The parent anion of the RGD tripeptide: Photoelectron spectroscopy and quantum chemistry calculations

Xiang Li,[1] Haopeng Wang,[1] Kit H. Bowen,[1,a)] G. Grégoire,[2] F. Lecomte,[2] Jean-Pierre Schermann,[2,a)] and Charles Desfrançois[2,a)]
[1]*Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, USA*
[2]*Laboratoire de Physique des Lasers, Institut Galilée, Université Paris13, UMR7538 CNRS, F-93430 Villetaneuse, France*

(Received 11 February 2009; accepted 26 April 2009; published online 1 June 2009)

The gas-phase conformation of the intact (parent) unprotected RGD^- peptide anion has been investigated using a combination of anion photoelectron spectroscopy and quantum chemistry calculations of its low-energy stable structures. The experimentally observed RGD^- species correspond to a conformation in which the guanidinium group is protonated, the C-terminus is neutral, the aspartic acid carboxyl is deprotonated, and the anion's excess electron orbital is localized on the protonated guanidinium. This structure is reminiscent of the RGD loop, which is the peptide motif recognized by *trans*-membrane integrins. The parent RGD^- radical anion was generated using a unique infrared desorption-photoemission-helium jet ion source, whose ability to produce radical anions of peptides may also have analytical mass spectrometric implications.
 [DOI: 10.1063/1.3137095]

I. INTRODUCTION

The search for new drugs benefits from the application of powerful methods such as x-ray crystallography, nuclear magnetic resonance (NMR), and electron imaging, all of which are capable of handling large biomolecular systems such as proteins in condensed phase environments. In gas phase (free) environments, however, electrospray and MALDI mass spectrometry have become mainstays of proteomics. While mass spectrometric methods cannot determine structural conformations in such systems alone, they can provide insight into the intrinsic structural and even reactive properties of crucial molecular regions that have been identified as being responsible for specific bioactivities and thus may be targets for structure-based drug design. For example, the presence of a specific terminal monosaccharide, such as L-fucose in glycoproteins or glycolipids, plays a fundamental role in blood groups, and the conformational space of L-fucose has been recently studied in the gas phase.[1] Another example of a small size molecular system is the peptide RGD sequence considered here. The RGD (L-arginyl-glycil-L-aspartic acid or arg-gly-asp) amino acid sequence and integrins, which serve as its receptors and transmit signals across cell membranes, constitute a major recognition system for cell adhesion. Integrins are *trans*-membrane proteins linking the extracellular matrix to the intracellular cytoskeleton of cells. They specifically recognize the short motif, RGD peptide, considered herein.[2] Each amino acid of the RGD sequence plays a role in the recognition process. Arginine (R) interacts with two aspartic acids (D) in one unit (α) of the integrins, while aspartic acid interacts with a metal cation in the other unit (β). Glycine (G) establishes weak $C-H\cdots O=C$ hydrogen bonds.[3] This specific binding can initiate cell-signaling processes and control several processes including cell proliferation, cellular transduction, and tumorgenesis. The integrin RGD binding site is thus a pharmaceutical target.[4]

The RGD peptide itself and peptides containing the RGD sequence have been the subject of theoretical[5,6] and mass-spectrometric[7] studies. The solution structure of a snake venom protein, γ-bungarotoxin, containing the RGD sequence has been determined by means of NMR.[8] As in several other proteins, the RGD sequence is located at the apex of a flexible loop.[9] The crystal structure of an integrin extracellular segment, complexed with a RGD-containing ligand, has been determined by x-ray crystallography.[10] It was observed that the overall structure is not appreciably modified by the ligand binding. This fact is attributed to the crystal constraints, which do not allow large rearrangements. The NMR study emphasizes the absence of rigidity and the x-ray study suggests that the removal of any constraint in a gas phase study may be of interest.

