Computer Simulation of Collision Induced Dissociation and Isolobal Analogy: The Case of Biotin and its Analogues

Yanghune Ha[1], Riccardo Spezia[2,*] and Kihyung Song[1,*]

[1]Department of Chemistry, Korea National University of Education, Cheongju, Chungbuk, Republic of Korea
[2]Sorbonne Universite, CNRS, Laboratoire de Chimie Théorique, LCT, 4, Place Jussieu, Paris, 75252 Cedex 05, France

August 13th 2020

*Correspondence to: riccardo.spezia@sorbonne-universite.fr, ksong@knue.ac.kr

ABSTRACT

We have studied how collision induced dissociation (CID) products and associated mechanisms change when a chemical group is modified by isolobal substitution, and in particular the sequence S, O, NH and CH_2. At this end, we have considered protonated biotin (vitamin B_7) and corresponding oxybiotin, N-biotin and C-biotin, which have the same structures except for one chemical group (the S in biotin which is substituted with the aforementioned isolobal ones). Collisional simulations with Ar were performed to model CID fragmentations and to have direct access to related mechanisms. Simulations show that the CID fragmentation of the four compounds are similar and the resulting fragments involve in a similar way the isolobal groups. Details on the mechanisms obtained from simulations are reported and discussed. This result shows that it is possible to predict, in principle and with a reasonable confidence, mass spectra of unknown molecules based on mass spectrum of a known one when isolobal modifications are done.

Keywords: theoretical mass spectrometry; chemical dynamics simulations; isolobal analogy; biotin

1. INTRODUCTION

The concept of isolobal analogy was introduced years ago by Hoffmann as a practical tool to understand and easily predict organic and organometallic reactions.[1] In the original publication, discussing the properties of $M(CO)_3$ and $M(C_6H_6)$ (where M is a metal) they "mean to imply that the number, symmetry properties, extent in space and energy of the frontier orbitals of the fragments are similar – not identical, but similar".[2] Thus, two compounds, where one chemical group is replaced by another with similar properties in the frontier orbitals, are expected to behave similarly. In mass spectrometry, this concept could be used to predict fragmentation patterns of molecules which are derived from a known one where a group is replaced by an isolobal one.

Understanding and even better predicting fragmentation patterns in collision induced dissociation (CID) is crucial in both fundamental and practical applications of mass spectrometry. One approach consists in library searching, for example using the extended NIST library.[3] Of course, libraries need experiments on all the species of interest and mechanisms can be proposed only for data for which experiments are available. Another approach uses machine learning algorithms to guess fragmentation patterns, like in the competitive fragmentation model (CFM) by Allen et al. which uses a probabilistic generative model.[4]

Another possibility to predict CID spectra independently from simulation is to use molecular simulations.[5] This approach, pioneered by Hase and co-workers,[6,7] was then extended in last 10 years to a variety of biological and organic molecules.[8] The same approach can be successfully used to model and predict surface induced dissociation and soft landing.[9,10] It is based on two alternative physical representations of the collisional activation: (1) by providing an excess energy equally distributed through the vibrational modes of the fragmenting molecule, (2) by direct simulating the collision with the inert gas.[11,12] Since it is based on the generation of an ensemble of reactive trajectories, the chemical dynamics approach provides at the same time the fragmentation products and the related mechanisms. The most relevant drawback is that it requires significant computational effort, and thus for relatively large systems highly accurate quantum chemistry methods cannot be used. However, semi-empirical Hamiltonians have shown to be able to provide relatively good fragmentation patterns and mechanisms,[13,14,15,16] in

particular for organic and rigid molecules, like testosterone[17] or methyl-guanine.[18] In this last study, a comparison with CFM machine learning prediction was done, showing that the chemical trajectory approach is comparable with it, providing at the same time more information on reaction mechanisms and physically ground results.

Here, we have investigated the CID fragmentation in relation with the isolobal analogy. In particular we have considered protonated biotin (also called vitamin B_7 or H) and then substituted the S atom with O (oxybiotin), NH (N-biotin) and CH_2 (C-biotin), forming a set of compounds shown in **Figure 1**. These atoms and fragments fulfill the "isolobal analogy" prescription since their frontier orbitals have similar energy, shape and number of electrons.[1]
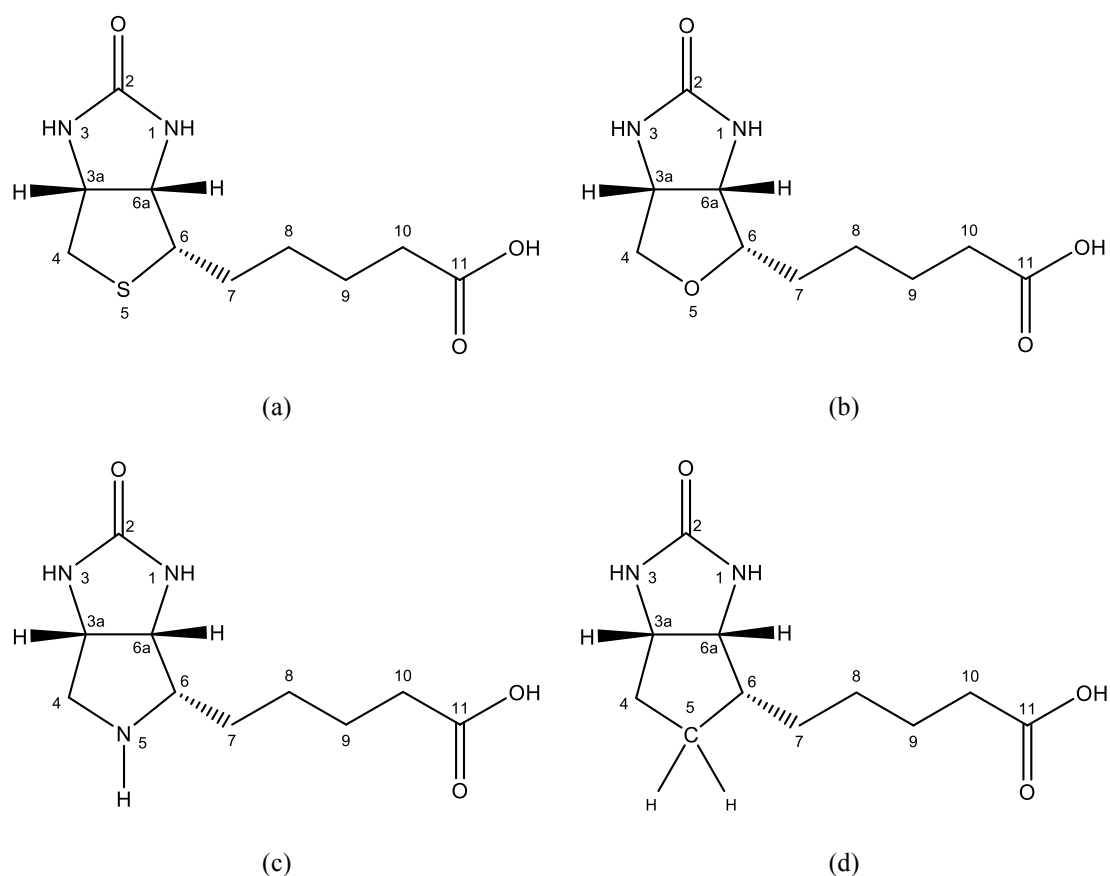
Figure 1. *Chemical structures of neutral (a) biotin, (b) oxybiotin, (c) N-biotin and (d) C-biotin.*

Other than a chemical rationale in modifying the S atom of biotin, there is also a biological interest on it. In fact, biotin acts as a coenzyme for carboxylase enzymes, involved in the synthesis of fatty acids, isoleucine, and valine, and in gluconeogenesis.[19,20]

If there is insufficient biotin in a higher organism, oxybiotin is the only analog capable of replacing it.[21] In general, single site substitution is common in biology and it can be useful to predict mass spectra of possible analogs independently from experiments. We have thus considered these four molecules, activated by modeling explicit collision with Ar atom and analyzed product ions and mechanisms. In particular, we focused on how the fragmentation spectra is (eventually) modified in the isolobal analogs and which are the related mechanisms. This will provide a guide to possibly predict the CID of different isolobal species from the spectrum of a known one.

