Neutral Zwitterions

DOI: 10.1002/ange.200906503

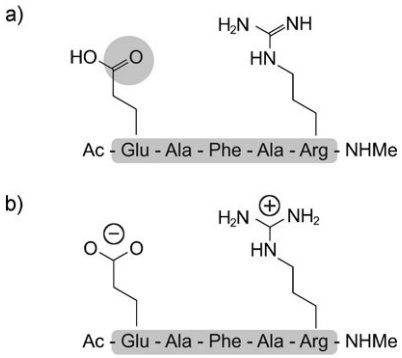
Internal Proton Transfer Leading to Stable Zwitterionic Structures in a Neutral Isolated Peptide**

Anouk M. Rijs, Gilles Ohanessian,* Jos Oomens, Gerard Meijer, Gert von Helden, and Isabelle Compagnon**

Decades of gas-phase spectroscopy of small biomolecules have enabled some of the intrinsic physical and chemical properties of the building blocks of life to be unraveled.[1] Bridging the gap towards an understanding of the biological function of these biomolecules has become an essential issue. For the processes that take place in the active sites of functional proteins at the molecular level to be understood, two critical aspects must be taken into account: 1) interactions with the biological environment (protons, electrons, metal ions, water molecules) and 2) the specific organization of a few significant amino acid residues nested in the well-defined local environment shaped by the entire protein. In this context, it is important to pursue a bottom-up approach, whereby elements of the environment can be introduced step-by-step in a controlled fashion until the biological function emerges.

A crucial discrepancy between the gas-phase structure of isolated amino acids and peptides and their biologically relevant counterparts is the transition from the canonical to the zwitterionic form. Whereas neutral, isolated amino acids[2] and peptides[3] have always been found in their canonical form, studies on ionic complexes have shown that zwitterionic forms may be stabilized by the addition of a proton,[4] an electron,[5] a metal cation,[6] or a metal dication[7] or by microsolvation.[8] In the case of overall-neutral complexes, the canonical-to-zwitterionic transition was observed upon the stepwise addition of solvent molecules.[9]

We report herein the first observation of an "autozwitterion" formed by intramolecular proton transfer between nearby residues in a neutral, isolated peptide. We specifically designed the pentapeptide Ac-Glu-Ala-Phe-Ala-Arg-NHMe (EAFAR; Scheme 1) with an appropriate structure for potential internal proton transfer between residues in a well-defined local environment. The residues intended for participation in proton transfer are the most acidic residue, glutamic acid (Glu), and the most basic residue, arginine (Arg), which were placed at the N and C termini of the peptide, respectively. To avoid interactions of the peptide extremities with the Glu and Arg side chains, and to make the vibrational signature of the Glu side chain unambiguous, we protected the N and C termini with CH_3CO (Ac) and $NHCH_3$ (NHMe) groups, respectively. In the gas phase, the Glu and Arg amino acids are in the canonical form;[2b,10] however, they are charged (deprotonated and protonated, respectively) in the biological environment as a result of proton exchange with the environment. The central part of the peptide consists of the Ala-Phe-Ala sequence, which was shown to adopt a 3_{10}-helical structure:[11] an ideal local architecture to bring the basic and acidic side chains into close proximity. The phenylalanine residue also served as the UV chromophore for IR–UV ion-dip spectroscopy (IR IDS) in this study.[12]



Scheme 1. a) Canonical structure of the capped EAFAR peptide. b) Zwitterionic structure that would result from proton transfer from the acidic side chain of glutamic acid to the basic side chain of arginine. The IR probes used in this study (the peptide C=O groups and side-chain carboxylic acid C=O group) are highlighted in gray.

