Exploring the Effects of Methylation on the CID of Lysine: A Combined Experimental and Computational Approach

Kenneth Lucas[a],[†] Amy Chen[b],[†] Megan Schubmehl,[†] Kristopher J. Kolonko,[‡] and George L. Barnes[*],[†]

[†]*Department of Chemistry and Biochemistry*
Siena College
515 Loudon Road
Loudonville, NY 12211
[‡]*Stewart's Advanced Instrumentation and Technology (SAInT) Center*
Siena College
515 Loudon Road
Loudonville, NY 12211

E-mail: gbarnes@siena.edu

[a]Presently at SUNY Upstate Medical University; College of Medicine;766 Irving Avenue; Syracuse, NY 13210
[b]ACS Project SEED High School Student; Presently pursing an undergraduate degree at Northwestern University

Abstract

We report the results of both experiments, simulations, and DFT calculations that focus on describing the reaction dynamics observed within the collision-induced dissociation of L-lysine and its side-chain methylated analogues, N_ε-Methyl-L-lysine (Me$_1$-lysine), N_ε,N_ε-Dimethyl-L-lysine (Me$_2$-lysine), and N_ε,N_ε,N_ε-Trimethyl-L-lysine (Me$_3$-lysine). There is good qualitative agreement between simulations and experiments. The major pathways observed in the experimental measurements were *m/z* 130 and 84, with the former dominant at low collision energies and the latter at intermediate to high collision energies. The *m/z* 130 peak corresponds to loss of H_2CO_2 while *m/z* 84 has the additional loss of $N(CH_3)_nH_{3-n}$. Within the time frame of the direct dynamics simulations, *m/z* 130 and 101 were the most populous peaks, with the latter identified as an intermediate to *m/z* 84. The simulations allowed for the determination of several reaction pathways that result in these products, and a graph theory analysis enabled the elucidation of the major structures that compose each peak. Methylation results in an increase in the preferential loss of the side-chain amide group and a reduced occurrence of cyclic structures within the population of the *m/z* 84 peak in simulations.

1 Introduction

Post-translational modifications (PTMs) are both common and expand the functionality of proteins and peptides.[1–3] PTMs increase chemical complexity, which in turn modifies the interactions present within these species resulting in important biological implications.[4] Methylation and acetylation of lysine, in particular, are predominately observed on histone proteins and can function as a biomarker for gene activation.[5–8] Methylation can occur at either the N-terminus or the side-chain and has three distinct forms, namely mono-, di-, and tri-methylation. Identifying methylation via tandem mass spectrometry (MS2), specifically collision-induced dissociation (CID), is a frequently occurring task. Indeed, MS2 spectroscopy

is a well-used tool for studying biological systems using both experimental[9–12] and theoretical, direct dynamics based studies.[13–18] That said, trimethylation and acetylation of lysine result in a mass difference of just ∼0.036 Da and requires high-resolution instruments to distinguish. Direct dynamics studies of methylation and acetylation can provide insight into the dynamics taking place during CID.

Direct dynamics simulations have shown great utility in elucidating atomic level information regarding dissociation mechanisms within MS2 spectroscopy.[16,17] Typically, these simulations make use of semi-empirical treatments of the intramolecular potential energy surface and yet still yield good agreement with experiment. Moreover, the results from an ensemble of direct dynamics simulations performed using an appropriate sampling method provide an unbiased means of exploring the reaction pathways of interest and often identify unexpected reaction mechanisms. Subsequently, the energetics of these reaction pathways can be obtained via higher-level DFT or *ab initio* calculations, and hence provide significant insight into the dynamics taking place within experiment.

Although PTMs are common, there have been only a few direct dynamics based studies focusing on their role and influence on the mechanisms seen in CID. In addition, the PTM systems studied experimentally are often too large for a direct dynamics computational study. In this work, we seek to elucidate the impact of methylation on the observed CID mechanisms. To accomplish this goal, we performed both direct dynamics simulations and experimental measurements on L-lysine (lysine), N_ε-Methyl-L-lysine (Me-lysine), N_ε,N_ε-Dimethyl-L-lysine (Me$_2$-lysine), and N_ε,N_ε,N_ε-Trimethyl-L-lysine (Me$_3$-lysine) to allow for a direct comparison between results.

As mentioned above, direct dynamics simulations are a valuable tool with a significant history of providing information regarding MS2 systems. Recently, several groups have started to use graph theory to analyze these simulations.[14,19–22] The application of graph theory to chemistry is well documented,[23] and its application within direct dynamics simulations typically begins with the definition of the adjacency matrix, which is closely related to what one of the authors termed the connectivity matrix in a previous work.[24] There are two primary means of obtaining the adjacency matrix from simulation data, namely the use of distance parameters to determine bonding[20–22] or the direct use of the bond order obtained from the quantum mechanical wavefunction used within the dynamics.[14,19] With the adjacency matrix in hand, a graph analysis allows for a determination of the number and types of fragments obtained, which we employ here to analyze the diversity of structures within each mass peak. This classification further simplifies the identification of the relevant mechanisms within the simulations.

An outline for the remainder of the paper is as follows: in Section 2, we provide an overview of our computational and experimental methods; in Section 3, we present our results and compare the CID pathways between lysine and Me$_n$-lysine; and in Section 4, we provide an overview.

2 Methods

Below we present out computational and experimental approach to studying L-lysine (lysine), N_ε-Methyl-L-lysine (Me$_1$-lysine), N_ε,N_ε-Dimethyl-L-lysine (Me$_2$-lysine), and N_ε,N_ε,N_ε-Trimethyl-L-lysine (Me$_3$-lysine).

2.1 Computational Approach

The literature contains several examples of the use of direct dynamics simulations to investigate MS2 systems. Moreover, there have been recent perspectives written on its use[13,16,25] as well as a tutorial review on the methods involved.[16] We refer the interested reader to those works for specific details regarding the approach and provide specific information for our systems of interest below.

