Journal of Molecular Graphics and Modelling 49 (2014) 1–10

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Quantum mechanical investigations on the role of C-terminal residue in influencing the structural features of dipeptides containing N-terminal proline

Gunajyoti Das*, Shilpi Mandal

Department of Chemistry, North Eastern Hill University, Shillong 793022, India

ARTICLE INFO

Article history:
Accepted 30 December 2013
Available online 13 January 2014

Keywords:
Proline
Dipeptide
Dipeptide–water complex
Solvation effects
Vibrational frequency
HOMO–LUMO energy gap

ABSTRACT

This study investigates the influence of the side chain moiety of C-terminal residue on the structural and molecular properties of seven dipeptides having proline at their N-terminal positions. The C-terminal component of the dipeptides is varied with seven different combinations viz. Ala, Leu, Asp, Thr, Asn, Arg and Sec. The calculations are carried out using B3LYP/6-311++G(d,p) level of theory in gas and implicit aqueous phase. Effects of explicit aqueous environment on the dipeptide structures are also investigated for two systems. The results furnished by this DFT study provide valuable information regarding the role of the side chain groups of C-terminal residues in determining the structural features of the amide planes, values of the ψ and ϕ dihedrals, geometry about the α-carbon atoms, theoretical IR spectra as well as the number and type of intramolecular H-bond interactions existing in the dipeptides, and extend a fine corroboration to the earlier theoretical and experimental observations. In aqueous phase the dipeptide geometries exhibit larger values of total dipole moments, greater HOMO–LUMO energy gaps and enhanced thermodynamic stability than those in gas phase. The explicit water molecules are found to modify the geometrical parameters related to the amide planes and vibrational spectra of the dipeptides.

1. Introduction

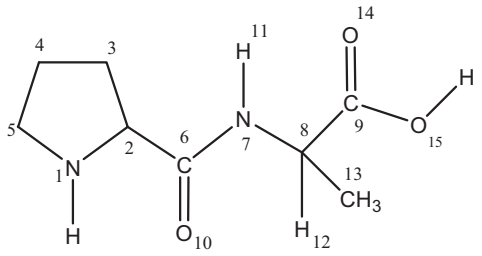
Small amino acid sequences are attractive targets for structural studies by modern methods of electronic structure theory due their immense chemical, biochemical and biological significance. Investigating and understanding the structural and molecular properties of dipeptides are of special interest because they play a major role in determining the functional specificity of proteins and polypeptides. Over the last few decades, dipeptides as well as their analogs have served as important model systems in elucidating some of the most important structural features of the protein backbone and as a result such efforts have been termed as the determination of the second half of the genetic code [1]. Dipeptides are also known to play numerous key biological roles in various life-supporting processes [2–6].

Gas phase computational investigations carried out on dipeptides [7–10] have revealed that steric hindrance of the side chain groups of the C-terminal residues and their ability to form intramolecular H-bonds play crucial role in determining the structural and molecular properties of dipeptides. Such gas-phase studies are useful in providing valuable information regarding their intrinsic properties free from the solvent or crystal phase effects. However, since dipeptides and other bioactive molecules operate in aqueous environment, it is of basic importance to consider the solvent effects of aqueous phase as well. Moreover, proteins and polypeptides contain many charged or polar groups which interact strongly with polar solvents like water, and as a result the solvent effects of aqueous phase often play critical roles in modifying the gas phase structural features of the dipeptides. The effects of solvation on the conformations and energies of dipeptides have been well discussed in the literature [11–18]. These studies have analyzed the energetics and structural features of the dipeptides in gas and in several different solvent phases to understand the effect of the surrounding environment on the stabilities and conformational preferences of the dipeptides. Often, in a strong polar solvent like water the interactions among the nearest-neighbor residues of the dipeptides are dramatically modified as compared to those in gas phase, which consequently affects the Ramachandran dihedrals (ψ, ϕ) [17–20] conferring markedly different conformations to the dipeptides in aqueous phase.

Among the 23 genetically encoded amino acids, proline (Pro) is unique in that its α-amino group is secondary and is fixed within a pyrrolidine ring. This distinctive cyclic structure of its side chain gives it an exceptional conformational rigidity compared to the

* Corresponding author. Tel.: +91 0364 2722615.
E-mail address: guna_das78@yahoo.co.in (G. Das).

1093-3263/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.12.009



$\Psi = N_1–C_2–C_6–N_7$
$\Phi = C_6–N_7–C_8–C_9$
$\Omega = C_6–N_7–C_8–C_{13}$

Fig. 1. Schematic representation of Pro–Ala showing atom numbering.

other canonical amino acids [21]. Compared to the other naturally occurring amino acid residues, the Pro residues are found to occur at a much higher than average frequency in proteins [22] and play an important role in conferring the structures of proteins [22,23] and peptides [24]. A number of computational investigations have been carried out both in conjunction with experimental studies and independently to gaining insights into the presence and biological functions performed by proline in the macromolecular context of real life systems [25–33].

This theoretical study, performed in gas and aqueous phase using PCM procedure [34], investigates the effects of steric interactions of the side chain functional group/groups of the varying C-terminal residue on the structural features of the peptide planes, geometry about the α-carbon atoms, values of the ψ and ϕ dihedrals, dipole moments, rotational constants, theoretically predicted vibrational spectra as well as the number and types of intramolecular H-bond interactions that may play crucial roles in determining the structure and stability of a set of seven dipeptides containing proline. The PCM procedure describes only the long-range effects and cannot take into account short-range interactions like possible solute-solvent intermolecular hydrogen bonds. For this reason we also investigate the effects of explicit aqueous environment on the structural and molecular properties of two dipeptides by considering three explicit water molecules surrounding the molecular geometries of the two dipeptides. It is expected that in this way both short and long-range interactions would be taken into account.