2. COMPUTATIONAL DETAILS

2.1 Reactant structures

Biotin has an imidazolidinone structure and thiolane ring (**Figure 1a**). The isolobal analogs considered here are: oxybiotin, N-biotin and C-biotin. They have the same structure as biotin but S is substituted by O, NH, CH_2 (also shown in **Figure 1**). Although N-biotin and C-biotin are hypothetical molecules, these two compounds provide optimal examples to study isolobal analogy on CID products and mechanisms.

Protonation can occur on different basic sites: six for biotin, oxybiotin and N-biotin and five for C-biotin. We define the different protonation sites as follows: (a) protonated on oxygen in the carbonyl group of the ring; (b) protonated on the nitrogen atom of amide group located in the opposite direction of the carboxylic group; (c) protonated on the other amide group; (d) protonated on S, O or N for biotin, oxybiotin and N-biotin, respectively (this site does not exist by definition in C-biotin); (e) protonated on the carbonyl of the carboxyl group and (f) protonated on the hydroxyl group. In **Figure 2** we show the resulting different tautomers for biotin, while in **Figures S1-S3** of the supporting material we report the same for oxybiotin, N-biotin and C-biotin.
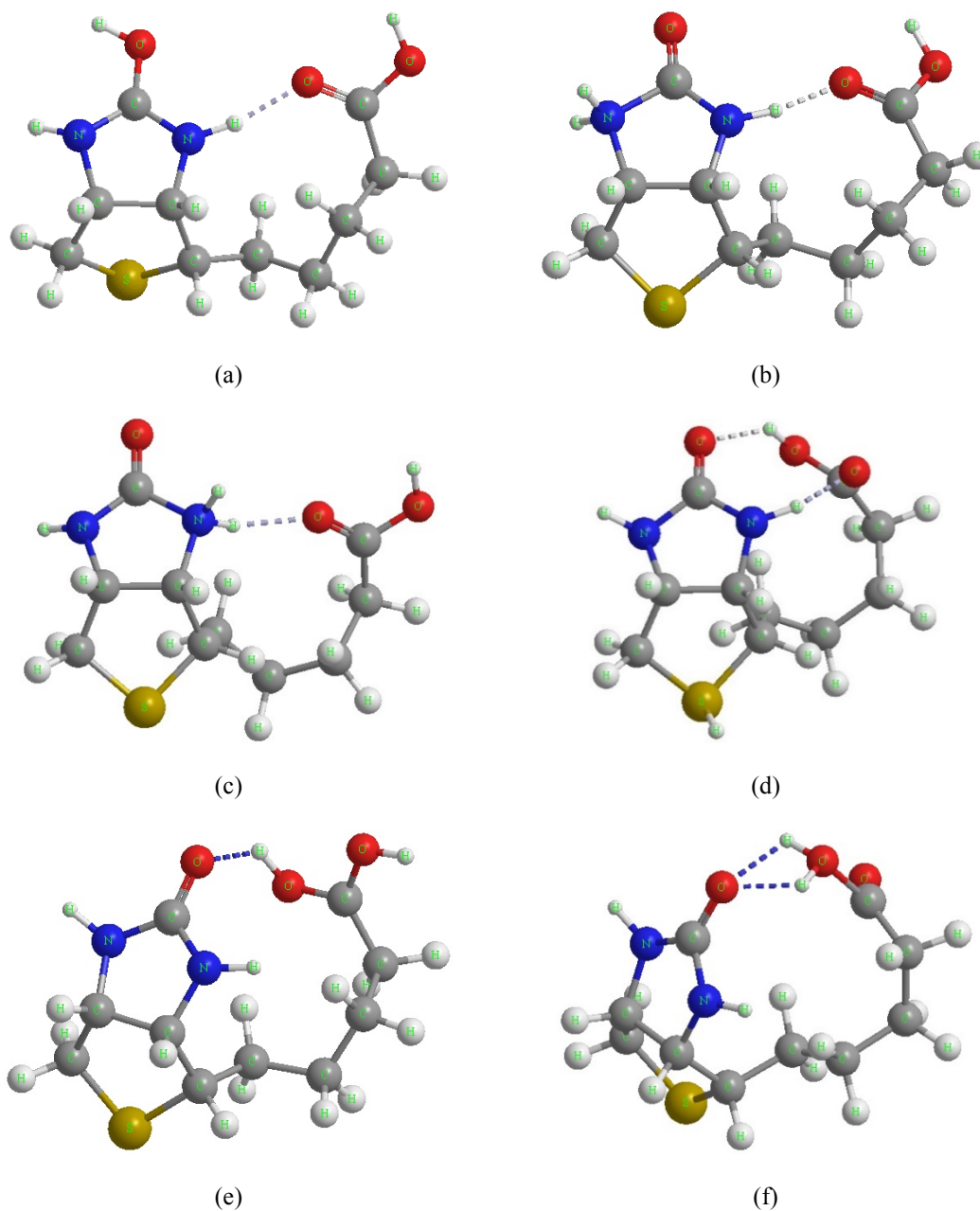
Figure 2. Structures of the different tautomers of protonated biotin.

The geometries of the different structures were optimized at Hartre-Fock (HF) and B3LYP[22,23] levels of theory with the 6-31G(d,p) and 6-311++G(d,p) basis sets and with a number of semi-empirical Hamiltonians, namely: MNDO-d,[24,25] RM1,[26] PM3,[27] PM6[28] and PM7.[29] The resulting optimized geometries of each tautomeric form are reported in **Figures 2** and **S1-S3**. From these calculations, it was possible to obtain the proton affinities of all the structures considered at different levels of theory.

HF and B3LYP calculations were performed using Gaussin 09 Rev D.01 software,[30] while semi-empirical Hamiltonians calculations with MOPAC 2016.[31]

2.2 Ion-Ar interaction potential

The potential energy for the collisional system, consisting of the ion and the neutral Ar atom (the collision gas) is described by:

$$V = V_{ion} + V_{Ar-ion} \qquad (1)$$

where V_{ion} is the intramolecular potential of the protonated molecule and V_{Ar-ion} is the intermolecular potential between Ar and the ion. This last term is as a sum of two-body interactions between Ar and the atoms of the ion. We used the expression proposed by Meroueh and Hase :[32]

$$V_{Ar-ion} = \sum_i a_{Ar-i} \exp(-b_{Ar-i} \times r_{Ar-i}) + \frac{c_{Ar-i}}{r_{Ar-i}^9} \qquad (2)$$

where i runs over all atoms of the ion and the parameters a_{Ar-i}, b_{Ar-i} and c_{Ar-i} are always positive as reported in the original publication and found in next studies. This interaction potential was able to correctly simulate the CID of a series of systems, like protonated urea, [33] protonated *N*-formylalanylamide[13] and testosterone.[17] We used parameters of H, C, O, N as from the original works of Meroueh and Hase[32] and Ortiz et al.[34] Parameters for S were obtained previously for sulfate group,[34] but the electron density of sulfur in biotin is clearly different and thus we have re-parametrized the potential for this new atom type. We have used at this aim the interaction between Ar and thiolane molecule to obtain parameters for S atom. The thiolane molecule containing thioether group (R-S-R') is a good model to represent the short-range repulsive interaction between Ar and S of biotin. The interaction potential energy between Ar and S was calculated from the shortest possible distance up to 3.5 Å, and the parameters a, b and c of Eq. (2) were obtained by fitting *ab initio* interaction energy curves carried out on the QCISD(T)/6-31++G(d,p) level of theory, which is the same level of theory used in previous parametrizations. The fit was performed by means of the least squares method as in Kaleidagraph software.[35]

2.3 Trajectory simulations

Direct dynamics simulations of the ion-Ar collision were performed for the most stable precursor tautomer of each molecule (biotin, oxybiotin, N- and C-biotin). As we will show in the results section, this will correspond to the tautomer (a) where the excess proton is on the oxygen of the carbonyl group of the imidazolidinone group.

