Eur. Phys. J. D **20**, 551–555 (2002)
DOI: 10.1140/epjd/e2002-00142-y

THE EUROPEAN PHYSICAL JOURNAL D

A DFT study of electron or hole localization in a peptide containing asparagin

C. Houée-Levin[1,a] and J. Bergès[2]

[1] Laboratoire de Chimie Physique[b], centre universitaire, 91405 Orsay Cedex, France
[2] Laboratoire de Chimie Théorique[c], Université Paris VI, 4 place Jussieu, 75005 Paris, France

Received 21 January 2002
Published online 13 September 2002 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2002

Abstract. The mechanisms of protein degradation induced by ionisation are of great interest for radiobiology, improvement of mass spectroscopy and industrial processes such as radio sterilisation. Sequences containing asparagin are very sensitive especially if surrounded by glycine. Very few techniques allow a satisfying understanding of the processes induced by creation of an anionic or cationic site in a peptide. We used the methods of quantum chemistry (DFT/B3LYP with 6-31G* basis set) to characterise the geometry modifications induced in the cations or in the anions derived from peptide Gly Asn Gly. The cationic sites are localised mostly close to the first peptidic bond and induce a lengthening of the Ca–C(O) bond. Conversely the anionic sites are localised on a carbonyl function. Implications are discussed considering the radiolytic products and the proposed mechanisms.

PACS. 31.15.Ar Ab initio calculations – 87.15.-v Biomolecules: structure and physical properties

1 Introduction

Protein sequences containing asparagin are known for their susceptibility to get degraded. The main degradation pathway is deamidation. This reaction was extensively studied and is well documented (see for instance [1] and references therein). It causes biologically and structurally important alterations in protein structures. Asparagin degradation is involved in ageing process [2,3]. This process is somewhat sequence-specific and the sequence Asn-Gly is especially sensitive [4], probably because steric hindrance from Glycine is minimum and conformation flexibility is maximum [5] as shown by recent *ab initio* calculations [6]. In the gas phase, sequences Gln-Gly are also a dominant dissociation site in protein fragmentation [7].

Oxidative damage enhances asparagin instability [8], which indicates a free radical route in asparagin degradation. Also, irradiation of proteins in the solid state, either lyophilised or in frozen aqueous solution, induces peptidic bond breakage leading to non-random fragmentations [9]. It was shown that sequences containing asparagin or aspartate groups are prone to bond cleavage leading this residue either as a C-terminal or at a N-terminal. Despite their importance in many biological (radioprotection, radiotherapy) and industrial processes (radio-sterilization of food and drugs), these processes are poorly understood. The aim of this work is thus to improve the understanding of free radical-induced degradation in an asparagin-containing peptide. The chosen sequence was Gly Asn Gly because of its sensitivity and also because it allows comparison of the roles of glycine at C- or at N-terminal.

[a] e-mail: chantal.houee-levin@lcp.u-psud.fr
[b] UMR 8000 du CNRS
[c] UMR 7616 du CNRS

2 Computational procedure

Three peptides were constructed. In the largest one called GNG (Glycine Asparagin Glycine peptide), the carboxylic function was replaced by an aldehyde (Fig. 1), to evaluate the role of the terminal carbonyl function. To precise the roles of carbonyl and amine terminal groups, two more truncated peptides were also considered: without the carbonyl group –CO of C-terminal (GNGa), and without both –CO and $-NH_2$ of terminal amine (GNGb). Without solvation, peptides terminal groups are neutral and not zwitterionic. The starting point was the geometry of this sequence in hen egg-white lysozyme, taken from the Protein Data Bank (2lym) (PDB).

All isolated species were optimised using DFT (B3LYP) with basis sets (6-31G*) using the program Gaussian 98 [10]. This method is commonly used for localised radical structures [11]. Some problems with DFT theory and negative ions have been suggested. However it is now clear that for bound systems reasonable estimates of electron affinities can be obtained with DFT theory [12]. The use of this relatively small basis set is

Table 1. Energies B3LYP/6-31G* of optimised tripeptides and of their related cations and anions.

