Structure and fragmentation of glycine, alanine, serine and cysteine radical cations. A theoretical study

Sílvia Simon\textsuperscript{a,*}, Adrià Gil\textsuperscript{b}, Mariona Sodupe\textsuperscript{b}, Juan Bertran\textsuperscript{b,*,1}

\textsuperscript{a}Institut de Química Computacional, Departament de Química, Universitat de Girona, Girona 17071, Spain
\textsuperscript{b}Departament de Química, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain

Received 1 December 2004; accepted 21 December 2004
Available online 28 June 2005

Abstract

The structure of glycine, alanine, serine and cysteine radical cations, as well as their C\textsubscript{x}–R fragmentations have been studied at the B3LYP/6-31 + G(d,p) level of theory. For all systems the loss of COOH\(^+\) is found to be the most favourable process. However, when R\textsubscript{1} size increases, fragmentations that lead to R\textsubscript{1} or R\textsubscript{C} start become competitive. The energy of a C\textsubscript{x}–R fragmentation can be related to the ionization potential of the two generated fragments, in such a way that the preferred process is the one that leaves the positive charge in the fragment with a lower ionization potential.

\(\text{© 2005 Elsevier B.V. All rights reserved.}\)

Keywords: Fragmentation; Radical cation; Glycine; Alanine; Serine; Cysteine

1. Introduction

Protein and peptide radicals are of great biological interest. The effect of oxidative damage in proteins, which is implicated in pathological disorders [1,2], is mainly due to the reactions that occur in amino acids. The knowledge of their structure and reactivity is also important to understand the role of transient species involved in protein radical catalysis [3]. Moreover, recent studies have shown that radical cations of some amino acids and oligopeptides can be produced by collision induced dissociation of [Cu\textsuperscript{II} (dien)M]\(^{2+}\) complex ion [4] and that their dissociation is very rich and differs considerably from that of protonated amino acids and peptides, which make them very attractive for peptide sequentiation. Because of that, in the past few years, the properties of different amino acid derived radicals have attracted considerable attention, both from an experimental [5–14] and theoretical point of view [4,15–36].

Glycine is the simplest amino acid and consequently an important model compound, which has been the subject of many experimental and theoretical investigations [8,9,11,15,17,24–29]. However, most of the performed studies have focused their attention on the structure and magnetic properties of the C-centered glycy1 [NH\(_2\)CHCOOH]\(^+\) radical, one of the radiation products of glycine in solution. Recently, glycy1 radical has also been generated in gas phase [37,38] by collisional neutralization of the stable glycy1 cation [NH\(_2\)CHCOOH]\(^+\), which is obtained by dissociative ionization of several amino acids such as phenylalanine or serine. Unimolecular decompositions are then studied by reionization mass spectrometry experiments. Moreover, photoion mass spectrometry studies of different amino acids in the 6–22 eV photon energy region have provided new information about their dissociative ionization products [13,14]. It has been shown that for glycine radical cation, the most intense peak is due to the aminomethyl radical ion, NH\(_2\)CH\(_2\), in completely agreement with our previous study [27], where the loss of the COOH radical was calculated to be the lowest energy ion fragmentation. This result has been confirmed by Lu et al. [29].

In the present work, we report density functional calculations for the unimolecular decomposition of glycine,
alanine, serine and cysteine radical cations. Although some experimental and theoretical reports can be found for alanine, cysteine and serine\cite{7,13,14,16,18,23,30–36}, there are only few studies that focus on the radical form \[\text{NH}_2\text{CHR}_1\text{COOH}\]^+\text{--R} and to our knowledge no computational studies have been performed for their radical cations. Different fragmentation processes derived from C--R breaking of \[\text{NH}_2\text{CHR}_1\text{COOH}\]^+\text{--R} will be considered. A relation between the most favourable fragmentation and ionization potential of different fragment will be discussed.