The use of IR IDS in combination with quantum-chemical calculations is a well-established technique for examining the secondary structure of peptides.[13] Besides the amide I

[*] Dr. A. M. Rijs, Prof. Dr. J. Oomens
FOM Institute for Plasma Physics "Rijnhuizen"
Edisonbaan 14, 3439 MN Nieuwegein (The Netherlands)
Fax: (+31) 30-603-1204
E-mail: a.m.rijs@rijnhuizen.nl
Prof. Dr. G. Ohanessian
Laboratoire des Mécanismes Réactionnels
Département de Chimie, Ecole Polytechnique, CNRS
91128 Palaiseau Cedex (France)
Fax: (+33) 1-6933-4801
E-mail: gilles.ohanessian@polytechnique.fr
Prof. Dr. G. Meijer, Dr. G. von Helden
Fritz-Haber-Institut der Max-Planck-Gesellschaft
Faradayweg 4–6, 14195 Berlin (Germany)
Dr. I. Compagnon
Université de Lyon, Université Lyon 1, CNRS, UMR 5579, LASIM
43 boulevard du 11 Novembre 1918, 69622 Villeurbanne (France)
Fax: (+33) 4-7243-1507
E-mail: compagnon@lasim.univ-lyon1.fr

[**] Financial support from the Dutch–French van Gogh program is gratefully acknowledged. A.M.R. acknowledges the Netherlands Organization for Scientific Research (NWO) for a VENI postdoctoral fellowship. We acknowledge the support of the FELIX staff, in particular Dr. B. Redlich and Dr. A. F. G. van der Meer.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200906503.

(mainly C=O stretching) and amide II (mainly CNH bending) vibrations, we used an additional IR probe to address the question of the charge state of the Glu side chain: the presence of a free acid C=O stretching vibration in the 1740–1800 cm^{-1} region offers an unambiguous signature of a carboxylic acid versus a carboxylate group. Note that the IR signature of the amide C=O stretching vibration, with vibrational frequencies ranging from 1640 to 1720 cm^{-1}, is distinct from that of a carboxylic acid C=O stretching vibration.

The peptides were brought into the gas phase by laser desorption and cooled by seeding into a supersonic expansion of argon. First, we recorded the excitation spectra by resonance-enhanced two-photon ionization (R2PI) to locate the positions of the UV resonances. As observed for various large molecules,[14] the UV spectrum of EAFAR is unresolved. Subsequently, ground-state IR spectra of mass-selected molecules were recorded by IR–UV ion-dip spectroscopy. The IR spectra were obtained in the mid-infrared region (1850–1000 cm^{-1}) by using the free electron laser FELIX.[15]

The main band centered around 1680 cm^{-1} and shaded pink in the IR spectrum of EAFAR (Figure 1a) is assigned to the amide I mode and is mainly due to unresolved combinations of the six local amide I bands (the presence of additional arginine-side-chain modes in this region, as observed in the calculated spectra, is discussed in more detail below). The broad feature around 1525 cm^{-1} corresponds to the amide II mode. The clear absence of the carboxylic acid C=O stretching vibration from the 1740–1800 cm^{-1} region shaded blue in Figure 1 indicates that the glutamic acid side chain is deprotonated in the local peptide environment and thus reveals the presence of a zwitterionic structure.

Further experimental evidence for this conclusion is provided by IR spectra of the two glutamic acid derivatives Z-Glu-OH and Z-Glu-NH_2. Both molecules contain a carbobenzyloxy group (Z) to facilitate the use of IR IDS (Figure 1b,c).[16] The IR spectrum of the major conformer of Z-Glu-OH (Figure 1b) exhibits two partially resolved bands at 1761 and 1775 cm^{-1}, which correspond to the side-chain and C-terminal carboxylic acid C=O stretching vibrations. The band at 1729 cm^{-1} was assigned to the amide C=O stretching mode. In the case of Z-Glu-NH_2 (Figure 1c), the carboxylic acid group of the side chain is unique, and was unambiguously identified at 1747 cm^{-1}. The intense band at 1713 cm^{-1} corresponds to the two overlapping amide I bands. The observation of carboxylic acid C=O stretching bands above 1740 cm^{-1} in these two spectra enabled us to determine a benchmark signature of the glutamic acid side chain in its canonical form. The clear absence of a peak in this region of the EAFAR spectrum (Figure 1a) confirms the formation of a zwitterionic structure through deprotonation of the Glu residue in the specific local environment of the peptide.

