



HAL
open science

Fragmentation Spectra Prediction and DNA Adducts Structural Determination

Andrea Carrà, Veronica Macaluso, Peter W Villalta, Riccardo Spezia, Silvia
Balbo

► **To cite this version:**

Andrea Carrà, Veronica Macaluso, Peter W Villalta, Riccardo Spezia, Silvia Balbo. Fragmentation Spectra Prediction and DNA Adducts Structural Determination. *Journal of The American Society for Mass Spectrometry*, 2019, 30 (12), pp.2771-2784. 10.1007/s13361-019-02348-7. hal-02385066

HAL Id: hal-02385066

<https://hal.sorbonne-universite.fr/hal-02385066>

Submitted on 28 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fragmentation Spectra Prediction and DNA Adducts Structural Determination

Andrea Carrà,¹ Veronica Macaluso,³ Peter W. Villalta,¹ Riccardo Spezia^{2*} and Silvia Balbo^{1*}

¹ *Masonic Cancer Center, University of Minnesota, 2231 6th Street SE, 55455 Minneapolis MN, USA*

² *Sorbonne Université, CNRS, Laboratoire de Chimie Théorique, LCT, F. 75005 Paris, France*

³ *Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement, Université d'Evry, CEA, CNRS, Université Paris Saclay, Bd. F. Mitterrand, 91025 Evry Cedex, France.*

4/10/2019

Correspondence to: riccardo.spezia@sorbonne-universite.fr , balbo006@umn.edu

Abstract

In this work, chemical dynamics simulations were optimized and used to predict fragmentation mass spectra for DNA adduct structural determination. ⁶O-methyl-guanine (⁶O-Me-G) was used as a simple model adduct to calculate theoretical spectra for comparison with measured high-resolution fragmentation data. An automatic protocol was established to consider the different tautomers accessible at a given energy and obtain final theoretical spectra by insertion of an initial tautomer. In the work reported here, the most stable tautomer was chosen as the initial structure, but in general any structure could be considered. Allowing for the formation of the various possible tautomers during simulation calculations was found to be important to getting a more complete fragmentation spectrum. The calculated theoretical results reproduce the experimental peaks such that it was possible to determine reaction pathways and product structures. The calculated tautomerization network was crucial to correctly identifying all the observed ion peaks, showing that a mobile proton model holds not only for peptide fragmentation but also for nucleobases. Finally, first-principles results were compared to simple machine learning fragmentation models.

Keywords: Adductomics; Collision Induced Dissociation; Fragmentation Prediction; Mobile Proton Model.

1. Introduction

In the era of genome-wide association studies and data-rich -omics analyses, complementary new technologies are needed to investigate the genotoxic effects resulting from complex exposures.[1] In this context, DNA adductomics is emerging as a powerful new tool for the characterization of DNA damage. This method allows the simultaneous detection of all covalently modified nucleosides (DNA adducts) in hydrolyzed DNA samples. The approach is based on the observation that collisional induced dissociation (CID) of protonated nucleosides leads to the fragmentation of the glycosidic bonds, resulting in the constant neutral loss (CNL) of the 2'-deoxyribose moieties and the formation of the corresponding protonated modified nucleobase ions. Our adductomic approach differs from conventional DNA adductomic approaches, and other data dependent methodologies, by the inclusion the triggering of high resolution/accurate mass (HRAM)-MS³ fragmentation upon the observation of the exact CNL (5 ppm). The presence of an MS³ event indicates detection of a DNA adduct with the corresponding fragmentation spectrum providing structural information regarding the detected adduct.[2]

DNA adductomics is still in its infancy and consequently, unlike other -omic techniques, it is not supported by automated data analysis tools or databases for structural identification of the observed DNA adducts. Fragmentation data is available for only a very limited number of DNA adducts and often only at the MS² level for nucleoside adducts, which is dominated by the loss of the deoxyribose moiety, and therefore not useful for adduct identification.[3] There are likely countless DNA adducts resulting from xenobiotic exposures and probably many endogenous adducts that have yet to be observed. The lack of DNA fragmentation data limits the ability of adductomic methodologies to confirm or determine the identity of the DNA adducts that have been reported.

In this work, direct chemical dynamics simulations were used to understand fragmentation pathways, corresponding to MS³ of DNA adducts. This method predicts products and fragmentation pathways independently from experiments and without pre-imposing any reaction channel. In this way, it is possible to obtain a theoretical CID fragmentation spectrum for any given structure resulting from the physical evolution of an activated system.[4] This approach was pioneered by Hase and co-workers,[5,6,7,8] and applied to many systems from small organic molecules[9] to peptides[10,11] and sugars.[12,13] Recently, two investigations of protonated[14] and cationized[15] uracil have been performed using the same direct dynamics approach and are the only two studies reported for nucleobases. Most chemical dynamic studies are aimed at analysis of data collected using triple-quadrupole instrumentation with conditions close to the single collision limit, and therefore consider explicit collisions between the molecular ion and an inert gas atom. Recently, the approach was extended to multiple collision activation (which is more appropriate for ion trap-type fragmentation) by activating the ion at a given internal energy that is randomly distributed among the different vibrational normal modes following a microcanonical statistical distribution, and used to study the fragmentation of small peptides.[16,17,18,19] Chemical dynamics is growing in popularity for the reproduction of mass spectrometric data, as indicated by its extension to surface induced dissociation of peptides [20,21,22] and electron ionization mass spectra of organic and inorganic molecules.[23,24,25]

In the present work, the fragmentation of the aglycone ion (O⁶-methyl-guanine, O⁶-Me-G) of O⁶-Me-dG, one of the most relevant DNA adducts, was modeled by means of chemical dynamics simulation with internal energy activation. This DNA adduct derives from guanine methylation, resulting from exposure of DNA to methylating agents, including those deriving from the metabolism, via cytochrome P450 oxidation of the tobacco specific nitrosamine and potent lung carcinogen: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK).[26] It is a well-established marker for NNK-derived DNA damage and plays a key role in the NNK-mediated lung carcinogenesis process.[27]

In the present work, a computational approach is proposed based on a coupled exploration of proton transfer (linking together the possible tautomers of O⁶-Me-G) and fragmentation, allowing the creation of a tautomerization network which is strictly related to the fragmentation processes. More generally, this semi-automated approach could be extended to all the systems in which CID reactivity is related to proton mobility, which is especially suited to DNA adductomic analysis due to the multiple tautomers of the nucleobases, and is the first step to building *in silico* fragmentation libraries.

2. Materials and Methods

2.1 Fragmentation Spectra Acquisition

The fragmentation spectra acquisition was performed by direct infusion with a syringe pump at 5·μL·min⁻¹ of O⁶-Me-dG (1 pmol·μL⁻¹) in H₂O:CH₃OH (1:1) with 0.01%_{v/v} formic acid into the electrospray ion source of an Orbitrap Fusion mass analyzer (Thermo Scientific). Ionization was performed in positive ion mode with a spray voltage of 3 kV, ion transfer tube temperature of 250 °C, and an S-Lens voltage of 60 V with the spectra acquired at a resolution of 60,000. The overall MS data acquisition consisted of three events: 1) the precursor ion (*m/z* 282) was isolated by the front quadrupole (isolation width of 1.5 *m/z*), 2) fragmented of the precursor ion was performed in the ion trap (CID level of 50%), 3) and the resulting fragment ion, O⁶-Me-G (*m/z* 166), formed upon cleavage of the glycosidic bond (MS² fragmentation), was further fragmented in the HCD cell at different energy levels (from 0 % to 100 %). The molecular formulas of the MS³ spectra ions were calculated using the Xcalibur software package (ver. 3.0.63 Thermo Scientific).

2.2 Electronic Structure Calculations

Geometries of the reactant ion (and some product ions) were determined by geometry optimization at different levels of theory. For O⁶-Me-G, six tautomers, corresponding to the protonation of the six heteroatoms (N¹, N², N³, O⁶, N⁷ and N⁹), as shown in **Figure 1** were considered. For each possible protonation site, 40 conformations of O⁶-Me-G with six tautomers (N¹, N², N³, O⁶, N⁷ and N⁹) were used as the starting points. The optimized geometry and energy (corrected by zero-point energy) of the minimum energy structure of each tautomer are then reported for comparison and used in subsequent calculations. Density functional theory (DFT) was used with the B3LYP functional [28,29] and the 6-311+G(d,p) basis set. This was considered as the reference calculation to evaluate the best semi-empirical Hamiltonian (SEH) for use in the reaction dynamics calculations. Several SEHs were considered, in particular: AM1,[30] AM1-D,[31] RM1,[32] RM1-D, PM3,[33] PM3-D, PM6 [34] and PM6-D, where D stands for dispersion correction with the method of Grimme [35]. For each minimum, the vibrational frequencies were calculated to verify that the calculated geometry is a real energetic minimum and to correct the electronic energy by the zero-point energy (ZPE). The optimized structures were used as initial structures for the chemical dynamic simulations (see section 2.3).