2. Methods

It is well known that amino acids can exist in a large number of conformations due to many single-bond rotamers. Glycine possible conformations and their radical cations have been studied previously\cite{17,26,34}. However, fewer studies exist for the other amino acids due to the larger number of degrees of freedom and possible conformers.

In order to find the lowest energy conformers of different amino acid we have applied the following strategy. We have started from the four most stable glycine conformers\cite{26} showed in Scheme 1 (I–IV), and for each amino acid we have performed a Monte Carlo Multiple Minimum (MCMM) conformational search\cite{39} with the MMFF94s force field\cite{40} allowing only the internal rotations of the side chain. Subsequently, full geometry optimizations at the B3LYP level\cite{41} with the 6-31\text{CC}G(d,p) basis set were carried out for all minima found in the previous conformational search. Radical cation structures were obtained after ionization and reoptimization of the B3LYP/6-31\text{CC}G(d,p) minima found for conformation IV. We have only considered the structures derived from this conformation because it was found to be the most stable one for glycine radical cation\cite{26}. All energy values reported in this paper include zero point energy correction and correspond to the lowest structure found at the B3LYP/6-31\text{CC}G(d,p) level for either system, neutral amino acid and radical cation. The nature of the stationary points, amino acids and their fragments, has been checked by vibrational frequency calculations.

We have chosen B3LYP as level of theory because previous results showed that MP2 and B3LYP geometries were quite similar\cite{26}, and CCSD(T) relative energies calculated at the MP2 and B3LYP geometries differed by less than 0.4 kcal/mol. Moreover, fragmentation energies of the B3LYP level were found to be in quite good agreement with CCSD(T) ones.

Net atomic charges and spin densities have been obtained using the natural population analysis of Weinhold et al.\cite{43}. All calculations have been performed with the GAUSSIAN 98 package\cite{44}.

3. Results and discussion

3.1. Neutral amino acids

Fig. 1 presents the optimized geometry parameters of the lowest conformers of neutral glycine, alanine, serine and cysteine at the B3LYP/6-31\text{CC}G(d,p) level of theory and including zero point energy. At this level of theory the lowest conformer of neutral glycine corresponds to structure I of Scheme 1. This structure shows a bifurcated hydrogen bond between the NH$_2$ group and the carbonylic group. The next conformer in energy, labelled II in Scheme 1, lies 0.715 kcal/mol above and presents a OH--N hydrogen bond. These results are in complete agreement with previous calculations at different levels of theory\cite{26}.

For neutral alanine structures I and II are almost degenerate. If no zero point corrections are taken into account structure II is the most stable conformer. However, the inclusion of zero point vibrational correction stabilizes structure I over II, the former becoming 0.288 kcal/mol lower than the latter. It should be noted that MP2/6-31\text{CC}G(d,p) calculations also give structure I as the most stable one, even if the zero-point energy is not considered. The nearly degeneracy of these two structures has already been noticed before by different authors\cite{30–32}, their relative stability depending on the level of theory used\cite{31}. Recently, an experimental study by Blanco et al. has confirmed that structure I is the most stable conformer for alanine\cite{35}.

Scheme 1.
While neutral glycine and alanine present similar intramolecular interactions, when R₁ is substituted by CH₂OH (in serine), extra intramolecular hydrogen bonds appear, which increases the number of stable conformers and may change their relative stability. It can be observed in Fig. 1 that, in addition to the bifurcated hydrogen bond between the NH₂ group and the carboxylic oxygen, the lowest conformer has also a hydrogen bond between the hydroxyl group of the side chain and the N atom of the amino group. This structure results to be only 0.156 kcal/mol more stable than the one derived from structure II of Scheme 1 which, in addition to the OH → N, presents a hydrogen bond between one H of the NH₂ group and the O of the side chain. These results are in completely agreement with previous studies [30,34] as well as with recent theoretical-experimental results by Lambie et al. [33].

For cysteine the situation is analogous to serine. The difference is that the additional hydrogen bonds formed with the side chain thiol group, CH₂SH, are weaker because thiol is a poorer hydrogen bond donor or acceptor than the hydroxyl group. This fact results in a better stabilization of structure II, which lies 0.766 kcal/mol below I. All results agree with previous studies, which predict the same global minimum [30,34] at different levels of theory.

