J Mol Model (2004) 10:198–203
DOI 10.1007/s00894-004-0184-y

ORIGINAL PAPER

Paranthaman Selvarengan · Ponmalai Gounder Kolandaivel

Molecular modeling of dipeptide and its analogous systems with water

Received: 14 November 2003 / Accepted: 26 January 2004 / Published online: 16 March 2004

Abstract Intermolecular hydrogen-bond interactions in the monohydrated complexes of formamide, *N*-methylacetamide and glycylglycine have been studied using ab initio and DFT methods. The geometries were optimized using second-order Møller–Plesset perturbation theory and the B3LYP DFT functional with the 6-311++G** basis set. It is observed that hydrogen-bond interactions at the carbonyl group of the peptide moiety are stronger than those at the amino group of the formamide and *N*-methylacetamide molecules. Because of the presence of cyclic hydrogen-bonding interactions in glycylglycine, the interaction at the amino group is higher than at the carbonyl. The ^{13}C and ^{15}N NMR shielding values were calculated for the non-hydrated and monohydrated complexes. Condensed Fukui functions have also been calculated for non-hydrated formamide, *N*-methylacetamide and glycylglycine molecules at the B3LYP/6-311++G** level of theory, and the results are discussed.

Keywords Formamide · *N*-Methylacetamide · Glycylglycine · Condensed Fukui functions

Introduction

The interaction of water with biomolecules is of great interest in biology. The network of hydrogen bonds involving water molecules with amino acids stabilizes the structure of a protein. [1, 2] Water molecules bound to a protein molecule establish hydrogen bonds and act as donors to C=O groups, and acceptors to N–H groups of the peptide chain. These hydrogen bonds play an important role in the determination of structure and activity of amino acids in proteins.

P. Selvarengan · P. G. Kolandaivel (✉)
Department of Physics,
Bharathiar University,
Coimbatore, 641 046, India
e-mail: ponkvel@hotmail.com
Fax: +91-422-2422387

The interactions of water molecules with peptide moieties have been studied extensively, both experimentally [3, 4, 5, 6, 7, 8] and theoretically. [9, 10, 11, 12, 13, 14, 15, 16, 17] Weir et al. [11] have studied the force field and vibrational frequencies for the dipeptide alanine–alanine in water using the supermolecule approach. Nemukhin et al. [12] have studied the hydrogen-bonded complex of the dipeptide *N*-acetyl-alanine-*N'*-methylamide with water using a QM/MM approach, and they concluded that QM/MM is also a suitable method for describing the conformational properties of dipeptides. Fu et al. [13] have studied the different conformations and interactions of formamide–water complexes using DFT methods and reported that a cyclic double hydrogen-bonded structure is most stable. Recently, Lecomte et al. [14] have reported that *N*-methylacetamide (NMA) is the best molecule for modeling hydration of proteins using Rydberg electron-transfer spectroscopy and quantum chemical methods. The main problem in choosing the molecules for the peptide model is that they do not have the same conformational flexibility as amino acids. To extrapolate from these small model peptide systems to water–protein systems, the study of interactions between glycylglycine and water is significant. This will provide valuable information regarding the hydration of proteins. Sieler et al. [6] have measured the spectra for the amide band of glycylglycine and *N*-acetylglycine in water, and studied the dynamics of atoms in an aqueous environment. Kameda et al. [7] have found that the glycylglycine-hydrate crystal has three water molecules per two molecules of dipeptide. The intermolecular hydrogen-bonding interactions in glycylglycine dimer have been studied by Abramov et al. [18] Chaudhuri et al. [19] have investigated the importance of electron-correlation effects in the peptide bond formation in glycylglycine using ab initio and density functional theory methods.

With these backgrounds, we have started the present study to gain more knowledge about the hydration of peptide model systems. Ab initio and density functional theory methods have been used to study the nature of the intermolecular interactions between water and the peptide

groups of formamide, NMA and glycylglycine within the supermolecule approach. Special emphasis has also been given to predicting the active binding sites for the above molecules using condensed Fukui functions.