It has been shown in structural analyses of γ-bungarotoxin[11] and other RGD-containing proteins that the amino acid residues flanking the RGD motif control the conformation of the RGD loop.[11] Thus, it is interesting to perform structural studies of the RGD peptide alone in order to determine its own intrinsic properties in the absence of any external influence, still keeping in mind that unprotected species allow for the possible formation of intramolecular bonds that may not exist when the RGD sequence is embedded in a larger peptide. We recently investigated the gas-phase structures of the protonated unprotected RGD peptide by means of mass selection and infrared multiphoton dissociation (IRMPD).[12] Depending on the proton location, four main

[a)]Authors to whom correspondence should be addressed. Electronic addresses: kbowen@jhu.edu, scherman@lpl.univ-paris13.fr, and desfranc@lpl.univ-paris13.fr.

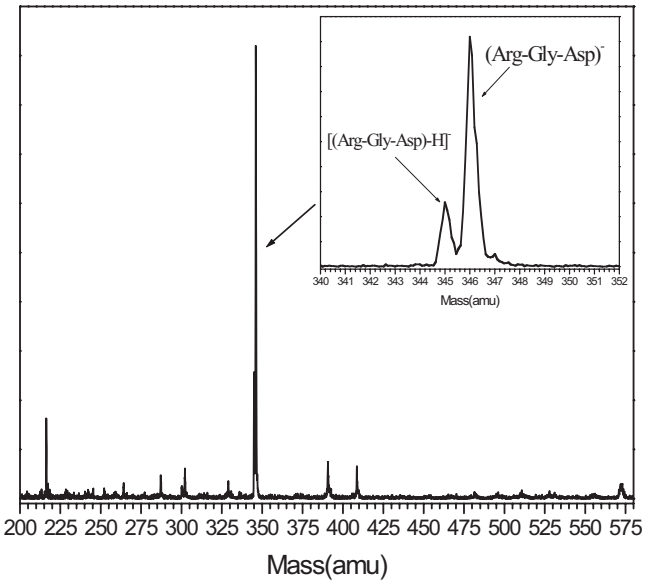

FIG. 1. Negative ion mass spectrum of RGD tripeptide obtained from our laser-desorption/photoemission supersonic ion source. The main peak corresponds to the parent intact RGD^- anion, at mass 346 amu, while a weaker peak corresponds to the deprotonated RGD species.

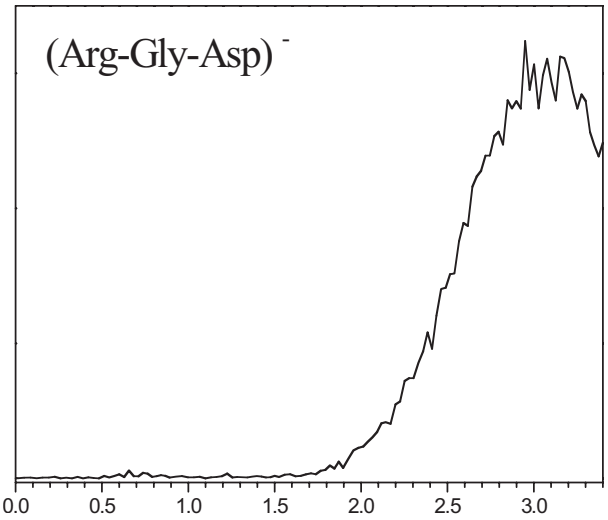

FIG. 2. Photoelectron spectrum of the parent intact RGD^- radical anion recorded with 3.493 eV detaching photons. The peak maximum at ~3 eV corresponds to the vertical detachment energy VDE. The spectral onset occurs at ~1.9 eV.

conformer families were calculated although only one was experimentally observed at 300 K. In the present case we studied the intact (parent) RGD radical anion, i.e., without either protonation or deprotonation. It is probably more subtle since the excess electron orbital may be localized in the vicinity of a positive charge or partially delocalized over an extended region. Moreover, to our knowledge, gas-phase structures of acidic amino acid residues have not been reported. Peptides containing acidic amino acids (glutamic acid, E, and aspartic acid, D) play an important role in molecular recognition among immune system[13] and have, until now, received little attention in gas-phase studies.[14]

Here, we report the formation of the gas phase RGD^- parent anion, the measurement of its anion photoelectron spectrum, and the calculation of its low energy geometric structures. Synergy between theory and experiment provides insight into the structural conformation of the RGD^- anion in the gas phase (in isolation).