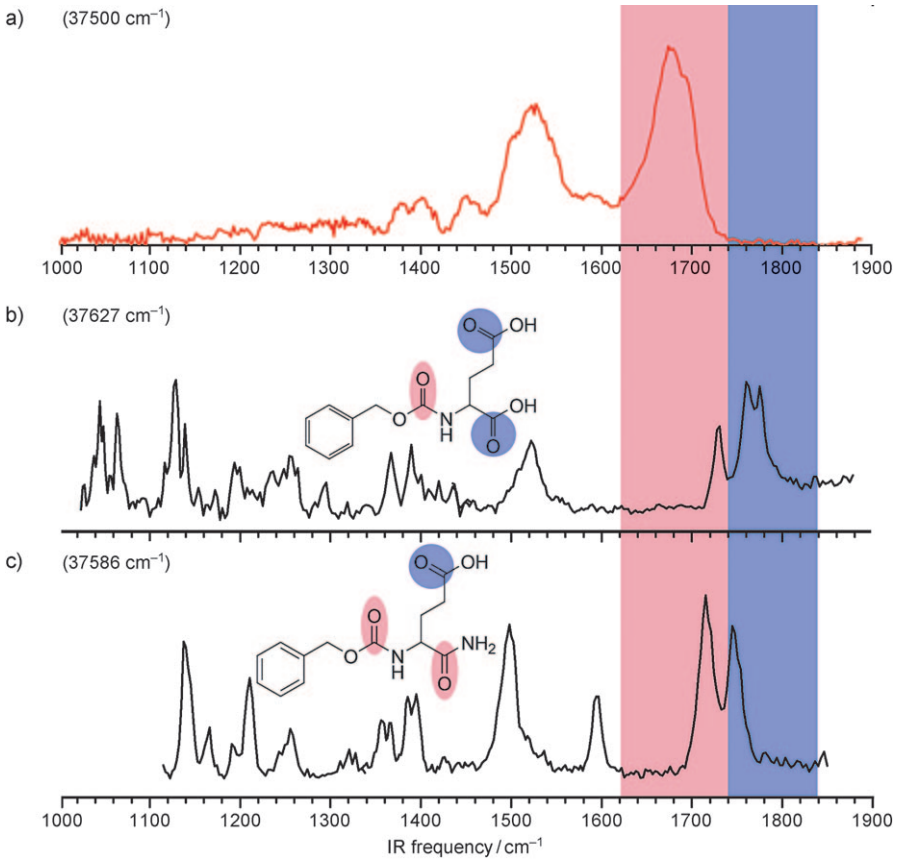


Figure 1. IR spectra of a) capped EAFAR, b) Z-Glu-OH, and c) Z-Glu-NH_2. The region for the peptide C=O stretching vibration is shaded pink, and the region for the carboxylic acid C=O stretching vibration is shaded blue. The energy of the UV photon is given in brackets.

We further examined the secondary structure of the peptide with the help of high-level quantum-chemical calculations. As an exhaustive exploration of the potential-energy surface of a molecule of the size and complexity of EAFAR is highly challenging, we devised the following modeling strategy. First, the peptide was simplified from Ac-Glu-Ala-Phe-Ala-Arg-NHMe to Ac-Glu-Gly-Gly-Gly-Arg-NHMe (EGGGR). For the conformational search, a series of typical starting structures was constructed: α-helical, 3_{10}-helical, distorted-β-strand, and strongly folded structures. In each case, both the canonical and the salt-bridge isomers were examined. A total of 17 structures were initially optimized at the HF/6-31G(d) level and reoptimized at the B3LYP/6-31+G(d) level. The energetics were verified at the MP2/6-311+G(2d,2p) level, without significant modifications to the lowest-energy structures. Finally, five EAFAR structures were rebuilt from the five most stable EGGGR structures and reoptimized at the B3LYP/6-31+G(d) level. Their vibrational frequencies were computed at the same level.