Computational procedures

The non-hydrated structure (ST1) and the monohydrated structures (ST2, ST3) of formamide, NMA, and glycylglycine have been optimized using second order Møller–Plesset perturbation theory [20] and the B3LYP [21, 22] DFT method with the 6-311++G** basis set. The optimized gas-phase structures of monohydrated complexes of formamide have been reoptimized using the polarized continuum model (PCM) of self-consistent reaction field theory (SCRF) to study the system in liquid phases. The interaction energy of the complexes studied were corrected for basis set superposition error (BSSE) using the Boys and Bernardi's counterpoise (CP) method, [23, 24]

$$E_{int}^{CP}(AB) = E_{AB}^{AB}(AB) - E_{AB}^{AB}(A) - E_{AB}^{AB}(B)$$

where $E_{AB}^{AB}(AB)$ is the energy of the supersystem and $E_{AB}^{AB}(A)$ and $E_{AB}^{AB}(B)$ represent the energies of the subsystems with complex geometry. Moreover, the geometry relaxation effects also contribute to the calculated interaction energies. Thus, to determine the contribution of the geometry-relaxation energy E_R in the interaction energy, we have used the equation $E_R = \sum_i E(i) - nE_m$, where $E(i)$ is the energy of the monomer in the complex geometry and E_m is the energy of the optimized monomer. [25] ^{13}C, ^{15}N NMR shielding tensors for carbon and nitrogen atoms in the peptide bond have been calculated using the gauge-independent atomic orbital (GIAO) method. [26, 27] Condensed Fukui functions $f_k(r)$ [28] have been calculated using the atomic charges q_K calculated by the Mulliken population analysis scheme (MPA),

$$f_K^+ = q_K(N+1) - q_K(N)$$

$$f_K^- = q_K(N) - q_K(N-1)$$

$$f_K^0 = {}^1\!/_2[q_K(N+1) - q_K(N-1)]$$

where N refers to the number of electrons in the system K, and condensed Fukui functions f_k^+, f_k^- and f_k^0 represent the nucleophilic, electrophilic, and radical attacks, respectively. All the computations were performed using Gaussian 98W. [29] All the figures were drawn using Molden. [30]

Results and discussion

The optimized geometrical parameters (selected values) calculated at the MP2/6-311++G** and B3LYP/6-311++G** levels of theory for the non-hydrated (ST1) and monohydrated (ST2, ST3) structures of formamide, NMA and glycylglycine are shown in Table 1 along with the experimental values. [7, 14] The generalized molec-

Table 1 Selected geometrical parameters of non-hydrated (ST1), monohydrated complexes (ST2, ST3) of formamide, N-methylacetamide and glycylglycine at the MP2/6-311++G** and B3LYP/6-311++G** levels of theory; distances (r) in Å and dihedral angles (θ) in degrees

Compounds		r(C1=O3)	r(N2–H4)	r(C1–N2)	r(C$^\alpha$–C)	r(NH...O)	r(CO...H)	(NCCN)	(HNCC)	(H4N2C1O3)
Formamide										
ST1	MP2	1.217	1.006	1.364	–	–	–	–	–	179.99
	B3LYP	1.212 (1.219)[a]	1.007 (1.002)	1.361 (1.352)	–	–	–	–	–	180.00
ST2	MP2	1.226	1.007	1.356	–	–	1.952	–	–	−170.75
	B3LYP	1.224	1.007	1.354	–	–	1.920 (2.03)	–	–	−178.99
ST3	MP2	1.220	1.012	1.360	–	2.005	–	–	–	169.97
	B3LYP	1.216	1.013	1.355	–	1.996 (2.1)	–	–	–	−179.99
NMA										
ST1	MP2	1.225	1.007	1.369	1.518	–	–	–	−0.007	180.00
	B3LYP	1.221	1.007	1.368	1.519	–	–	–	−0.004	−179.99
ST2	MP2	1.234	1.007	1.359	1.513	–	1.876	–	8.611	−173.14
	B3LYP	1.230	1.007	1.357	1.514	–	1.861	–	2.116	−177.65
ST3	MP2	1.229	1.011	1.362	1.518	2.025	–	–	8.134	−173.08
	B3LYP	1.225	1.011	1.359	1.520	2.040	–	–	0.001	−179.99
Glycylglycine										
ST1	MP2	1.239	1.014	1.347	1.522	–	–	13.064	−0.917	−179.24
	B3LYP	1.235 (1.221)	1.014 –	1.344 (1.333)	1.527 (1.529)	–	–	9.209	1.155	−177.42
ST2	MP2	1.246	1.016	1.340	1.521	–	1.951	11.175	−1.345	−179.80
	B3LYP	1.244	1.015	1.338	1.525	–	1.926	8.087	0.279	−178.38
ST3	MP2	1.241	1.020	1.346	1.523	1.870	–	29.893	4.009	−174.36
	B3LYP	1.238	1.020	1.344	1.527	1.904 (2.110)	–	25.094	5.082	−173.24