II. METHODS AND RESULTS

A. Experimental

We here used a source for generating intact (parent) anions of involatile molecules. This source previously allowed us to obtain intact nucleobase,[15] nucleoside,[16] and nucleotide[17] anions in the gas phase. We produced for the first time the intact tripeptide molecular parent anions. Figure 1 shows its negative ion mass spectrum with a strong intensity of the intact RGD^- anion and a relatively weaker intensity fragment anion, which has lost a hydrogen atom from the parent anion. Among the attributes of the mass spectrum of this species is its remarkable "cleanliness." Several other fragment anions of RGD were also observed, although at much weaker intensities.

This source solves the longstanding problem of producing intact (parent) anions of biomolecules in the gas phase without fragmenting them. Details of this source have been published elsewhere.[16] Briefly, low-power infrared laser pulses (1064 nm, 1.17 eV/photon) from a neodymium-doped yttrium aluminum garnet (Nd:YAG) laser were used to desorb neutral biomolecules from a slowly moving graphite bar, which had been thinly coated with sample. Almost simultaneously, electrons were generated by visible laser pulses (another Nd:YAG laser operated at 532 nm, 2.33 eV/photon) striking a rotating yttrium oxide disk. Since yttria's work function of ~2 eV is slightly below the photon energy of the visible laser, low energy electrons were produced,[18] and this process is critical to the formation of intact biomolecular ions. At the same time a pulsed gas valve provides a collisionally cooling jet of helium to carry away excess energy and stabilize the resulting parent anions. This novel source provides a pathway for generating parent radical anions of peptides and this development may be of interest in analytical mass spectrometry, where ion generation without the need to rupture bonds is an advantage.

The photoelectron spectrum was recorded by crossing a beam of RGD^- parent anions with a fixed-frequency photon beam (355 nm or 3.493 eV/photon). The resultant photodetached electrons were energy analyzed using a magnetic bottle energy analyzer with a resolution of 35 meV at EKE=1 eV. Photodetachment of electrons is governed by the energy-conserving relationship, $h\nu$=EBE+EKE, where EBE is the electron binding (transition) energy, EKE is the measured electron kinetic energy, and $h\nu$ is the photon energy.

The photoelectron spectrum of the RGD anion is presented in Fig. 2. It exhibits a broadband with an onset at ~1.9 eV and a peak maximum at 3.0 eV. The broadness of this band suggests significant structural differences between the equilibrium structure of RGD and its anion RGD^-. The peak maximum provides the vertical detachment energy (VDE), which is the energy difference between the anion and its neutral counterpart at the equilibrium geometry of the anion. Thus, the measured VDE of RGD is 3.0 eV. The onset

TABLE I. Relative energies (ΔE), VDEs, and "local" adiabatic electron affinities (EA_a) at three different levels of fully optimized calculations. All values are in eV (1 eV=23.06 kcal/mol=96.5 kJ/mol).

Anion structures	G1	G2	G3	P1	P2
ΔE (AM1)	0	0.36	0.59	1.11	1.21
ΔE (B3LYP/6-31+G*)	0	0.14	0.23	0.31	0.65
ΔE (RI-MP2/def-TZVP)	0	0.10	0.35	0.18	0.47
VDE (AM1)	3.78	3.59	4.21	2.16	1.80
VDE (B3LYP/6-31+G*)	2.74	2.96	3.76	2.74	1.23
VDE (RI-MP2/def-TZVP)	2.76	2.80	3.97	2.67	1.66
EA_a (B3LYP/6-31+G*)	0.59	0.43	0.30	0.46	0.33

of the photoelectron signal may reflect the adiabatic electron affinity (EA_a), which is the energy difference between the ground vibronic state of the neutral and that of its corresponding anion. However, if there was to be poor Franck–Condon overlap between the lowest vibrational levels of the anion and its corresponding neutral, the actual EA_a value may be lower than the onset at 1.9 eV. On the other hand, if vibrational hot bands in the anion were to come into play, the actual EA_a value might lie at an EBE value, which is higher than the onset.

