Indian J. Phys. **76B** (4), 577–580 (2002)

I J P B
— an international journal

Correlation effects on peptide molecules

Puspitapallab Chaudhuri and Sylvio Canuto
Instituto de Fisica, Universidade de São Paulo, CP 66318, 05315-970 São Paulo, SP, Brazil
E-mail : puspita@ift.unesp.br

Abstract : *Ab initio* methods are used to study the four possible dipeptides that can be formed from alanine and glycine in gas phase. We calculate the geometry, binding energy and rotational spectra of alanylalanine, alanylglycine, glycylalanine and glycylglycine. The geometries are optimized using Hartree-Fock (HF) and second-order Moller-Plesset perturbation theory (MP2). Single-point energy calculations with high-level CCSD(T) are performed on MP2 geometries. Both electron correlation and zero-point vibrational energy corrections are found to be very important and increase the calculated binding energy. Most of the correlation effects are obtained in second-order. The inclusion of zero-point correction is found to be essential.

Keywords : Correlation effects, peptide molecules.

PACS No. : 33.15.Ry

1. Introduction

Peptide bonds are of enormous scientific interest as it is the primary requirement for the structure of all the proteins and consequently, for the origin and continuation of life. With the recent detection of more than 100 species of amino acids in interstellar space [1] and with the continuing searches for more, there is always a possibility to find proteins in the cool interstellar space. In this situation the study of peptide bond formation from neutral amino acids becomes of interest.

The formation of a peptide bond is an example of a condensation reaction. The two molecules of amino acids join covalently to form the bond together with the accompanying removal of a molecule of water. Since glycine and alanine are the two most simple amino acids, the peptide bonds involving alanine or glycine are the most widely studied ones by the theoreticians [2-6]. In the present paper we investigate the structure and energetics of the four peptide molecules that may be formed from the mixture of alanine and glycine in gas phase. When they split out water to form a dipeptide, either could be the donor for either a hydrogen or a hydroxyl group and hence, theoretically, it gives rise to four possibilities for the product (see Figure 1a) :

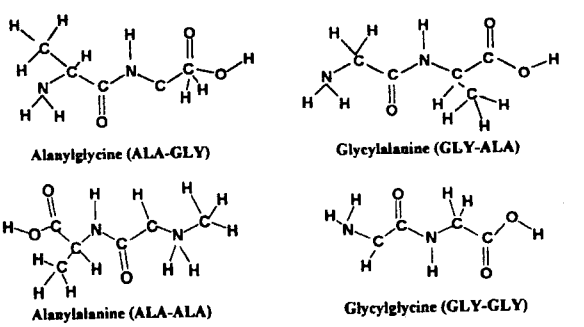

Figure 1a. The four peptide molecules considered in the present study.

(1) Alanyl glycine (ALA-GLY) :
CH_3-CH(NH_2)-CO-NH-CH_2-COOH

(2) Glycyl alanine (GLY-ALA) :
H-CH(NH_2)-CO-NH-CH (CH_3)-COOH

(3) Alanyl alanine (ALA-ALA) :
CH_3-CH(NH_2)-CO-NH-CH (CH_3)-COOH

(4) Glycyl glycine (GLY-GLY) :
H-CH(NH_2)-CO-NH-CH_2-COOH

We use high-level *ab initio* quantum mechanical methods to model all the four molecules mentioned above. The basic interest of the present study is to analyze the effect of higher-order electron correlation on the binding energy. However rotational properties are also considered.

2. Method of calculation

Fully unconstrained geometry optimizations are performed for ALA-GLY, GLY-ALA, ALA-ALA, GLY-GLY, Alanine, Glycine and water using Hartree-Fock (HF) and second-order Moller-Plesset theory (MP2) using gradient techniques [7] and the 6-31G* basis set. After obtaining the optimized geometry at the MP2 level, single-point calculations are performed using many-body perturbation/coupled-cluster theories (MBPT/CC) [8] with the same 6-31G* basis set. These include the more sophisticated coupled-cluster models such as CCSD [8] and CCSD(T) [9]. Thus at the highest level, geometry optimization of the seven molecules mentioned above are performed at the MP2/6-31G* level and single-point calculations are further made up to the CCSD(T)/6-31G* level. Comparison between these results allows a systematic analysis of the electron correlation effects on the binding energies. The binding energies for all the four peptides are calculated with all the above theoretical models, with and without the zero-point energy correction. Thus we can also see the effect of zero-point energy on the binding energies.

To make sure that all optimized geometries are true minimum the geometry optimizations are followed by a calculation of the hessian. These hessians are then used to calculated the harmonic vibrational frequencies are therefrom the zero-energy vibration. All the calculations are performed using the GAUSSIAN/98 program [10].

3. Results and discussions

The peptide bond structure :

It is well known [1] that both glycine and alanine can have several structural conformations depending on the internal rotational degrees of freedom. The present optimized structure of glycine corresponds closely to the structure IVn of Császár [11], whereas that of alanine corresponds to the structure IIb [12].