[a] Values in parenthesis are experimental results taken from [7, 14]

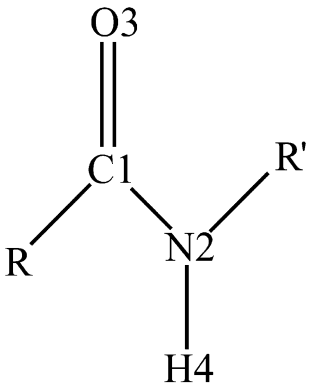

Fig. 1 Molecular structure of the compounds studied and numbering of the atoms involved in the peptide bond. R=H and R′=H: for formamide. R=CH_3 and R′=CH_3: for NMA. R=NH_2CH_2 and R′=CH_2COOH: for glycylglycine

ular structure with attached groups and atom numbering is given in Fig. 1.

In the present study, we have considered two intermolecular interactions. The first is the interaction of a water molecule with the oxygen atom of the carbonyl group of the peptide moiety (ST2); the second is that of a water molecule with the hydrogen atom of the amino group of the peptide moiety (ST3). The optimized structures (ST1, ST2, and ST3) of formamide, NMA and glycylglycine are shown in Fig. 2a–c. Extensive structural studies have been reported for formamide and NMA, [31] and our present calculated values agree very well with those values. The calculated C1=O3...H and N2-H4...O hydrogen-bond length values are compared with the experimental results, [7, 14] and it is found that the hydrogen-bond lengths calculated with MP2 are much closer to the experimental values. The interaction of a water molecule with the peptide moiety changes the structural parameters of the complex molecule. Comparison with the non-hydrated (ST1) structure of all complexes reveals that there is a slight increase in the N2-H4 bond length in the ST3 structure. The hydrogen-bonding interaction between the C1=O3 bond and the water molecule in the ST2 structure elongates the C1=O3 bond and shortens the C1-N2 bond. The same trend is also observed in the ST3 structure. Because of the elongation of the C1=O3 bond due to the interaction with water molecule, the electrons are delocalized in the bond. Thus, the nearby nitrogen atom gains more charge from carbon, which decreases the C1-N2 bond length. The hydrogen-bonding interaction with the N2-H4 group, does not influence the C1=O3 bond, because of the rigidity of the C1-N2 bond.

In formamide, when a water molecule interacts with the carbonyl oxygen, the peptide bond deviates from planarity by approximately 10° at the MP2 level of theory. The same trend is also observed for NMA. The presence of multiple intermolecular hydrogen-bond interactions in glycylglycine produces different trends in the planarity of the peptide bond. The C1-N2 bond in the ST2 structure of glycylglycine is strong, which restricts the deviation from planarity, but in the ST3 structure, the interaction with the N2-H4 group increases the flexibility of the C1-N2 bond, which causes a deviation (~5°) from planarity. This can be acceptable, since Arthur et al. [32] reported on the basis of crystallographic data that substantial deviations from planarity could be tolerated with a standard deviation up to 6°.