To determine the best SEH for use in the reaction dynamic calculations, the different SEHs were benchmarked with respect to B3LYP calculations. First, the relative energies of the different tautomers and their geometries were considered. The calculated root mean square deviation of atomic positions (RMSD) was used to quantify the geometric distortion induced by the use of different computational approaches. Secondly, the formation energies of some fragmentation products (chosen based upon examination of the experimental data) were calculated. The comparison between SEH and B3LYP calculations was done by considering the mean absolute errors (MAE) for the energies and the root mean square deviations of atomic positions (RMSD) for the geometries. In both cases B3LYP results were considered as reference values. DFT calculations were done with the Gaussian09

package,[36] while SEH calculations were performed with the MOPAC-version 5.022mn software.[37]

2.3 Chemical Dynamics Simulations

Direct dynamics simulations were performed using the VENUS package [38] coupled with MOPAC-version 5.022mn software. Energies and forces were calculated on-the-fly at PM6-D and AM1 levels of theory, which were determined to be the best performing semi-empirical methods (see Results Section 3.2). The unrestricted formalism was used in the quantum chemistry calculations. Note that given the number of atoms (and electrons) of the system and the number of trajectories, chemical dynamics simulations at the DFT level are currently computationally too expensive and the best performing SEHs were used for simulations. Note that the use of these SEHs was reported to be in agreement with experimental mass spectra in several systems.[4]

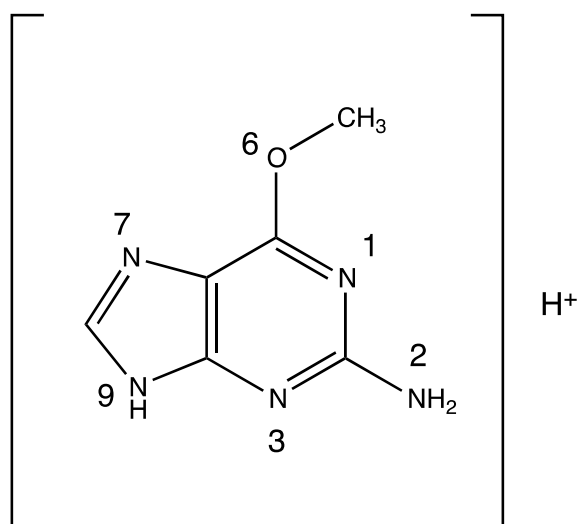


Figure 1. *O*⁶-Me-G structure with numbering indicating different protonation sites.

The minimum energy geometry of each tautomer was used as the initial structure and activated at given internal energies (E_{int}) using a classical microcanonical normal mode sampling.[39] The initial rotational energy was sampled from a 300 K Boltzmann distribution. Two sets of simulations were performed: (i) $E_{\text{int}} = 1254$ kJ/mol given to the energetic minimum of each tautomer such that they all have the same amount of internal energy and different total energies, (ii) $E_{\text{int}} = E_{\text{T}} + \Delta$, where $E_{\text{T}} = 1254 + \Delta$ kJ/mol for the energetically most stable tautomer (N^7) and Δ is the energy difference between N^7 and the other tautomers. The energy value was chosen in order to have enough reactive trajectories on timescales available for chemical dynamics simulations.

For each internal energy and each initial tautomer, 1000 trajectories were run, each 10 ps long. Newton's equations of motion were numerically integrated using a sixth-order symplectic algorithm,[40,41] with a step size of 0.2 fs which assures good internal energy conservation. In total, six tautomers per energy and method were investigated for a total of approximately 18000 trajectories and 1800 ns of simulated reaction time. From a computational point of view, each trajectory needed approximately 2.5 hours on a single processor of the local cluster for a total computational time of 45000 hours.

2.4 Automatic determination of tautomerization network and fragmentations

Fragmentation trajectory calculations were done in Cartesian coordinates, and thus the outputs are pure xyz files without any chemical information. By using an in-house code based on graph

theory,[16,42] product ion or tautomer formation from a given trajectory could be automatically determined as could charge position and Lewis structure from Natural Bond Orbital (NBO) analysis[43]. For each trajectory, there are three simulated outcomes: fragmentation, tautomerization, or the unchanged parent ion and tautomer. A four-stage simulation procedure was developed and is shown in **Figure 2**, with the first step being determination of the starting point as the most stable tautomer of the precursor ion at a given internal energy (see Section 2.3). Next, the trajectory simulation was performed at a given internal energy (see Section 2.3), followed by the analysis of simulation result by in-house software to automatically determine the final fragment ion structure or the tautomer formed. Upon formation of a tautomer, a new trajectory is initiated, keeping track of which tautomers resulting in which fragment ions. The final theoretical mass spectrum of a given energy and SEH is obtained by summing up the results of the 1000 trajectories run for each initial protonation site.

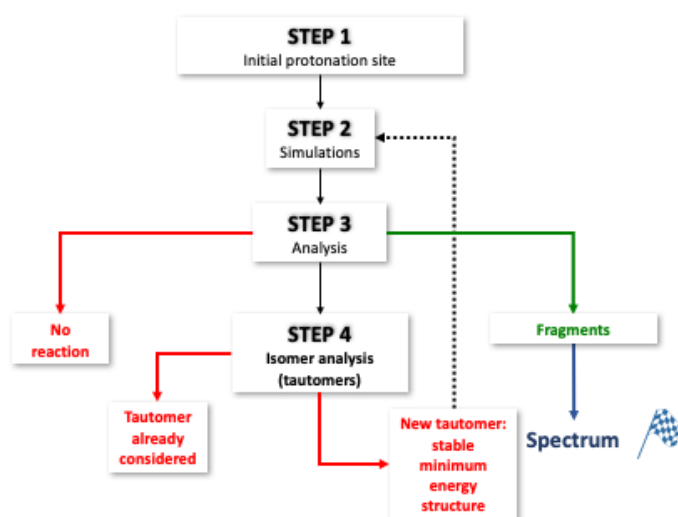


Figure 2. Schematic representation of the procedure used to automatically build the tautomerization network and fragmentation products.

3. Results and discussion

3.1 Experimental Spectrum

MS³ spectra recorded at increasing collision energies showed different fragmentation profiles. The largest number of fragment ions was observed at an HCD level of 50 % (**Figure 3A**). Collision energies of higher and lower values provided no or low intensity ions in the high and low mass regions, respectively (**Figure 3B and C**).

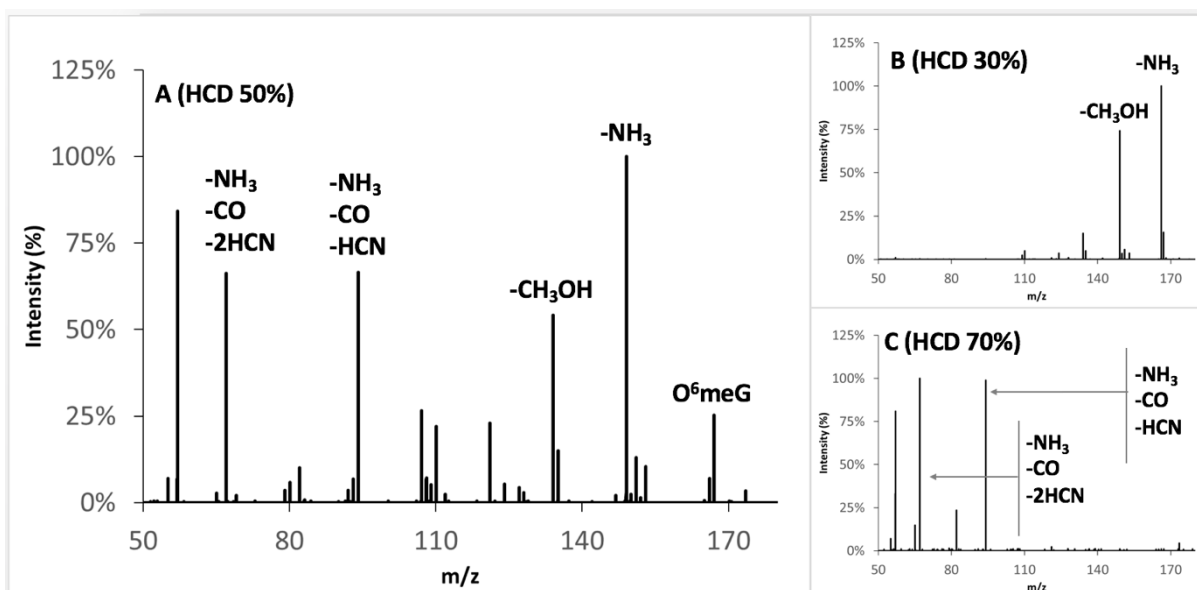


Figure 3. Fragmentation of O^6 -Me-G at different HCD collision levels. Panels A, B and C show fragmentation spectra recorded at HCD level of 30 %, 50 % and 70 %, respectively.

The MS³ fragmentations, observed with the HCD level of 30%, were dominated by the neutral losses of NH₃ and CH₃OH (m/z 149.0461 and 134.0464, respectively). On the other hand, the low mass region of the 70% HCD spectrum is dominated by the combined neutral losses of HCN, NH₃ and CO.

3.2 Benchmark of Semi-Empirical Methods.

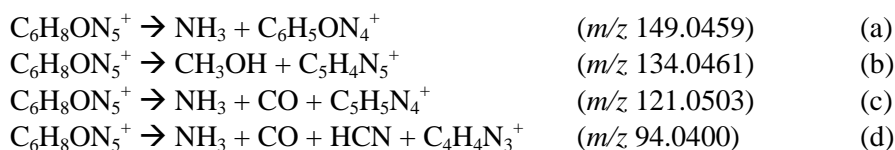
The results of the SEH benchmark calculations on the energies and geometries of the six tautomers of the precursor ion are shown in **Table 1**. The results show that PM6-D provides the smallest MAE and RMSD, and therefore is deemed to be the most accurate SEH for reproducing the relative energies and geometries of the precursor ion tautomer. This is important for providing a reliable tautomerization network.