B. Computational

In order to interpret our experimental photoelectron spectrum, we performed quantum chemistry calculations of the possible low-energy isomers of the RGD^- peptide anion. As a starting point, we considered the low-energy conformers of the protonated species that we previously studied.[12] From those protonated RGD structures, we removed one proton, either on the N-terminus, the guanidinium, or aspartic acid residue, or on the C-terminus. We added an extra electron on these neutral structures and then fully optimized those anions at the semiempirical AM1 level, using Hyperchem.[19] During these optimizations, proton transfer sometimes occurred between the guanidine side chain and the carboxylic group of either the acid aspartic residue or the C-terminus. From this first stage of calculations, we obtained only two sets of conformers: one family (G anions) in which the excess electron is located on the protonated guanidinium residue, one carboxylic group being deprotonated, and the other family (P anions) in which the extra electron lies on the peptide chain, all residues and termini being neutral. The G anion structures appeared to be about 1 eV more stable than the P ones and their VDE values were found to be about 3.8 eV, compared with about 2 eV for the P anions.

The three lowest-energy structures of the G anion family (G1, G2, and G3) and the two lowest-energy structures of the P anion family (P1, P2) were then fully optimized at the B3LYP/6-31+G* level, using the GAUSSIAN 03 suite of programs.[20] In order to obtain the VDE, the energies of the corresponding neutrals were calculated at the geometries of the ground state anion. Next, the neutrals were also fully optimized so that the energies of the corresponding neutral equilibrium structures can be obtained. The energy differences between the anion and the local neutral equilibrium structures lead to values of the "local" adiabatic electron affinities. Those values are only the estimates of the true absolute adiabatic electron affinity since we did not perform a systematic search for the absolute minima neither for the anion nor for the neutral RGD structures. All computational results for the relative anion energies (ΔE), the VDE, and the local EA_a are displayed in Table I. During the second stage of calculations, the resulting equilibrium geometries did not evolve very much, except for the P1 conformer in which a proton transfer occurred between the C-terminus carboxylic acid and the neighboring peptide chain carbonyl on which the excess electron localized. This conformer was not found in AM1 calculations, but it led here to a more stable anion and to a larger VDE, as compared to the P2 conformer in which all residues and termini remain neutrals.

In order to ascertain the energetic values, we also performed resolution-of-identity second-order Møller-Plesset method (RI-MP2) calculations with triple zeta plus polarization (def-TZVP) basis sets on the C, N, and O atoms and with double zeta plus polarization (def-SVP) basis sets on the H atoms, using the TURBOMOLE suite of programs.[21–26] Those basis sets were used because they are similar to Pople's 6-311G* and 6-31G* basis sets, while the 6-31+G* basis set is not available in the TURBOMOLE package. Again, full optimizations of the anions were performed starting from the B3LYP equilibrium conformers. At this third stage of calculations, no important structural changes occurred during the optimizations. Figures 3 and 4 display the highest occupied molecular orbital (HOMO) of the G1 and G3 and P1 and P2 anions calculated at the RI-MP2 level, respectively.