The dipeptides are constructed by keeping proline as a fixed component in the N-terminus whereas the component in the C-terminus is varied with seven different combinations. The seven different amino acids chosen for the C-terminus position are alanine (Ala), leucine (Leu), aspartic acid (Asp), threonine (Thr), asparagine (Asn), arginine (Arg) and selenocysteine (Sec). All these amino acid residues are taken as neutral (non-ionic) species. The C-terminal residues are chosen on the basis of the propensity of their side chain moieties to be in contact with a polar solvent like water. The side chain moieties of Ala and Leu are non-polar (hydrophobic), polar acidic for Asp, polar uncharged for Thr and Asn while polar basic for Arg. On the other hand, Sec is structurally similar to cysteine; the only difference is that in Sec a selenium-containing selenol group replaces the sulfur-containing thiol group of cysteine [18]. It is also known that the p*K*a of free Sec is significantly lower than that of free Cys [35]. The standard three letter abbreviation is used to represent an amino acid while a particular dipeptide is named by listing the N-terminal residue first. Thus, Pro–Ala dipeptide corresponds to a structure in which proline is in the N-terminal position and alanine is in the C-terminal position. Fig. 1 schematically represents the atom numbering assigned to the seven dipeptides considering Pro–Ala as an example. The C_6—N_7 is the peptide linkage of a given dipeptide structure while C_2 and C_8 are the α-carbon atoms of the N- and C-terminal residues respectively. This DFT study, performed in gas phase as well as in implicit and explicit aqueous environment, is expected to provide valuable insights regarding the role of C-terminal residues in influencing the structural and molecular properties of dipeptides at an atomic level, which in turn may help us to understand the structure–function relationships of proteins and enzymes containing Pro residues.

2. Computational methodology

The B3LYP/6-311++G(d,p) level of theory [36,37] of Gaussian 03 and 09 packages [38,39] was used to carry out full geometry optimization and vibrational frequency calculations on all the dipeptide geometries. The efficacy of B3LYP/6-311++G(d,p) in studying conformational behavior and various other properties of amino acids has been explained in literature [40]. All the computations were conducted in gas as well as in simulated aqueous phase using a polarizable continuum model (PCM). The two intermolecular H-bonded complexes Pro–Ala–W_3 and Pro–Thr–W_3 were also optimized at B3LYP/6-311++G(d,p) level using PCM. The accuracy of self-consistent reaction field (SCRF) model in predicting the structure and energetics of dipeptides has already been justified in the literature [41]. Absence of imaginary frequency value in the vibrational frequency calculations proved the optimized geometries of all the systems studied as true minima. Zero point energy (ZPE) corrections were applied to the total energies of all the dipeptides using a scaling factor 0.9877 [42]. The vibrational frequencies below 1800 cm^{-1} were scaled with 1.01 and for those above 1800 cm^{-1} a correction factor 0.9679 was used [42]. Use of diffuse functions is important to take into account the relative diffuseness of lone pair of electrons when a molecule under investigation contains the same [43] while polarization functions are useful in studying the conformational aspects where stereoelectronic effects may play an important role [44].

3. Results and discussion

Fig. 2 depicts the optimized geometries of all the seven dipeptides in aqueous and gas phase while those of two dipeptide–water complexes are portrayed in Fig. 3. Table 1 assembles the DFT data on solvation enthalpies (E_{Sol}), solvation free energies (G_{Sol}), zero point vibrational energies (ZPVE), rotational constants, dipole moments and HOMO/LUMO energy gaps of the dipeptides calculated in gas and aqueous phase using B3LYP/6-311++G(d,p) level. Table 2 collects the geometric parameters for some structurally significant intramolecular H-bonding interactions that play crucial roles in the energetics and in conferring the observed conformations to the dipeptides in both the phases. Table 3 assembles some of the characteristic frequency and intensity values (given in parentheses) of the dipeptides calculated at the B3LYP/6-311++G(d,p) level of theory. The predicted total electronic energies (E) and Gibbs free energies (G) as well as their respective ZPVE corrected values, E_{corr} and G_{corr} respectively, of all the dipeptides in both the phase are listed in Table S1 of Supplementary Information (SI). For a particular dipeptide, the E_{Sol} and G_{Sol} are calculated by subtracting the gas-phase energies (using E_{corr} and G_{corr} values) from its solvent-phase energies [45]. The HOMO/LUMO energies of all the systems studied are also listed in Table S1. The bond length and bond angle values related to the amide planes of the dipeptides are listed in Tables S2 and S3 of the SI respectively (the gas phase values are given in parentheses). The planarity of the peptide planes of the dipeptides are monitored by considering four dihedral angles, viz. C_2—C_6—N_7—C_8, O_{10}—C_6—N_7—H_{11}, C_2—C_6—N_7—H_{11} and O_{10}—C_6—N_7—C_8 while the backbone structural features of the dipeptides are monitored by taking into account the

two well known Ramachandran dihedral angles ψ (N_1—C_2—C_6—N_7) and ϕ (C_6—N_7—C_8—C_9). The orientations of the side-chain moieties of C-terminal residues are specified by the dihedral angle Ω (C_6—N_7—C_8—C_{13}). The gas and aqueous phase values of these dihedral angles are presented in Table S4 of the SI. Table S5 of the SI represents the gas and aqueous phase data on the geometrical parameters considered to examine the geometry around the α-carbon atoms. The theoretical IR spectra of all the systems calculated in gas and aqueous phase (scaled with a correction factor 0.9679) are reported in the SI while Table S6 presents a comparison of the gas phase IR assignments furnished in this study with those obtained previously by various gas phase theoretical and experimental studies.