Table 1. Comparison of performances of the SEHs for calculations of both tautomers and reactions, where B3LYP/6-311++G(d,p) calculations are used as references.

Method	Tautomers		Reactions
	MAE (kJ/mol)	RMSD (Å)	MAE (kJ/mol)
PM6-D	14.9	0.076	66
PM6	15.1	0.081	65
PM3-D	32.3	0.101	148
PM3	24.2	0.087	65
RM1-D	29.0	0.079	88
RM1	23.9	0.079	75
AM1-D	28.6	0.113	81
AM1	24.4	0.119	57

The relative energies of the different tautomers using B3LYP are shown in **Figure 4A**, and all the relative energies obtained from SEHs are reported in **Table S1**. The same stability order among the tautomers was observed independent of the level of theory used with N⁷ being the lowest in energy. Based on a Boltzmann population, tautomer N⁷ is the only populated protonation state before activation at the experimental conditions (\approx 300 K). The geometries of the minima are reported in **Table S2** of the supporting information.

The performance of the SEHs were evaluated using several experimentally observed fragment ions, i.e. by simple qualitative consideration of suspected ion structures. The reactions considered are:



The MAEs of the fragment processes obtained from the different methods (B3LYP is also considered here as the reference) are reported in **Table 1** and the details of the energies are reported in **Table S3**. Of note, AM1 provides the best agreement, but PM6-D (as well as PM6) provides very similar values. All the SEHs overestimate the reaction energies for reactions (a)-(c) while they underestimate the energetics for reaction (d). While such MAEs (57 and 66 kJ/mol for AM1 and PM6-D, respectively) are relatively large with regards to the quantitative calculation of reaction barriers and related kinetics, they could be reasonable for the qualitative description of the fragmentation processes, in particular when a much larger activation energy is used in simulations. These values are in fact about 5% of the activation energy and differences of this scale have been shown to be useful in understanding other systems such as the reactivity of deprotonated diproline [18] and L-cysteine sulfate anion [19].

Overall, based on the results on tautomers (relative energy and geometry) and reaction products, the PM6-D and AM1 Hamiltonians were used in the chemical dynamics simulations. Note that PM6-D was recently used with success to study other gas phase reactions.[19,42,44]

3.3 Tautomerization network

The activation energy must be defined in chemical dynamics simulations. Since it cannot be determined from experimental parameters in the present case (e.g. the HCD value does not directly relate to the amount of energy that the ion gets), the activation energy was set empirically based on the results of the preliminary fragmentation runs. Energy values in a reasonable range (800-1460 kJ/mol) were tested and it was determined that a value of 1254 kJ/mol provides enough reactivity to have statistically converged results in the computationally allowed simulation time-length. Two sets of energies were then considered: the same E_{int} (= 1254 kJ/mol) for all tautomers or a different value as a function of their relative energy ($E_{\text{int}} + \Delta$, as described in Section 2.3). However, tautomerization and fragmentation results are relatively unaffected by these energy differences and therefore only the last case (i.e. $E_{\text{int}} + \Delta$) is discussed for simplicity. Note that these values are much higher than the fragmentation barriers, being about 5% of the activation energy, so that errors in the SEHs should have a little effect on the fragmentation processes.

Simulations indicated the interconnectedness of all six tautomers, as graphically described in **Figure 4**, with the lowest energy tautomer (N^7) isomerizing directly to N^1 and N^9 , which further isomerize to N^2 (or O^6) and N^3 , respectively. This tautomerization is consistent with the well-known mobile proton model which is relevant for peptide fragmentation,⁴⁵ with proton transfer occurring most likely between protonation sites close in space. In this sense, the N^1 tautomer plays a key role because it is very low energetically and spatially close to the N^7 position, thus it can easily receive the proton. Moreover, it is geometrically close to the energetic N^2 and O^6 positions. Thus, the presence of protonation sites with low relative energy, like N^1 (+10 kJ/mol) and N^3 (+30 kJ/mol), facilitates the transfer of the proton, making accessible the higher energy tautomers which are spatially distant (or sterically hindered) from the minimum energy proton position.

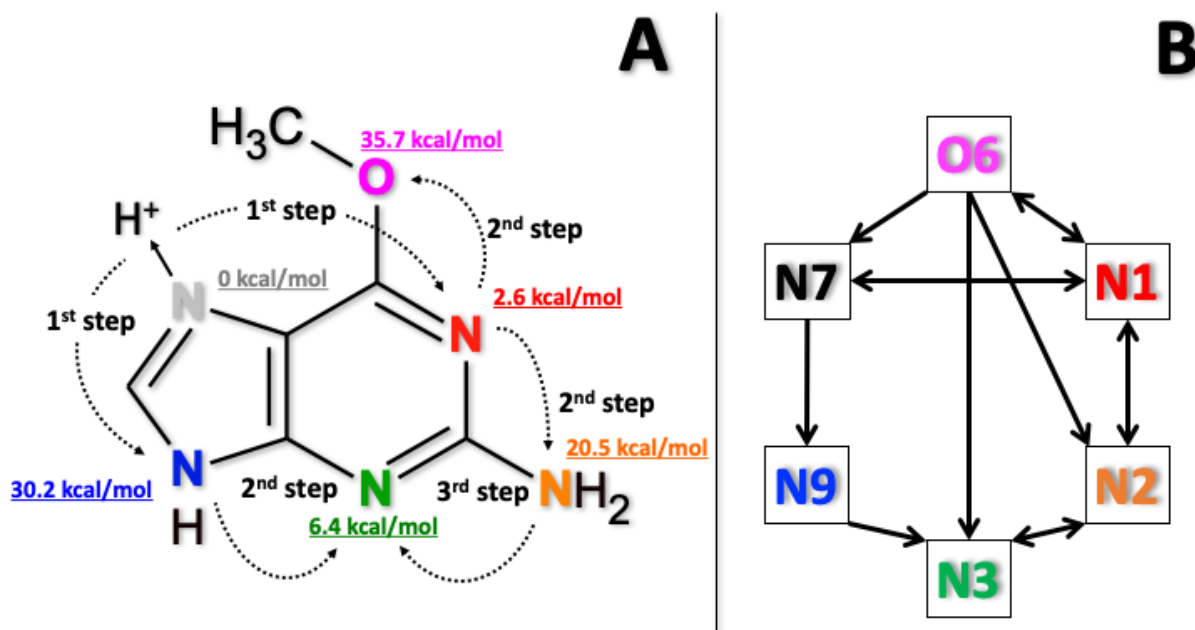


Figure 4. Proton transfer steps and schematic isomerization network observed in *O*⁶-Me-G chemical dynamics simulations. Panel A: proton motions observed from simulations with relative energy of different tautomers. Panel B: schematic isomerization network.

3.4 Fragmentation products

Once the tautomerization connectivity was clarified, which basically shows that all the tautomers are accessible at the energy (and time-length) of the simulation conditions, the fragmentation products could be determined. The product ions (as *m/z*) resulting from PM6-D simulations are summarized in **Figure 5** (similar results are obtained from AM1 simulations and are reported in **Figure S1**). The fragmentation spectra of the tautomers N¹, N², N³, O⁶, N⁷ and N⁹ are shown in panels A, B, C, D, E and F, respectively, each plotted opposite the experimental fragmentation spectrum of *O*⁶-Me-G at an HCD level of 50%. A visual inspection of all fragmentation profiles reveals varying amounts of agreement between the different tautomers and the experimental spectrum. As can be seen from fragmentation spectra simulations, the experimentally abundant fragment ions *m/z* 149.0459 and 134.0461, open chain fragments deriving from the neutral loss of NH₃ and CH₃OH, were preferentially formed from the protonated N² and O⁶ sites, strengthening the importance of proton mobility in opening up the fragmentation pathways.