Initial conditions for the ions were obtained from most stable structures, and initial positions and momenta obtained from vibrational normal mode Boltzmann sampling at 300 K.[36] Rotational energy was also added corresponding to the same temperature. The collisional system was set as follows: (1) the ion was randomly rotated about its Euler angles; (2) the Ar was placed at 13 Å distance (ensuring no ion-Ar interaction at the beginning); (3) an initial relative velocity was impressed to Ar and ion, with velocity vectors connecting the two centers of mass, corresponding to a relative energy of 30 eV in the center of mass framework; (4) the Ar atom was shifted by a randomly obtained impact parameter in the 0-5.0 Å range. This maximum value was estimated by the maximum length of the protonated systems which was about 9.2 Å. These collisional simulation initial conditions were set as described in details by Hase et al.[37]

Simulations were then run and stopped when the Ar-ion distance exceeds 300 Å and in any case with a maximum simulation time of 40 ps. Equations of motions were integrated using velocity Verlet algorithm with a time step of 0.2 fs. For each system, we performed between 10,000 and 20,000 trajectories in order to have enough reactive events for the given collision energy and simulation time. Namely, we have between 288 (for C-biotin) and 1003 (for biotin) reactive trajectories, as reported in **Table S1** of the supporting information. The relative energy of 30 eV ensures also enough fragmentation to a have a statistical consistent number of products and it is what successfully used recently to model protonated testosterone fragmentation.[17]

3. RESULTS AND DISCUSSIONS

3.1 Ar-S interaction potential parameters

The interaction potential between Ar and S bound to two carbon atoms (C-S-C) in thiolane was obtained by performing QCISD(T)/6-31++G(d,p) calculations along the Ar…S C_{2v} axis of thiolane. The resulting *ab initio* energy curve is plotted in **Figure 3**

together with the result of fitted Equation 2. Corresponding parameters for Ar-S interaction are reported in **Table 1**.
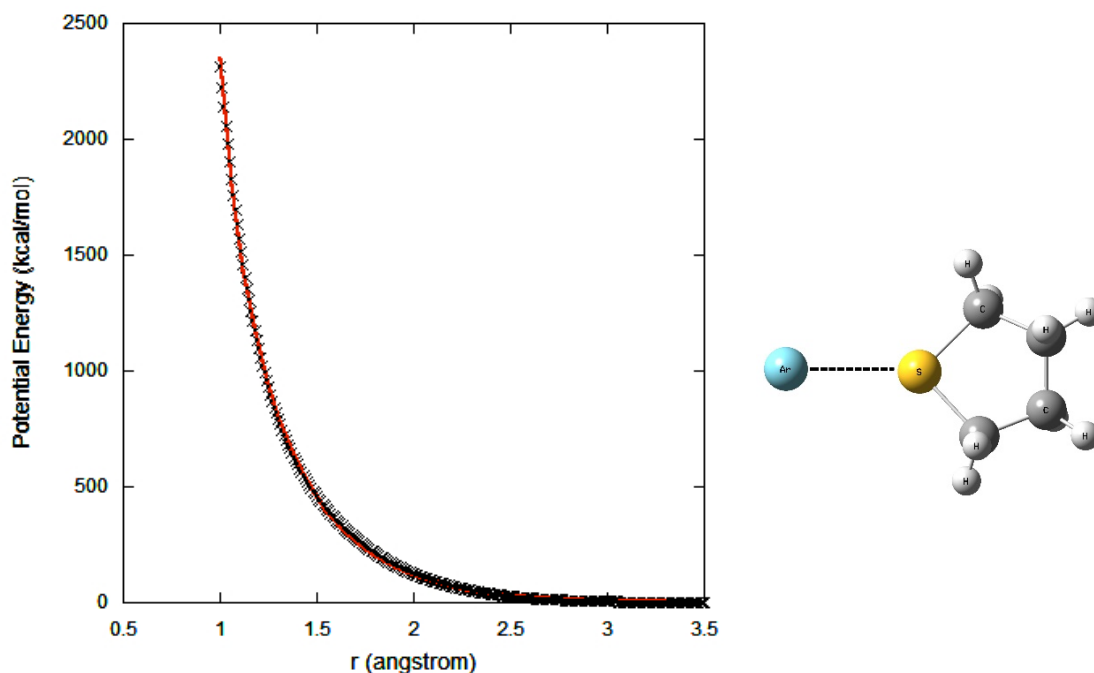
Figure 3. Potential energy for Ar/S(CH2)4 as obtained from QCISD(T)/6-31++G(d,p) calculations (crosses) and fitted Equation 2.

Table 1. Intermolecular potential parameters of sulfur atom

potential	a [1]	b [2]	c [3]
Ar/S(CH$_2$)$_4$(ArS)	21634.0	2.5991	739.11

[1] Units are in kcal/mol. [2] Units are in Å^{-1}. [3] Units are in Å^9 kcal/mol.

3.2 Proton affinities

Proton affinities are calculated for each stable tautomer of biotin and isolobal analogs. In this way, it is possible to determine the most stable pronation state which will be used as starting structure for the following chemical dynamics simulations. Results are reported in **Tables 2-5**. In all systems, the (a) protonation site (on the carboxyl oxygen of imidazolidinone ring) is the most favorable one. Values are very similar for the different structures, which is reasonable since the protonation site is relatively far from the

modified group. Interestingly, for N-biotin the proton affinity of N-analog atom – site (d) – is much closer to site (a) than in biotin and oxybiotin, such that it becomes slightly higher when calculated at B3LYP/6-311++G(d,p), RM1 or PM7 levels of theory.

Table 2. Proton affinities for biotin [a]

	ion type (a)	ion type (b)	ion type (c)	ion type (d)	ion type (e)	ion type (f)
HF / 6-31G(d,p)	<u>237.827</u>	217.820	223.635	208.919	230.635	224.321
HF / 6-311++G(d,p)	<u>235.860</u>	217.365	223.029	209.893	212.692	222.274
B3LYP / 6-31G(d,p)	<u>233.257</u>	220.055	224.987	206.991	229.003	222.569
B3LYP / 6-311++G(d,p)	<u>229.203</u>	216.975	221.778	206.434	223.698	218.001
MNDO-d	<u>168.318</u>	145.957	148.263	161.202	157.071	137.742
RM1	<u>182.809</u>	177.649	181.017	158.401	181.173	176.824
PM3	<u>202.598</u>	191.167	190.578	182.970	182.814	172.221
PM6	155.682	156.357	<u>159.775</u>	149.318	155.038	150.902
PM7	<u>152.531</u>	149.646	151.436	142.566	151.382	149.319

[a] Units are in kcal/mol.

Table 3. Proton affinities for oxybiotin [a]

	ion type (a)	ion type (b)	ion type (c)	ion type (d)	ion type (e)	ion type (f)
HF / 6-31G(d,p)	<u>240.516</u>	218.196	224.224	214.428	231.060	228.388
HF / 6-311++G(d,p)	<u>238.161</u>	216.888	193.723	213.397	213.943	222.894
B3LYP / 6-31G(d,p)	<u>235.175</u>	226.335	225.271	204.091	229.577	223.728
B3LYP / 6-311++G(d,p)	<u>230.962</u>	216.836	221.573	185.463	227.396	218.707
MNDO-d	<u>167.439</u>	145.622	147.426	139.601	155.568	136.016

RM1	<u>183.149</u>	177.279	180.386	147.046	180.745	175.970
PM3	<u>204.633</u>	191.597	192.426	154.288	183.683	172.915
PM6	153.074	154.990	<u>158.907</u>	137.826	154.223	149.126
PM7	<u>151.894</u>	149.218	151.678	124.786	151.427	148.708

[a] Units are in kcal/mol.