3.1. Structure and stability of the dipeptides

As evident from Table 1 the total dipole moments of all the dipeptides are larger in aqueous phase (3.415–7.206 D) than those in the gas phase (2.344–4.264 D). This indicates that the dipeptide structures possess greater polar character and have greater affinity to polar solvents. It has been generally accepted that large dipole moments often lead to an extra stabilization in aqueous solution. This fact is well supported by the data on the total electronic energies and Gibbs free energies of the dipeptides calculated at B3LYP/6-311++G(d,p) level of theory (reported in Table S1 of the SI). The negative values of E_{Sol} (range from −9.08 to −17.48 kcal/mol) and G_{Sol} (range from −9.74 to −18.24 kcal/mol) suggest that the

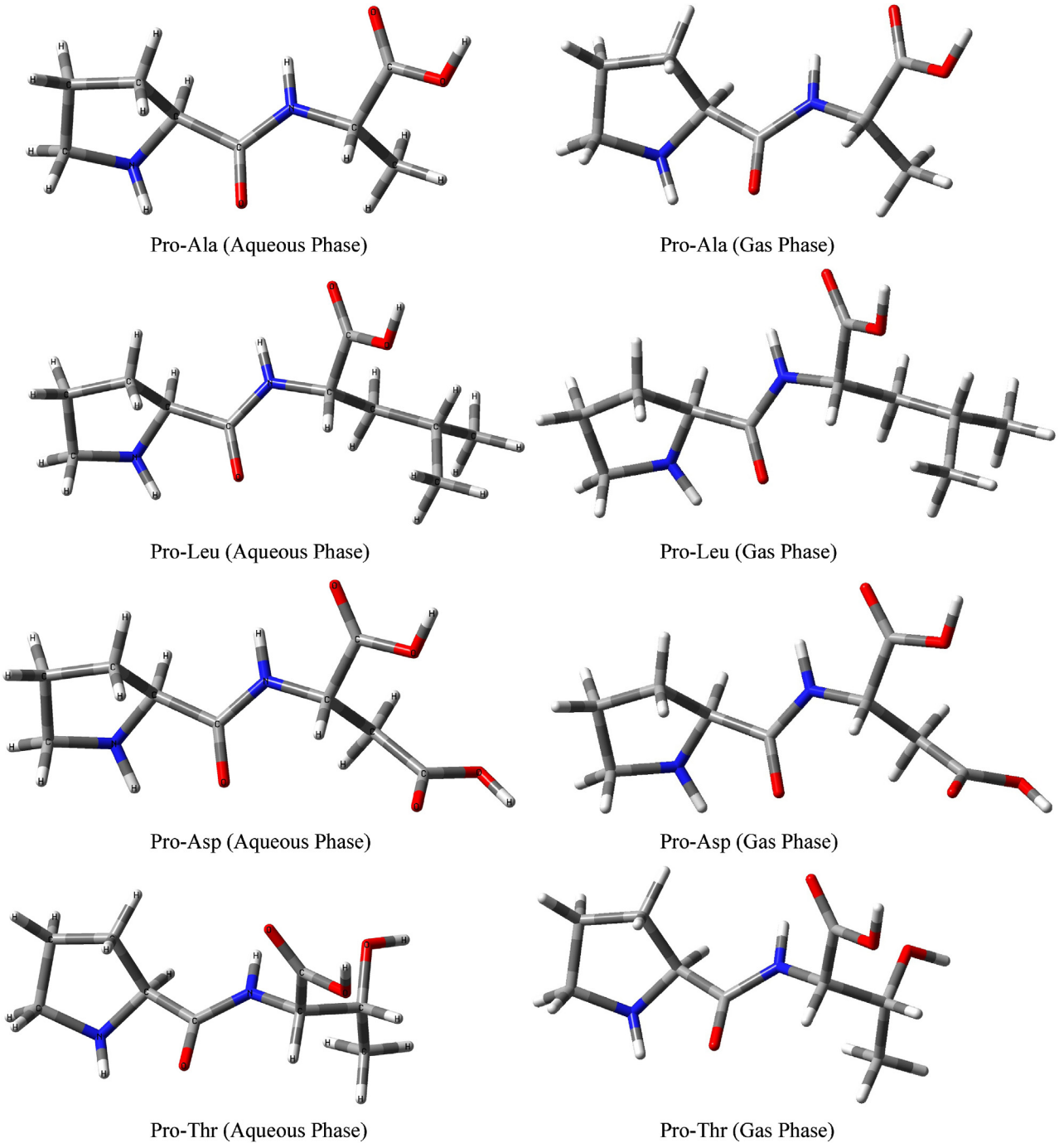


Fig. 2. Optimized geometries of all the dipeptides in aqueous and gas phase.