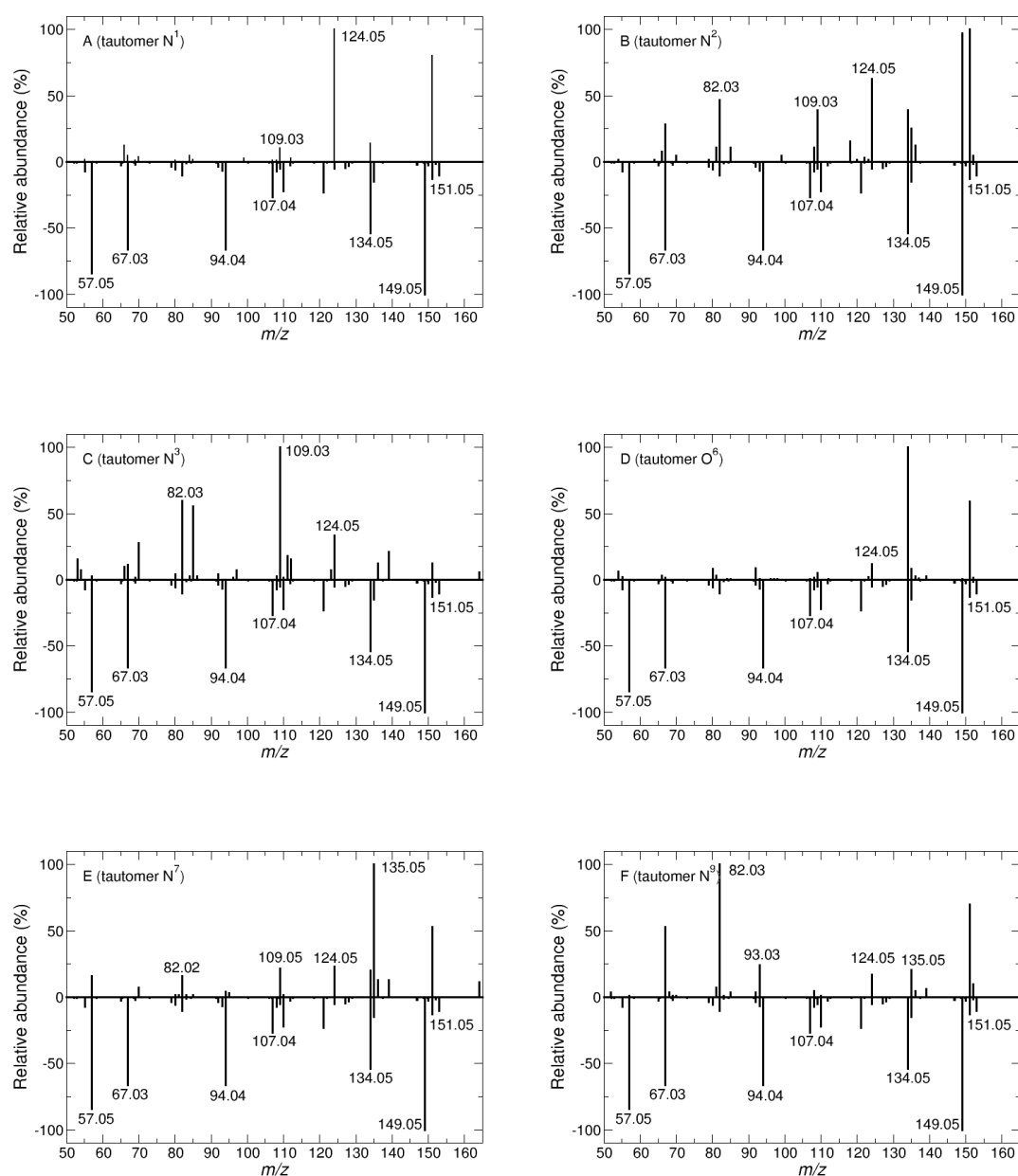


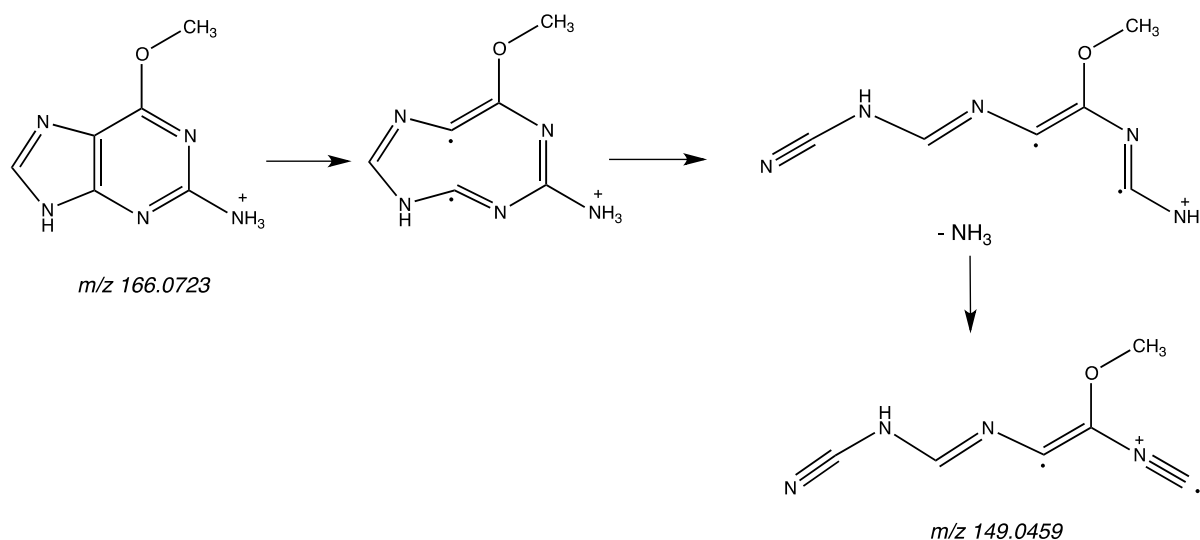
Figure 5. Comparison between computational (using PM6-D and activation energies corresponding to $1254 + \Delta kJ/mol$ from the most stable structure) and experimental (HCD 50%) fragmentation spectra of O^6 -Me-G. The experimental spectrum is reported with negative intensities whereas the computational spectra are reported with positive intensities. Panels A, B, C, D, E and F correspond to the fragmentation of N^1 , N^2 , N^3 , O^6 , N^7 and N^9 tautomeric precursors, respectively.

3.5 Fragmentation mechanisms

Using chemical dynamics simulations, both final product ions and the corresponding fragmentation mechanisms can be determined. As discussed previously, a critical first step is the transfer of the proton between different tautomer positions. Unfortunately, there is no experimental evidence as to the populated tautomer in the gas phase, and thus the relative energies were determined using DFT calculations. Protonation of the N^7 position was found to be the most stable tautomer and therefore used as the starting point for the simulations; however, all the tautomers are connected when the ion

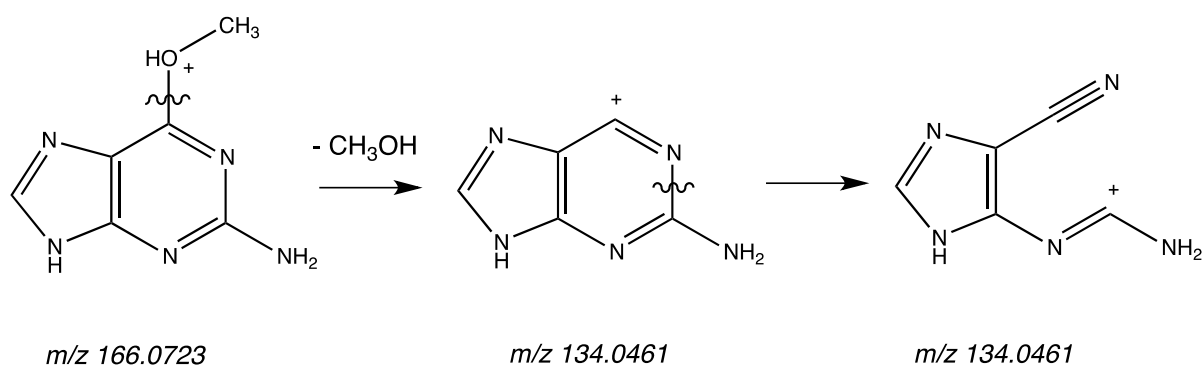
gets sufficiently energetic (as reported in previous section), such that the overall simulated fragmentation is not strongly dependent upon the first tautomer used in the iterative procedure described in **Figure 2**. The mechanisms corresponding to the formation of the six most abundant ions observed in the experimental spectra, i.e.: m/z 149.0459, 134.0461, 94.0400, 67.0291, 107.0357 and 57.0452, are discussed below.

The neutral loss of NH_3 results in the production of m/z 149.0459 and the simulations indicated that the production of the ion requires the migration of the proton to the N^2 site prior to fragmentation. The most direct way (i.e. least number of proton transfers from the most stable tautomer, N^7) consists first of a proton transfer to the N^1 position followed by a second transfer to the N^2 position and then neutral loss of ammonia. Note, the N^2 tautomer can also be formed from the N^3 and O^6 tautomers. This fragmentation pathway is summarized in **Scheme 1**.