2.1.1 Structures and Simulation Method

We generated initial structures for lysine and Me$_n$-lysine (n=1-3) with the methylation occurring on the side-chain nitrogen using Avogadro.[26] The charge of each species was +1, which was accomplished by placing the excess proton on the side-chain for lysine and Me$_{1-2}$-lysine. Me$_3$-lysine has a +1 charge without an additional proton. These

species are of sufficient size that it would be haphazard to trust human intuition for finding the global potential energy minimum. Hence, a simulated annealing approach was employed. In particular, the species were imported into GROMACS[27] where an initial geometry optimization was performed, followed by a 300 K equilibration. The simulated annealing procedure then commenced using six heat-cool cycles in which the temperature was ramped to 1000 K over 100 ps held at 1000 K for 100 ps and then cooled to 0 K over 100 ps. The resulting 0 K structures were optimized at the RM1 level as implemented in Mopac2016.[28] The molecular mechanical potential for GROMACS was based on the Gromos 54A7[29] potential and then modified using the Automatic Topology Builder (ATB)[30–32] to accommodate both the methylation and charged nature of each species.

Direct dynamics simulations were performed using the RM1 semi-empirical method[33] to calculate the lysine potential energy surface. RM1 has been shown to produce good results for numerous previous simulations involving protonated peptides.[18,24,25,34,35] Although the potential energy parameters for Ar CID given in the literature[36] could have been used for the PTM systems, we instead choose to make use of the microcanonical sampling scheme to impart a statistical distribution of internal energy.[37] Good agreement with experiment was achieved, which suggests that sudden or fast fragmentation events[19] are not that significant for these systems. A range of internal energies was considered for each system with lysine, Me_1-lysine, and Me_2-lysine having values of 250, 300, and 350 kcal/mol, and Me_3-lysine having values of 300, 350, 400 kcal/mol. We performed 400 trajectories with randomly selected initial conditions for each specified internal energy. The particular internal energy values were chosen to allow for significant reactivity to explore the chemical space without producing too many low-mass products. During the analysis, the products are considered either collectively or as a function of internal energy. In general, reactivity increased with increasing internal energy within the time frame of the simulations.

Hamilton's equations of motion were solved using a 6th order symplectic integration scheme,[38] for a maximum simulation time of 50 ps with a 1 fs step size and output written every 50 fs using our in-house simulation package tightly coupled with Mopac2012.[39] As we are focused on the charged fragments, neutral fragments were removed from the simulation if they were at least 15 Å away from any charged fragments. In addition, a trajectory is halted early if the final charged fragment had an $m/z \leq 60$. Conservation of energy was excellent for all trajectories.

2.1.2 Theoretical Mass Spectra and Reaction Mechanisms

As described previously,[14,40] our in-house simulation software obtains the bond order between all QM atoms as a function of time throughout the simulation. These bond orders are averaged over a 5 fs window to reduce occurrences of momentary fluctuations. This averaged bond order matrix is converted to an adjacency matrix and recorded every 50 fs for post-analysis. On-the-fly, connectivity analysis allows for an initial preview of the reactive events. In post analysis, fragmentation events that quickly recombine are removed, which allows for straightforward and clear analysis of when fragmentation occurs. This information can be correlated with proton motion to determine which proton migrations influence bond cleavage events.[40] Combining this information with a graph theory analysis[20,21] of each product enables the automatic calculation of adjacency matrix weighted by atomic number, the lowest eigenvalue of which identifies unique structures within each product. Hence, the direct dynamics simulations combined with connectivity analysis and graph theory allow for easy identification of the most important products and the mechanistic steps taken to reach them. With this information on hand, structures and energetics of relevant minima, intermediates, and transition states were obtained for the lysine system at the ωB97X-D/aug-cc-pVTZ level of theory as implemented in the Gaussian09 software package.[41] We note here that we expect that potential energy profiles for the PTM analogs are similar to lysine. In addition, it was common for proton transfers to occur via an uphill path within the lysine system, and often a TS could not be located.

2.2 Experimental Approach

High purity samples of L-lysine (lysine), N_ε-Methyl-L-lysine (Me_1-lysine), N_ε,N_ε-Dimethyl-L-lysine (Me_2-lysine), and N_ε,N_ε,N_ε-Trimethyl-L-lysine (Me_3-lysine) were obtained from Sigma Aldrich and used without further purification. Deuterium oxide ($\geq$99.96 atom % in D) and L-lysine (α-^{15}N) was obtained from Cambridge Isotope Laboratories (Cambridge, MA). All samples were prepared in mass spectrometry grade, 1:1 mixture of acetonitrile and water purchased from Sigma Aldrich. High-resolution mass spectrometry data was acquired on a Bruker Maxis Impact HD (Quadrupole-Time of Flight) spectrometer using sodium formate solution as the mass calibrant and the standard heated electrospray ionization (ESI) source. Pseudo MS3 experiments were performed utilizing in-source fragmentation. Typical collision energies considered were 5, 10, 15, 20, 25, and 30 eV using nitrogen within the CID cell. Additional experimental details are provided in the supporting information.

3 Results and Discussion

The results from the ESI high-resolution MS and the simulations are in good qualitative agreement and provide complementary insight. This good agreement is highlighted in Figure 1 which contains example experimental/simulated mass spectra and experimental break-down curves for lysine and Me_n-lysine. Additional experimental data is provided in the supporting information. The most prominent experimental peaks for each system are provided in Table 1, all but four of which are observed in the direct dynamics simulations. At low and intermediate collision energies, a mixture of significant peaks is observed. At high collision energy, the m/z 84 peak ($C_5H_{10}N$) is dominant in all species. From this experimental observation, it seems likely that a common mechanistic pathway exists for lysine and its methylated analogues that lead to the m/z 84 peak, which is consistent with the findings from our direct dynamics simulations.