3.2. Radical cations

As mentioned, only those conformers derived from structure IV of Scheme 1 have been considered as starting points in the optimization of ionized species. Structure IV was found to be the most stable one for glycine radical cation at the MP2 and CCSD(T) levels of theory and is expected to be also the lowest one for alanine, serine and cysteine. It should be noted that at the B3LYP level, structure III(+) of glycine radical cation was found to be lower in energy [26]. However, this isomer presents a two center-three electron interaction between NH₂ and OH groups and these structures are known to be overstabilized by present density functionals due to a bad cancellation of the self interaction error by the exchange functional [42].

Optimized geometries for the lowest conformers derived from structure IV are given in Fig. 2. It can be observed that in all cases the initially pyramidalized –NH₂ group becomes more planar in the radical cation species, in agreement with the fact that –NH₂ acquires an important –NH₂⁺ character upon removal of one electron (see Table 1). Thus, ionization of these amino acids increases the –NH₂ acidity which favors the intramolecular hydrogen bond interaction. Glycine and alanine radical cations present a shorter –NH → O hydrogen bond (2.10 Å) than serine (2.27 Å) or cysteine (2.16 Å), because for the formers the charge and spin density mainly lie at the –NH₂ group, whereas for serine and cysteine they are more delocalized. Accordingly, the open shell orbital in glycine and alanine radical cations has an important contribution of the px orbital of N, whereas for serine and cysteine the open shell orbital becomes more delocalized at the side chain, especially for cysteine (see Fig. 3).

In addition to the strengthening of intramolecular hydrogen bond, other major geometry changes occur upon ionization. These changes can be related to the nodal properties of the open shell orbital or the HOMO orbital from which the electron is removed. It can be observed in Fig. 3 that for glycine, alanine and serine this orbital shows an antibonding character between C and N and thus, it is not surprising that the C–N distance decreases upon ionization. In contrast, this distance is almost unaffected for cysteine, in agreement with the nature of this orbital which does not...
present any contribution at the C atom (see Fig. 3). On the other hand, it is observed that the C–R1 distance in alanine and serine increases upon ionization, in agreement with the fact that the HOMO orbital presents an important bonding character between the two linked atoms. Particularly striking is the increase observed for serine which is about 0.2 Å. However, for cysteine the C–R1 distance remains again almost unaltered (~1.54 Å) due to the properties of the open shell orbital. The most remarkable change for cysteine is the formation of a two center-three electron hemibond between NH2 and SH of the side chain. This kind of bond is quite common for –SH groups and, probably, is responsible for the stabilization of this conformer. It should be noted that although such species are known to over stabilize by DFT methods, this overstabilization is much smaller when elements of third period are involved. In fact, MP2 single point calculations confirm that this hemibond structure is clearly the lowest energy conformer derived from structure IV. For serine, we have not been able to locate such a structure with a two center — three electron interaction.

3.3. Unimolecular decompositions

Table 2 gathers the energy corresponding to different Cα fragmentation processes for glycine, alanine, serine and cysteine. Four different Cα–R bond cleavages can be considered: Cα–COOH, Cα–R1 (with R1 equal to H, CH3, CH2OH or CH2SH for glycine, alanine, serine and cysteine, respectively), Cα–H and Cα–NH2. Such cleavage can be produced in two different ways: that is, by losing a neutral radical (COOH–, R1–, H–, NH2–) or a cation (COOH+–, R1+–, H+–, NH2+–). Thus, eight different reactions can be considered, as they are collected in Table 2.

First of all let us compare the two set of reactions depending on the loss of a neutral radical or a cation fragment. It can be observed that for all four amino acids the reactions corresponding to the loss of neutral radicals (G+→G+ + R–) are more favourable than those corresponding to the loss of the cationic fragments (GR+→G+ + R+–).