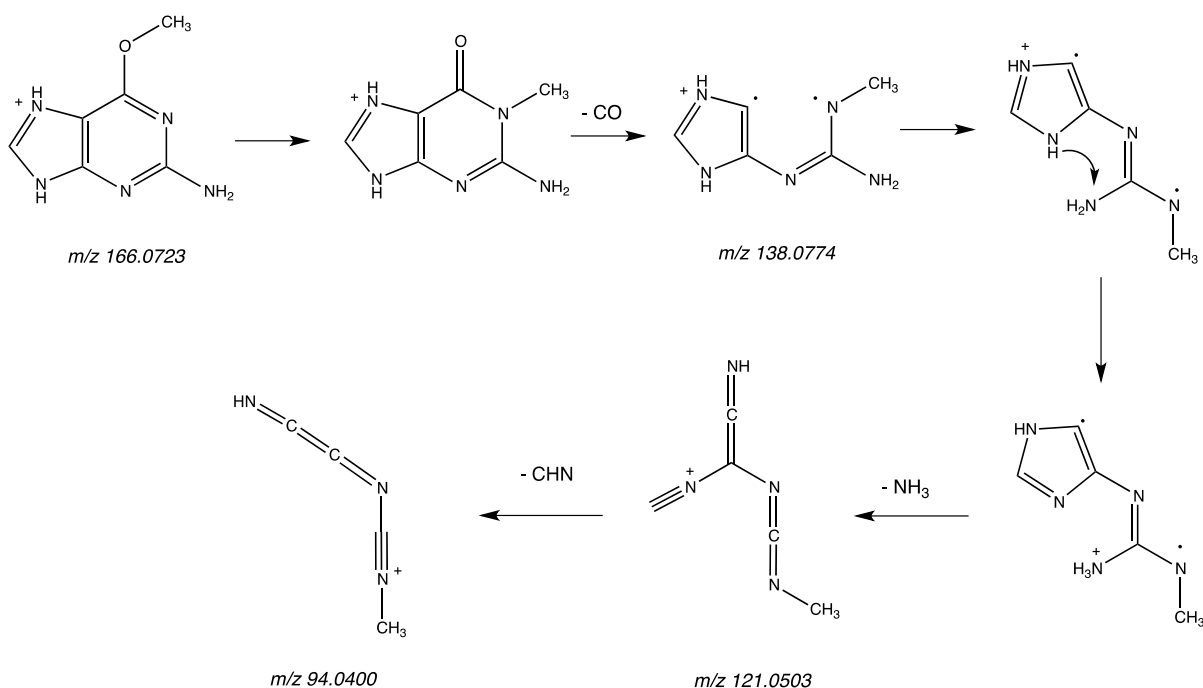
Scheme 1. Formation mechanism of ion m/z 149.0459 from the tautomer N^2 as obtained from chemical dynamics simulations.

Fragment ion m/z 134.0461 results from the neutral loss of CH_3OH , and the fragmentation mechanism as observed in simulations is reported in **Scheme 2**. The mechanism involves a proton transfer to the lone pairs of electrons of the O^6 position and the subsequent heterolytic cleavage of the σ bond between C^6 and O^6 . The positive charge then localized on the C^6 induces the opening of the purinic ring between N^1 and C^2 . Finally, the positive charge is localized on C^2 and a nitrile group is formed with the N^1 terminal atom. The resulting ion is also obtained from simulations starting from different tautomers, but in every case the proton has to migrate to oxygen atoms. Of course, when starting from the O^6 tautomer, the loss of CH_3OH happens immediately.



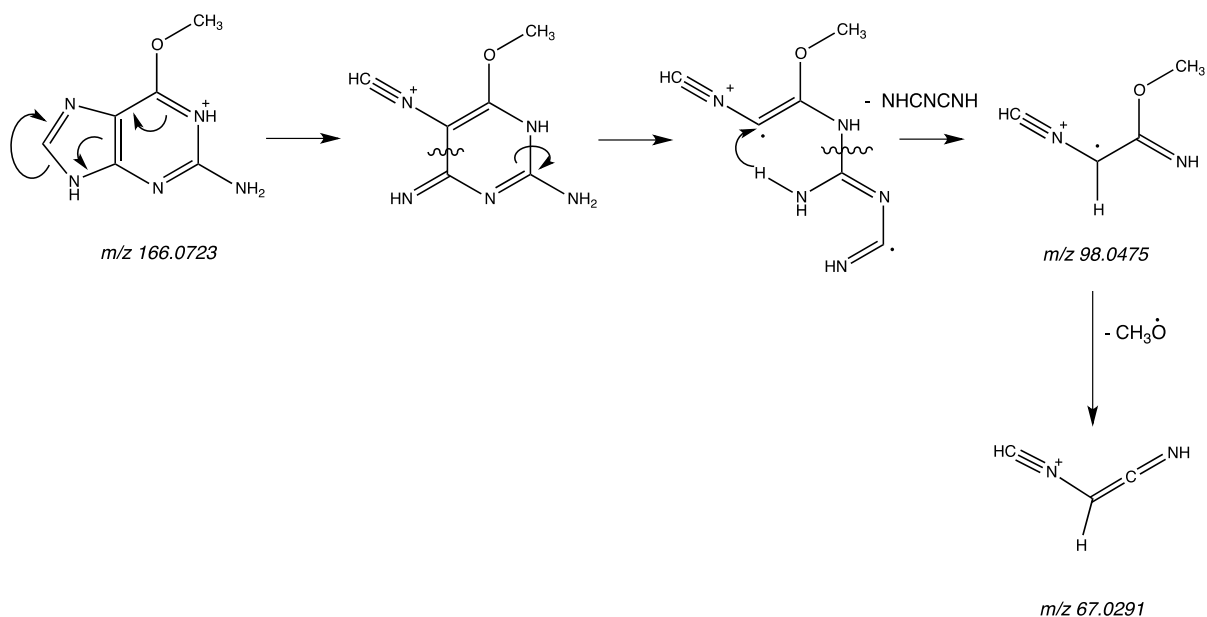
Scheme 2. Formation mechanism of ion m/z 134.0461 from the tautomer O^6 as obtained from chemical dynamics simulations.

Ion m/z 94.0400 is only obtained directly from tautomer N^7 , mainly with the mechanism shown in **Scheme 3** (another less abundant mechanism is reported in **Scheme S1** of the supporting information). The fragmentation is triggered by the transfer of the methyl group to the adjacent N^1 atom, forming an unstable species with cleavage of C-O and N-C bonds forming unstable intermediate which then loses neutral CO. The formed ion is also unstable and it sequentially loses NH_3 and CHN resulting in the final product ion m/z 94.0400. Note that the intermediates shown in **Scheme 3** (and all schemes issued from chemical dynamic simulations) are not stationary points but just structures observed in trajectories. Thus, the surviving structures could further react on longer time-scales.



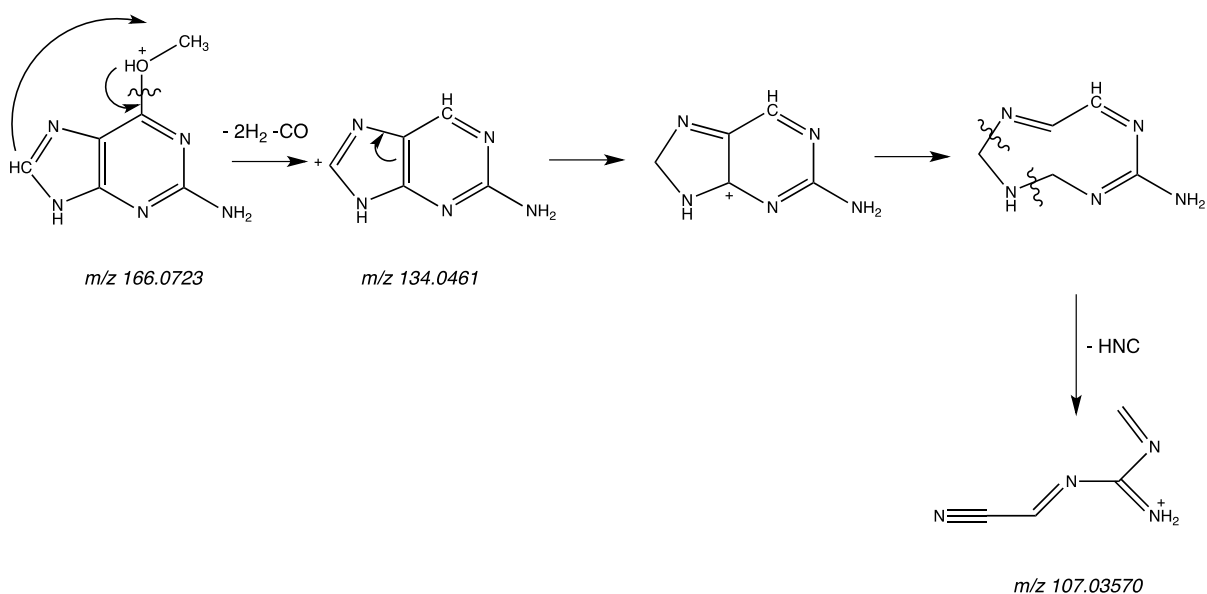
Scheme 3. Formation mechanism leading to ion m/z 94.0400 from tautomer N^7 as obtained from chemical dynamics simulations.

Ion m/z 67.0291 is obtained from tautomers N^1 , N^2 , N^3 , O^6 and N^9 . The most probable pathway is the one shown in **Scheme 4**. The first step consists of forming the tautomer N^1 (directly from the most stable tautomer N^7 , but also from N^2 and O^6) and then the opening of the 5-membered ring occurs followed by the opening of the 6-membered ring. The resulting species rearranges and the N^1 -C bond breaks, forming an intermediate ion m/z 98.0475 which then loses CH_3O forming the final species m/z 67.0291.



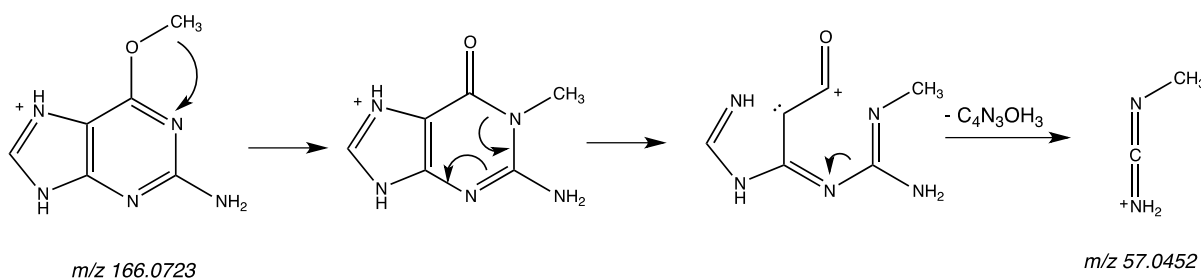
Scheme 4. Formation mechanism leading to ion m/z 67.0291 starting from N^1 tautomer as obtained from PM6-D chemical dynamics simulations. Charges, bonds and unpaired electrons are shown as resulting from NBO calculations.