The Figure 1a illustrates the four peptide molecules considered in this paper. The bond lengths and angles of the peptide unit, C^{α}-CO-NH-C^{α}, in alanylglycine (ALA-GLY) as obtained by the two methods are shown in Figure 1b and in Table 1 where they are compared with the experimental parameters [13-15].

Table 1. Geometry of the peptide unit in the peptide molecule, ALA-GLY. Bond lengths are given in Angstrom and the bond angles in degree.

Geometry	HF/ 6-31G*	MP2/ 6-31G*	Ref. [13]	Ref. [14]	Ref. [15]
C-N	1.349	1.359	1.32	1.335	1.33
C^{α}-C	1.531	1.528	1.53	1.522	1.52
N-C^{α}	1.436	1.441	1.47	1.449	1.45
N-H	0.995	1.014	1.00	0.96	1.00
C-O	1.203	1.237	1.24	1.229	1.23
< C^{α}-C-N	116.09	116.26	114	116.6	116
< C^{α}-C-O	121.88	121.79	121	120.4	121
< N-C-O	122.03	122.05	125	122.9	123
< H-N-C^{α}	117.35	117.26	114	118.4	118
< C^{α}-N-C	121.28	120.55	123	121.9	122
< H-N-C	121.37	122.29	123	119.8	120

In comparing our results among themselves and with experimental values one should recall that experimental measurements on the structures of peptides are done in the condensed phase that contain both bonded and non-bonded interactions. However, comparing with the three sets of experimental results, we note an overall good agreement between theory and experiment. Both HF and MP2 produce almost the same values for the bond angles. However the HF bond lengths are generally smaller than that of MP2. The C-N bond length is always longer than that of experiment at all theoretical levels as has already been observed before [16].

Rotational spectra :

The rotational constants are very useful in the radio-astronomy identification of molecules in the interstellar space. The dipole moment, on the other hand, is important as it is related to the intensity of the rotational lines. These rotational constants can be evaluated directly from the moments of inertia. In Table 2 we compare the rotational constants and dipole moments obtained by HF and MP2 methods. We note that $A > B \cong C$ in all cases, indicating that all the four molecules can be represented by a near-prolate approximation in a rigid model.

The binding energy :

Table 3 shows the calculated binding energies (BE) obtained by these two methods along with the zero-point energy corrections to it. BE are calculated as a difference between the sum of the equilibrium energy of the final channel (peptide molecule + water) and that of the individual amino

Table 2. The rotational constants (A, B and C) in the units of GHz and dipole moments (□) in debye of the four peptide molecules, alanylglycine, glycylalanine, alanylalanine and glycylglycine.

	Alanylglycine (GLY-GLY)		Glycylalanine (GLY-ALA)		Alanylalanine (ALA-ALA)		Glycylglycine (GLY-GLY)	
	HF	MP2	HF	MP2	HF	MP2	HF	MP2
A	3.2463	3.1872	3.0393	3.0127	2.2159	2.2022	5.2066	4.9885
B	0.5771	0.5706	0.6404	0.6389	0.5622	0.5569	0.6722	0.6691
C	0.5291	0.5261	0.5709	0.5655	0.4914	0.4887	0.6015	0.5972
□	2.9188	3.1738	2.8092	3.0111	2.8633	3.0930	2.8595	3.0847

acid components in the initial channel. Table 3 shows that both electron correlation effects and zero-point vibrational corrections (ZPC) increase the BE The electron correlation effect obtained in second-order MP2 gives a substantial contribution to BE. Among the four molecules the effect is maximum in case of GLY-GLY where BE increases by 3.15 kcal/mol and minimum in ALA-ALA where the increase is only 0.64 kcal/mol. ZPC are found to be more important for GLY-GLY at the HF level but for ALA-ALA at the correlated MP2 level. Considering the MP2 results, the most bound system is GLY-GLY with a BE of 6.34 kcal/mol. Composite systems such ALA-GLY and GLY-ALA are intermediate between the ALA-ALA and GLY-GLY systems. At this stage it is appropriate to include higher level correlation corrections. Thus, single-point full fourth-order MP4 and CCSD(T) calculations are performed for all molecules in their MP2 geometries. The results are shown in Table 4. There are some interesting aspects related to the higher-order correlation contributions. First, all calculated BE are decreased at the CCSD(T) level compared to the previous MP2 results. Comparison between the results obtained at the MP4(SDQ) and MP4(DQ) shows that single excitations at the fourth order decrease the BE but comparison between MP4(SDQ) and CCSD shows that high-order single and double excitations increase BE for all systems. Additionally, the contribution of triple excitations in fourth-order, as seen in MP4(SDTQ), and higher-order as seen in CCSD(T), do not have a defined sign. They increase BE for GLY-GLY and GLY-ALA, decrease for ALA-ALA and have a minor effect for ALA-GLY. If one considers the highest level of calculation, CCSD(T), and ZPC, at the MP2 level, we find that the best result considered here gives BE of 4.86 kcal/mol, 5.09 kcal/mol, 5.61 kcal/mol and 5.89 kcal/mol for ALA-ALA, ALA-GLY, GLY-ALA and GLY-GLY, respectively.