The five most stable structures of EAFAR (up to 77 kJ mol^{-1}) are pre-

sented in Figure 2. Interestingly, all α-helical starting structures collapsed to 3_{10} helices during optimization, which is consistent with the findings of Mons and co-workers.[11] The three most stable structures are zwitterionic (ZW-1, ZW-2, and ZW-3 in Figure 2), whereas the two high-energy conformers, which lie at or above 50 kJ mol^{-1}, have a neutral, canonical structure (N-4 and N-5 in Figure 2). The lowest-energy structure is a 3_{10} helix with two H bonds (CO···HN) that stabilize a salt-bridge structure between the carboxylate and guanidinium groups. The structure of second-lowest energy is another 3_{10} helix stabilized by a single H bond (CO···HN). A compact structure consisting of three consecutive turns (γ, γ, β) with two H bonds is found at higher energy. Structure N-4 is the lowest-energy canonical form. It has a 3_{10} conformation and exhibits an H-bond (COH···N) motif between the carboxylic acid OH group and the guanidino functionality of the canonical Glu and Arg side chains. Thus, this structure is the canonical counterpart of structure ZW-2. The last conformer, N-5, has a canonical compact structure. It contains an H bond of the type CO···HN between the carbonyl oxygen atom of the Glu side chain and the Arg side chain (see the Supporting Information for further discussion of this particular binding motif).

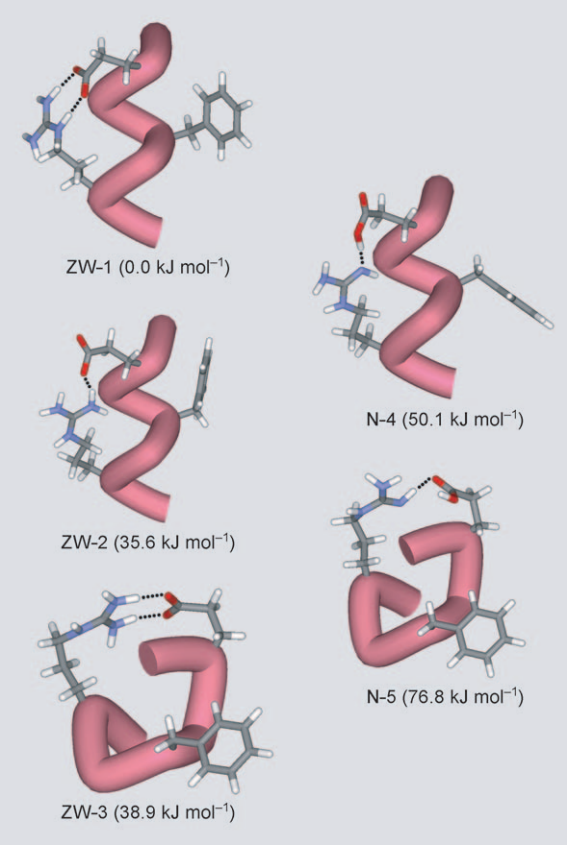


Figure 2. Low-energy conformations of EAFAR. Relative MP2 energies are given in brackets. H bonds involving Arg and Glu side chains are indicated with dotted lines.

The computed spectra of the five EAFAR structures described above are shown in Figure 3 in comparison with the experimental spectrum (top). The spectrum of the lowest-energy structure ZW-1 is overall in very good agreement with the experimental spectrum. The following features of the spectra deserve particular attention: 1) The free carboxylic acid C=O stretching band, observed at 1760 cm^{-1} in the spectrum of N-4, is absent from the experimental spectrum as well as from all calculated spectra of zwitterionic species. 2) The amide I feature in the 1640–1700 cm^{-1} region arises from unresolved combinations of the six local amide I bands. For the zwitterions, additional modes involving vibrations of the protonated arginine side chain are observed in the same region (1680–1730 cm^{-1}). These modes result in a broadening of the amide I peak of ZW-1, which is consistent with the experimental observations. 3) All calculated structures exhibit a series of delocalized backbone vibrations between 1420 and 1480 cm^{-1}, which account for the distinct feature observed at 1450 cm^{-1} in the experimental spectrum. 4) The specific vibrational modes of carboxylate ions are known to