The total energies calculated for the non-hydrated and monohydrated complexes of formamide, NMA and glycylglycine at the MP2 and DFT levels are given in Table 2. In all cases, structure ST2 is more stable than ST3, but in glycylglycine, both levels of theory predict structure ST3 to be more stable than ST2. This is due to the formation of cyclic double hydrogen-bonded interactions between the water molecule and the N-terminus of glycylglycine and the amino group of the peptide moiety. The hydrogen-bonded complexes (ST2, ST3) of formamide in polar medium (water) were studied using the PCM of self-consistent reaction field theory at the B3LYP/6-311++G** level of theory. It is predicted that the total energy of the complexes ST2 and ST3 of formamide are −246.44378 and −246.44422 Hartrees, respectively, which shows that structure ST3 is more stable than ST2 in liquid phase. This can be confirmed by the dipole moment of the complexes. In the gas phase, the dipole moment of ST3 is higher than ST2, and this order is reversed for the complexes in the liquid phase. This study is restricted to formamide complexes, due to the inadequacy if the memory of our computer system. However, one can expect from the dipole moment that the order of stability is reversed for NMA in liquid phase, but at the same time the order of stability is the same for both gas and liquid phases for the glycylglycine complex, since the dipole moment and total energy of ST3 are higher than ST2.

The interaction energies for the monohydrated complexes calculated using the MP2 and DFT methods are shown in Table 2. It has been observed that the structures ST2 of formamide and NMA have higher interaction energies than ST3. The difference between the interaction energies of the two structures ST2 and ST3 is ~2 kcal mol^{-1} for formamide and NMA complexes, and for glycylglycine it is ~4 kcal mol^{-1}. The interaction energy of ST3 is greater than that of ST2 because of cyclic hydrogen bonding-interactions in the ST3 structure of glycylglycine. The contribution of geometry-relaxation energies to the interaction energies of all the complexes have also been calculated and the values are given in Table 2. The above energies are small for all the structures except the ST3 structure of glycylglycine, where it is found to be 2.59 and 2.80 kcal mol^{-1} at the MP2 and DFT levels, respectively. The dipole moment of all the structures computed at both levels of theory are shown in Table 2. The dipole moment of the formamide-water (ST2) complex is found to be low compared with the NMA-water and glycylglycine-water complexes because the water dipoles in ST2 of formamide point almost in opposite directions. However, in the α-helix, it is noted