In all cases the most favourable process is the one corresponding to the loss of COOH– radical. This fact was previously observed for glycine [27], and is also true for alanine, serine and cysteine. These results are in good agreement with mass spectrometry studies for glycine and alanine, which show that the most intense peaks are (m/z = 30) and (m/z = 44), respectively, which can be assigned to the NH2CH2+ and NH2CH2CH+ ions formed by loss of the COOH– radical. For serine and cysteine no experimental results are found. When the entropic effects are taken into account, all reaction energies are lowered about 11 kcal/mol (except for those which lose an hydrogen atom or a proton); that is, the loss of COOH– is still the most probable process, being a spontaneous process for alanine (∆G298 K = −5.0 kcal/mol).

When the positive charge stays at the fragment that does not contain the Cα atom (reactions 5,6,7,8) the loss of COOH+ is the most favourable process only for glycine and alanine. For serine and cysteine the loss of the side chain,

![Fig. 2. B3LYP/6-31++G(d,p) optimized geometries of the lowest conformers of glycine, alanine, serine and cysteine radical cations.](image)
CH$_2$OH$^+$ and CH$_2$SH$^+$, respectively, and formation of glycyl radical is clearly preferred.

To get a deeper insight into the C$_a$–R fragmentation preferences let us consider the following thermodynamic cycles:

$$
\begin{align*}
G^- + R^- & \rightarrow D_0(-R^-) \\
G^- + R & \rightarrow -IP(G) + IP(R^-) \\
G + R & \rightarrow D_0(G) + IP(R^-) \\
G + R^- & \rightarrow D_0(G) - IP(R^-)
\end{align*}
$$

$D_0(GR)$ correspond to the homolytic dissociation energy of the neutral amino acid, $D_0(-R^C)$ and $D_0(-R^%)$ the dissociation energy of the radical cation leading to $R^C$ and $R^%$, respectively, and IP(GR) the adiabatic ionization potential of the considered amino acid. IP(R$^-$) and IP(G$^-$) correspond to the ionization potential of each fragment. From this scheme it can be observed that the dissociation energies leading to $R^+$ and $R^-$ can be decomposed as

$$
\begin{align*}
D_0(-R^-) &= -IP(GR) + D_0(GR) + IP(G^-) \\
D_0(-R^+) &= -IP(GR) + D_0(GR) + IP(R^-)
\end{align*}
$$

depending on the preference for the loss of neutral radical or a cation fragment. Table 3 shows the adiabatic ionization potential (with zero point correction included) for each amino acid (first row) and fragments. Neutral amino acid C$_a$–R dissociation energies are given in Table 4.

From this energy decomposition scheme it can be noted that for each amino acid the preference in the loss of the neutral radical fragment (R$^-$) or cation fragment (R$^+$) only depends on the relative ionization potentials of IP(R$^-$) and IP(G$^-$). It can be observed in Table 3 that the ionization energy of the fragment that contains C$_a$ (IP(G)) is lower and, because of that, the loss of R$^-$ is always the preferred process. For example, for alanine the ionization potential of NH$_2$CHCH$_3$% (IP $= 133.2$ kcal/mol) is lower than that of COOH% (IP $= 192.4$ kcal/mol), which makes the loss of

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Fragmentation energies including zero point correction for glycine, alanine, serine and cysteine radical cations (in kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHR$_1$]$^+$ + [COOH]$^-$</td>
</tr>
<tr>
<td>(2)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHCOOH]$^+$ + R$_1^-$</td>
</tr>
<tr>
<td>(3)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHR$_1$COOH]$^+$ + H$^+$</td>
</tr>
<tr>
<td>(4)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [CHR$_1$COOH]$^+$ + [NH$_2$]$^+$</td>
</tr>
<tr>
<td>(5)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHR$_1$]$^+$ + [COOH]$^+$</td>
</tr>
<tr>
<td>(6)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHCOOH]$^+$ + R$_1^-$</td>
</tr>
<tr>
<td>(7)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [NH$_2$CHR$_1$COOH]$^+$ + H$^+$</td>
</tr>
<tr>
<td>(8)</td>
<td>[NH$_2$CHR$_1$COOH]$^+$ - [CHR$_1$COOH]$^+$ + [NH$_2$]$^+$</td>
</tr>
</tbody>
</table>

(a) R$_1$=H, CH$_3$, CH$_3$OH, CH$_3$SH for Gly, Ala, Ser and Cys, respectively.
neutral COOH\(^-\) the most favourable process. The same conclusion can be reached for each pair of fragmentation of all four amino acids.