Tripeptide	Molecule	Anion		Cation	
	Energy (a.u.)	Energy (a.u.)	Electronic affinity (eV)	Energy (a.u.)	Ionisation energy (eV)
GNG	−833.20674	−833.20595	−0.02	−832.92294	7.72
GNGa	−719.89751	−719.86595	−0.86	−719.60804	7.96
GNGb	−664.51730	−664.51600	−0.45	−664.25513	7.13

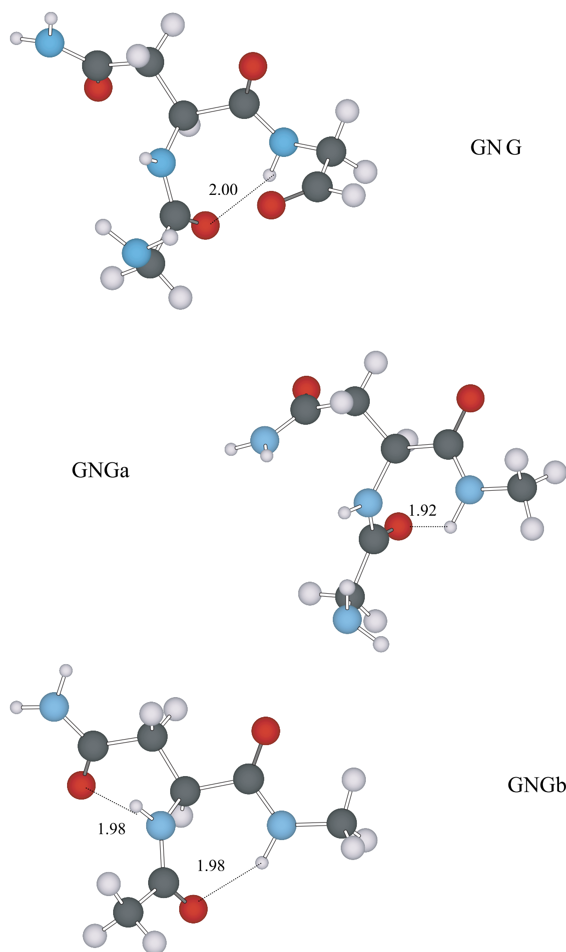

Fig. 1. Optimised geometries of peptides. Dotted lines: hydrogen bonding with interatomic distances less than *ca.* 2 Å.

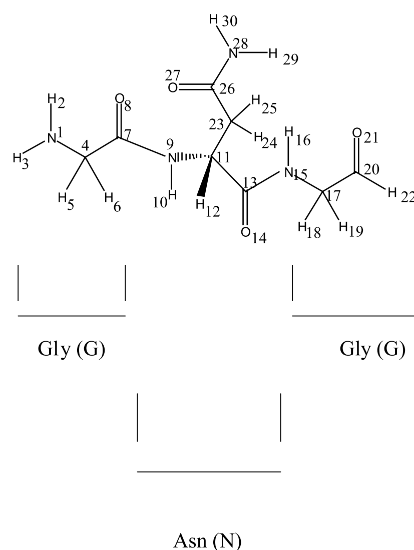

Fig. 2. Numbering of atoms.

justified by the number of geometry optimisations that had to be performed, and by the fact that DFT methods are known to be little basis set dependent [13]. In addition, our previous works demonstrated that such basis sets are sufficient [14,15]. Due to the size and the flexibility of these peptides it was not possible to perform full optimisations with larger basis sets. However single point calculations were done with B3LYP/6-31+G* and B3LYP/6-311+G(2d, 2p).

Calculations were performed on the Nec SX-5 of IDRIS (Orsay, France). One full optimisation of each peptide required more than 60 h cpu.

3 Results and discussion

All species, molecules or free radicals, were fully optimised. Their energies are gathered in Table 1. Molecular structures are compared in Figures 1 and 2 gives the numbering of atoms.

In lysozyme, the sequence corresponds to a turn thus it has many degrees of freedom and few hydrogen bonds. In optimised structures, hydrogen bonds are created for stabilisation. The network is similar in all peptides (Fig. 1, dotted lines). In the three molecules, intramolecular H-bonds create links between the same atoms of the peptidic bonds: H_{16} and O_8. Carbonyl and amine terminals are not directly involved.

Analysis of the energies of the cation and anion free radicals shows that electronic affinities are weak and negative, especially for GNGa, whereas ionisation energies are high, as it should be for an organic molecule. The values of ionisation energies are rather close, the differences do not seem to be significant. We are aware that values of electronic affinities may be somewhat imprecise because of the lack of diffuse orbitals. However we have noticed that they lack the self-consistency even with diffuse orbitals (verified by us on these peptides) [14,16]. Nevertheless, it seems that for bound systems and atom-centred radicals reasonable estimates of electronic affinities can be obtained with DFT theory.