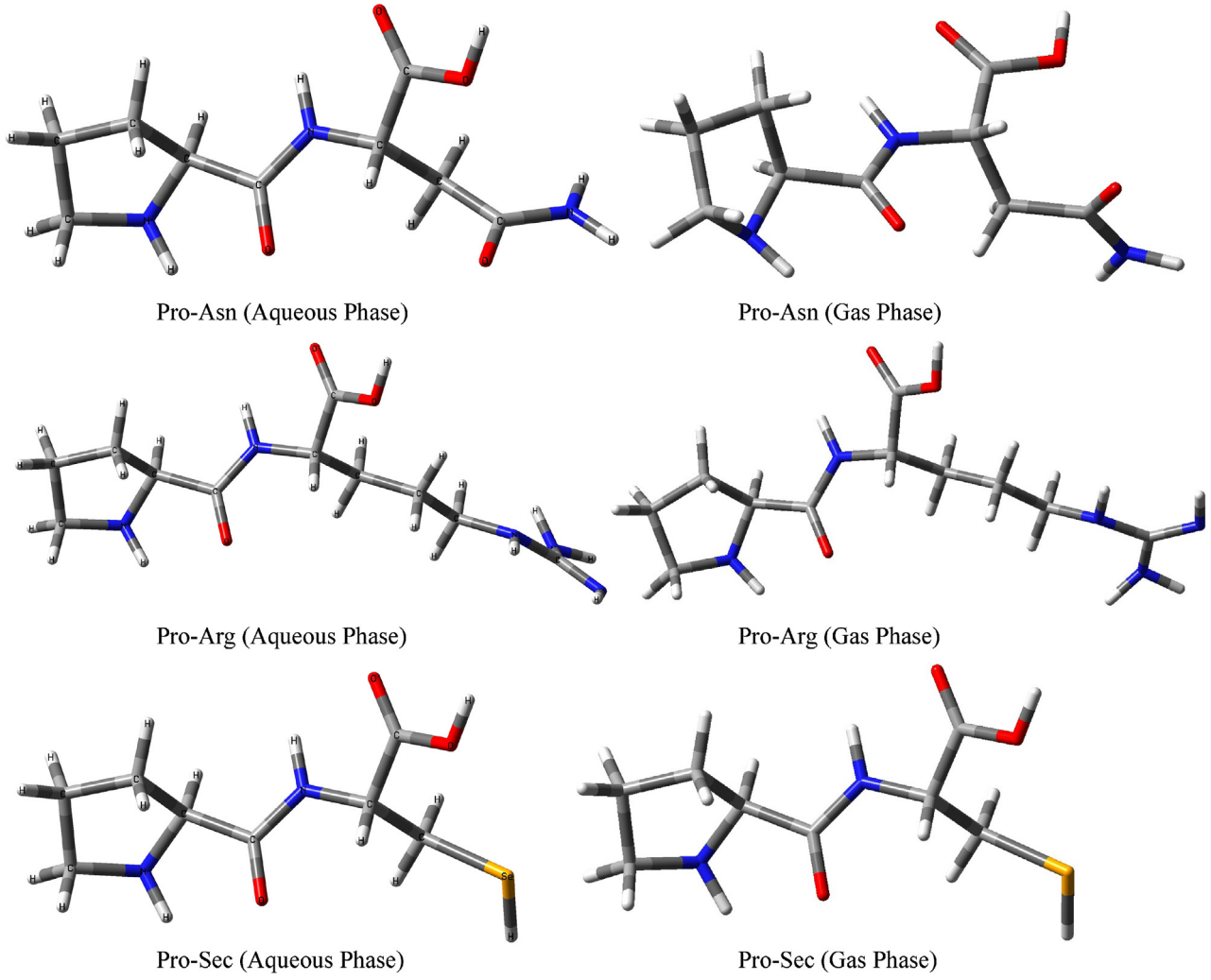


Fig. 2. (*Continued*)