Table 4. Proton affinities for N-biotin [a]

	ion type (a)	ion type (b)	ion type (c)	ion type (d)	ion type (e)	ion type (f)
HF / 6-31G(d,p)	<u>240.469</u>	224.928	227.665	240.371	234.100	229.235
HF / 6-311++G(d,p)	<u>238.829</u>	210.509	226.882	238.758	216.545	227.104
B3LYP / 6-31G(d,p)	<u>235.884</u>	227.754	228.457	235.132	232.183	226.714
B3LYP / 6-311++G(d,p)	231.122	224.040	224.890	<u>232.102</u>	226.507	221.886
MNDO-d	<u>168.188</u>	146.629	147.445	160.405	155.386	136.159
RM1	186.119	182.625	183.743	<u>190.428</u>	183.195	179.376
PM3	<u>205.433</u>	194.979	194.711	192.415	184.334	174.060
PM6	156.195	160.384	161.702	<u>166.231</u>	156.058	152.228
PM7	153.257	152.973	152.803	<u>155.431</u>	152.232	150.253

[a] Units are in kcal/mol.

Table 5. Proton affinities for C-biotin [a]

	ion type (a)	ion type (b)	ion type (c)	ion type (d)	ion type (e)	ion type (f)
HF / 6-31G(d,p)	<u>242.297</u>	227.125	229.229	-	235.556	230.259
HF / 6-311++G(d,p)	<u>240.450</u>	226.376	228.566	-	233.123	228.259
B3LYP / 6-31G(d,p)	<u>237.473</u>	229.051	229.942	-	233.574	227.890

B3LYP / 6-311++G(d,p)	<u>233.422</u>	225.434	226.586	-	228.130	223.305
MNDO-d	<u>171.302</u>	150.931	152.705	-	157.584	138.550
RM1	<u>188.189</u>	185.302	186.140	-	184.884	181.569
PM3	<u>207.837</u>	197.899	197.084	-	185.424	175.551
PM6	160.042	163.928	<u>164.976</u>	-	159.247	155.918
PM7	<u>156.033</u>	155.932	155.617	-	154.596	153.030

[a] Units are in kcal/mol.

The semi-empirical methods globally underestimate the proton affinity, but they show similar trend as from HF and B3LYP calculations, except for PM6 which provides ion type (c) as the most stable in contrast with HF and B3LYP results. Given such poor performance of PM6 in proton affinity, we did not add any empirical dispersion, which is however implicit in PM7 parametrization. Between different methods tested, PM3 seems to better reproduce these values. We have thus used it in the following chemical dynamics simulations. Note that the same semi-empirical Hamiltonian was successfully used to study CID of other molecules, like notably galactose sulfate.[34]

3.3 Fragmentation yields and products

Reactivity yield is about 10% for biotin, but then it decreases when moving to oxybiotin (7.8%), N-biotin (4.2%) and C-biotin (2.4%), which roughly reflects the increasing in C-X bond strength from S to C (where X = S, O, N and C). Details are reported in **Table S1** of the supporting material.
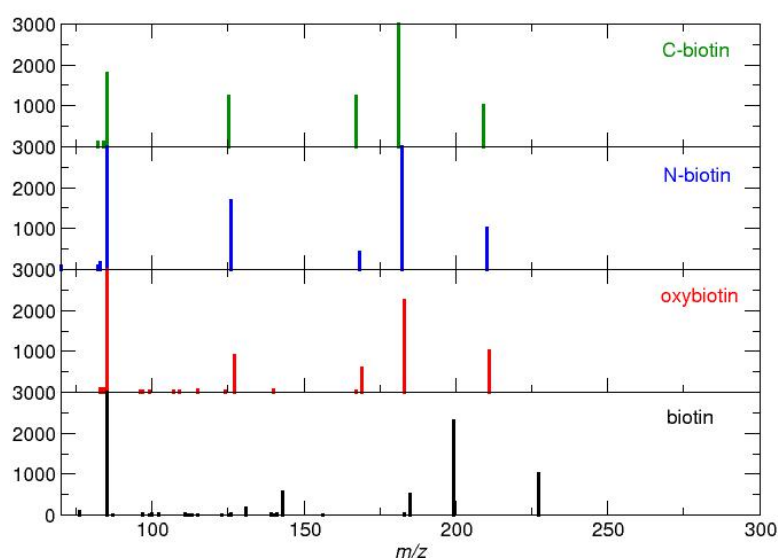
Fragmentation products of biotin are compared with what reported from experiments[38] in the MassBank database.[39] The two fragmentation spectra are shown in **Figure 4**. From simulations, the theoretical MS/MS spectrum is obtained simply by counting the product ions at the end of each trajectory and normalizing the intensities with respect to the ion *m/z* 227.08 which is the same normalization reported in the experiments.
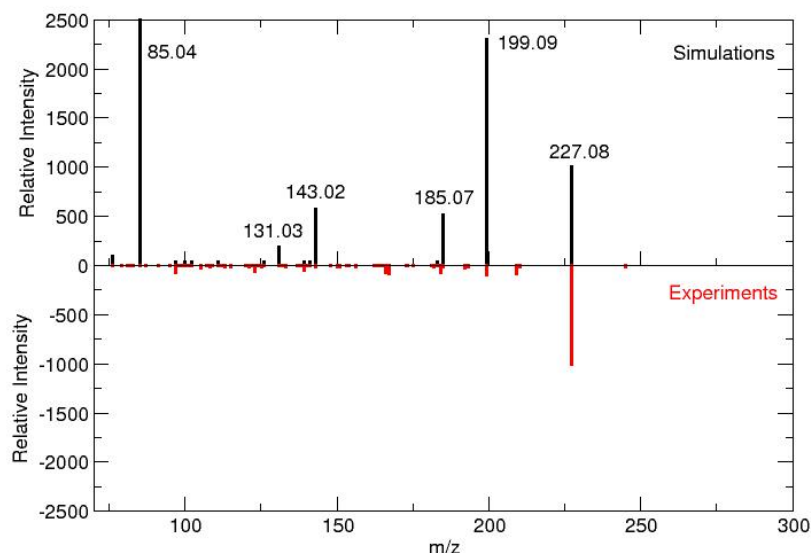
Figure 4. *CID spectra of biotin as obtained from simulations (upper panel, in black) and reported experimentally (lower panel, in red; the sign is inversed). Both spectra are normalized with respect to the first peak, m/z 227.08. Precursor ion is not reported.*

Globally, the match between experiments and simulations is quite good: the main experimental peaks are found also in simulations with some differences in relative intensity. In particular, experimentally, the most intense peak is *m/z* 227.08 which is found less intense in simulations. Simulations report ion *m/z* 199.09 more abundant, while it is found less in experiments. A possible explanation can be that in the experiments once that ion *m/z* 199.09 is formed it can further react. We did not have any direct evidence from trajectories, since our simulation time-length was not enough to observe a secondary fragmentation. We will discuss some possible secondary products in next section. The largest difference concerns peak *m/z* 85.04 which is very abundant in simulations while very small in experiments. This is one of the lightest peak observed and thus the difference in abundance can be due to several reasons. One possibility is that the energy in simulations is higher than in experiments (a direct comparison on energetics cannot be done since experiments were not done in a triple-quad or QQTof instrument), thus smallest fragments are more probable. Furthermore, we do not have information on how the detection range is set in experiments, which can impact the observed number of ions

in the low mass range of the instrument if it was optimized for higher *m/z* peaks. In any case, this low mass peak is not a characteristic peak. Details on most important mechanisms will be reported in section 3.4, here we mostly remark that the characteristic peak is obtained (*m/z* 227.08) and the global pattern is in reasonable agreement with experiments. Note that collisional conditions are also different, so we do not expect a perfect match, in particular concerning intensities.

Figure 5. Comparison of fragmentation products obtained from CID simulations of biotin, oxybiotin, N-biotin and C-biotin. In y-axis relative intensity, where the normalization was set on the first fragmentation peak of each species. Precursor ions are not reported.