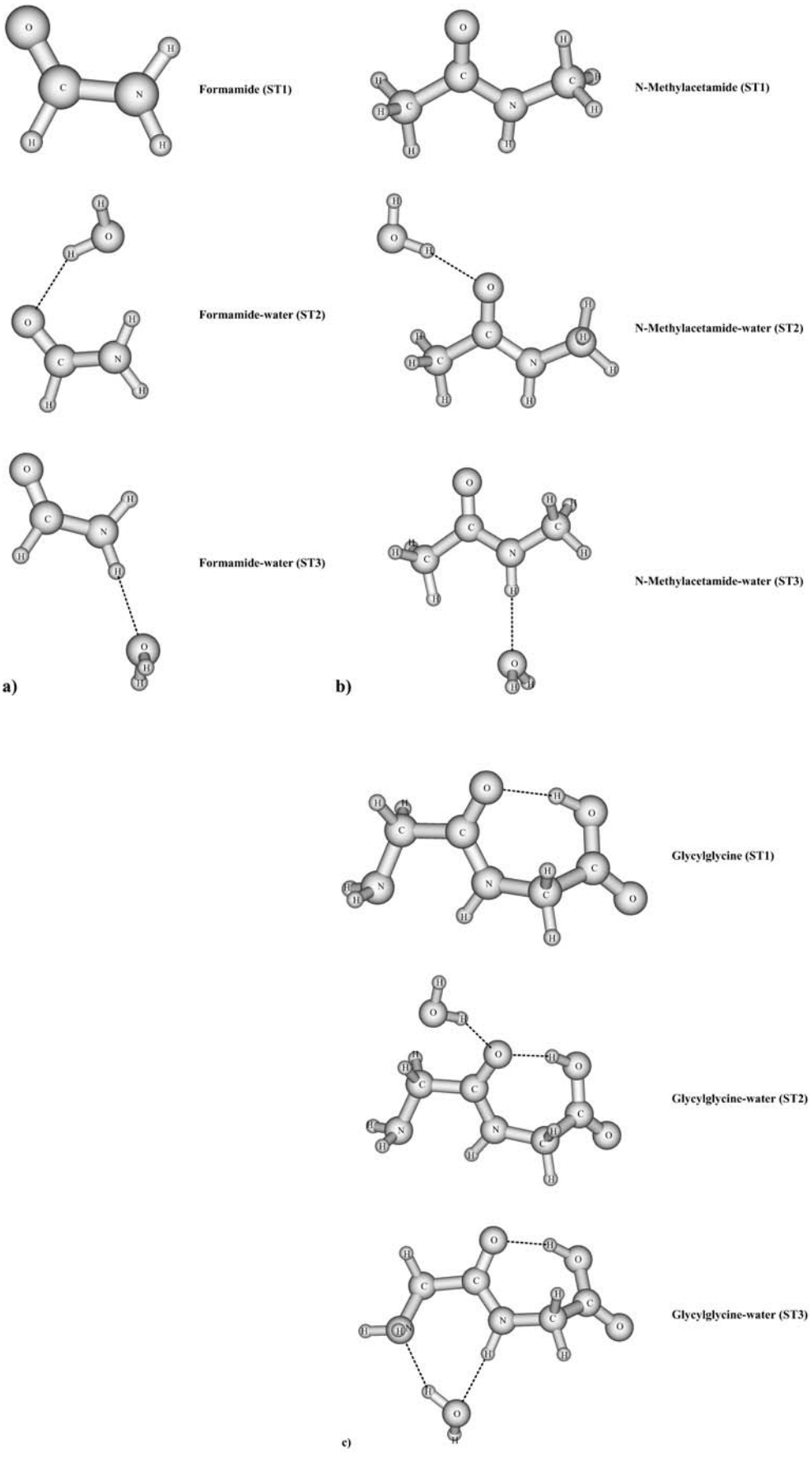

Fig. 2 Optimized non-hydrated and monohydrated structures of **a** formamide, **b** NMA, **c** glycylglycine

Table 2 Total energies (E_{tot}, in hartrees), interaction energies (E_{int}, in kcal mol^{-1}), relaxation energies (E_R, in kcal mol^{-1}), dipole moments (μ, in Debye), ^{13}C and ^{15}N NMR shielding values[a] (in ppm) at the MP2/6-311++G** and B3LYP/6-311++G** levels of theory

Compounds		E_{tot}	E_{int}	E_R	μ	^{13}C	^{15}N
Formamide							
ST1	MP2	−169.50540	–	–	4.381	–	–
	B3LYP	−169.95485	–	–	4.044 (3.73)[b]	21.943	145.124
ST2	MP2	−245.79584	−9.94	0.20	2.764	–	–
	B3LYP	−246.42832	−9.83	0.46	2.742	17.032	135.623
ST3	MP2	−245.79043	−6.31	−0.04	6.888	–	–
	B3LYP	−246.42251	−5.85	0.13	6.731 (6.20)	20.576	139.333
SCRF							
ST2		−246.44378	–	–	4.445	–	–
ST3		−246.44422	–	–	7.273	–	–
NMA							
ST1	MP2	−247.89553	–	–	4.204	–	–
	B3LYP	−248.60602	–	–	3.909 (3.80)	12.246	136.309
ST2	MP2	−324.18387	−8.23	−0.19	4.397	–	–
	B3LYP	−325.07643	−7.78	0.33	4.648 (4.33)	6.295	135.773
ST3	MP2	−324.18128	−6.48	−0.31	7.213	–	–
	B3LYP	−325.07273	−5.29	0.16	6.855 (6.56)	10.160	134.917
Glycylglycine							
ST1	MP2	−491.31328	–	–	8.372	–	–
	B3LYP	−492.60691	–	–	7.682	2.134	128.975
ST2	MP2	−567.60031	−7.78	0.18	6.380	–	–
	B3LYP	−569.07627	−7.13	0.35	5.706	−1.960	125.211
ST3	MP2	−567.60181	−11.13	2.59	9.117	–	–
	B3LYP	−569.07686	−9.96	2.80	8.283	4.105	121.894

