-energy collision-induced dissociation spectra of peptides. *Anal. Chem.* **76**, 3908–3922 (2004)
19. Lu, T.F., Law, C.K.: Toward accommodating realistic fuel chemistry in large-scale computations. *Prog. Energy Combust. Sci.* **35**, 192–215 (2009)
20. Khan, S.S., Zhang, Q.Z., Broadbelt, L.J.: Automated mechanism generation. Part 1: mechanism development and rate constant estimation for voc chemistry in the atmosphere. *J. Atmos. Chem* **63**, 125–156 (2009)
21. Harper, M.R., Van Geem, K.M., Pyl, S.P., Marin, G.B., Green, W.H.: Comprehensive reaction mechanism for *n*-Butanol pyrolysis and combustion. *Combust. Flame* **158**, 16–41 (2011)
22. Anslyn, E.V., Dougherty, D.A.: *Modern Physical Organic Chemistry*. University Science Books, Sausalito, California (2006)
23. Savitski, M.M., Kjeldsen, F., Nielsen, M.L., Garbuzynskiy, S.O., Galzitskaya, O.V., Surin, A.K., Zubarev, R.A.: Backbone carbonyl group basicities are related to gas-phase fragmentation of peptides and protein folding. *Angew. Chem. Int. Ed.* **46**, 1481–1484 (2007)
24. Sztáray, J., Memboeuf, A., Drahos, L., Vékey, K.: Leucine enkephalin —a mass spectrometry standard. *Mass Spectrom. Rev.* **30**, 298–320 (2011)
25. Stearns, J.A., Seaiby, C., Boyarkin, O.V., Rizzo, T.R.: Spectroscopy and conformational preferences of gas-phase helices. *Phys. Chem. Chem. Phys.* **11**, 125–132 (2009)
26. Strittmatter, E.F., Lemoff, A.S., Williams, E.R.: Structure of cationized glycine, $Gly{\cdot}M^{2+}$ (M=Be, Mg, Ca, Sr, Ba), in the gas phase: intrinsic effect of cation size on zwitterion stability. *J. Phys. Chem. A* **104**, 9793–9796 (2000)
27. Wyttenbach, T., Bushnell, J.E., Bowers, M.T.: Salt bridge structures in the absence of solvent? The case for the oligoglycines. *J. Am. Chem. Soc.* **120**, 5098–5103 (1998)
28. *Macromodel*; Schrödinger, LLC: Portland, Oregon, 2009.

29. Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., Still, W.C.: Macromodel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **11**, 440–467 (1990)
30. Jorgensen, W.L., Maxwell, D.S., Tirado-Rives, J.: Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996)
31. Jorgensen, T.J.D., Gardsvoll, H., Ploug, M., Roepstorff, P.: Intramolecular migration of amide hydrogens in protonated peptides upon collisional activation. *J. Am. Chem. Soc.* **127**, 2785–2793 (2005)
32. Berger, A., Loewenstein, A., Meiboom, S.: Nuclear magnetic resonance study of the protolysis and ionization of *N*-Methylacetamide. *J. Am. Chem. Soc.* **81**, 62–67 (1959)
33. Schulman, J.M., Disch, R.L.: A simple formula for the zero-point energies of hydrocarbons. *Chem. Phys. Lett.* **113**, 291–293 (1985)
34. Grice, M.E., Politzer, P.: Use of molecular stoichiometry to estimate vibrational-energy. *Chem. Phys. Lett.* **244**, 295–298 (1995)
35. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery Jr., J.A., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A.: *Gaussian 03*. Gaussian, Inc., Wallingford CT (2004)
36. *Q-Chem*, ver. 3.2; Q-Chem, Inc.: Pittsburgh, 2009.