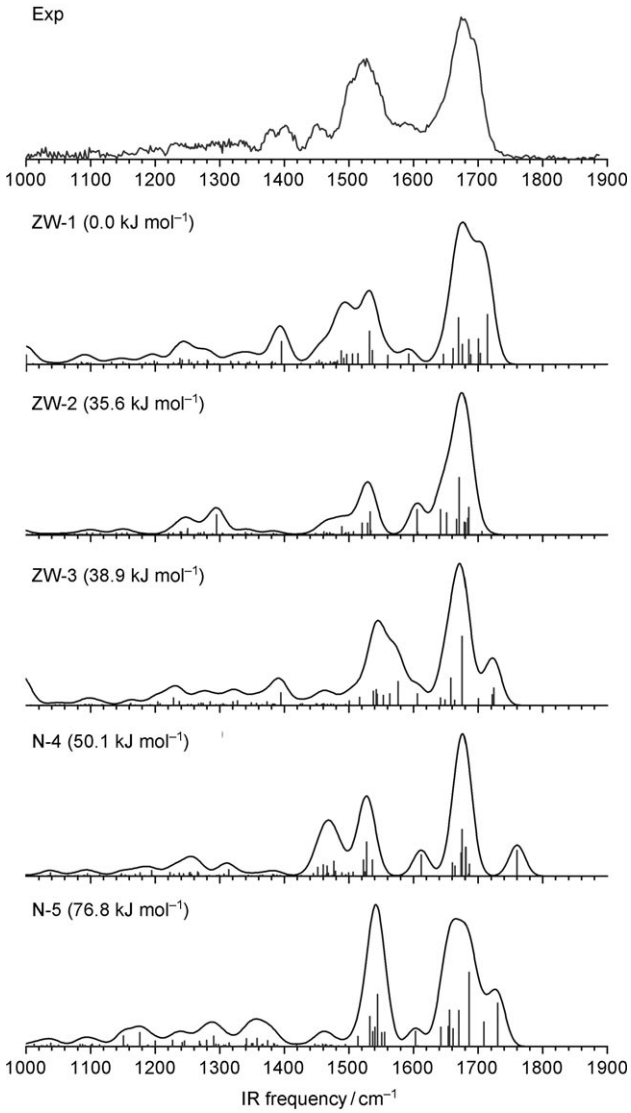


Figure 3. Comparison of the IR spectrum of the capped EAFAR peptide with the calculated IR spectra (scaling factor: 0.97) of the low-energy conformers.

be very sensitive to the environment[17] and thus constitute a promising probe of the molecular arrangement of the zwitterions. In the present case, for ZW-1, ZW-2 and ZW-3, the COO^- asymmetric modes range from 1560 to 1660 cm^{-1} and correspond to a mildly active region of the experimental spectrum. However, as the canonical forms exhibit pure arginine-side-chain modes in the same region, carboxylate asymmetric stretch modes can not be used for direct structural probing. 5) In contrast, the symmetric carboxylate mode appears to offer a diagnostic probe of the H-bond motif between the Glu and Arg side chains. Indeed, in the case of a single H bond (in ZW-2), a significant band appears around 1300 cm^{-1}; this band is blue-shifted to 1390 cm^{-1} when two H bonds are formed between the two side chains (in ZW-1 and ZW-3). The presence of this band at 1390 cm^{-1} in the experimental spectrum confirms the existence of the low-energy 3_{10}-helix structure stabilized by two H bonds.

In conclusion, we have demonstrated that "autozwitterionization" can occur in the local environment of a neutral, isolated peptide without additional interactions with the external biological environment. The transition from the canonical form of individual amino acids to the observed zwitterionic peptide is initiated by internal proton transfer between glutamic acid and arginine residues. The symmetric carboxylate stretching vibration can serve as a complementary probe that overcomes the intrinsic limitations that arise from the congestion of the amide bands in the spectra of large peptides. This diagnostic probe may facilitate the exploration of zwitterionic structures in neutral, isolated peptides.