3.1 Lysine

The experimental measurements find that in lysine ($C_6H_{15}N_2O_2$), the two most significant peaks are m/z 84 ($C_5H_{10}N$) and 130 ($C_6H_{12}NO_2$). This finding is in agreement with previous measurements made by Zhang *et al.*[42] as well as Dookeran *et al.*[43] Our direct dynamics simulations also find that m/z 84 and 130 are common products. However, we find that m/z 101 ($C_5H_{13}N_2$) is not only important but, on the timescale of our simulations, the peak with the largest population for internal energies of 300 and 350 kcal/mol.

Table 1: All experimental peaks above a fraction of 0.01 normalized intensity. Peaks resulting from less common isotopes, such compounds containing ^{13}C, are removed. All peaks are observed in the simulations, unless indicated.

Peak	Fraction	Loss
Lysine[a]		
84.0807	0.135	$NH_3 + H_2CO_2$
129.1024[b]	0.021	H_2O
130.0864	0.434	NH_3
147.1130	0.328	None
Me_1-lysine[a]		
84.0806	0.159	$N(CH_3)H_2 + H_2CO_2$
98.0965	0.023	$NH_3 + H_2CO_2$
116.1073[b]	0.022	CH_3NO
130.0863	0.347	$N(CH_3)H_2$
144.1020	0.023	NH_3
161.1286	0.336	None
Me_2-lysine[a]		
84.0805	0.165	$N(CH_3)_2H + H_2CO_2$
130.0861	0.299	$N(CH_3)_2H$
130.1222[b]	0.063	CH_3NO
131.1258	0.005	CO_2
158.1173	0.016	NH_3
175.1439	0.353	None
Me_3-lysine[a]		
84.0807	0.380	$N(CH_3)_3 + H_2CO_2$
130.0865	0.336	$N(CH_3)_3$
144.1385[b]	0.042	CH_3NO
189.1601	0.140	None

[a] Same experimental collision energies as in Figure 1.
[b] Not observed in simulations.

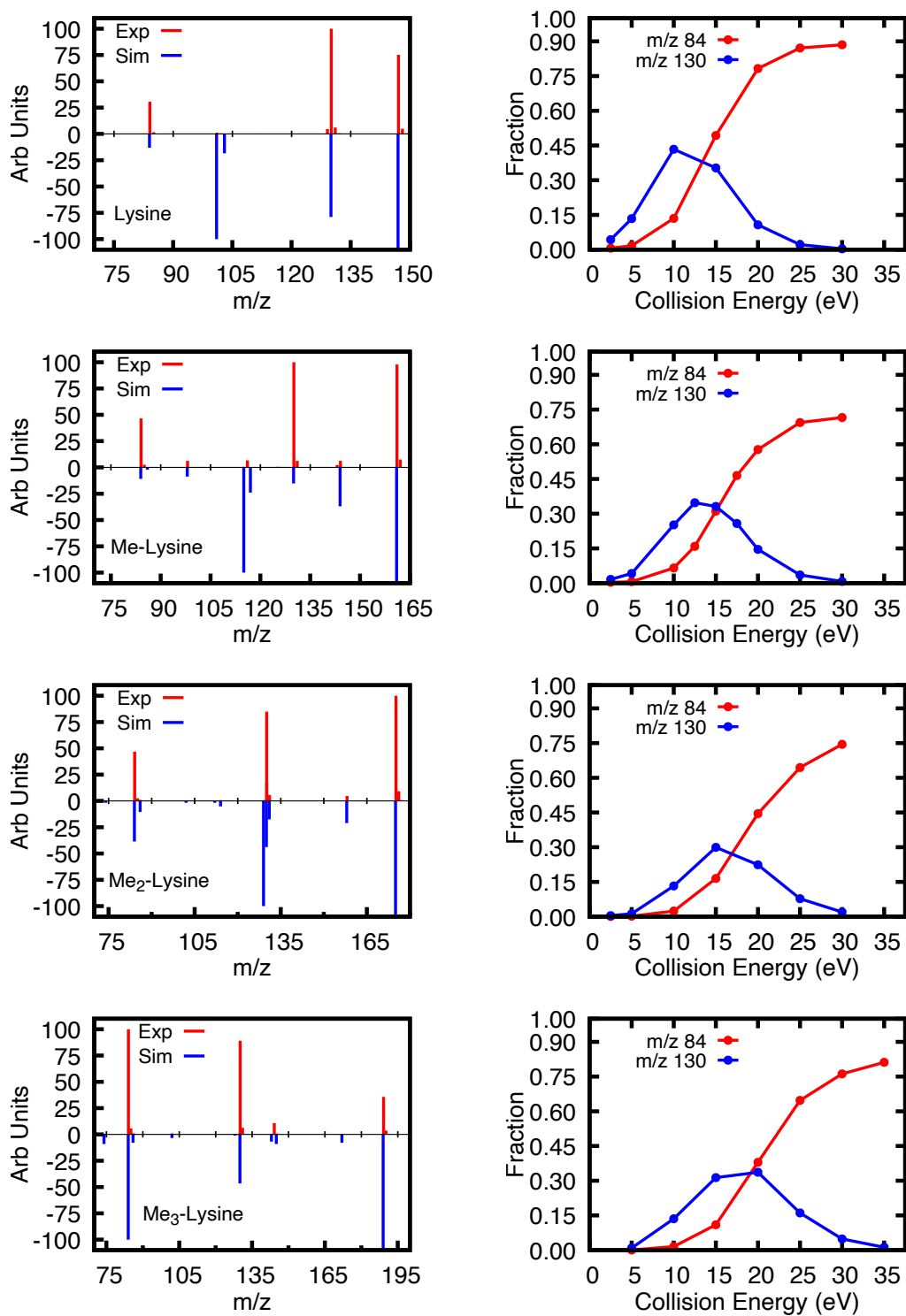The m/z 84 peak corresponds to the loss of NH_3

Figure 1: Experimental/simulated mass spectra are presented in the first column with experimental break-down curves provided in column 2. The experimental *m/z* spectrum intensity in the first column is scaled to the most intense peak for each given spectrum. In contrast, the simulated spectrum is scaled to the most intense fragment peak to ease comparison. The experimental collision energies are 10, 12.5, 15, and 20 eV, while the simulated internal energies are 250, 300, 350, and 400 kcal/mol, respectively. In column 2, the experimental break-down curves are relative to the total ion signal collected rather than the most intense peak.