Because the IP(\text{GR}) term is the same for a given amino acid, differences on the C\(_2\)-R fragmentations are due to the differences on the neutral C\(_2\)-R dissociation energies, D\(_0\)(GR), and to the ionization energies of the different fragments. However, it can be observed in Table 4 that the homolytic dissociation energies of C\(_2\)-COOH, C\(_2\)-NH\(_2\), C\(_2\)-H, bonds of neutral amino acids, D\(_0\)(GR), are quite similar and, thus, the differences between the fragmentation energies mainly arise from the changes on the IP(\text{G‘}) or IP(\text{R‘}) ionization energies and also on the C\(_2\)-R\(_1\) dissociation energy, which for alanine, serine and cysteine is about 10–25 kcal/mol smaller than the C\(_2\)-COOH, C\(_2\)-NH\(_2\), C\(_2\)-H ones. Thus, reactions corresponding to the loss of R\(_1\) fragment (radical or cation), become more favourable. It is remarkably that the reaction energy corresponding the loss of R\(_1\) decreases significantly for serine and cysteine. In spite of this, still the dominant factor comes from the ionization energy of the fragments; that is, the preferred process is the one that leaves the positive charge in the fragment with lower ionization energy. Because for all four amino acids the [NH\(_2\)CHR\(_1\)]\(^-\) fragment is the one with a lower ionization energy, the [NH\(_2\)CHR\(_1\)COOH]\(^-\) + [COOH]\(^-\) fragmentation is the most favourable unimolecular decomposition in all cases. The second most favourable process is the lost of R\(_1\) to lead the glycyl cation because the C\(_2\)-R\(_1\) bond is weaker. It is interesting to note that while for glycine and alanine this second fragmentation is about 15 kcal/mol more costly than the loss of COOH\(^-\), for serine and cysteine, the energy difference between the two processes decreases to only 5 kcal/mol. Thus, for amino acids with larger side chain and weaker C\(_2\)-R\(_1\) bonds one could expect that formation of glycyl cation [NH\(_2\)CHR\(_1\)COOH]\(^+\) \(\rightarrow\) NH\(_2\)CHCOOH\(^+\) + R\(_1\) becomes competitive with the loss of COOH or even more favourable.

Finally, let us discuss about the formation of glycyl radical, which takes place in reaction (6). As it was pointed out, the formation of glycyl radical is not energetically favourable from glycine (179.8 kcal/mol) or alanine (82.9 kcal/mol), while for serine and cysteine the formation of glycyl starts being a competitive process. This fact is mainly due to the side chain (R\(_1\)) ionization potential. It can be observed from Table 3 that ionization potential from CH\(_2\)OH and CH\(_2\)SH is half the ionization potential of H atom. It can be expected that amino acids which present R\(_1\) with low energy ionization will be possible candidate to the formation of glycyl radical. Thus the ionization energy can be a parameter to predict the possible fragmentation process and help to interpret the mass spectrometry experiments.

### 4. Conclusions

In this work, the structure of glycine, alanine, serine and cysteine radical cations as well as their fragmentations have been studied at the B3LYP/6-31 + + G(d,p) level of theory. Geometry changes upon ionization have been interpreted through the nodal properties of the open shell orbital. Among all different C\(_2\)-R fragmentation processes, the one that loses COOH\(^-\) is the most favourable one. Nevertheless, for amino acids with increasing R\(_1\) size, fragmentations leading to R\(_1\) or R\(_1\)\(^+\) start being competitive. This is important because both processes can lead to the formation of glycyl radical, indirectly or directly, respectively. A thermodynamic cycle has shown that the formation of glycyl radical is connected with the ionization potential of R\(_1\) in such a way that the more ionizable is the side chain the easier would be the formation of glycyl radical.