To obtain fragment ion m/z 107.0357, tautomer O^6 must first be generated, followed by the transfer of the proton to the adjacent carbon atom, inducing the breaking of the C-O bond. The leaving OCH_3 group is highly reactive: it first loses H_2 followed by the transfer of an H atom from the ring, resulting in the formation of H_2 and CO molecules. The ring is now highly unstable, and the breaking of two C-N bonds leads to the formation of an intermediate metastable ion. This metastable ion loses HNC forming the final m/z 107.0357 ion. The mechanism is shown in **Scheme 5**. Note that, as detailed in section 3.2, the proton transfer step is crucial: the O^6 tautomer, in fact, can be generated from the following sequence (assuming the initial tautomer being the most stable one): $N^7 \rightarrow N^1 \rightarrow O^6$ or $N^7 \rightarrow N^9 \rightarrow N^3 \rightarrow N^2 \rightarrow N^1 \rightarrow O^6$.



Scheme 5. Formation mechanism leading to ion m/z 107.0357 from tautomer O^6 as obtained from chemical dynamics simulations.

Finally, ion m/z 57.0452 can occur from tautomers N^7 , N^3 and N^9 . The mechanism obtained from the most stable tautomer is shown in **Scheme 6**. It is initiated by the migration of CH_3 to N^1 which induces the cleavage of the N^1 - C^6 bond (opening of 6-membered ring), followed by breakage of the N^9 - C^4 bond (opening of 5-membered ring). The N^3 - C^2 bond of the resulting ion breaks leading to the formation of the m/z 57.0452 ion. Simulations generated product ions with masses of m/z 57.0452, corresponding to the $\text{C}_2\text{H}_5\text{N}_2^+$ stoichiometry, but other trajectories generated ion m/z 57.0214, with less abundance, corresponding to $\text{C}_2\text{H}_3\text{ON}^+$ stoichiometry. The experimentally measured mass indicates that the $\text{C}_2\text{H}_5\text{N}_2^+$ stoichiometry is the one formed, illustrating the value of accurate mass measurements for use in efforts to develop *in silico* approaches to simulate mass spectra as well as for use in fragmentation database searching.



Scheme 6. Formation mechanism leading to ion m/z 57.0452 from tautomer N^7 as obtained from chemical dynamics simulations.

3.6 Comparison with *in silico* CFM-ID results

The chemical dynamic simulation results and experimental spectra are compared here with the MS^2 spectra predicted with the competitive fragmentation model by Allen et al. [46] using the CFM-ID web interface [47]. Note that in this approach the spectra are provided by using a probabilistic generative model for the fragmentation process, while in chemical dynamics simulations fragmentation products are directly obtained from dynamical evolution of a molecular system to which excess energy is injected. The CFM-ID results using three energies (10, 20 and 40 V) are shown in **Figure 6** along with the 50% HCD experimental spectrum for comparison.

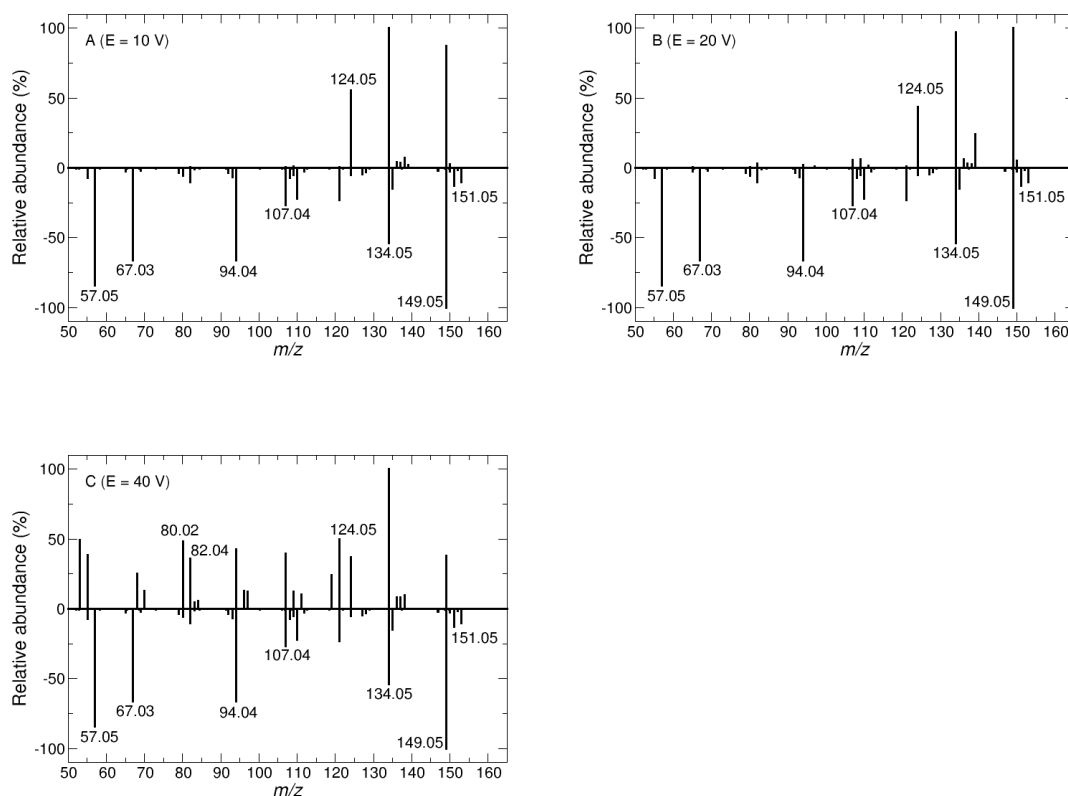


Figure 6. Comparison of the CFM-ID prediction with experimental data (50% HCD) for three energy values.

An overall quantitative comparison between CFM-ID, chemical dynamics simulations and experimental spectra was done by means of a similarity (S) function, defined as:

$$S = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_i^n A_i B_i}{\sqrt{\sum_i^n A_i^2} \sqrt{\sum_i^n B_i^2}}$$

where A_i and B_i are two vectors composed by the logarithm of the signal of each m/z value. Note that, similarly to what is done in spectroscopy,[48,49] the logarithm of the intensity is considered in order to enhance the weight of the peak position and reduce the role of differences in intensity. Results using the different tautomers as initial structure and summing up all the chemical dynamics products are shown in **Table 2** along with the CFM-ID results. First, it should be noted that CFM-ID provides a similarity value of 45.5% at 20 and 40 V and 6.2% at 10 V. PM6-D simulations starting from the most stable tautomer, N^7 , provide a similar value (42.5%) and a slightly better value (46.1%) when starting from N^2 . Less similarity is observed when starting from other tautomers. Summing up all the MS^2 spectra obtained by chemical dynamics simulations results in a similarity of 46.6% which is similar to CFM-ID values at 20 and 40 V. The two activation energies provide globally the same results (see Table 2). AM1 simulations provide similarity values which are much less in agreement. This is probably due to the poorer performances of AM1 in reproducing proton transfer which is a key step in fragmentation mechanisms as shown in previous section.

Table 2. Similarity between calculated and experimental spectra. Values shown for N^1 , N^2 , N^3 , O^6 , N^7 and N^9 are the results for the chemical dynamic simulations of the corresponding individual tautomers. The value for “Total simulations” is the result corresponding to all the chemical dynamics simulations products. CFM-ID results are also shown for the three energies.

System	Similarity (%)		
	PM6-D		AM1
E (kJ/mol)	1254 + Δ	1254	1254
N^1	24.9	37.1	20.9
N^2	46.1	43.0	41.8
N^3	33.8	25.6	27.5
O^6	31.9	31.0	20.4
N^7	42.5	31.7	25.1
N^9	37.9	37.9	33.5
Total simulations	46.6	42.9	27.8
CFM-ID (10 V)	6.2		
CFM-ID (20 V)	45.5		
CFM-ID (40 V)	45.5		

To compare the CFM-ID and simulations results in detail, the final structures proposed by CFM-ID were considered with those obtained from simulations, focusing on the most abundant ion peaks for which the mechanisms were reported previously. In **Figure 7** the structures obtained with the two approaches are summarized. Product ion structures for m/z 149.0459 and 134.0461 are very similar, however in simulations the rings are not closed. This is expected since dynamics are performed in the microcanonical ensemble, and thus the system will not dissipate the excess energy needed for the rings to form[4]. Product ion structures for m/z 94.0400 and 107.0357 are more different due to both the lack of ring structures as well as the large differences in the relative positions of the nitrogen and carbon atoms. In the case of ion m/z 94.0400, some trajectories end up with a 5-membered ring which is different that proposed by CFM-ID. For ions m/z 67.0291 and 57.0452, CFM-ID did not assign intensities but did propose structures which are very different from those calculated by the simulation. Finally, the product ion structures for m/z 57.0452 were found to be very similar in both methods with a difference only in the localization of the excess proton.