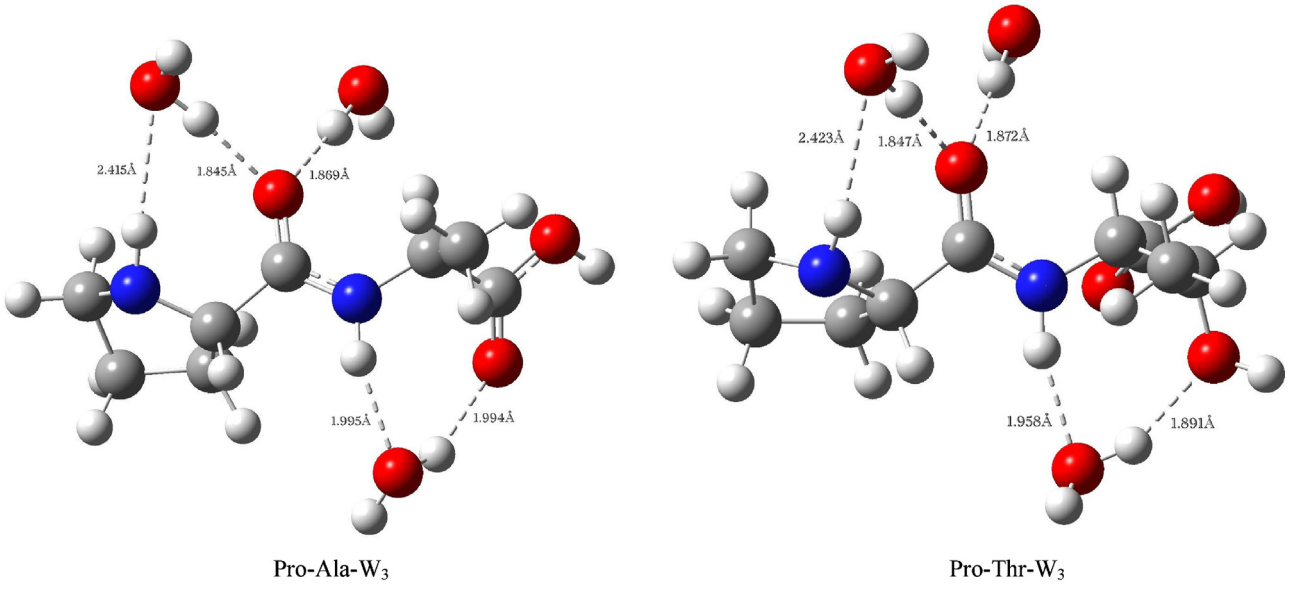


Fig. 3. Optimized geometries of Pro–Ala–W_3 and Pro–Thr–W_3 in aqueous phase.

Table 1
Calculated solvation enthalpies[a] (E_{Sol}), solvation free energies (G_{Sol}), zero point vibrational energies (ZPVE; in Hartrees), rotational constants (GHz), dipole moments (Debye) as well as HOMO and LUMO energy gaps (eV) of the dipeptides in gas and aqueous phase.

Dipeptides	Phases	ZPVE	E_{sol}	G_{sol}	Rotational constants			Dipole Moments	HOMO–LUMO energy gap
					A	B	C		
Pro–Ala	Aqueous	0.227062	−9.08	−9.74	1.8714	0.3601	0.3227	3.856	5.7225
	Gas	0.227422			1.8481	0.3667	0.3260	2.796	5.2695
Pro–Leu	Aqueous	0.311408	−9.35	−10.08	0.9359	0.2397	0.2099	3.853	5.6660
	Gas	0.312104			0.9128	0.2457	0.2117	2.735	5.1726
Pro–Asp	Aqueous	0.242157	−13.17	−13.73	1.0323	0.2456	0.2117	4.545	5.5666
	Gas	0.242751			1.0065	0.2492	0.2131	2.811	5.2080
Pro–Thr	Aqueous	0.259652	−11.7	−11.95	1.0524	0.3117	0.2844	5.099	5.4390
	Gas	0.259948			1.0618	0.3166	0.2784	2.805	5.1277
Pro–Asn	Aqueous	0.254167	−16.10	−16.52	1.0479	0.2452	0.2125	7.206	5.6918
	Gas	0.254414			1.0107	0.2499	0.2140	3.992	5.2899
Pro–Arg	Aqueous	0.340844	−17.48	−18.24	0.6563	0.1205	0.1070	7.086	5.5465
	Gas	0.341415			0.6335	0.1255	0.1081	4.264	5.1836
Pro–Sec	Aqueous	0.225020	−10.30	−11.10	1.0565	0.2031	0.1768	3.415	5.8039
	Gas	0.225518			1.0490	0.2049	0.0887	2.344	5.4175
Pro–Ala–W_3	Aqueous	0.299174			0.5352	0.3023	0.2374	8.299	5.5367
Pro–Thr–W_3	Aqueous	0.332218			0.4422	0.2441	0.2155	9.645	5.7021

[a] Energies in kcal/mol (ZPVE corrected and scaled with a correction factor 0.9877).

Table 2
Structural parameters for the intramolecular H-bond interactions detected (distances[a] in angstrom) in the dipeptides of proline in gas and aqueous phase (values of the A—H···B angles are given in parentheses).

Dipeptides	Phases	$O_{10}\cdots H—C_8$	$O_{10}\cdots H—N_1$	$O_{14}\cdots H—N_7$	$O_{SC}\cdots H—C_8$	$O_{SC}\cdots H—N_7$	$Se\cdots H—C_8$
Pro–Ala	Aqueous	2.557 (92.2)	2.245 (112.2)	2.315 (101.7)	*abs*	*abs*	*abs*
	Gas	2.566 (90.5)	2.249 (110.9)	2.278 (103.6)	*abs*	*abs*	*abs*
Pro–Leu	Aqueous	2.477 (97.0)	2.236 (112.5)	2.408 (98.6)	*abs*	*abs*	*abs*
	Gas	2.531 (92.8)	2.249 (111.2)	2.312 (102.3)	*abs*	*abs*	*abs*
Pro–Asp	Aqueous	2.588 (89.5)	2.244 (112.1)	2.278 (102.9)	2.532 (101.2)	*abs*	*abs*
	Gas	2.711 (86.4)	2.256 (110.8)	2.199 (106.4)	2.485 (101.9)	*abs*	*abs*
Pro–Thr	Aqueous	2.436 (99.1)	2.234 (112.4)	*abs*	*abs*	2.457 (98.5)	*abs*
	Gas	2.344 (103.9)	2.256 (110.7)	2.604 (87.3)	*abs*	2.750 (85.4)	*abs*
Pro–Asn	Aqueous	2.571 (90.8)	2.239 (112.3)	2.284 (102.6)	2.513 (102.1)	*abs*	*abs*
	Gas	2.750 (85.4)	2.247 (111.2)	2.183 (107.0)	2.411 (103.7)	*abs*	*abs*
Pro–Arg	Aqueous	2.504 (95.2)	2.238 (112.3)	2.361 (100.2)	*abs*	*abs*	*abs*
	Gas	2.545 (91.3)	2.247 (110.9)	2.273 (103.6)	*abs*	*abs*	*abs*
Pro–Sec	Aqueous	2.547 (91.9)	2.241 (112.2)	2.295 (102.3)	*abs*	*abs*	3.034 (86.6)
	Gas	2.598 (88.7)	2.254 (110.9)	2.244 (104.5)	*abs*	*abs*	3.035 (86.6)
Pro–Ala–W_3	Aqueous	2.384 (103.1)	2.501 (101.0)	2.706 (85.1)	*abs*	*abs*	*abs*
Pro–Thr–W_3	Aqueous	2.377 (102.7)	2.512 (100.4)	*abs*	*abs*	2.740 (92.1)	*abs*