+ H_2CO_2, and our graph theory analysis reveals that nine different structures correspond to this peak. However, three of those structures make up 84.5% of the population within the peak. These three most important structures, accounting for 53.2%, 18.1%, and 13.3% of the peak population, respectively, are shown in Figure 2 along with reaction mechanisms that can produce them. The energetics of these mechanisms are shown in Figure 3.

The simulated m/z 84 pathways illustrate that this product can be formed by losing NH_3 as the first or second step with the nitrogen originated from either the side-chain or the N-terminus. Although loss is possible from either location, most simulated reaction pathways show loss of the side-chain, which is the preferred product for all internal energies considered. When considering all internal energies simulated, 84.6% of the m/z 84 population obtained shows loss of the side-chain. This finding is in good agreement with our experimental measurements of lysine (α-^{15}N labeled), which allows for an experimental determination of side-chain vs. N-terminus loss. Examining the m/z 84 and 85 populations within the ^{15}N labeled experiments shows a loss of the side-chain ~90% of the time. After the loss of the side-chain, the ion must then lose H_2CO_2 to reach the final m/z 84 product. In simulations, this most commonly occurs as $C(OH)_2$ as depicted in the first three pathways in Figure 2. However, some trajectories show loss of either H_2O + CO or HCO_2H. It is likely that in the long time limit, $C(OH)_2$ will rearrange to one of these two more stable neutral species. Moreover, the three structures shown for m/z 84 can interconvert between each other, and at the ωB97X-D/aug-cc-pVTZ level of theory, the six-member ring structure has the lowest energy.

Simulations also revealed a unique m/z 84 formation pathway that shows the initial loss of $C(OH)_2$ followed by NH_3. This pathway is the sole source of the ~15% of trajectories that lose the N-terminus and begins with a proton transfer from the side-chain nitrogen to the N-terminus. This type of proton transfer was commonly observed both in the present simulations and in our previous works. [14,44] No distinct TS between these two protonation sites has been found, but rather the side-chain and N-terminus form short, strong

hydrogen bonds. [45] After the excess proton is located on the N-terminus, there are two possible reactive pathways open to the ion. One pathway begins by transferring a proton from the N-terminus nitrogen to an oxygen of the C-terminus, resulting in either $C(OH)_2$ or H_2O + CO loss depending on the receiving oxygen. The product of this loss has m/z 101 and can subsequently rearrange to the m/z 84 product through a ring-closure reaction or from a mixture of protonated cis/trans 1-Pentene-1,5-diamine, which are long-lived on the timescale of our simulations. We note that the final products can also be reached via a direct transfer of the proton from the side chain nitrogen to the oxygen of the C-terminus. These branching pathways show that the m/z 101 population within the simulations is easily explained as an intermediate on the way to m/z 84. Our graph theory analysis shows that nine different structures correspond to m/z 101. However, 59.3% of the products are given by the boxed structure shown in Figure 2 with the other structures representing either intermediates along this pathway or structures that can interconvert with the structure displayed.

In addition to addressing the presence of m/z 101, these pathways show that m/z 130 is also an intermediate to m/z 84. While the graph theory analysis finds 14 unique structures for the m/z 130 peak, eight of these structures occur in just a single trajectory and are likely fleeting intermediates. Only two structures show sizable populations, as shown in Figure 2, with ~59.0% arising from the pathway that branches between m/z 130, m/z 101, and m/z 84. Experimentally, as seen in Figure 1, m/z 130 starts as the most dominant peak but is supplanted by m/z 84, which is consistent with the intermediate nature of m/z 130.