In **Figure 5** we show the theoretical MS/MS spectrum obtained as results of all collisional simulations. Details and corresponding formula are listed in **Table S2** where in each raw we report *m/z*, formula and abundance of the ions of the different isolobal anolog fragmentation products. Globally, the fragmentation patterns are very similar, in particular concerning the characteristic peaks of biotin. The intensities show some differences. In particular, moving from biotin to C-biotin, the relative intensity of ion *m/z* 199.09 (which is *m/z* 183.11, 182.13 and 181.13 in oxybiotin, N- and C-biotin, respectively) becomes more intense, and also other peaks that are less intense in biotin become more relevant. As we have discussed previously, simulations of biotin report the

low *m/z* ion 85.04 with high intensity. This same *m/z* ion was found in other isolobal analogs, because the modified groups are in the neutral fragments. Details on the mechanisms observed in simulations are reported in next section.

3.4 Fragmentation mechanisms

We now analyze in details reaction mechanisms for the formation of the most relevant reaction products. We discuss the mechanisms observed in biotin together with the equivalent ones for the other systems, leading to isolobal analog fragments.

$C_{10}H_{15}N_2O_2S^+$ *(m/z=227.08) and isolobal analog ions.* Fragment *m/z* 227.08 is the most abundant peak in experiments and, as discussed previously, obtained also in simulations, but with less relative abundance. Analog fragments *m/z* 211.11, 210.12 and 209.13 are obtained in simulations of oxybiotin, N-biotin and C-biotin, respectively. The product ion corresponds to water loss with three possible mechanisms, shown in **Scheme 1**, where for simplicity we report only what was observed in biotin, the others being similar given the isolobal substitution. The water loss can be obtained from the two oxygen atoms, from carbonyl in the imidazolidinone ring or from the carboxylic tail. In the first case, the mechanism is rather complex, while for the latter the OH can pick a hydrogen from either the NH group of imidazolidinone or a vicinal CH. The last mechanism is the most abundant (about 67%) while the first two have lower abundance (12 and 21%, respectively). We should note that the product ion *m/z* 209.07, observed experimentally in biotin but not in simulations, corresponds to the subsequent water loss from the product ion *m/z* 227.08. Another ion, *m/z* 192.05 observed experimentally, but not in simulations, is likely coming from further ammonia loss from *m/z* 209.07. Collisional simulations are limited in time-length and this is clearly at the origin of the missing of these secondary and tertiary product ions in simulations.
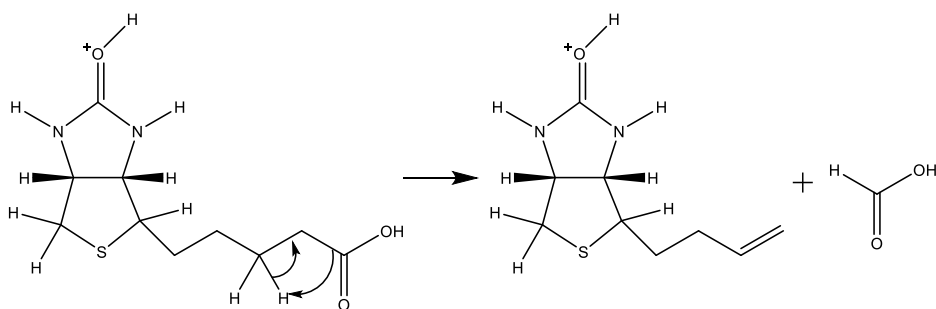
$C_9H_{15}N_2OS^+$ *(m/z=199.09) and isolobal analog ions.* The peak *m/z* 199.09 corresponds to HCOOH lost and is observed in both simulations and experiments as an intense peak. The mechanisms are the same for the different species, leading to ions *m/z* 183.11, 182.13 and 181.13 for oxybiotin, N-biotin and C-biotin, respectively. For simplicity, we report in **Scheme 2** only the mechanism observed in biotin simulations. As discussed in section 3.2, this ion was much more abundant in simulations than in experiments. This can be

due to secondary fragmentations, which unfortunately we did not observe from trajectories which are too short to allow such further reactions. Based on the structure of the product ion we can suggest a H_2O loss providing ion *m/z* 181.09 which is observed in experiments but not in trajectories.
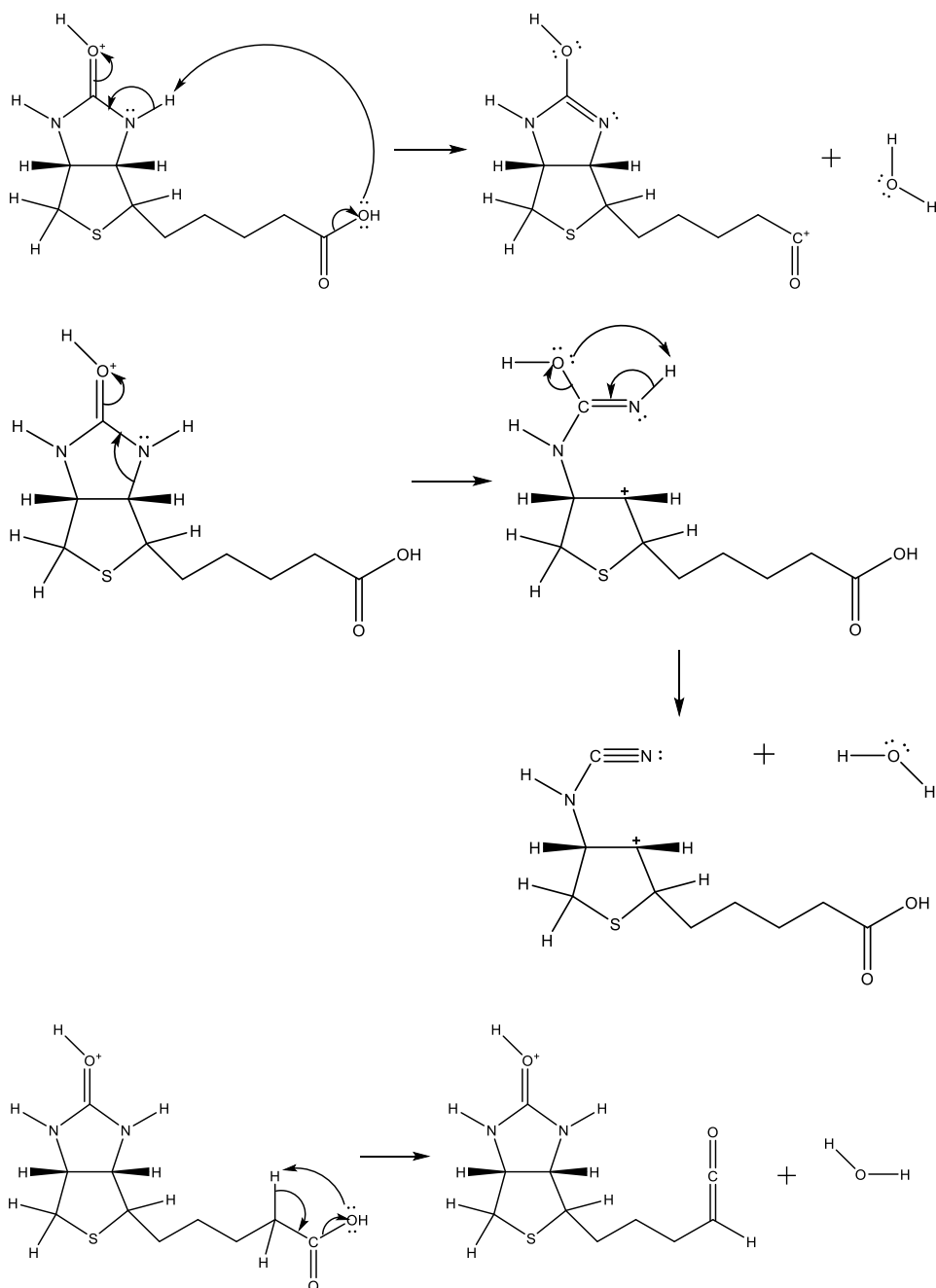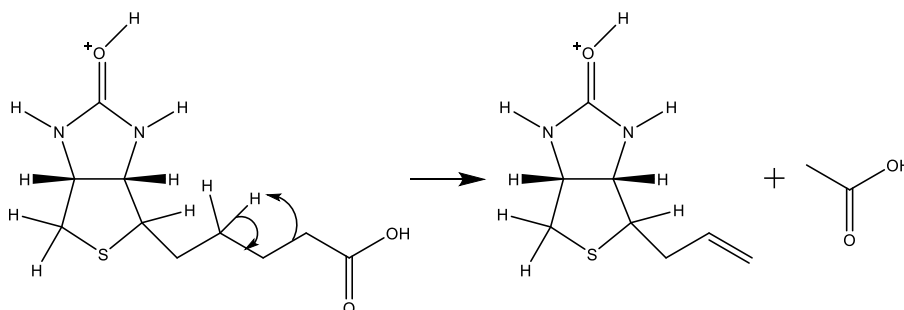