Received: November 18, 2009
Published online: February 28, 2010

Keywords: conformation analysis · hydrogen transfer · IR–UV spectroscopy · neutral zwitterions · peptide design

[1] a) J. P. Simons, *Phys. Chem. Chem. Phys.* **2004**, *6*, E7; b) M. S. de Vries, P. Hobza, *Annu. Rev. Phys. Chem.* **2007**, *58*, 585.
[2] a) R. D. Suenram, F. J. Lovas, *J. Mol. Spectrosc.* **1978**, *72*, 372; b) C. J. Chapo, J. B. Paul, R. A. Provencal, K. Roth, R. J. Saykally, *J. Am. Chem. Soc.* **1998**, *120*, 12956.
[3] R. Antoine, M. Broyer, P. Dugourd, G. Breaux, F. C. Hagemeister, D. Pippen, R. R. Hudgins, M. F. Jarrold, *J. Am. Chem. Soc.* **2003**, *125*, 8996.
[4] P. D. Schnier, W. D. Price, R. A. Jockusch, E. R. Williams, *J. Am. Chem. Soc.* **1996**, *118*, 7178.
[5] a) S. J. Xu, W. J. Zheng, D. Radisic, K. H. Bowen, *J. Chem. Phys.* **2005**, *122*, 091103; b) M. Gutowski, P. Skurski, J. Simons, *J. Am. Chem. Soc.* **2000**, *122*, 10159.
[6] a) J. M. Talley, B. A. Cerda, G. Ohanessian, C. Wesdemiotis, *Chem. Eur. J.* **2002**, *8*, 1377; b) N. C. Polfer, J. Oomens, *Mass Spectrom. Rev.* **2009**, *28*, 468; c) C. Kapota, J. Lemaire, P. Maître, G. Ohanessian, *J. Am. Chem. Soc.* **2004**, *126*, 1836; d) R. A. Jockusch, W. D. Price, E. R. Williams, *J. Phys. Chem. A* **1999**, *103*, 9266.
[7] R. C. Dunbar, N. C. Polfer, J. Oomens, *J. Am. Chem. Soc.* **2007**, *129*, 14562.
[8] a) R. A. Jockusch, A. S. Lemoff, E. R. Williams, *J. Am. Chem. Soc.* **2001**, *123*, 12255; b) S. J. Xu, M. Nilles, K. H. Bowen, *J. Chem. Phys.* **2003**, *119*, 10696; c) M. F. Bush, J. S. Prell, R. J. Saykally, E. R. Williams, *J. Am. Chem. Soc.* **2007**, *129*, 13544.
[9] M. N. Blom, I. Compagnon, N. C. Polfer, G. von Helden, G. Meijer, S. Suhai, B. Paizs, J. Oomens, *J. Phys. Chem. A* **2007**, *111*, 7309.
[10] The existence of glutamic acid derivatives in their canonical form in the gas phase was shown in this study (Figure 1 b,c).
[11] W. Chin, F. Piuzzi, J. P. Dognon, L. Dimicoli, B. Tardivel, M. Mons, *J. Am. Chem. Soc.* **2005**, *127*, 11900.
[12] R. J. Lipert, S. D. Colson, *J. Chem. Phys.* **1988**, *89*, 4579.
[13] W. Chin, F. Piuzzi, I. Dimicoli, M. Mons, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033.
[14] a) A. M. Rijs, I. Compagnon, J. Oomens, J. S. Hannam, D. A. Leigh, W. J. Buma, *J. Am. Chem. Soc.* **2009**, *131*, 2428; b) A. Abo-Riziq, B. O. Crews, M. P. Callahan, L. Grace, M. S. de Vries, *Angew. Chem.* **2006**, *118*, 5290; *Angew. Chem. Int. Ed.* **2006**, *45*, 5166; c) J. M. Bakker, C. Plützer, I. Hünig, T. Häber, I. Compagnon, G. von Helden, G. Meijer, K. Kleinermanns, *ChemPhysChem* **2005**, *6*, 120.
[15] D. Oepts, A. F. G. van der Meer, P. W. van Amersfoort, *Infrared Phys. Technol.* **1995**, *36*, 297.
[16] I. Compagnon, J. Oomens, J. Bakker, G. Meijer, G. von Helden, *Phys. Chem. Chem. Phys.* **2005**, *7*, 13.
[17] J. Oomens, J. D. Steill, *J. Phys. Chem. A* **2008**, *112*, 3281.