Experimental MS measurements show that at high collision energy ($\geq$ 15eV), m/z 84 is the dominant feature of the MS spectrum. The mechanisms presented along with the DFT energetics make it clear why this is the preferred product. Isomeric fragments of m/z 130 are common intermediates in each of these pathways that lead to m/z 84, and MS3 experiments confirm that fragment m/z 130 yields m/z 84. While both DFT calculated barriers and energy minimum for m/z 130 structures are lower than those for m/z 84, the formation of the m/z 84 fragment results from a non-
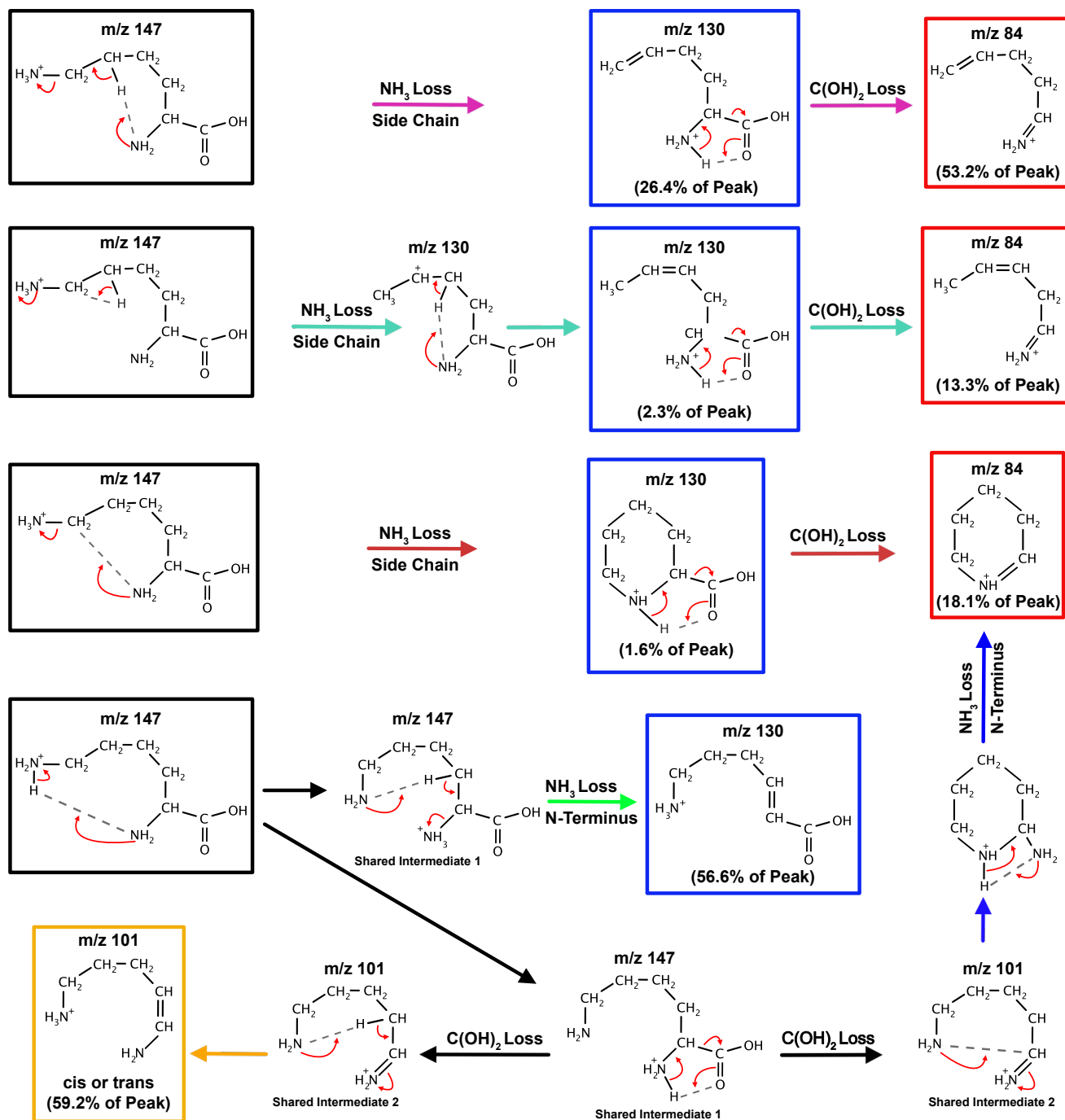
Figure 2: Representations of the frequently occurring reaction pathways observed in the direct dynamics simulations. The initial ion is boxed in black, *m/z* 84 in red, *m/z* 101 in orange, and *m/z* 130 in blue. The energetics are provided in Figure 3 with the reaction arrows above color coded to match that figure.3.
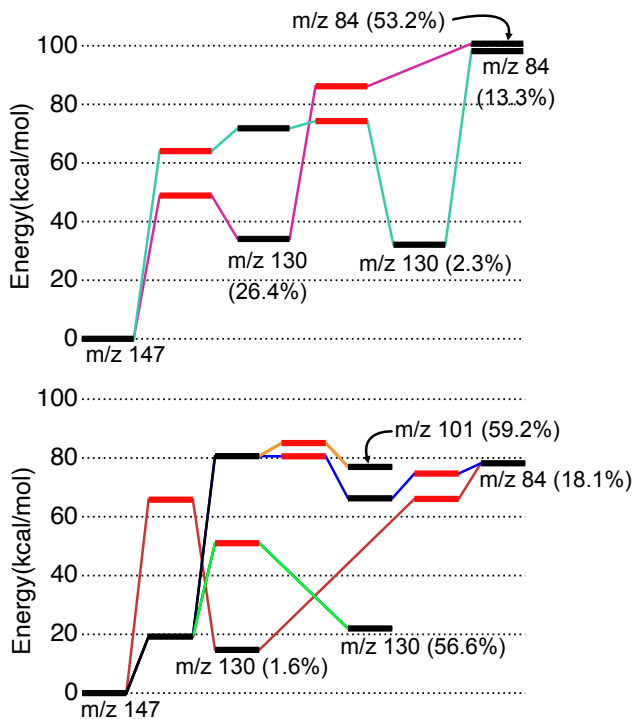
Figure 3: Energies for pathways shown in Figure 2 calculated at the ωB97-D/aug-cc-pVTZ level of theory are shown above. Minima are labeled with solid black lines, while TSs labeled with red. Pathways are displayed in two panels and color coded in agreement with Figure 2 for clarity.

reversible step. Thus, given enough time and energy, the system will eventually reach this product. The same is true for the simulated pathways that lead to m/z 101, which is confirmed via pseudo MS3 experiments on m/z 101. Indeed, m/z 101 has a slightly higher formation TS from the parent ion m/z 147 than the other products, but it also has a reversible barrier and given sufficient time, it will lead to m/z 84. The formation of the m/z 130 structure that accounts for 56.6% of the simulated peak population appears to be a low energy endpoint with a relatively low barrier to formation. However, it is easy to believe that this structure will ultimately also lose H_2CO_2, in some form, to result in m/z 84. Reaching m/z 84 from this m/z 130 structure will require rearrangements that were not observed within the time frame of our simulations.

Table 1 shows water loss as an experimental peak that is not observed in simulations. This reaction is slightly above the 0.01 cutoff and accounts for a fraction of 0.021. However, this product does

not reach the cutoff for the other systems. Moreover, while we do not see water loss in the lysine simulations, it is observed in simulations for Me$_2$-lysine, and it seems likely that if more lysine simulations were calculated, some water loss would also be observed.