Scheme 1. *Mechanisms observed for the formation of m/z 277.08 as from CID simulations of biotin.*

Scheme 2. *Mechanism observed for the formation of m/z 199.09 as from CID simulations of biotin.*

$C_8H_{13}N_2OS^+$ (m/z=185.07) and isolobal analog ions. Ion *m/z* 185.07 observed in biotin corresponds to CH_3COOH neutral loss and the same mechanism is reported for the other systems. As previously, the mechanism does not involve the S, O, NH or CH_2 isolobal modified groups and thus they are the same (in **Scheme 3** we report for simplicity only the case of biotin).
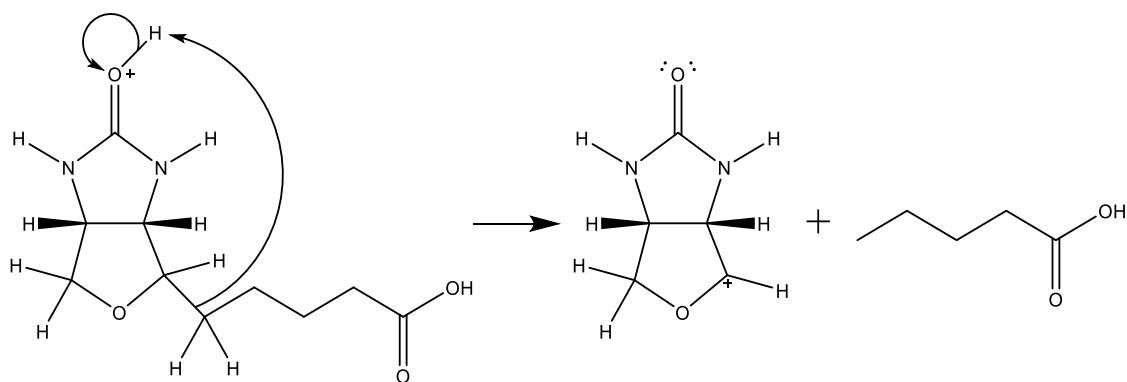

Scheme 3. *Mechanism observed for the formation of m/z 185.07 as from CID simulations of biotin.*

$C_5H_7N_2OS^+$ (m/z 143.02) and isolobal analog ions. The mechanism responsible for the formation of these ions involve the isolobal moieties. They have in common the cleavage of C-C bond, but then a hydrogen atom is taken from a different site. The mechanisms observed in the case of biotin are reported in **Scheme 4**: hydrogen is released from one C atom of the broken C-C bond, from NH in the imidazolidinone ring or from the CH group in common between the two rings. Similar mechanisms are obtained for the oxybiotin (product ion $C_5H_7N_2O_2^+$, *m/z* 127.05), with the exception of this last mechanism which is not observed, while a new one is. This corresponds to a proton transfer from the protonated carbonyl and it is reported in **Scheme 5**. For both biotin and oxybiotin the

different mechanisms are observed with about the same occurrence.
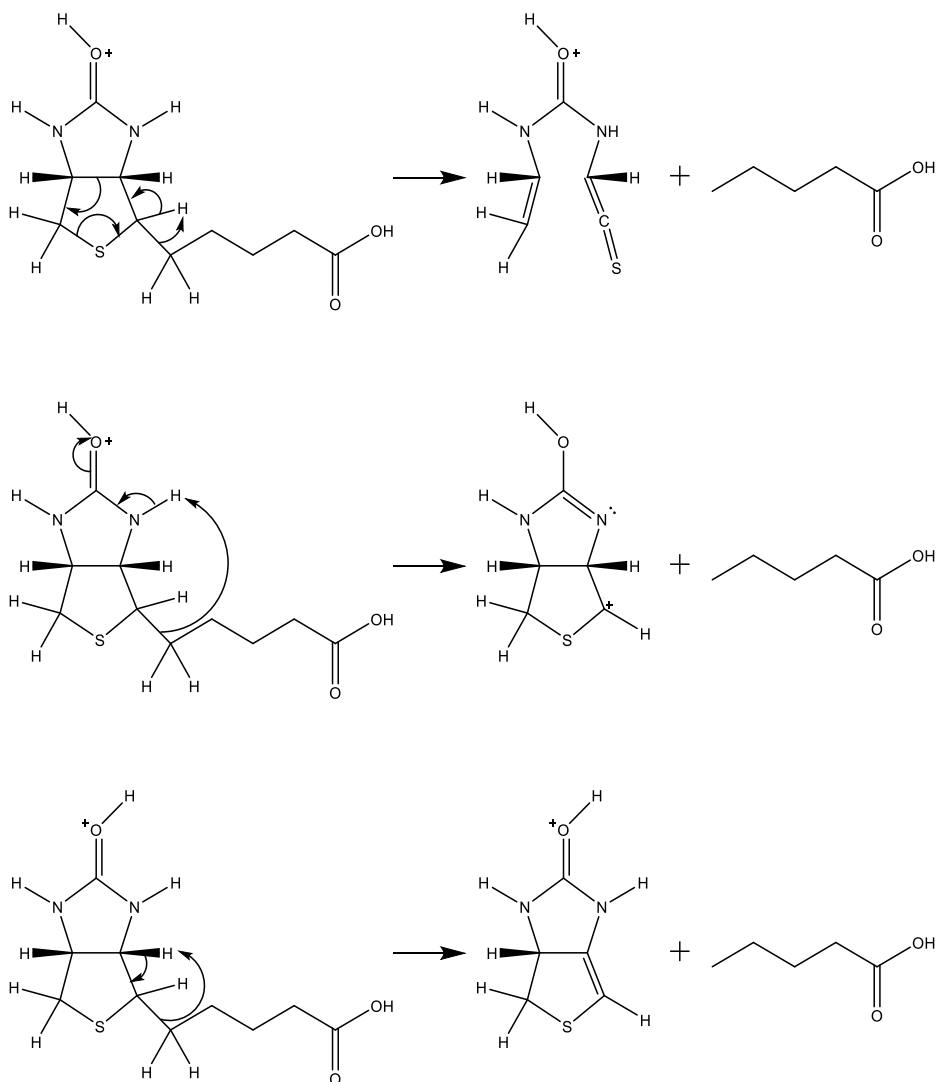
Scheme 4. *Mechanism observed for the formation of m/z 143.02 as from CID simulations of biotin.*