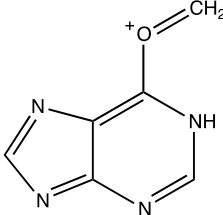
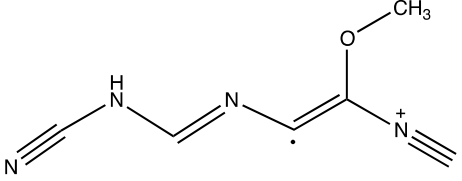
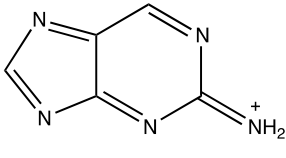
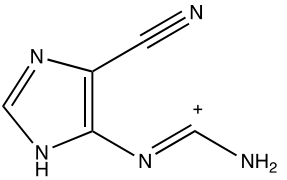
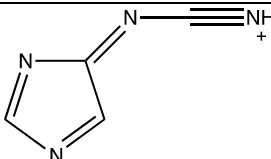
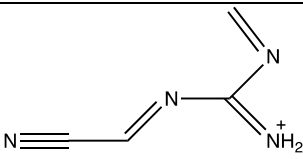
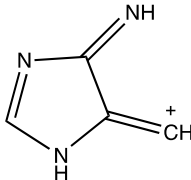
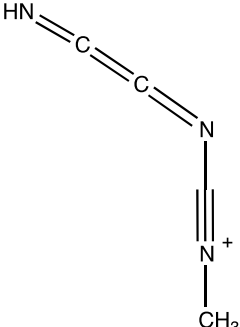
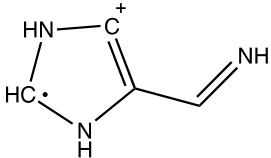
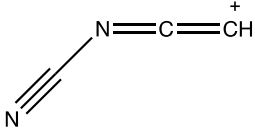
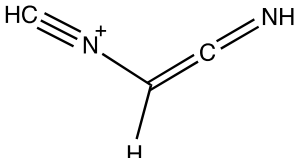
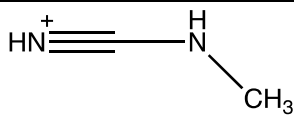
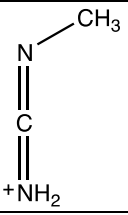
m/z	CFM-ID	CID chemical dynamics simulations
149.0459		
134.0461		
107.0357		
94.0400		
		
67.0291		
57.0452		

Figure 7. Chemical structures as obtained from CFM-ID spectra (left column) and chemical dynamics simulations (right column).

Table 3. Presence of main experimental peaks in PM6-D simulations (considering all the initial structures) and CFM-ID (at 40 V).

m/z	Experimental intensity ^a	Simulations PM6-D	Simulations AM1	CFM-ID
153.0401	m	-	-	-
151.0494	m	✓	✓	-
149.0459	h	✓	✓	✓
135.0306	m	✓	✓	-
134.0461	h	✓	✓	✓
124.0510	l	✓	✓	✓
121.0503	m	-	-	✓
110.0355	m	✓	-	-
109.0509	l	✓	✓	✓
108.0433	l	✓	✓	-
107.0357	h	✓	✓	✓
94.0400	h	✓	-	✓
93.0082	l	✓	✓	-
82.0399	m	✓	✓	✓
80.0246	l	✓	✓	✓
67.0291	h	✓	✓	-
57.0452	h	✓	✓	-
55.0291	l	✓	✓	✓

^aPeaks with intensities > 25% labeled as *h* (high), 10 - 25% labeled as *m* (medium), and 5-10% labeled as *l* (low), as obtained from experiments at 50% HCD.

The similarity function provides a good measure of the ability of both the chemical dynamic simulations and CFM-ID modeling to predict experimental spectra. Both proved to be non-quantitatively satisfactory with similarity values that are less than 50% and able to predict the appearance of the most abundant product ion masses. **Table 3** contains the most abundant peaks observed in the 50% HCD spectrum and indicates their presence (or absence) in chemical dynamics simulations (PM6-D and AM1) and CFM-ID modeling. Of note, only peaks at 153.0401 and 121.0503 m/z are missing in PM6-D simulations. AM1 simulations provide similar results, with two additional missing peaks, m/z 110.0355 and 94.0400, reflecting the smaller values of the similarity function. More peaks are absent in the CFM-ID modeling results. While qualitative (i.e. presence of the peak) agreement using simulations is often found, the modeling of intensities is more challenging and there are several factors which must be taken into account. Most importantly, short simulation time lengths due to computational practicalities limits the accuracy of intensity prediction, and also necessitates the use of relatively high activation energies. The use of semi-empirical Hamiltonians (which is computationally mandatory for the present systems) is another source of quantitative error in simulation results. Furthermore, the choice of internal energy in simulation cannot be related directly to the experimental HCD value used for fragmentation.

Summarizing, CFM-ID modeling and chemical dynamics simulations, while very different approaches, result in qualitative similar results with respect to experimental spectra. However, in contrast to probabilistic generative approaches like CFM-ID, chemical dynamics simulations provide detailed insight into the physical processes at work in the fragmentation processes as well as realistic gas phase product ion structures. This data and subsequent investigations will provide a fundamental understanding of gas phase fragmentation processes which can be used to guide both experimental fragmentation ion interpretation as well as drive future technological innovations. In addition, improvements to the chemical dynamics simulations performed here will be possible through the development of more accurate and faster theoretical methods for the calculation of energies and

forces, and/or through the use of longer simulation time-lengths. In addition, improvements in computational power will provide for more accurate *in silico* product ion spectra generation.

4. Conclusions

In this work, the fragmentation mechanisms of O⁶-Me-G have been studied using chemical dynamics simulations and compared the results with experimental mass spectra. The fragmentation simulation using internal energy activation of the ion at PM6-D level of theory provided qualitative agreement with experiments and indicates probable fragmentation mechanisms. Calculated spectra using AM1 semi-empirical Hamiltonians are in poorer agreement with experimental spectra.

Here, for the first time, chemical dynamics simulations were compared with high-resolution mass spectra, providing confident product ion structure determinations and related reaction mechanisms observed in experiments and simulations. For example, the combination of chemical dynamics simulations with high resolution experiments, allow the exact identification of the structure of ion *m/z* 57.0452.

From the observed fragmentation mechanisms responsible for the appearance of the different product ions, the mobility of the proton has been found to be a critical factor, similarly to what is well established in peptide fragmentation[45]. The agreement between chemical dynamics simulations and the experimental mass spectrum suggests the possibility of using such a theoretical approach to assist in the structural identification of uncharacterized adducts in the emerging field of DNA adductomics, for example in the determination of isomeric structures (e.g. different position of methyl group). More broadly, the prediction of major fragmentation pathways for classes of compounds for which there is little or no fragmentation data can assist in the development of screening assays for these compounds. We feel that through our future efforts and increases in the computational power available to us that the simulations needed for these goals will become increasingly more accurate and routine.

Acknowledgments

We thank ANR DynBioReact (Grant No. ANR-14-CE06-0029-01) and ASMS award for financial support. A.C. was partially supported by the 2017 American Society for Mass Spectrometry Postdoctoral Career Development Award. Mass spectrometry was carried out in the Analytical Biochemistry Shared Resource of the Masonic Cancer Center, University of Minnesota, funded in part by Cancer Center Support Grant CA-077598.

References

- ¹ Roukos, D.H.: Genome-wide association studies: how predictable is a person's cancer risk? *Expert Rev. Anticancer Ther.* 9, 389-392 (2009)
- ² Balbo, S., Hecht, S. S., Upadhyaya, P., Villalta, P. W.: Application of a High-Resolution Mass-Spectrometry-Based DNA Adductomics Approach for Identification of DNA Adducts in Complex Mixtures. *Anal. Chem.* 86, 1744-1752 (2014)
- ³ Balbo, S., Turesky, R. J., Villalta, P. W.: DNA Adductomics. *Chem. Res. Toxicol.* 27, 356-366 (2014)
- ⁴ Song, K., Spezia, R.: *Theoretical Mass Spectrometry*. De Gruyter, Berlin (2018)
- ⁵ de Sainte Claire, P., Hase, W. L.: Thresholds for the Collision-Induced Dissociation of Clusters by Rare Gas Impact. *J. Phys. Chem.* 100, 8190-8196 (1996)
- ⁶ Liu, J., Song, K., Hase, W. L., Anderson, S. L.: Direct dynamics study of energy transfer and collision-induced dissociation: Effects of impact energy, geometry, and reactant vibrational mode in H₂CO⁺-Ne collisions. *J. Chem. Phys.* 119, 3040-3050 (2003)