3.2 Me$_n$-lysine

All of the significant experimental peaks for the Me$_n$-lysine systems are provided in Table 1. Good qualitative agreement between experiments and simulations is seen for all systems. As shown in both this table and Figure 1, the Me$_n$-lysine series all have m/z 84 and 130 as the most critical peaks for the collision energies considered. These peaks are closely related to the same peaks in lysine and correspond to loss of $N(CH_3)_nH_{3-n}$ for m/z 130 and the additional loss of H_2CO_2 for m/z 84. Simulations indicated that similar mechanisms are observed for Me$_n$-lysine as for lysine. However, simulations reveal some mechanistic differences. The simulations for m/z 84 for lysine showed that a six-member ring structure would form through two different pathways. However, this structure is infrequently observed within the timeframe of the simulations for any of the methylated systems. The loss of the N-terminus in the methylated systems will not result in m/z 84, which reduces the population of this structure. However, the pathway that involves the direct loss of the side-chain followed by ring-closure is now only observed at 400 kcal/mol and in only five trajectories. That said, the methylated systems do still exhibit cyclic final products. In Me$_2$-lysine and Me$_3$-lysine, a five-member ring structure is the third most common, consisting of ∼10% of the m/z 84 population for each of those systems. An equivalent structure is also observed in Me$_1$-lysine, though it is not as prevalent. These differences aside, simulations predict that the most most common structure for m/z 84 in the methylated systems is the equivalent of the straight-chain structure that accounts for 53.2% of the lysine m/z 84 population as shown in Figure 2. The addition of methyl groups likely slows ring closure, making the observation of the equivalent structure in Me$_n$-lysine less common in simulations.

In considering the pathways that form the

m/z 130 peak, the most important structure seen in lysine involved loss of the N-terminus. The equivalent pathway in Me_n-lysine will no longer appear within the *m/z* 130 peak, and the overall population within the *m/z* 130 peaks is decreased compared to lysine. Comparing reactivity between systems directly is non-trivial since they have different internal degrees of freedom; however, for an internal energy of 300 kcal/mol, the total fraction of reactive simulations is 0.585 with *m/z* 130 contributing 0.14 to that total. In comparison, at an internal energy of 350 kcal/mol in Me_1-lysine, the total fraction of reactive simulations is 0.558, with *m/z* 130 contributing 0.068 to that total. When considering all internal energies, the major product seen in all systems for *m/z* 130 is the equivalent of the lysine structure that accounts for 26.4% of the peak. This structure accounts for 64.2%, 72.2%, and 64.6% of the populations in *m/z* 130 for the Me_n-lysine systems (n=1,2,3), respectively. Cyclic structures formed primarily from intramolecular reaction of the N-terminus and the side-chain carbons are observed for all systems and account for a larger fraction of the population in the peak than was seen for lysine. However, this observation may be due to the reduced overall population of the peak within the simulations.

While *m/z* 84 and 130 are the most important peaks and occur due to loss of the side-chain nitrogen, there are peaks that correspond to the loss of the N-terminus, namely *m/z* 98, 112, and 126. These peaks correspond to the loss of NH_3 + H_2CO_2 in Me_n-lysine, respectively, and would be found within *m/z* 84 for lysine. These first two peaks are observed in the simulations but are rare events. For Me_1-lysine, this product is observed in 22 trajectories across all internal energies considered, with the most common structure being analogous to the six-member ring structure observed in lysine. For Me_2-lysine the number drops to just three trajectories, all with unique structures, none of which correspond to that seen in lysine structure. Loss of the N-terminus alone is not observed for Me_3-lysine. In comparison, loss of the N-terminus occurred in roughly 15% of trajectories that resulted in *m/z* 84 for lysine. Methylation of the side-chain has effectively shutdown loss of the N-terminus since $N(CH_3)_nH_{3-n}$ is a better leaving group.

The equivalent *m/z* 130 peaks in Me_n-lysine are *m/z* 144, 158, and 172, respectively. None of these products are shown in Table 1, meaning that their experimental intensity is less than 1% of the total ion signal. Within the simulations with internal energies of 350, 350, and 400 kcal/mol, these peaks are observed with fractions of 0.05, 0.03, and 0.0175 for Me_n-lysine, respectively. In comparison, the fraction of *m/z* 130 for these systems is 0.0675, 0.0625, and 0.10. Although we note that the fractions in simulations depend on the time window calculated, the N-terminus loss peaks follow the trend of decreasing importance with increased methylation but are seen more commonly than in experiment.

In simulations of lysine, it was seen that the *m/z* 101 peak was the most common for intermediate internal energies and that this product was an intermediate to *m/z* 84. The analogous peaks in Me_n-lysine are *m/z* 115, 129, and 145. For Me_1-lysine and Me_2-lysine, these peaks are again the most populous peaks within the simulations. When considering all internal energies, the most common structure observed, making up 57.1% and 57.8% of the of *m/z* 115 and *m/z* 129 population, is the analogue of the boxed 101 structure shown in Figure 2. In the Me_2-lysine system, a six-member ring structure is the next most common structure with 14.5% of the population. However, it would be an intermediate to *m/z* 112, i.e., loss of the N-terminus, rather than *m/z* 84, side-chain loss pathway. For Me_3-lysine, the *m/z* 145 peak, which corresponds to the loss of CO_2 rather than H_2CO_2, is not the most populous peak. Rather at 400 kcal/mol, this peak makes up a fraction of just 0.02 of the calculated trajectories. In comparison, *m/z* 84 has a fraction of 0.22 and is the most significant fragment peak.

The simulations do not observe some of the peaks outlined in Table 1, namely *m/z* 116.1073 in Me_1-lysine, *m/z* 130.1222 in Me_2-lysine, and 144.1385 in Me_3-lysine. These peaks correspond to loss of CH_3NO with the nitrogen originating from the N-terminus. These peaks are minor contributors to the experimental spectra, though the contribution in the Me_2-lysine does reach ∼6% of the total ion signal. To investigate if these peaks originated from sudden fragmentation events following collision with the CID gas, Ar CID simula-

tions for Me_2-lysine were performed with collision energies ranging from 10 to 20 eV, but CH_3NO was not observed. A few trajectories produced CO due to the sudden collision with Ar followed by the formation of NH_4^+. If the excess proton had transferred back to the larger fragment, it would produce the experimentally observed m/z ratio, i.e., loss of $CO+NH_3$. However, that did not occur in the simulations, and these were rare events making it unlikely that a sudden event is leading to this peak.