37. Shao, Y., Molnar, L.F., Jung, Y., Kussmann, J., Ochsenfeld, C., Brown, S.T., Gilbert, A.T.B., Slipchenko, L.V., Levchenko, S.V., O'Neill, D.P., DiStasio Jr., R.A., Lochan, R.C., Wang, T., Beran, G.J.O., Besley, N. A., Herbert, J.M., Lin, C.Y., van Voorhis, T., Chien, S.H., Sodt, A., Steele, R.P., Rassolov, V.A., Maslen, P.E., Korambath, P.P., Adamson, R.D., Austin, B., Baker, J., Byrd, E.F.C., Dachsel, H., Doerksen, R.J., Dreuw, A., Dunietz, B.D., Dutoi, A.D., Furlani, T.R., Gwaltney, S.R., Heyden, A., Hirata, S., Hsu, C.-P., Kedziora, G., Khalliulin, R.Z., Klunzinger, P., Lee, A.M., Lee, M.S., Liang, W.-Z., Lotan, I., Nair, N., Peters, B., Proynov, E.I., Pieniazek, P.A., Rhee, Y.M., Ritchie, J., Rosta, E., Sherrill, C.D., Simmonett, A.C., Subotnik, J.E., Woodcock III, H.L., Zhang, W., Bell, A.T., Chakraborty, A.K., Chipman, D.M., Keil, F.J., Warshel, A., Hehre, W.J., Schaefer III, H.F., Kong, J., Krylov, A.I., Gill, P.M.W., Head-Gordon, M.: Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.* **8**, 3172–3191 (2006)
38. Mohammed, O.F., Jas, G.S., Lin, M.M., Zewail, A.H.: Primary peptide folding dynamics observed with ultrafast temperature jump. *Angew. Chem. Int. Ed.* **48**, 5628–5632 (2009)
39. Williams, S., Causgrove, T.P., Gilmanshin, R., Fang, K.S., Callender, R.H., Woodruff, W.H., Dyer, R.B.: Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* **35**, 691–697 (1996)
40. Vaden, T.D., de Boer, T.S.J.A., Simons, J.P., Snoek, L.C., Suhai, S., Paizs, B.: Vibrational spectroscopy and conformational structure of protonated polyalanine peptides isolated in the gas phase. *J. Phys. Chem. A* **112**, 4608–4616 (2008)
41. Hudgins, R.R., Ratner, M.A., Jarrold, M.F.: Design of helices that are stable in vacuo. *J. Am. Chem. Soc.* **120**, 12974–12975 (1998)
42. Hudgins, R.R., Mao, Y., Ratner, M.A., Jarrold, M.F.: Conformations of Gly_nH^+ and Ala_nH^+ peptides in the gas phase. *Biophys. J.* **76**, 1591–1597 (1999)
43. Kinnear, B.S., Hartings, M.R., Jarrold, M.F.: The energy landscape of unsolvated peptides: helix formation and cold denaturation in $Ac\text{-}A_4G_7A_4+H^+$. *J. Am. Chem. Soc.* **124**, 4422–4431 (2002)
44. Kohtani, M., Jones, T.C., Schneider, J.E., Jarrold, M.F.: Extreme stability of an unsolvated α-Helix. *J. Am. Chem. Soc.* **126**, 7420–7421 (2004)
45. Wieczorek, R., Dannenberg, J.J.: A-helical peptides are not protonated at the N-Terminus in the gas phase. *J. Am. Chem. Soc.* **126**, 12278–12279 (2004)
46. Åqvist, J., Luecke, H., Quiocho, F.A., Warshel, A.: Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl. Acad. Sci. USA* **88**, 2026–2030 (1991)
47. Hol, W.G.J., van Duijnen, P.T., Berendsen, H.J.C.: The α-Helix dipole and the properties of proteins. *Nature* **273**, 443–446 (1978)
48. Roos, G., Loverix, S., Geerlings, P.: Origin of the pK_a perturbation of N-Terminal cysteine in α- and 310-Helices: a computational DFT study. *J. Phys. Chem. B* **110**, 557–562 (2006)
49. Scheraga, H.A.: Effect of side chain-backbone electrostatic interactions on the stability of α-helices. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5585–5587 (1985)
50. Wada, A.: The α-Helix as an electric macro-dipole. *Adv. Biophys.* **9**, 1–63 (1976)
51. Wu, Q.Y., van Orden, S., Cheng, X.H., Bakhtiar, R., Smith, R.D.: Characterization of cytochrome c variants with high-resolution FTICR mass spectrometry—correlation of fragmentation and structure. *Anal. Chem.* **67**, 2498–2509 (1995)
52. Zhang, Z., Bordas-Nagy, J.: Peptide conformation in gas phase probed by collision-induced dissociation and its correlation to conformation in condensed phases. *J. Am. Soc. Mass Spectrom.* **17**, 786–794 (2006)