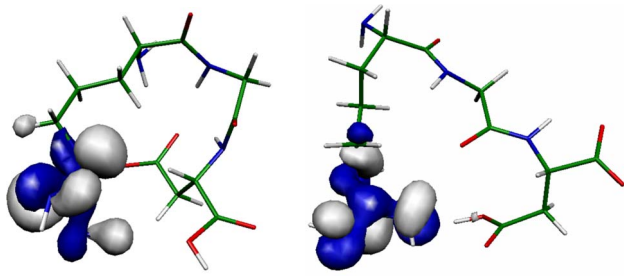

FIG. 3. (Color online) HOMO of the G1 (left) and G3 (right) anion structures calculated at the RI-MP2 level. The excess electron is localized on the guanidinium side of the protonated arginine residue, while the C-terminus is neutral and the aspartic acid residue is deprotonated (G1) or vice versa (G3). The guanidinium group is not any more planar. The calculated VDE is about 2.75 eV for G1, in close agreement with the experimental value of 3 eV, while it is about 3.9 eV for G3, probably because this structure is more far from the neutral equilibrium structures that are more folded.

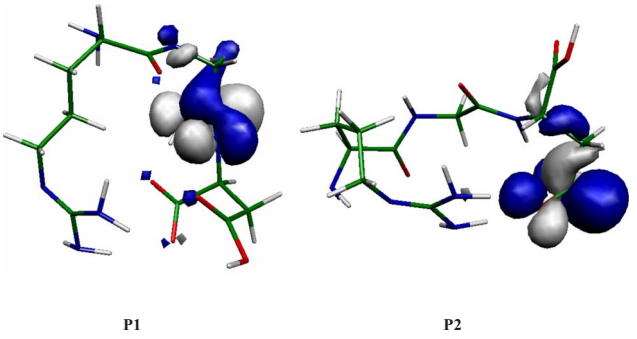

FIG. 4. (Color online) HOMO of the P1 (left) and P2 (right) anion structures calculated at the RI-MP2 level. The excess electron is localized either on the second protonated carbonyl group of the peptide chain (P1) or on the aspartic acid neutral residue (P2). The calculated VDE is then about 2.7 eV for P1, i.e., comparable to values of the G conformers, but only 1.2 eV (B3LYP) or 1.7 eV (RI-MP2) for P2, because the extra electron is now bound to a neutral group rather than to a protonated one.

As can be seen from Table I, the AM1 results lead to larger ΔE and VDE values as compared to B3LYP and RI-MP2 results, which themselves are surprisingly similar. As expected, the conformers that correspond to an excess electron bound to a protonated group, either on guanidine (G1, G2, G3) or on a peptide chain carbonyl (P1), possess a larger VDE as compared to the conformer (P2), in which the extra electron is only bound to a neutral (acid aspartic) group. Both G conformers possess similar HOMOs. G1 (shown in Fig. 3 and also in the supplementary material[27]) and G2 (data given in the supplementary material[27]) structures are similar, with a deprotonated acid aspartic residue that interacts strongly with several N–H bonds. On the other hand, G3 (Fig. 3) is less folded and its deprotonated C-terminus points away from the guanidine group, leading to less internal H-bonding between C=O and N–H groups. This is probably the reason why its VDE is larger, because this geometry differs more from its calculated corresponding neutral geometry, which itself is more folded. We also note that, as the result of electron binding, the guanidinium group of the G conformers and the amide group of the P1 conformer are no longer planar, with the carbon atom becoming slightly tetrahedral.

III. DISCUSSION

Given the present experimental anion formation conditions (laser desorption and attachment of low-energy photoemitted electrons in a cooling supersonic beam), it is expected that the anion population is comparable to a thermal distribution at a temperature of 200–300 K. Although we did not calculate the relative free energies, the total energies calculated here, together with the measured VDE value, provide insight into the possible conformers that are observed in the experiment. Due to its high relative energy (0.5–0.6 eV) and low VDE (1.2–1.7 eV), conformer P2 must be discarded. If conformer G3, whose VDE is about 4 eV, would have been noticeably populated in the beam, the EBE maximum in our photoelectron spectrum would not have occurred within our experimental energy window, which it did. This is consistent with its high relative energy. With about the same relative energy of 0.2–0.3 eV, conformer P1 is then also unlikely to be noticeably present, although its VDE fits well with the experimental value. Thus only the G1 and G2 conformers, whose relative energies are about two times lower, i.e., 0.1–0.15 eV, remain viable candidates. They both possess VDE values that are in good agreement with the experimental data and they may both contribute to the experimental spectrum. The corresponding calculated "local" electron affinities, i.e., 0.4–0.6 eV, are much lower than the observed threshold in the PES spectrum, 1.9 eV, but, as outlined above, this may be due to the large structural changes between the anions and the neutrals and poor Franck–Condon overlap.