4   Summary

We have presented results from experiments and direct dynamics simulations aimed at exploring the reaction pathways that occur within lysine and its side-chain methylated analogues, Me_n-lysine with n=1 - 3. Experimentally, for the lower collision energies considered, the m/z 130 peak ($C_6H_{12}NO_2$) is the most important, with the m/z 84 peak ($C_5H_{10}N$) becoming dominant as collision energy increases. These two peaks correspond to loss of H_2CO_2 and H_2CO_2 + $N(CH_3)_nH_{3-n}$ + H_2CO_2, respectively.

Direct dynamics simulations are in good qualitative agreement with the experiment. Multiple reaction pathways were identified from the simulations and quantified through ωB97X-D/aug-cc-pVTZ single point calculations to identify relevant minima and transition states. Simulations show that in the lysine system, m/z 84 can be formed from loss of nitrogen at either the side-chain, which is preferred, or the N-terminus, and occurring as either the first or second step. Simulations of lysine show that the m/z 101 peak ($C_5H_{13}N_2$) is the most important within the time frame considered. In addition, the simulations, the DFT calculations, and pseudo MS3 experiments show that m/z 101 is an intermediate to the formation of m/z 84. Three primary structures are found for m/z 84, all of which can interconvert, with the six-member ring structure being the energetic minimum. The m/z 130 is an intermediate along most of the reaction pathways described. In fact, the DFT calculations reveal that the formation of m/z 130 has the lowest energy barriers, which is consistent with the experimental results.

The experiments and simulations of the Me_n-lysine systems allow for differentiation of the side-chain from the N-terminus in the major peaks. It is seen that side-chain loss is dominant in both experiments and simulations, which is consistent with $N(CH_3)_nH_{3-n}$ being a better leaving group than NH_3. The structures observed in simulations of Me_n-lysine are analogous to those in lysine, though the m/z 84 shows a de-emphasis of the six-member ring structure.

Our graph theory analysis allowed for an automatic and in-depth view of the structures that made up each peak observed in the simulated mass spectra. These insights allowed for the targeted DFT based calculations that characterized the energy landscape for the lysine system. We expect that the lysine energy landscape is similar to that of the methylated analogues. Moreover, we also expect that using graph theory based post-simulation analysis will continue to be developed and become an invaluable tool within the community due to the insight that it provides.

5   ACKNOWLEDGEMENTS

References

(1) Kumar, A.; Narayanan, V.; Sekhar, A. Characterizing Post-Translational Modifications and Their Effects on Protein Conformation Using NMR Spectroscopy. Biochemistry **2019**, _59_, 57–73.

(2) Ryšlavá, H.; Doubnerová, V.; Kavan, D.; Vaněk, O. Effect of posttranslational modifications on enzyme function and assembly. J. Proteomics **2013**, _92_, 80–109.

(3) Boscher, C.; Dennis, J. W.; Nabi, I. R. Glycosylation, galectins and cellular signaling. Curr. Opin. Cell Biol. **2011**, _23_, 383–392.

(4) Mann, M.; Jensen, O. N. Proteomic analysis of post-translational modifications. Nat. Biotechnol. **2003**, 21, 255–261.

(5) Hyun, K.; Jeon, J.; Park, K.; Kim, J. Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. **2017**, 49, e324.

(6) Zhang, X.; Huang, Y.; Shi, X. Emerging roles of lysine methylation on non-histone proteins. Cell. Mol. Life Sci. **2015**, 72, 4257–4272.

(7) Zhang, X.; Wen, H.; Shi, X. Lysine methylation: Beyond histones. Acta Biochim. Biophys. Sin. (Shanghai). **2012**, 44, 14–27.

(8) Ivanov, G. S.; Ivanova, T.; Kurash, J.; Ivanov, A.; Chuikov, S.; Gizatullin, F.; Herrera-Medina, E. M.; Rauscher, F.; Reinberg, D.; Barlev, N. A. Methylation-Acetylation Interplay Activates p53 in Response to DNA Damage. Mol. Cell. Biol. **2007**, 27, 6756–6769.

(9) Cooks, R. G.; Ast, T.; Pradeep, T.; Wysocki, V. H. Reactions of Ions with Organic Surfaces. Acc. Chem. Res. **1994**, 27, 316–323.

(10) Papayannopoulos, I. A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. Mass Spectrom. Rev. **1995**, 14, 49–73.

(11) Williams, E. R. Tandem FTMS of Large Biomolecules. Anal. Chem. **1998**, 70, 179A–85A.

(12) Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L. Surface-induced dissociation of small molecules, peptides, and non-covalent protein complexes. J. Am. Soc. Mass Spectrom. **2008**, 19, 190–208.

(13) Martin Somer, A.; Macaluso, V.; Barnes, G. L.; Yang, L.; Pratihar, S.; Song, K.; Hase, W. L.; Spezia, R. Role of Chemical Dynamics Simulations in Mass Spectrometry Studies of Collision-Induced Dissociation and Collisions of Biological Ions with Organic Surfaces. J. Am. Soc. Mass Spectrom. **2020**, 31, 2–24.

(14) Lucas, K.; Barnes, G. L. Modeling the Effects of O-Sulfonation on the CID of Serine. J. Am. Soc. Mass Spectrom. **2020**, 31, 1114–1122.

(15) Barnes, G. L.; Shlaferman, A.; Strain, M. Fast fragmentation during surface-induced dissociation: An examination of peptide size and structure. Chem. Phys. Lett. **2020**, 754, 137716.