[a] NMR shielding values and SCRF calculations using B3LYP/6-311++G** level of theory
[b] Values in parentheses are experimental results taken from [14]

that there is a cooperative effect that will increase the dipole moment in the peptide. In NMA, the cooperative effect is detected, which indicates the increase in dipole moment than the non-hydrated structure. However, this effect is not observed in glycylglycine. It is concluded that, among the three investigated molecules, NMA is the most suitable for modeling the hydration of proteins.

The ^{13}C and ^{15}N shielding values were calculated for the structures ST1, ST2 and ST3 of all molecules at the B3LYP/6-311++G** level of theory and the values are shown in Table 2. Comparisons have been made between the nitrogen and carbon atoms of ST2, and ST3 with the reference to the structure ST1 of the respective compounds. The calculated values show that a significant change is observed in the shielding values of ^{13}C in all the complexes after interaction with water molecules (ST2, ST3). But it is noted that the ^{15}N NMR values in these structures are slightly different, which is due to the presence of lone pair electrons on nitrogen, which restrict the deshielding of the nitrogen atom and gain more charge from nearest neighbors. The strength of the hydrogen bonding in ST3 is weaker than the interaction with the carbonyl group. During the formation of hydrogen bonding (ST2), the oxygen atom in the CO group aquires more charge from the nearest hydrogen and carbon atoms. Thus, the carbon atom is less shielded than in the structure (ST1). In structure ST3, the oxygen atom in the water molecule attracts more charge from the neighboring hydrogen atoms leaves nitrogen more shielded. Thus, we conclude that the structures with stronger hydrogen-bonding interactions are less shielded. The shielding values of ST3 are higher than those for ST2, so ST2 has stronger hydrogen-bonding interactions.

Table 3 Calculated condensed Fukui functions at the B3LYP/6-311++G** level of theory

Compound		C=O*	N–H*
Formamide	f_k^+	0.522	0.095
	f_k^0	0.289	0.657
	f_k^-	0.056	1.219
NMA	f_k^+	0.287	0.093
	f_k^0	0.165	0.426
	f_k^-	0.042	0.759
Glycylglycine	f_k^+	0.152	0.007
	f_k^0	0.105	0.086
	f_k^-	0.059	0.166

The recent study [33] on halomethane molecules reveals that condensed Fukui functions successfully predict the reactive sites. The binding sites of amino acids have also been studied theoretically using ab initio methods for the glycine-water complex. [34] In the present study, condensed Fukui functions have been calculated from the atomic charges of the hydrogen and oxygen atoms in the peptide for the non-hydrated structure (ST1) of formamide, NMA and glycylglycine at the B3LYP/6-311++G** level of theory and are listed in Table 3. The calculated condensed Fukui function values f_k^+, f_k^0, and f_k^-, for C=O* are 0.522, 0.289, 0.056 in

formamide, 0.287, 0.165, 0.042 for NMA and 0.152, 0.105, 0.059 for glycylglycine, which show that oxygen is the favorable reactive site for nucleophilic and the hydrogen atom in the amino group is the most favorable for electrophilic attack, because f_k^- is greater than f_k^0 and f_k^+ values for all the molecules.