[a] Only the (B···H) distances are listed; *abs* = absent; O_{SC} = oxygen atom belonging to —SC group of C-terminal residues; Angle = degrees.

Table 3
Frequencies[a] (in cm^{-1}) and IR intensities (in km/mol) of various vibrational modes[b] obtained from the theoretical vibrational spectra of the proline dipeptides in gas and aqueous phase. Intensities are given in parentheses.

Dipeptides	Phases	$\nu(O_{15}—H)$	$\nu(C_6{=}O_{10})$	$\nu(N_7—H_{11})$	$\delta(N_7—H_{11})$	$\nu(C_6—N_7)$	$\nu(N_1—H)$	$\nu(C_8—H)$	$\nu(C_2—H)$
Pro–Ala	Aqueous	3618 (142)	1697 (599)	3478 (115)	1561 (395)	1255 (50)	3405 (38)	2975 (10)	2910 (86)
	Gas	3635 (87)	1744 (250)	3480 (61)	1553 (239)	1244 (34)	3419 (28)	2960 (7)	2881 (9)
Pro–Leu	Aqueous	3607 (132)	1695 (619)	3482 (102)	1561 (418)	1250 (56)	3407 (41)	3016 (9)	2906 (107)
	Gas	3626 (81)	1741 (233)	3484 (56)	1552 (284)	1239 (44)	3419 (29)	3000 (16)	2878 (76)
Pro–Asp	Aqueous	3615 (162)	1700 (538)	3477 (125)	1555 (393)	1258 (40)	3410 (39)	2969 (9)	2906 (20)
	Gas	3635 (100)	1745 (230)	3472 (73)	1548 (280)	1251 (15)	3421 (28)	2943 (7)	2879 (107)
Pro–Thr	Aqueous	3618 (133)	1703 (509)	3491 (96)	1556 (471)	1266 (107)	3407 (40)	2994 (19)	2908 (137)
	Gas	3639 (83)	1748 (219)	3505 (40)	1555 (275)	1261 (92)	3421 (27)	2982 (19)	2881 (4)
Pro–Asn	Aqueous	3615 (150)	1698 (376)	3477 (122)	1557 (393)	1255 (99)	3409 (40)	2978 (10)	2906 (16)
	Gas	3636 (93)	1745 (232)	3470 (74)	1550 (290)	1246 (95)	3418 (29)	2937 (13)	2880 (92)
Pro–Arg	Aqueous	3608 (135)	1696 (622)	3481 (110)	1560 (431)	1252 (57)	3406 (40)	2997 (18)	2907 (76)
	Gas	3625 (85)	1741 (232)	3480 (61)	1549 (290)	1239 (46)	3419 (29)	2966 (6)	2880 (21)
Pro–Sec	Aqueous	3610 (157)	1700 (443)	3472 (124)	1548 (426)	1244 (99)	3410 (40)	2979 (4)	2907 (21)
	Gas	3630 (97)	1745 (218)	3475 (65)	1545 (266)	1235 (67)	3421 (28)	2966 (3)	2878 (87)
Pro–Ala–W_3	Aqueous	3609 (126)	1665 (511)	3311 (587)	1608 (504)	1301 (98)	3431 (52)	3011 (16)	2919 (60)
Pro–Thr–W_3	Aqueous	3618 (139)	1674 (396)	3291 (684)	1611 (598)	1306 (60)	3431 (48)	2998 (20)	2923 (57)

[a] The frequencies below 1800 cm^{-1} are scaled with 1.01 and for those above 1800 cm^{-1} a correction factor 0.9679 is used.
[b] Vibrational modes: ν = stretching; δ = amide II (bending mode).

aqueous phase dipeptide geometries possess more thermodynamic stability than those in gas phase. Furthermore, large dipole moment values may be helpful in detecting these dipeptides in the microwave spectrum since the microwave transition intensities are proportional to the square of the dipole moments. On the other hand, efficiency of DFT method in predicting the rotational constants of biologically important molecules has been discussed in literature [46]. In the absence of any experimental data on rotational constants these theoretically predicted values may assist experimentalists in determining the other conformers of the dipeptides using the rotational spectroscopy. Table 1 also represents the DFT data on the energy gaps between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies of the dipeptides in both the phases. As mentioned in the earlier literature [47], the data furnished by this DFT study also suggest that the HOMO–LUMO energy gaps of the dipeptides increase in the presence of a solvent with high dielectric constant compared to those in gas phase. In aqueous phase the predicted HOMO–LUMO energy gaps of the dipeptides are (range from 5.4390 to 5.8039 eV) always higher than their respective gas phase values (range from 5.1277 to 5.4175 eV) which is also indicative of the increased stability of the dipeptides in aqueous phase.