53. Park, C., Goddard III, W.A.: Stabilization of alpha-helices by dipole-dipole interactions within alpha-helices. *J. Phys. Chem. B* **104**, 7784–7789 (2000)
54. Wu, Y.D., Zhao, Y.L.: A theoretical study on the origin of cooperativity in the formation of 3_{10}- and α-helices. *J. Am. Chem. Soc.* **123**, 5313–5319 (2001)
55. Ireta, J., Neugebauer, J., Scheffler, M., Rojo, A., Galvan, M.: Density functional theory study of the cooperativity of hydrogen bonds in finite and infinite alpha-helices. *J. Phys. Chem. B* **107**, 1432–1437 (2003)
56. Koch, O., Bocola, M., Klebe, G.: Cooperative effects in hydrogen-bonding of protein secondary structure elements: a systematic analysis of crystal data using Secbase. *Proteins Struct. Funct. Bioinf.* **61**, 310–317 (2005)
57. Tsemekhman, K., Goldschmidt, L., Eisenberg, D., Baker, D.: Cooperative hydrogen bonding in amyloid formation. *Protein Sci.* **16**, 761–764 (2007)
58. Marqusee, S., Robbins, V.H., Baldwin, R.L.: Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5286–5290 (1989)
59. McLuckey, S.A., Cameron, D., Cooks, R.G.: Proton affinities from dissociations of proton-bound dimers. *J. Am. Chem. Soc.* **103**, 1313–1317 (1981)
60. Stevenson, D. P.: Ionization and Dissociation by Electronic Impact - the Ionization Potentials and Energies of Formation of Sec-Propyl and Tert-Butyl Radicals—Some Limitations on the Method. *Disc. Faraday Soc.* **10,** 35–45 (1951)
61. Barsnes, H., Eidhammer, I., Martens, L.: A global analysis of peptide fragmentation variability. *Proteomics* **11**, 1181–1188 (2011)
62. Anders, L.R., Beauchamp, J.L., Dunbar, R.C., Baldeschwieler, J.D.: Ion-cyclotron double resonance. *J. Chem. Phys.* **45**, 1062–1063 (1966)
63. Jue, A.L., Racine, A.H., Glish, G.L.: The effect of ion trap temperature on the dissociation of peptide ions in a quadrupole ion trap. *Int. J. Mass Spectrom.* **301**, 74–83 (2011)
64. Klassen, J.S., Kebarle, P.: Collision-induced dissociation threshold energies of protonated glycine, glycinamide, and some related small peptides and peptide amino amides. *J. Am. Chem. Soc.* **119**, 6552–6563 (1997)
65. El Aribi, H., Rodriquez, C.F., Almeida, D.R.P., Ling, Y., Mak, W.-W. N., Hopkinson, A.C., Siu, K.W.M.: Elucidation of fragmentation mechanisms of protonated peptide ions and their products: a case study on glycylglycylglycine using density functional theory and threshold collision-induced dissociation. *J. Am. Chem. Soc.* **125**, 9229–9236 (2003)
66. Armentrout, P. B.; Heaton, A. L. Thermodynamics and Mechanisms of Protonated Diglycine Decomposition: A Guided Ion Beam Study. *J. Am. Soc. Mass Spectrom.* **2011,** doi:10.1007/s13361-011-0225-6.

67. Cordero, M.M., Houser, J.J., Wesdemiotis, C.: The neutral products formed during backbone fragmentations of protonated peptides in tandem mass-spectrometry. *Anal. Chem.* **65**, 1594–1601 (1993)
68. Laskin, J. Denisov, E. Futrell, J.: Comparative Study of Collision-Induced and Surface-Induced Dissociation. 2. Fragmentation of Small Alanine-Containing Peptides in FT-ICR MS. *J. Phys. Chem. B* **105**, 1895–1900 (2001)
69. Yeh, R.W., Grimley, J.M., Bursey, M.M.: Collisionally induced fragmentation of protonated oligoalanines and oligoglycines. *Biol. Mass Spectrom.* **20**, 443–450 (1991)
70. Pu, D., Clipston, N.L., Cassady, C.J.: A comparison of positive and negative ion collision-induced dissociation for model hepta-peptides with one basic residue. *J. Mass Spectrom.* **45**, 297–305 (2010)