### Table 3
**Adiabatic ionization potential (kcal/mol) of neutral aminoacids and radical fragments (zero point corrections included)**

<table>
<thead>
<tr>
<th></th>
<th>GLY</th>
<th>ALA</th>
<th>SER</th>
<th>CYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH(_2)CHR(_1)COOH</td>
<td>208.6</td>
<td>203.8</td>
<td>200.4</td>
<td>195.5</td>
</tr>
<tr>
<td>[NH(_2)CHR(_1)COOH](^+)</td>
<td>167.4</td>
<td>154.8</td>
<td>153.0</td>
<td>155.4</td>
</tr>
<tr>
<td>[CHR(_1)COOH](^-)</td>
<td>209.8</td>
<td>199.0</td>
<td>164.2</td>
<td>168.9</td>
</tr>
<tr>
<td>[NH(_2)CHR(_1)](^-)</td>
<td>147.4</td>
<td>133.2</td>
<td>140.4</td>
<td>138.2</td>
</tr>
<tr>
<td>[NH(_2)CHCOOH]</td>
<td>167.4</td>
<td>167.4</td>
<td>167.4</td>
<td>167.4</td>
</tr>
<tr>
<td>H(^+)</td>
<td>314.8</td>
<td>314.8</td>
<td>314.8</td>
<td>314.8</td>
</tr>
<tr>
<td>[NH(_2)](^+)</td>
<td>293.2</td>
<td>293.2</td>
<td>293.2</td>
<td>293.2</td>
</tr>
<tr>
<td>[COOH](^+)</td>
<td>192.4</td>
<td>192.4</td>
<td>192.4</td>
<td>192.4</td>
</tr>
<tr>
<td>R(_1)(^-)</td>
<td>314.8</td>
<td>228.5</td>
<td>177.3</td>
<td>177.1</td>
</tr>
</tbody>
</table>

### Table 4
**Zero point corrected dissociation energies of neutral glycine, alanine, serine and cysteine (kcal/mol)**

<table>
<thead>
<tr>
<th>GR(\rightarrow)G‘ + R‘ (\rightarrow) (a)</th>
<th>GLY</th>
<th>ALA</th>
<th>SER</th>
<th>CYS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH(_2)CHR(_1)COOH (\rightarrow) [NH(_2)CHR(_1)](^-) + [COOH](^-)</td>
<td>79.0</td>
<td>76.8</td>
<td>77.9</td>
<td>75.4</td>
</tr>
<tr>
<td>NH(_2)CHR(_1)COOH (\rightarrow) [NH(_2)CHCOOH](^+) + R(_1)(^+)</td>
<td>73.6</td>
<td>58.3</td>
<td>56.3</td>
<td>52.2</td>
</tr>
<tr>
<td>NH(_2)CHR(_1)COOH (\rightarrow) [NH(_2)CHR(_1)](^+) + H(^+)</td>
<td>73.6</td>
<td>70.6</td>
<td>72.6</td>
<td>72.3</td>
</tr>
<tr>
<td>NH(_2)CHR(_1)COOH (\rightarrow) [CHR(_1)COOH](^+) + [NH(_2)](^-)</td>
<td>74.3</td>
<td>69.1</td>
<td>73.2</td>
<td>70.7</td>
</tr>
</tbody>
</table>

(a) R\(_1\) = H, CH\(_3\), CH\(_2\)OH, CH\(_2\)SH for Gly, Ala and Cys, respectively.
Acknowledgements

Financial support from MCYT and FEDER (project BQU2002-04112-C02), DURSI (project 2001SGR-00182), and the use of the computational facilities of the Catalonia Supercomputer Center (CESCA) are gratefully acknowledged.

References