Understanding the effects of solvation and functional group/groups present in side chain moiety of the C-terminal residue on the geometrical parameters, viz. bond lengths, bond angles and dihedral angles, related to dipeptide structures is of fundamental importance to accurately determine the structural aspects of large peptide chains. Table S2 presents the gas and aqueous phase bond length values of the five bonds of the amide planes i.e. C_2—C_6, C_6=O_{10}, C_6—N_7, N_7—H_{11} and N_7—C_8. As reported in previous studies [17,18], this DFT study on the dipeptides containing N-terminal proline residues also predicts very little variance in the bond length values of the amide planes as the identity of the C-terminal residue of the dipeptides changes. The five bonds remain basically unaltered and show a maximum deviation of 0.008 Å in gas phase and 0.007 Å in aqueous phase from their respective average values. In aqueous phase the three bonds C_2—C_6, C_6=O_{10} and C_6—N_7 show maximum deviations up to 0.005, 0.008 and 0.009 Å respectively from their corresponding gas phase values, whereas the other two bonds N_7—H_{11} and N_7—C_8 do not change much and show little deviation up to only 0.003 Å. The elongation of the C_6=O_{10} bonds in aqueous phase (elongated up to 0.008 Å) and shortening of the C_6—N_7 bonds in aqueous phase (by a maximum magnitude of 0.009 Å) are in good agreement with the previous theoretical results [12]. The gas and aqueous phase values of the six bond angles of the amide planes i.e. C_2—C_6—O_{10}, C_2—C_6—N_7, O_{10}—C_6—N_7, C_6—N_7—C_8, C_6—N_7—H_{11} and H_{11}—N_7—C_8 are presented in Table S3. With maximum variation of 1.0° in gas phase and 0.7° in aqueous phase these results too suggest very little changes in the bond angle values of the amide planes as the C-terminal residue of the dipeptides changes. In aqueous phase the three bond angles C_2—C_6—O_{10}, C_2—C_6—N_7 and O_{10}—C_6—N_7 show maximum deviations up to 0.7, 0.6 and 0.6° respectively from their corresponding gas phase values, whereas the other three bond angles C_6—N_7—C_8 C_6—N_7—H_{11} and H_{11}—N_7—C_8 are deviated up to 1.0°.

The limited usefulness of protein structure prediction based on a simplified model where only the Ramachandran angles are considered has been well discussed in the literature [8,9]. It is generally accepted that the amide planes of small dipeptides are not completely planar and the peptide bond geometry is sensitive to the steric bulk of the amino acid side chains. Since the protein structures usually contain thousands of amino acid residues, the small variations in geometrical parameters around the peptide bonds can lead to large errors in predicting the overall structures of large peptide chains and proteins. The gas and aqueous phase values of the four dihedral angles, viz. C_2—C_6—N_7—C_8, O_{10}—C_6—N_7—H_{11}, C_2—C_6—N_7—H_{11} and O_{10}—C_6—N_7—C_8, considered to monitor the planarity of the peptide planes of the dipeptides are listed in Table S4. Evidently, if the amide plane of a particular dipeptide has to be planar then the two dihedral angles C_2—C_6—N_7—C_8 and O_{10}—C_6—N_7—H_{11} should be close to 180° and the other two i.e. C_2—C_6—N_7—H_{11} and O_{10}—C_6—N_7—C_8 should be close to 0°. The data presented in Table S4 shows that in aqueous phase the values of the four dihedral angles deviate up to a maximum value of 5.7° from the expected value whereas in gas phase the maximum deviation observed is 8.1°. Thus, the four dihedral angles do not exhibit large variations from their expected values in both the phases, however, the extent of deviations clearly indicates that the geometry of the amide planes are not completely planar regardless of whether the systems are in gas phase or in strong polar solvents like water. Out of the seven dipeptides studied here, only in three systems, viz. Pro–Asp, Pro–Thr and Pro–Asn, the amide planes show enhanced planarity in aqueous phase as compared to gas phase. This suggests that solvation effects are not the only factors that can influence the planarity of the amide planes. Therefore, it is reasonable to assume that the amide plane geometry of a given dipeptide principally depends on two factors – (i) size and type of functional group/groups present in the side chain part of the C-terminal residue (—SC group) and (ii) intramolecular H-bond forming ability of the H- and O-atoms of the amide plane with their adjacent moieties belonging to the C- and N-terminal residues. The intramolecular H-bond interactions that play crucial role in dictating the structural features of the amide planes and in imparting the observed conformations to the dipeptides in gas and aqueous phase are listed in Table 2 and a discussion on these interactions is also offered in a succeeding section of this paper.