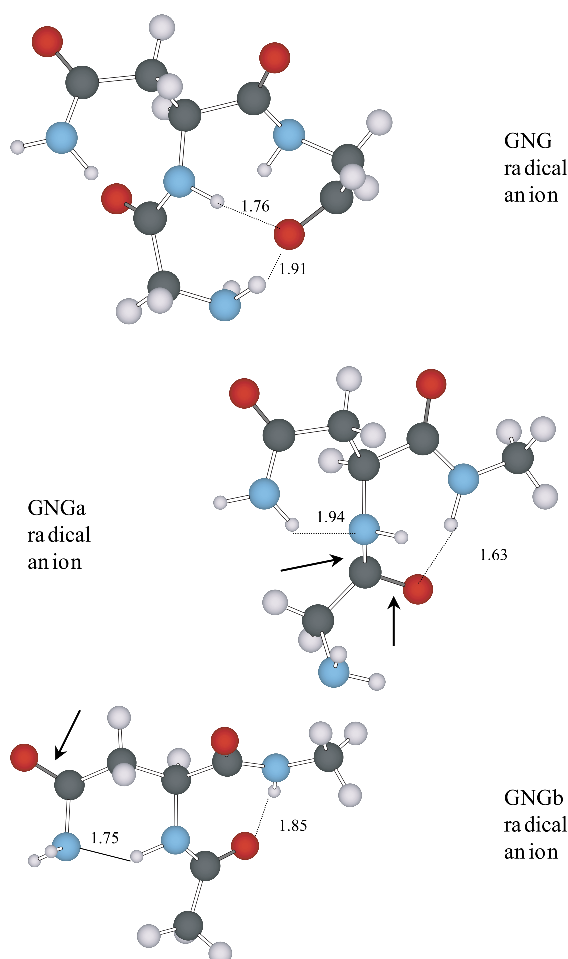

Fig. 3. Optimised geometries of anions. Dotted lines: hydrogen bonding with interatomic distances less than *ca.* 2 Å. Arrows indicate the bonds whose length is modified compared to that in the molecule.

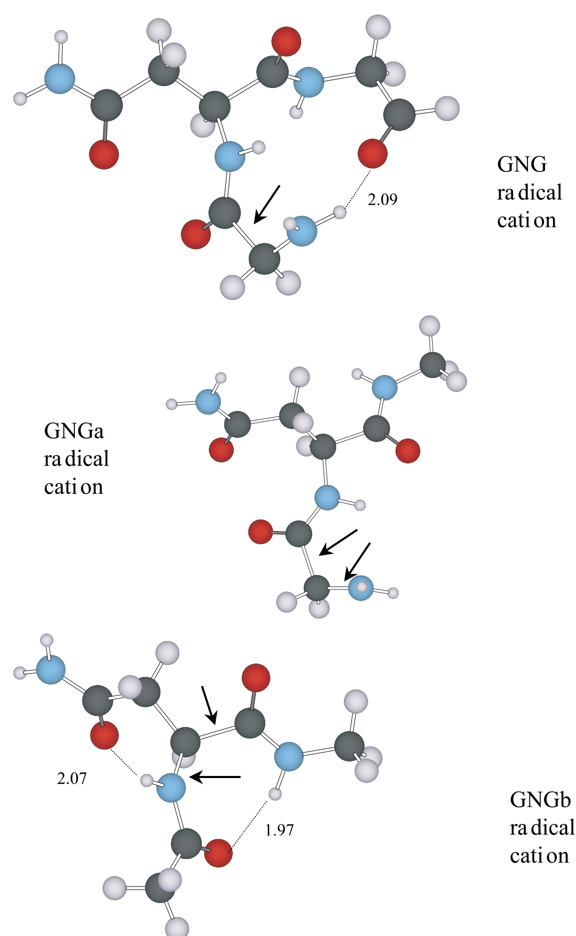

Fig. 4. Optimised geometries of cations. Dotted lines: hydrogen bonding with interatomic distances less than *ca.* 2 Å. Arrows indicate the bonds whose length is modified compared to that in the molecule.

Optimised structures of anions and cations are represented in Figures 3 and 4 respectively. In GNG radicals cations as well as anions, the terminal amine and carbonyl are hydrogen bonded. The intramolecular hydrogen bond O_8–H_{16} observed in the molecular peptides remains in GNGb cation and in GNGa and GNGb anions. In these three entities, it is reinforced by one more hydrogen bond between the first glycine and asparagin moieties. Furthermore hydrogen bonds are very short creating a tight network in each anion.