Table 3. Calculated binding energies (kcal/mol) for the peptide molecules using the basis set 6-31G* at the HF and MP2 level. BE and ZPC denote the binding energy and the correction due to the difference in zero-point vibrational energies on the binding energy.

Molecule	BE(HF)	After ZPC	BE(MP2)	After ZPC
ALA-GLY	1.34	3.82	3.05	5.62
GLY-ALA	1.47	3.95	3.71	6.11
ALA-ALA	1.79	4.58	2.64	5.46
GLY-GLY	1.02	3.18	4.17	6.34

Table 4. Calculated binding energies using MP2 geometries for the peptide bonds, with 6-31G* basis set in different methods. ZPC is the calculated MP2 zero-point energy correction.

METHOD	ALA-GLY	GLY-ALA	ALA-ALA	GLY-GLY
MP2	3.05	3.71	2.64	4.17
MP2 + ZPC	5.62	6.11	5.46	6.34
MP3	2.96	3.52	2.59	3.92
MP4(D)	2.67	3.28	2.35	3.62
MP4(DQ)	2.64	3.18	2.48	3.36
MP4(SDQ)	2.34	2.94	1.91	3.15
MP4(SDTQ)	2.22	3.01	1.95	3.32
CCSD	2.51	3.08	2.37	3.25
CCSD(T)	2.51	3.21	2.03	3.72
CCSD(T) + ZPC	5.09	5.61	4.86	5.89

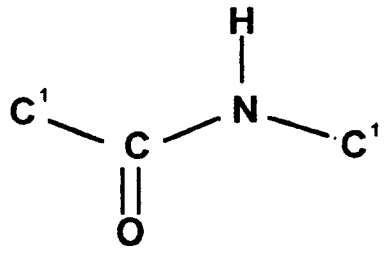

Figure 1b. The structure of the peptide unit taken from ALA-GLY to be compared with the experiments in Table 1. The symbol C^1 stands for the standard notation $C^{\square}$ used in the text and table.

4. Conclusions

The present *ab initio* calculations show the effects of electron correlation on the binding energy of the peptide molecules considering the four possible combination of

alanine and glycine to form dipeptides in gas phase. HF and MP2 methods are used to optimize geometries and to calculate the total energies and vibrational frequencies of the four molecules : ALA-GLY, GLY-ALA, ALA-ALA and GLY-GLY. The effect of electron correlation is found to be significant as it increases the binding energies nearly duplicating the HF value in all cases. Most of the correlation effect is obtained already in second-order but higher-order effects slightly decrease the binding energies (around 0.5 kcal/mol). The zero-point energy corrections are found to be crucial and they also increase the binding energy. Among the four molecules considered, GLY-GLY is found to have the highest binding energy (5.9 kcal/mol) while ALA-ALA has the lowest one (4.9 kcal/mol).

References

[1] A G Császár *Prog. in Biophys. and Mol. Biol.* **71** 243 (1999)

[2] W D Cornell, I R Gould and P A Kollman, *J. Mol. Struc. (Theochem)* **392** 101 (1997)

[3] I K Roterman *et al Biomol. Struc. Dynam.* 7 421 (1989)

[4] I R Gould, W D Cornell and I H Hillier *J. Am. Chem. Soc.* **116** 9250 (1994)

[5] T Head-Gordon *et al J. Am. Chem. Soc.* **113** 5989 (1991)

[6] H Bohm and S Brode *J. Am. Chem. Soc.* **113** 7129 (1991)

[7] H B Schlegel *J. Comp. Chem.* **3** 214 (1982)

[8] R J Bartlett *J. Phys. Chem.* **93** 1997 (1989)

[9] K Raghavchari *Ann. Rev. Phys. Chem.* **42** 615 (1991)

[10] M J Frisch *et al Gaussian 98, Revision A.7, Pittsburgh P A,* (1998)

[11] A G Császár *J. Am. Chem. Soc.* **114** 9568 (1992)

[12] A G Császár *J. Phys. Chem.* **100** 3541 (1995)

[13] R B Corey and L Pauling, *Proc. Roy. Soc.* (London) **B141** 10 (1953)

[14] E Benedetti *Peptides* eds. M Goodman and J Meinenhofer (New York : Wiley) (1977)

[15] R E Marsh and J Donohue *Adv. Protein Chem.* **22** 249 (1967)

[16] J N Scarsdale *et al J. Am. Chem. Soc.* **105** 3438 (1983)

This work has been partially supported by CNPq and FAPESP, Brazil.