We again emphasize that we did not try to systematically explore the full anion potential energy surface, so that other low-energy conformers may exist and may contribute to the experimental spectrum. However, the present results suggest that the observed low-energy RGD^- anion conformers correspond to structures in which the anion excess electron orbital is localized on the protonated guanidinium, with the C-terminus being neutral and the aspartic acid carboxyl being deprotonated.

In addition, the RGD sequence exhibits a large propensity for the formation of a loop when it is embedded in proteins.[28] This secondary structure is crucial for the recognition of the RGD sequence by integrins. For example, the flexible RGD loop structure[11] in dendroaspin (PDB entry 1DRS) corresponds to residues 43–45 and is located on the surface, maintained by disulfide bridges.[29] A quantitative structure-activity relationship (QSAR) study of RGD-containing peptides has been conducted in a platelet aggregation assay using NMR for structural determinations.[30] This study stressed that the major QSAR criteria are the respective positions of the two major recognition sites, i.e., the charged sidechains of arginine and aspartic acid. Those positions are defined as the distance between the respective C^β atoms of R and D and the pseudodihedral defining the R and D sidechain orientation. This angle is given by the respective positions of the C^ζ and C^α atoms of arginine and the C^α and C^δ atoms of aspartic acid. Platelet aggregation representing biological activity is inhibited only if this pseudodihedral angle is comprised in between −45° and +45°. Once this criterium is fulfilled, biological activity increases when the distance between the charged centers and/or between the C^β atoms of R and D increases within a range in between 4.4

TABLE II. Calculated (RI-MP2) distance and angle parameters, relevant for bioactivity of the RGD sequence, as a function of the different anion conformers. See text for definition and discussion.

RGD conformation	G1	G2	G3	P1	P2
Distance between the C^β atoms (Å)	4.568	5.65	7.74	6.99	7.852
Pseudodihedral angle (deg)	78	−20	11	−69	44

and 9 Å. As shown in Table II, the distances between the respective C^{β} atoms of R and D in the five studied configurations of the gas-phase RGD anion agree with the range of distances measured in the studied bioactive compounds of reference.[30] However, while the pseudodihedral angle criterion is fulfilled for configurations G2, G3 and P2, only the biologically relevant conformations G1 and G2 were observed experimentally. It is interesting to note that although the studied gas-phase anions are not constrained by any protein backbone, a sizeable number of their configurations retain structures compatible with bioactivity.

ACKNOWLEDGMENTS

This material is based in part (KHB) upon work supported by the U.S. National Science Foundation under Grant No. CHE-0809258.