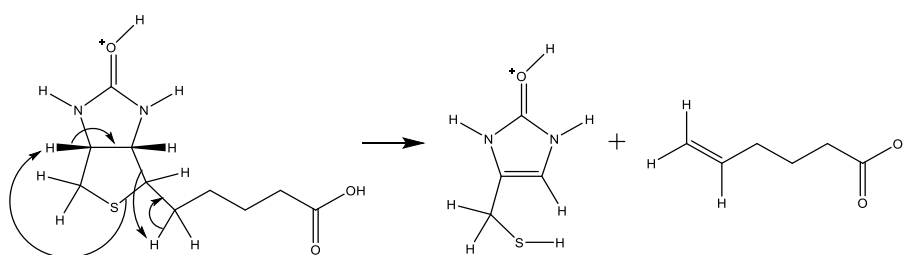**Scheme 5.** *Mechanism observed for the formation of m/z 127.05 specific of oxybiotin as from CID simulations.*

In the case of N-biotin the corresponding product ion $C_5H_8N_3O^+$(m/z 126.07) is formed with a different mechanism, where the hydrogen atom on the NH isolobal group is abstracted, as reported in **Scheme 6**. Other two mechanisms in common with biotin and oxybiotin (abstraction of H from the breaking C-C bond and from the other NH group) are also observed. For C-biotin the product ion $C_6H_9N_2O^+$ (m/z 125.07) is also formed with a mechanism in which the hydrogen is taken from the isolobal modified group (see **Scheme 6**). These new mechanisms typical of N- and C-biotin cannot be observed by definition in biotin and oxybiotin, but, since the proton can be abstracted also from other sites, finally we do not observe any difference in the product ions.
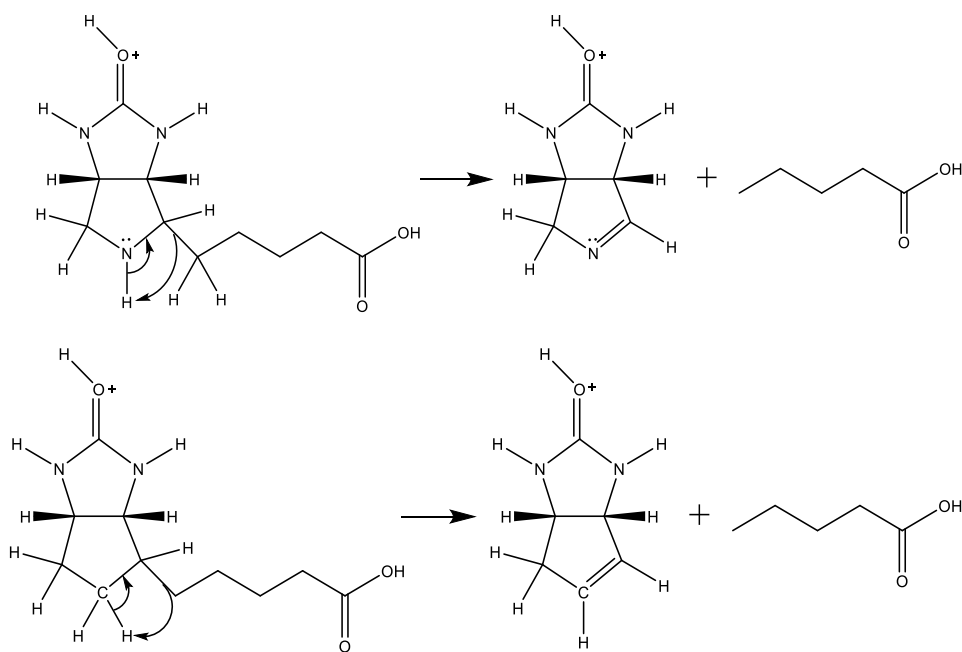
Scheme 6. *Mechanism observed for the formation of m/z 126.07 and 125.07 specific*

of N-biotin (upper) and C-biotin (lower), respectively, as from CID simulations.

$C_4H_7N_2OS^+$ *(m/z 131.03) and isolobal analog ions*. The formation of *m/z* 131.03 in biotin involves the breaking of S-C bond, as reported in **Scheme 7**. The analog mechanism is found also in oxybiotin (ion $C_4H_7N_2O_2^+$, *m/z* 115.05), but not in N-biotin and C-biotin. The final product for biotin and oxybiotin contains SH and OH group, respectively. Together with the difference in C-X bond (X = S, O, N, C) this can be at the origin of the absence of this product ion in simulations of N- and C-biotin.
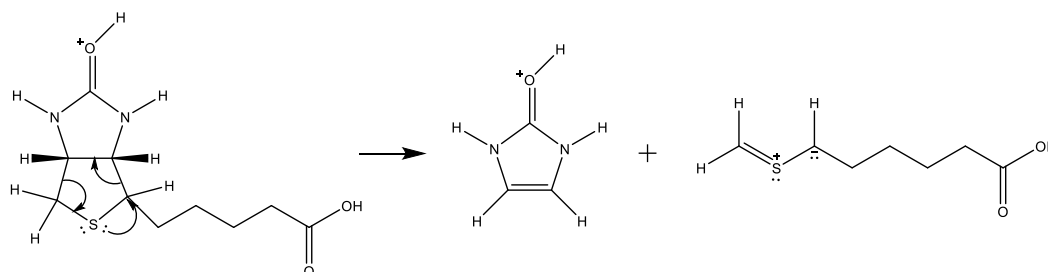

Scheme 7. *Mechanism observed for the formation of m/z 131.03 as from CID simulations of biotin.*

Differences in the reactivity of isolobal analogs for this and other products are minor since they concern ions for which we obtained only few reactive trajectories, so any conclusion is not statistically ground. They are reported for completeness in the supporting information. Moreover, they are in a region of the spectrum which shows experimentally very low signals.

$C_3H_5N_2O^+$ *(m/z 85.04) product ion*. This product was observed in all simulations and with a very high intensity. It is the most abundant product in biotin, oxybiotin and N-biotin, while it is not reported in experiments. The mechanism is reported in **Scheme 8**. In some biotin and oxybiotin simulations, the neutral further breaks in two molecules ($CH_2X + C_6H_{10}O_2$, where X = S or O). The product ions are the same in all the systems, since the isolobal analog groups are in the neutral molecule after fragmentation. The abundance of this ion in simulations and discrepancy with respect to experiment can have different origins. We have mentioned two possible reasons in section 3.2: (i) simulations are performed at relatively high energy and the direct collision of the Ar on the linear

chain can enhance this pathway; (ii) the experimental detector can be optimized for ions of higher *m/z* values and the low side of the experimental range can be underestimated. These reasons are, based on our experience, the most probable source of discrepancy but we should mention also two other possible reasons: (iii) the product ion can further react and we do not observe these secondary reactions due to limitations in simulation time-length; (iv) an artifact due to the semi-empirical Hamiltonian.

Scheme 8. *Mechanism observed for the formation of m/z 85.04 as from CID simulations of biotin.*

4. CONCLUSIONS

In this work, we have studied how protonated biotin and isolobal analogs oxybiotin, C-biotin and N-biotin fragments in CID simulations. Experiments are available only for biotin and simulations have shown a good agreement, being able to elucidate the different fragmentation mechanisms. The isolobal analogs seem to react in a similar way, in particular for the characteristic peaks. Even when the mechanisms involve the isolobal analog part, the final fragmentation products are often the same, because of the variety of possible mechanisms. The few differences in reactivity which induce a difference in fragmentation ion products (and thus in the MS/MS signal) concern reaction pathways with low abundance and which are not fingerprint peaks. Unfortunately, there are no experiments on isolobal analogs, and likely this work will stimulate new experimental and theoretical studies: the former to verify how MS/MS spectra are modified by isolobal substitution, the latter to determine with advanced electronic structure calculations the potential energy surfaces starting from what suggested by trajectory simulations.

Concluding, this work shows that it is possible to use the isolobal analogy to guess with a certain confidence the presence many fragmentation products, but for this the detailed mechanisms should be known for at least one system. This result paves the way of using

in the future chemical dynamics simulations on one species and use simple mechanistic considerations to propose fragments of isolobal analogs. In the future, it will be thus possible to predict mass spectrum of unknown molecules based on mass spectrum of the known ones with the help of computer simulations also with the aim of building a theoretical MS/MS database which can support experiments and complement existent databases.