Conclusions

The non-hydrated (ST1) and monohydrated structures (ST2, ST3) of formamide, NMA and glycylglycine have been optimized using the MP2/6-311++G** and B3LYP/6-311++G** levels of theory. The planarity of the peptide bond depends on the rigidity of the C-N bond. The calculated interaction energies show that the hydrogen-bonding interaction with the carbonyl group is stronger than with the amino group. Among the three molecules investigated, NMA is the most suitable for modeling protein hydration due to the presence of a cooperative effect. Condensed Fukui functions calculated at the B3LYP/6-311++G** level of theory show that the oxygen atom in the carbonyl group is the favorable reactive site for nucleophilic attack and hydrogen atom is the favorable reactive site for electrophillic attack.

Acknowledgements The authors thank Department of Science and Technology, Government of India, (DST) for financial assistance in the form of project (SP/S1/H-27/99) to complete this work.

References

1. Jarrold ME (1999) Acc Chem Res 32:360–367
2. Daune M (1999) Molecular biophysics. Oxford University Press, Oxford
3. Lovas FJ, Suenram RD, Fraser GT (1988) J Chem Phys 88:722–729
4. Fraser GT, Suenram RD, Lovas FJ (1988) J Mol Struct 189:165–172
5. Engdahl A, Nelander B, Astrand PO (1993) J Chem Phys 99:4894–4907
6. Sieler G, Schweitzer-Stenner R (1997) J Am Chem Soc 119:1720–1726
7. Kameda T, Takeda N, Ando S, Ando I, Hashizume D, Ohashi Y (1998) Biopolymers 45:333–339
8. Manas ES, Getahun Z, Wright WW, de Grado WF, Vanderkooi JM (2000) J Am Chem Soc 122:9883–9890
9. Dixon DA, Dobbs KD, Valentini JJ (1994) J Phys Chem 98:13435–13439
10. Besley NA, Hirst JD (1999) J Am Chem Soc 121:8559–8566
11. Weir AF, Lowery AH, Williams RW (2001) Biopolymers 58:577–591
12. Nemukhin AV, Grigorenko BL, Bochenkova AV, Topol IA, Burt SK (2002) J Mol Struct (Theochem) 581:167–175
13. Fu A, Du D, Zhou Z (2003) J Mol Struct 623:315–325
14. Lecomte F, Lucas G, Gregoire G, Schermann JP, Desfrancois C (2002) Eur Phys J D 20:449-457
15. Tazaki K, Shimizu K (1998) J Phys Chem B 102:6419–6424
16. Apostolakis J, Ferrara P, Caflisch A (1999) J Chem Phys 110:2099–2108
17. Kalko SG, Guardia E, Padro JA (1999) J Phys Chem B 103:3935–3941
18. Abramov YA, Volkov A, Wu G, Coppens P (2000) J Phys Chem B 104:2183–2188
19. Chaudhuri P, Canuto S (2002) J Mol Struct (Theochem) 577:267–279
20. Møller C, Plesset MS (1934) Phys Rev 46:618–622
21. Becke AD (1993) J Chem Phys 98:5648–5652
22. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
23. Boys SF, Bernardi F (1970) Mol Phys 19:553–566
24. Sedano PS (2001) Doctoral dissertation. Universitat de Girona, Spain
25. Kulkarni SA, Bartolotti LJ, Pathak RK, (2003) Chem Phys Lett 372:620–626
26. Ditchfield R (1974) Mol Phys 27:789–807
27. Wolinski K, Hinton JF, Pulay P (1990) J Am Chem Soc 112:8251–8260
28. Yang W, Mortier WJ (1986) J Am Chem Soc 108:5708–5711
29. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery Jr JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels JD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Rega N, Salvador P, Dannenberg JJ, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2001) Gaussian 98, Revision A.11.2. Gaussian, Pittsburgh, PA
30. Schaftenaar G, Noordik JH (2000) J Comput-Aided Mol Des 14:123–134
31. Scheiner S (1997) Hydrogen bonding: a theoretical perspective. Oxford University Press, New York
32. MacArthur MW, Thorton JM (1996) J Mol Biol 264:1180–1195
33. Senthilkumar K, Kolandaivel P (2002) Mol Phys 100:3817–3822
34. Basch H, Stevens WJ (1990) Chem Phys Lett 169:275–280