-
- ⁷ Martínez-Núñez, E., Fernández-Ramos, A., Vázquez, S. A., Marques, J. M. C., Xue, M., Hase, W. L.: Quasiclassical dynamics simulation of the collision-induced dissociation of $\text{Cr}(\text{CO})_6^+$ with Xe. *J. Chem. Phys.* 123 (15), 154311 (2005)
- ⁸ Meroueh, S. O., Wang, Y., Hase, W. L.: Direct Dynamics Simulations of Collision- and Surface-Induced Dissociation of N-Protonated Glycine. Shattering Fragmentation. *J. Phys. Chem. A* 106, 9983-9992 (2002)
- ⁹ Spezia, R.; Salpin, J.-Y.; Gaigeot, M.-P.; Hase, W. L.; Song, K., Protonated Urea Collision-Induced Dissociation. Comparison of Experiments and Chemical Dynamics Simulations. *J. Phys. Chem. A* 113, 13853-13862 (2009)
- ¹⁰ Spezia, R., Martens, J., Oomens, J., Song, K.: Collision-induced dissociation pathways of protonated Gly_2NH_2 and Gly_3NH_2 in the short time-scale limit by chemical dynamics and ion spectroscopy. *Int. J. Mass Spectrom.* 388, 40-52 (2015)
- ¹¹ Spezia, R., Lee, S. B., Cho, A., Song, K.: Collision-induced dissociation mechanisms of protonated penta- and octa-glycine as revealed by chemical dynamics simulations. *Int. J. Mass Spectrom.* 392, 125-138 (2015)
- ¹² Molina, E. R., Eizaguirre, A., Haldys, V., Urban, D., Doisneau, G., Bourdreux, Y., Beau, J.-M., Salpin, J.-Y., Spezia, R.: Characterization of Protonated Model Disaccharides from Tandem Mass Spectrometry and Chemical Dynamics Simulations. *ChemPhysChem* 18, 2812-2823 (2017)
- ¹³ Ortiz, D., Salpin, J.-Y., Song, K., Spezia, R.: Galactose-6-Sulfate collision induced dissociation using QM+MM chemical dynamics simulations and ESI-MS/MS experiments. *Int. J. Mass Spectrom.* 358, 25-35 (2014)
- ¹⁴ Molina, E. R., Ortiz, D., Salpin, J.-Y., Spezia, R.: Elucidating collision induced dissociation products and reaction mechanisms of protonated uracil by coupling chemical dynamics simulations with tandem mass spectrometry experiments. *J. Mass Spectrom.* 50, 1340-1351 (2015)
- ¹⁵ Rodriguez-Fernandez, R., Vazquez, S. A., Martinez-Nunez, E.: Collision-induced dissociation mechanisms of $[\text{Li}(\text{uracil})]^+$. *Phys. Chem. Chem. Phys.* 15, 7628-7637 (2013)
- ¹⁶ Spezia, R., Martin-Somer, A., Macaluso, V., Homayoon, Z., Pratihari, S., Hase, W. L.: Unimolecular dissociation of peptides: statistical vs. non-statistical fragmentation mechanisms and time scales. *Faraday Discuss.* 195, 599-618 (2016)
- ¹⁷ Homayoon, Z., Pratihari, 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 $\text{TIK}(\text{H}^+)_2$ tripeptide. Mechanisms and Kinetic Parameters. *J. Phys. Chem. A* 120, 8211-8227 (2016)
- ¹⁸ Martin-Somer, A., Martens, J., Grzetic, J., Hase, W.L., Oomens, J., Spezia, R.: Unimolecular Fragmentation of Deprotonated Diproline $[\text{Pro}_2\text{-H}]^-$ Studied by Chemical Dynamics Simulations and IRMPD Spectroscopy. *J. Phys. Chem. A* 122, 2612-2625 (2018)
- ¹⁹ Macaluso, V., Scuderi, D., Crestoni, M. E., Fornarini, S., Corinti, D., Dalloz, E., Martinez-Nunez, E., Hase, W. L., Spezia, R.: L-Cysteine Modified by S-Sulfation: Consequence on Fragmentation Processes Elucidated by Tandem Mass Spectrometry and Chemical Dynamics Simulations. *J. Phys. Chem. A* 123, 3685-3696 (2019)
- ²⁰ Park, K., Deb, B., Song, K., Hase, W. L.: Importance of Shattering Fragmentation in the Surface-Induced Dissociation of Protonated Octaglycine. *J. Am. Soc. Mass Spectrom.* 20, 939-948 (2009)
- ²¹ 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.* 131, 17185-17193 (2009)
- ²² Gregg, Z., Ijaz, W., Jannetti, S., Barnes, G. L.: The Role of Proton Transfer in Surface-Induced Dissociation. *J. Phys. Chem. C* 118, 22149-22155 (2014)
- ²³ Bauer, C. A., Grimme, S.: Elucidation of Electron Ionization Induced Fragmentations of Adenine by Semiempirical and Density Functional Molecular Dynamics. *J. Phys. Chem. A* 118, 11479-11484 (2014)
- ²⁴ Bauer, C. A., Grimme, S.: How to Compute Electron Ionization Mass Spectra from First Principles. *J. Phys. Chem. A* 120, 3755-3766 (2016)
- ²⁵ Asgeirsson, V., Bauer, C. A., Grimme, S.: Quantum chemical calculation of electron ionization mass spectra for general organic and inorganic molecules. *Chem. Sci.* 8, 4879-4895 (2017)
- ²⁶ Hecht, S. S.: DNA adduct formation from tobacco-specific N-nitrosamines. *Mut. Res.-Fund. Mol. M.* 424, 127-142 (1999)

- ²⁷ Peterson, L. A., Hecht, S. S.: O⁶-Methylguanine Is a Critical Determinant of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone Tumorigenesis in A/J Mouse Lung. *Cancer Research*, 51, 5557-5564 (1991).
- ²⁸ Becke, A. D.: A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* 98, 1372-1377 (1993)
- ²⁹ Lee, C., Yang, W., Parr, R. G.: Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37, 785-789 (1988)
- ³⁰ Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., Stewart, J. J. P.: Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* 107, 3902-3909 (1985)
- ³¹ McNamara, J. P., Hillier, I. H.: Semi-empirical molecular orbital methods including dispersion corrections for the accurate prediction of the full range of intermolecular interactions in biomolecules. *Phys. Chem. Chem. Phys.* 9, 2362-2370 (2007)
- ³² 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.* 27, 1101-1111 (2006)
- ³³ Stewart, J. J. P.: Optimization of parameters for semiempirical methods I. *Method. J. Comput. Chem.* 10, 209-220 (1989)
- ³⁴ Stewart, J. J. P.: Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* 13, 1173-1213 (2007)
- ³⁵ Grimme, S.: Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* 25, 1463-1473 (2004)
- ³⁶ Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; 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.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
- ³⁷ Stewart, J. J. P., Fiedler, L. J., Zhang, P., Zheng, J., Rossi, I., Hu, W.-P., Lynch, G. C., Liu, Y.-P., Chuang, Y.-Y., Pu, J., Li, J., Cramer, C. J., Fast, P. L., Truhlar, D. G.: MOPAC 5.022mn; Department of Chemistry and Supercomputing Institute, University of Minnesota: Minneapolis, MN, 2015.
- ³⁸ Hase, W. L., Duchovic, R. J., Hu, X., Komornicki, A., Lim, K. F., Lu, D.-H., Peslherbe, G. H., Swamy, K. N., Vande Linde, S. R., Zhu, L., Varandas, A., Wang, H., Wolf, R.: VENUS96: A general chemical dynamics computer program. *QCPE Bull.* 16, 43 (1996)
- ³⁹ Hase, W. L., Buckowski, D. G.: Monte carlo sampling of a microcanonical ensemble of classical harmonic oscillators. *Chem. Phys. Lett.* 74, 284-287 (1980)
- ⁴⁰ Schlier, C., Seiter, A.: Symplectic Integration of Classical Trajectories: A Case Study. *J. Phys. Chem. A* 102, 9399-9404 (1998)
- ⁴¹ Schlier, C., Seiter, A.: High-order symplectic integration: an assessment. *Comp. Phys. Comm.* 130, 176-189 (2000)
- ⁴² Jeanvoine, Y., Largo, A., Hase, W. L., Spezia, R.: Gas Phase Synthesis of Protonated Glycine by Chemical Dynamics Simulations. *J. Phys. Chem. A* 122, 869-877 (2018)
- ⁴³ Carpenter, J. E., Weinhold, F.: Analysis of the geometry of the hydroxymethyl radical by the different hybrids for different spins natural bond orbital procedure. *J. Mol. Struct. (Theochem)*, 139, 41-62 (1988)
- ⁴⁴ Jeanvoine, Y., Spezia, R.: The formation of urea in space. II. MP2 vs PM6 dynamics in determining bimolecular reaction products. *Theor. Chem. Acc.* 138, 1 (2019)

-
- ⁴⁵ Dongré, A. R., Jones, J. L., Somogyi, Á., Wysocki, V. H.: Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: An evidence for the mobile proton model. *J. Am. Chem. Soc.* 118, 8365–8374 (1996)
- ⁴⁶ Allen, F., Greiner, R., Wishart, D.: Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics* 11, 98-110 (2015)
- ⁴⁷ Allen, F., Pon, A., Wilson, M., Greiner, R., Wishart, D.: CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res.* 42, W94–W99 (2014)
- ⁴⁸ Kempkes, L. J. M., Martens, J., Berden, G., Houthuijs, K. J., Oomens, J.: Investigation of the position of the radical in z_3 -ions resulting from electron transfer dissociation using infrared ion spectroscopy. *Faraday Discuss.* 217, 434-452 (2019).
- ⁴⁹ Fu, W., Hopkins, W. S.: Applying Machine Learning to Vibrational Spectroscopy. *J. Phys. Chem. A*, 122, 167–171 (2018)