[1] P. Çarçabal, T. Patsias, I. Hünig, B. Liu, C. Kaposta, L. C. Snoek, D. P. Gamblin, B. G. Davis, and J. P. Simons, Phys. Chem. Chem. Phys. **8**, 129 (2006).
[2] K. E. Gottschalk and H. Kessler, Angew. Chem. Int. Ed. **41**, 3767 (2002).
[3] J. Bella and M. J. Humphries, BMC Struct. Biol. **5**, 4 (2005).
[4] M. D. Pierschbacher and E. Ruoslathi, J. Biochem. **262**, 17294 (1987); J. Takagi, Biochem. Soc. Trans. **32**, 403 (2004).
[5] F. D. Suvire, A. M. Rodriguez, M. L. Mak, J. G. Papp, and R. D. Enriz, J. Mol. Struct.: THEOCHEM **540**, 257 (2001).
[6] J. C. P. Koo, G. A. Chass, A. Perczel, O. Farkas, L. L. Torday, A. Varro, J. G. Papp, and I. G. Csizmadia, J. Phys. Chem. A **106**, 6999 (2002).
[7] T. Solouki, R. C. Fort, A. Alomary, and A. Fattahi, J. Am. Soc. Mass Spectrom. **12**, 1272 (2001).
[8] M. J. Sutcliffe, M. Jaseja, E. I. Hyde, X. J. Lu, and J. A. Williams, Nat. Struct. Biol. **1**, 802 (1994).
[9] M. A. McLane, S. VijayKumar, C. Marcinkiewicz, J. J. Calvete, and S. Niewiarowski, FEBS Lett. **391**, 139 (1996).
[10] J. P. Xiong, T. Stehle, B. Diefenbach, R. G. Zhang, R. Dunker, D. L. Scott, A. Joachimiak, S. L. Goodman, and M. A. Arnaout, Science **294**, 339 (2001).
[11] J. H. Shiu, C. Y. Chen, L. S. Chang, Y. C. Chen, Y. H. Lo, Y. C. Liu, and W. J. Chuang, Proteins: Struct., Funct., Bioinf. **57**, 839 (2004).
[12] G. Grégoire, M. P. Gaigeot, D. C. Marinica, J. Lemaire, J. P. Schermann, and C. Desfrançois, Phys. Chem. Chem. Phys. **9**, 3082 (2007).
[13] H. L. Ploegh, J. Clin. Invest. **115**, 2077 (2005).
[14] M. E. Sanz, A. Lessari, J. C. Lopez, and J. L. Alonso (unpublished).
[15] M. Haranczyk, M. Gutowski, X. Li, and K. H. Bowen, Proc. Natl. Acad. Sci. U.S.A. **104**, 4804 (2007); J. Phys. Chem. B **111**, 51 (2007); X. Li, K. H. Bowen, M. Haranczyk, R. A. Bachorz, K. Mazurkiewicz, J. Rak, and M. Gutowski, J. Chem. Phys. **127**, 174309 (2007).
[16] S. T. Stokes, X. Li, A. Grubisic, Y. J. Ko, and K. H. Bowen, J. Chem. Phys. **127**, 084321 (2007).
[17] S. T. Stokes, A. Grubisic, X. Li, Y.-J. Ko, and K. H. Bowen, J. Chem. Phys. **128**, 044314 (2008).
[18] M. Mitsui and A. Nakajima, J. Chem. Phys. **117**, 9740 (2002).
[19] HYPERCHEM 7.0, Gainesville, FL, 2005.
[20] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Revision B.04, Gaussian Inc., Pittsburgh, PA, 2003.
[21] R. Ahlrichs, B. M. Haser, M. H. Horn, and C. Kolmel, Chem. Phys. Lett. **162**, 165 (1989).
[22] F. Weigend and M. Haser, Theor. Chem. Acc. **97**, 331 (1997).
[23] C. Hattig, A. Hellweg, and A. Kohn, Phys. Chem. Chem. Phys. **8**, 1159 (2006).
[24] A. Schaefer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).
[25] F. Weigend, M. Haser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).
[26] A. Schafer, H. Horn, and R. Ahlrichs, J. Chem. Phys. **97**, 2571 (1992).
[27] See EPAPS Document No. E-JCPSA6-130-012921 for optimized coordinates and structures of the G1 and G2 conformers at the B3LYP/6-31 +G* level. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[28] I. Y. Torshin, Med. Sci. Monit. **8**, BR301 (2002).
[29] D. Monleon, V. Esteve, H. Kovacs, J. J. Calvete, and B. Celda, Biochem. J. **387**, 57 (2005).
[30] S. Kostidis, A. Stavrakoudis, N. Biris, D. Tsoukatos, C. Sakarellos, and V. Tsikaris, J. Pept. Sci. **10**, 494 (2004).