To describe localisation of unpaired electrons, we focused on spin densities rather than on charge distribution modifications because charges are not only on heavy atoms but split between hydrogens and C or N. Spin densities and modified bond lengths are gathered in Table 2. For anions, spin density is always on CO bonds. As a consequence, there is an important increase of this bond length. However the localisation of this carbonyl function is dependent on the anion: in GNG, spin density is on the CO terminal whereas in GNGa, it is on the CO function of the first peptidic bond (C_7–O_8), and in GNGb anion, on the amide of asparagin. All these results are independent of the size of the basis and were confirmed by single point calculations with 6-31+G* and 6-311+G(2d, 2p) basis sets. In GNGa, the N_1–C_4 bond length does not change much (0.02 Å) but the peptidic bond is strongly elongated (Tab. 2). Also, the stereochemistry of the peptidic bond is modified: whereas it adopts a "trans" conformation in the molecule and in the cation, it is almost "cis" in the anion.

It was proposed that in aqueous solution, hydrated electron adds at the C=O bond of peptide C-terminal [17]. In simple amino acids, the resulting radical anion undergoes deamination [18]. The same reaction occurs with simple peptides like glycylglycine. However with longer peptides such as tri- or tetra-glycine in aqueous solution, electron localisation on the carbonyl function is said to lead also to deamidation [19], *i.e.* breaking of the N_9–C_{11} bond (Fig. 2). In the solid phase, electron localization would also occur on a carbonyl group. Our calculations support these hypotheses. In addition, differences in the CO function with neighbouring group would be in agreement with the variety of final compounds, which depends on the sequence.

Table 2. Bond lengths and spin densities in all entities.

	Molecule	Cation		Anion	
	Bond length (Å)	Spin densities	Modified bond lengths (Å)	Spin densities	Modified bond lengths (Å)
GNG	C_4–C_7 1.54	O_8 0.40	C_4–C_7 1.65		
		N_1 0.38			
		O_{27} 0.11			
	C_{20}–O_{21} 1.21			C_{20} 0.68	C_{20}–O_{21} 1.31
				O_{21} 0.21	
GNGa	C_4–C_7 1.52	O_8 0.46	C_4–C_7 1.63		C_7–O_8 1.31
	C_4–N_1 1.48	N_1 0.29	C_4–N_1 1.40		
		O_{27} 0.15			
	C_7–O_8 1.25			C_7 0.65	C_7–O_8 1.31
	C_7–N_9 1.33			O_8 0.26	C_7–N_9 1.49
GNGb	C_{11}–C_{13} 1.55	O_{14} 0.42	C_{11}–C_{13} 1.61		
	C_{11}–N_9 1.47	N_9 0.22	C_{11}–N_9 1.41		
		O_{27} 0.18			
	C_{26}–O_{27} 1.22			C_{26} 0.58	C_{26}–O_{27} 1.29
				O_{27} 0.32	

The geometrical changes could indicate that preferential cleavage might occur at the peptidic bond. No elongation of the N_9–C_{11} bond is seen, which would lead to deamidation. Thus mechanisms seem to be different with the presence of water as a proton donor.

As for cations, the spin density is localised on N1 of $–NH_2$ terminal for GNG and GNGa, and on the carbonyl group of the first peptidic bond. They move to N_9 and O_{14} in GNGb, hence both peptidic bonds may be affected. Thus suppression of C-terminal does not change spin densities, but that of amine does. Nevertheless, these atoms are always separated by a methylene group. It is associated with an important variation of bond lengths: elongation of the C–C bond adjacent to the concerned peptide bond (Ca–C(O) in the PDB terminology, here C_4–C_7 or C_{11}–C_{13}) (from 0.06 to 0.11 Å) and simultaneous decrease of the adjacent C–NH (from 0.06 to 0.08 Å). Part of the spin density is also on the amide group for the three peptides (O_{27}).

Experimental data on oxidation of polypeptides indicate that the most stable radicals formed after OH attack in aqueous solution would result form H abstraction from Ca (in the PDB notation) (here C_4, C_{11} or C_{17}). Examination of the C–H bond lengths of all Ca carbons shows that none of them is elongated. Thus this mechanism should proceed through one-electron oxidation followed by proton abstraction by the solvent. However it is difficult to identify products coming from the Ca–C(O) bond break since compounds coming from one-electron oxidation are very numerous [20] and propositions of mechanism are somewhat specious.