(16) Pratihar, S.; Barnes, G. L.; Hase, W. L.; Pratihara, S.; Barnes, G. L.; Hase, W. L. Chemical Dynamics Simulations of Energy Transfer, Surface-Induced Dissociation, Soft-Landing, and Reactive- Landing in Collisions of Protonated Peptide Ions with Organic Surfaces. Chem. Soc. Rev. **2015**, 45, 3595–3608.

(17) Pratihar, S.; Barnes, G. L.; Laskin, J.; Hase, W. L. Dynamics of Protonated Peptide Ion Collisions with Organic Surfaces. Consonance of Simulation and Experiment. J. Phys. Chem. Lett. **2016**, 7, 3142–3150.

(18) Frederickson, D.; McDonough, M.; Barnes, G. L. A Computational Comparison of Soft Landing of Alpha-Helical vs Globular Peptides. J. Phys. Chem. B **2018**, 122, 9549–9554.

(19) Barnes, G. L.; Shlaferman, A.; Strain, M. Fast fragmentation during surface-induced dissociation: An examination of peptide size and structure. Chem. Phys. Lett. **2020**, 754, 137716.

(20) Vázquez, S. A.; Otero, X. L.; Martinez-Nunez, E. A trajectory-based method to explore reaction mechanisms. Molecules **2018**, 23, 3156.

(21) Rodríguez, A.; Rodríguez-Fernández, R.; A. Vázquez, S.; L. Barnes, G.; J. P. Stewart, J.; Martínez-Núñez, E. tsscds2018: A code for automated discovery of chemical reaction mechanisms and solving the kinetics. J. Comput. Chem. **2018**, 39, 1922–1930.

(22) Rossich-Molina, E.; Salpin, J.-Y.; Spezia, R.; Martinez-Nunez, E. On the gas phase frag-

mentation of protonated uracil: a statistical perspective. Phys. Chem. Chem. Phys. **2016**,

(23) Trinajstic, N. Chemical Graph Theory, 2nd ed.; CRC Press, 1992; p 352.

(24) Barnes, G. L.; Young, K.; Yang, L.; Hase, W. L. Fragmentation and Reactivity in Collisions of Protonated Diglycine with Chemically Modified Perfluorinated Alkylthiolate-Self-Assembled Monolayer Surfaces. J. Chem. Phys. **2011**, 134, 094106.

(25) Barnes, G. L.; Hase, W. L. Energy Transfer, Unfolding, and Fragmentation Dynamics in Collisions of N-Protonated Octaglycine with an H-SAM Surface. J. Am. Chem. Soc. **2009**, 131, 17185–17193.

(26) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. J. Cheminform. **2012**, 4, 17.

(27) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX **2015**, 1-2, 19–25.

(28) Stewart, J. P. Mopac2016. 2016; `http://openmopac.net`.

(29) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; Van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. Eur. Biophys. J. **2011**, 40, 843–856.

(30) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated force field Topology Builder (ATB) and repository: Version 1.0. J. Chem. Theory Comput. **2011**, 7, 4026–4037.

(31) Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E. Testing and validation of the Automated Topology Builder (ATB) version

2.0: Prediction of hydration free enthalpies. J. Comput. Aided. Mol. Des. **2014**, 28, 221–233.

(32) Canzar, S.; El-Kebir, M.; Pool, R.; Elbassioni, K.; Mark, A. E.; Geerke, D. P.; Stougie, L.; Klau, G. W. Charge group partitioning in biomolecular simulation. J. Comput. Biol. **2013**, 20, 188–198.

(33) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. J. Comput. Chem. **2006**, 27, 1101–1111.

(34) Homayoon, Z.; Pratihar, S.; Dratz, E.; Snider, R.; Spezia, R.; Barnes, G. L.; Macaluso, V.; Martin Somer, A.; Hase, W. L. Model Simulations of the Thermal Dissociation of the $TIK(H^+)_2$ Tripeptide: Mechanisms and Kinetic Parameters. J. Phys. Chem. A **2016**, 120, 8211–8227.

(35) Gu, M.; Zhang, J.; Hase, W. L.; Yang, L. Direct Dynamics Simulations of the Thermal Fragmentation of a Protonated Peptide Containing Arginine. ACS Omega **2020**, 5, 1463–1471.

(36) Meroueh, O.; Hase, W. L. Collisional Activation of Small Peptides. J. Phys. Chem. A **1999**, 103, 3981–3990.

(37) Hase, W. L.; Buckowski, D. G. Monte carlo sampling of a microcanonical ensemble of classical harmonic oscillators. Chem. Phys. Lett. **1980**, 74, 284–287.

(38) Schlier, C.; Seiter, A. High-Order Symplectic Integration: An Assessment. Comput. Phys. Commun. **2000**, 130, 176–189.

(39) Stewart, J. P. Mopac2012. 2012; `http://openmopac.net`.

(40) Gregg, Z.; Ijaz, W.; Jannetti, S.; Barnes, G. L. The Role of Proton Transfer in Surface-Induced Dissociation. J. Phys. Chem. C **2014**, 118, 22149–22155.

(41) Frisch, M. J. et al. Gaussian 09, Revision D.01. 2009.

(42) Zhang, P.; Chan, W.; Ang, I. L.; Wei, R.; Lam, M. M.; Lei, K. M.; Poon, T. C. Revisiting Fragmentation Reactions of Protonated α-Amino Acids by High-Resolution Electrospray Ionization Tandem Mass Spectrometry with Collision-Induced Dissociation. Sci. Rep. **2019**, 9, 6453.

(43) Dookeran, N. N.; Yalcin, T.; Harrison, A. G. Fragmentation reactions of protonated α-amino acids. J. Mass Spectrom. **1996**, 31, 500–508.

(44) Shaikh, K.; Blackwood, J.; Barnes, G. L. The Effect of Protonation Site and Conformation on Surface-Induced Dissociation in a Small, Lysine Containing Peptide. Chem. Phys. Lett. **2015**, 637, 83–87.

(45) Humbel, S. Short strong hydrogen bonds: A valence bond analysis. J. Phys. Chem. A **2002**, 106, 5517–5520.