ACKNOWLEDGMENTS

REFERENCES

[1] R.Hoffmann. Building bridges between inorganic and organic chemistry (Nobel Lecture). *Angew. Chem. Int. Ed.* 21 (1982) 711-724.

[2] M.Elian, M.M.L. Chen, D.M.P.Mingos, R. Hoffmann. Comparative Bonding Study of Conical Fragments. Inorg. Chem. 15 (1976) 1148-1155.

[3] https://chemdata.nist.gov/

[4] F.Allen, R.Greiner, D.Wishart. Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. Metabolomics. 11 (2015) 98-110.

[5] K.Song, R.Spezia. Theoretical Mass Spectrometry: Tracing Ions with Classical Trajectories. Walter de Gruyter GmbH & Co , Berlin 2018.

[6] S.O.Meroueh, Y.Wang, W.L.Hase. Direct Dynamics Simulations of Collision- and Surface-Induced Dissociation of N-Protonated Glycine. Shattering Fragmentation. J. Phys. Chem. A 106 (2002) 9983-9992.

[7] J.Liu, K.Song, W.L.Hase, S.L.Anderson. Direct Dynamics Study of Energy Transfer and Collision-Induced Dissociation: Effect of Impact Energy, Geometry, and Reactant Vibrational Modes in H_2CO^+-Ne Collisions. J. Chem. Phys. 119 (2003) 3040-3050.

[8] A.Martin-Somer, V.Macaluso, G.L.Barnes, L.Yang, S.Pratihar, K.Song, W.L.Hase, R.Spezia. 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. 31 (2020) 2-24.

[9] S.Pratihar, G.L.Barnes, W.L.Hase. 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. 45 (2016) 3595-3608.

[10] S.Pratihar, G.L.Barnes, J.Laskin, W.L.Hase. Perspective: Dynamics of Protonated Peptide Ion Collisions with Organic Surfaces. Consonance of Simulation and Experiment. J. Phys. Chem. Lett. 7 (2016) 3142-3150.

[11] R.Spezia, A.Martin-Somer, V.Macaluso, Z.Homayoon, S.Pratihar, W.L.Hase. Unimolecular Dissociation of Peptides: Statistical vs Non-Statistical Fragmentation Mechanisms and Time Scales. Faraday Discuss. 195 (2016) 599-618.

[12] A.Martin-Somer, J.Martens, J.Grzetic, W.L.Hase, J.Oomens, R.Spezia. Unimolecular

Fragmentation of Deprotonated Diproline [Pro$_2$-H]$^-$ Studied by Chemical Dynamics Simulations and IRMPD Spectroscopy. J. Phys. Chem. A 122 (2018) 2612-2625.

[13] D.Ortiz, P.Martin-Gago, A.Riera, K.Song, J.-Y.Salpin, R.Spezia. Gas-Phase Collision Induced Dissociation Mechanisms of Peptides. Theoretical and Experimental Study of N-Formylalanylamide Fragmentation. Int. J. Mass Spectrom. 335 (2013) 33-44.

[14] H.Bednarski, K.Sohlberg, M.Domanski, J.Weszka, G.Adamus, M.Kowalczuk, V.Coran. A combined theoretical and experimental study of mechanisms of fragmentation active for PHB oligomers in negative-ion mode multistage mass spectrometry. Int. J. Mass Spectrom. 304 (2011) 15-24.

[15] K.Lucas, G.L.Barnes. Modeling the Effects of O-sulfonation on the CID of Serine. J. Am. Soc Mass Spectrom. 31 (2020) 1114-1122 (2020)

[16] R.Rodríguez-Fernández, S.A.Vázquez, E.Martínez-Núñez. Collision-induced dissociation mechanisms of [Li (uracil)]$^+$. Phys. Chem. Chem. Phys. 15 (2013), 7628-7637

[17] G.Lee, E.Park, H.Chung, Y.Jeanvoine, K.Song, R.Spezia. Gas phase fragmentation mechanisms of protonated testosterone as revealed by chemical dynamics simulations. Int. J. Mass Spectrom. 407 (2016) 40-50.

[18] A.Carrà, V.Macaluso, P.W.Villalta, R.Spezia, S.Balbo. Fragmentation Spectra Prediction and DNA Adducts Structural Determination. J. Am. Soc. Mass Spectrom. 30 (2019) 2771-2784.

[19] M.Obermayer, F.Lynen. Structure of biotin enzymes. Trends in Biochemical Sciences 1 (1976) 169-171.

[20] P.V.Attwood. The structure and the mechanism of action of pyruvate carboxylase. Int. J. Biochem. Cell Biol. 27 (1995) 231-249.

[21] M.Casutt, T.Koppe, M.Schwarz. Vitamins, 9. Biotin. Ullmann's Encyclopedia of Industrial Chemistry 2012.

[22] A.D.Becke. A new mixing of Hartree–Fock and local density-functional theories. J. Chem. Phys. 98 (1993) 1372-1377.

[23] C.T.Lee, W.T.Yang, R.G.Par, Development of the Colle–Salvetti correlation- energy formula into a functional of the electron density. Phys. Rev. B 37 (1988) 785–789.

[24] M.J.S.Dewar, W.Thiel. Ground states of molecules. 38. The MNDO method. Approximations and parameters. J. Am. Chem. Soc. 99 (1977) 4899.

[25] S.Grimme. Accurate description of van der Waals complexes by density functional theory including empirical corrections. J. Comput. Chem. 25 (2004) 1463−1473.

[26] G.B.Rocha, R.O.Freire, A.M.Simas, J.J.P.Stewart. RM1: a reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. J. Comput. Chem. 27 (2006) 1101−1111.

[27] J.J.P.Stewart. Optimization of parameters for semiempirical methods I. J. Comput. Chem. 10 (1989) 209−220.

[28] J.J.P.Stewart. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. J. Mol. Model. 13 (2007) 1173–1213.

[29] J.J.P.Stewart. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. J. Mol. Model. 19 (2013) 1-32.

[30] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Nor- mand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT, 2009.

[31] J.J.P.Stewart. MOPAC2016; Stewart Computational Chemistry: Colorado Springs, CO, 2016.

[32] O.Meroueh, W.L.Hase. Collisional activation of small peptides. J. Phys. Chem. A 103 (1999) 3981-3990.

[33] R.Spezia, J.-Y.Salpin, M.-P.Gaigeot, W.L. Hase, K.Song. Protonated urea collision-induced dissociation. Comparison of experiments and chemical dynamics simulations. J. Phys. Chem. A 113 (2009) 13853-13862.

[34] D.Ortiz, J.-Y.Salpin, K.Song, R.Spezia. Galactose-6-Sulfate collision induced dissociation using QM+MM chemical dynamics simulations and ESI-MS/MS

experiments. Int. J. Mass Spectrom. 358 (2014) 25-35.

[35] P.D.Kirsch, J.G.Ekerdt. KaleidaGraph: Graphing and Data Analysis. J. Am. Chem. Soc. 122 (2000) 11755-11755.

[36] G.H.Peslherbe, H.Wang, W.L.Hase. Monte Carlo Sampling for Classical Trajectory Simulations. Adv. Chem. Phys. 105 (2007) 171−201.

[37] W.L.Hase, D.G.Buckowski, K.N.Swamy. Dynamics of Ethyl Radical Decomposition. 3. Effect of Chemical Activation vs. Microcanonical Sampling. J. Phys. Chem. 87 (1983) 2754-2763.

[38] S.Beisken, M.Earll, C.Baxter, D.Portwood, Z.Ament, A.Kende, C.Hodgman, G.Seymour, R.Smith, P.Fraser, M.Seymour, R.M.Salek, C.Steinbeck. Metabolic differences in ripening of Solanum lycopersicum "Ailsa Craig" and three monogenic mutants. Sci. Data, 1 (2014) 140029.

[39] http://www.massbank.jp/RecordDisplay?id=ML002701&dsn=MetaboLights