4 Conclusion

In this work we have investigated the geometrical changes induced by ionisation, in a tripeptide based upon the sequence Gly Asn Gly.

DFT calculations show different patterns for electron localisation in radical anions and cations. For anions the electron is preferentially located on the CO bonds, but the concerned bond varies with the environment. It belongs to the first glycine moiety in GNG and GNGa and to asparagin if both terminal groups, COH and NH_2 are missing. For cations, in all peptides, spin density is partially on asparagin and is shared between CO and NH separated by a methylene group.

In general, the first peptidic bond (C_7–N_9) is more easily either reduced or oxidised than the second one (C_{13}–N_{15}). It would mean that in the gas phase, N-terminal glycine is more radio-sensitive than the C-terminal one. It suggests also that preferential bond cleavage would occur between this glycine and asparagin. The latter would then end up as N-terminal, as observed after irradiation of lysozyme in frozen aqueous solutions [8]. Asparagin is always the target in cations, in good agreement with instability of this amino acid in oxidative conditions [7]. In anions, this residue is a target only in truncated peptide GNGb. However this is a situation similar to that encountered in polypeptidic chain, in agreement with the sensitivity of this residue toward ionising radiation. Finally, to extrapolate these results to protein radio-sensitivity, one should also take into account the protein's constraints and dynamics which might influence the sequence behaviour. Development of the recent methods QM/MM would allow further studies.

We thank IDRIS for providing computer time (project 010268).

References

1. N.E. Robinson, B.A. Robinson, Proc. Natl. Acad. Sci. US **98**, 12409 (2001)
2. P.J. Groenen, M.J. van Dongen, C.E. Voorter, H. Boemendal, W.W. de Jong, FEBS Lett. **322**, 69 (1993)

3. H. Lindner, B. Sarg, H. Grunicke, W. Hellinger, J. Cancer Res. Clin. Oncol. **125**, 182 (1999)
4. L. Takemoto, T. Emmons, Curr. Eye Res. **10**, 865 (1991)
5. T. Geiger, S. Clarke, J. Biol. Chem. **262**, 785 (1987)
6. J.L. Radkiewicz, H. Zipse, S. Clarke, K.N. Houk, J. Am. Chem. Soc. **123**, 3499 (2001)
7. A.P. Jonnson, T. Bergman, H. Jornvall, W.J. Griffiths, Rapid. Commun. Mass Spectrom. **15**, 713 (2001)
8. A.R. Hopkiss, Biogerontology **2**, 173 (2001)
9. A. Filali-Mouhim, M. Audette, M. St Louis, L. Thauvette, L. Denoroy, F. Penin, X. Chen, N. Rouleau, J.-P. Le Caer, J. Rossier, M. Potier, M. Le Maire, Int. J. Radiat. Biol. **72**, 63 (1997)
10. *Gaussian 98, Revision A.9*, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1998
11. A. Rauk, D.A. Armstrong, J. Bergès, Can. J. Chem. **79**, 405 (2001)
12. X. Li, Z. Cai, M.D. Sevilla, J. Phys. Chem. A **106**, 1596 (2002)
13. B. Braida, P.C. Hiberty, A. Savin, J. Phys. Chem. A **102**, 7872 (1998)
14. J. Bergès, F. Fuster, J.-P. Jacquot, B. Silvi, C. Houée-Levin, Nukleonika **45**, 23 (2000)
15. S. Carles, C. Desfrançois, J.P. Schermann, J. Bergès, C. Houée-Levin, Int. J. Mass Spectrom. **205**, 227 (2001)
16. J. Bergès, E. Kassab, D. Conte, E. Adjadj, C. Houée-Levin, J. Phys. Chem. **101**, 7809 (1997)
17. W.M. Garrison, Curr. Top. Radiat. Res. **4**, 43 (1968)
18. W.M. Garrison, Free Radic. Biol. Med. **6**, 285 (1989)
19. J.B. D'Arcy, M. Sevilla, Radiat. Phys. Chem. **13**, 119 (1979)
20. C. Houée-Levin, C. Sicard-Roselli, in *Radiation Chemistry, Present status and future trends*, edited by C.D. Jonah, B.S.M. Rao (Elsevier, Amsterdam, 2001), pp. 553-584