Table S4 also lists the —SC groups of C-terminal residues of the seven dipeptides as well as the gas and aqueous phase values of their ψ, ϕ and Ω dihedral angles. The gas and aqueous phase values of ψ furnished by this DFT study suggest that little variance results in the values of these dihedral angles as the identity of the C-terminal residue of a given dipeptide changes. On the other hand, the influence of size and type of functional group/groups present in the —SC group on the structural features of the amide planes as well as on the values of ϕ dihedrals has been well documented in the literature [8,17,18]. A large sized —SC group requires more physical space to accommodate itself in between the amide plane and carboxylic group of the C-terminal residue of a given dipeptide which may consequently influence the value of ϕ and Ω as well as planarity of the amide plane. Further, the —SC group, depending on the type of functional group/groups present in it, may exert electrostatic repulsive or electrostatic attractive forces on its neighboring atoms belonging to the peptide plane and the carboxylic group of the C-terminal residue of a particular dipeptide which may also affect the planarity of the amide plane as well as the values of ϕ and Ω angles. Thus, the smaller values of ϕ observed in the case Pro–Thr, −118.7° in gas phase and −94.1° in aqueous phase, as compared to that of Pro–Ala, −155.0° in gas phase and −151.7° in aqueous phase, even though the —SC group of Thr is larger than that of Ala, can be explained on the basis of the type of functional group/groups present in the —SC groups of Thr and Ala [18]. The data collected in Table S4 suggests that the implicit solvation does not show dramatic influence on the conformational aspects of the dipeptides about their ψ, ϕ and Ω dihedrals. However, the effects of the explicit aqueous environment are quite prominent on the ϕ dihedrals of the two dipeptides Pro–Ala and Pro–Thr. Regarding the aqueous phase conformational propensities of a given dipeptide about its ψ and ϕ dihedrals, an earlier theoretical study on alanine dipeptide at

MP2 level [12] has revealed that the conformers $C7_{eq}$ ($\psi = 90.1^\circ$ and $\phi = -86.3^\circ$) and C5 ($\psi = 143.8^\circ$ and $\phi = -156.4^\circ$) are the dominant species in gas phase and nonpolar solvents extending a fine corroboration to the experimental observations [48,49]. However, in strong polar solvents like water the conformers C5 and β ($\psi = 142.1^\circ$ and $\phi = -64.0^\circ$) emerge as the most dominant species with β being marginally higher in energy than C5. Since the PCM procedure does not consider the explicit water–dipeptide interactions and there is no existing experimental evidence of C5 occurring as the most dominant specie in water, it is suggested that the effects of the explicit aqueous environment move the conformers C5 and β to an intermediate conformation. In view of the above conclusions our DFT results are in good agreement with the previous theoretical and experimental observations. Out of the seven dipeptides studied here, six of them, except Pro–Thr, are of C5 type conformers whose aqueous phase ψ values range from 152.6° to 158.0° while the ϕ values span from −144.0° to −156.0°. The aqueous phase geometry of the Pro–Thr dipeptide ($\psi = 155.2^\circ$ and $\phi = -94.1^\circ$) can be considered to be close to the β conformer on the energy surface while its gas phase conformation ($\psi = 160.2^\circ$ and $\phi = -118.7^\circ$) seems to be closer to the C5 conformation stabilized by the $O_{14}\cdots H{-}N_7$ interaction which is absent in its aqueous phase conformation (see later). On the other hand, the introduction of an explicit aqueous environment by considering three explicit water molecules significantly changes the conformational features about the ϕ dihedrals of Pro–Ala and Pro–Thr in the two water–dipeptide complexes Pro–Ala–W_3 ($\psi = 152.1^\circ$ and $\phi = -120.9^\circ$) and Pro-Thr-W_3 ($\psi = 152.1^\circ$ and $\phi = -106.3^\circ$), imparting them intermediate conformations between the C5 and β type conformers. Thus, the present study reveals that the solvent effects as well as the identity of the C-terminal residue play important roles in influencing the conformational behavior of a given dipeptide about its ϕ dihedral angle.

3.2. Geometry about the α-carbon atoms

$N_1{-}C_2{-}C_3$, $N_1{-}C_2{-}C_6$ and $C_3{-}C_2{-}C_6$ are the three bond angles considered to monitor the geometry around the C_2 α-carbon atoms of the dipeptides while $N_7{-}C_8{-}C_9$, $N_7{-}C_8{-}C_{13}$ and $C_{13}{-}C_8{-}C_9$ are the same for the C_8 α-carbon atoms. It is evident from Table S5, which collects the gas and aqueous phase data on the bond angles about the α-carbon atoms, that the geometries about the C_2 atoms of the dipeptides remain more or less unchanged with the change in the identity of the C-terminal residue. The three bond angles about the C_2 atoms show a maximum deviation of only 0.2° in gas phase and 0.3° in aqueous phase from their respective average values. This is expected because, the stereoelectronic effects of the varying –SC groups on the geometry of the C_2 atoms are very little as they reside at a distance of four bonds away from these α-carbon atoms. On the contrary, the three angles about the C_8 atoms show a maximum deviation of 2.2° in gas phase and 2.8° in aqueous phase from their respective average values. The observed deviations in the geometry of C_8 α-carbon atoms can be explained by invoking the two factors – size and the type of functional groups present in the –SC groups as previously mentioned while discussing the planarity of the peptide planes. Since the varying –SC group is situated adjacent to the C_8 atom the geometry around it is affected by the changing identity of the –SC group. Thus, the results furnished by this DFT study on the geometry of the α-carbon atoms of the seven dipeptides strongly point to the fact that the geometry about the α-carbon atoms is sensitive to the identity of the side chain moieties of a given amino acid sequence. This factor must be taken into account while considering the structures of larger peptides and proteins since small variations in the geometrical parameters around each α-carbon atoms can also lead to large errors in predicting the overall structure of large peptide chains and proteins. Regarding the solvation effects on the geometry of the C_2 and C_8 atoms, it is evident from Table S5 that the geometry about the C_8 atoms of the dipeptides (maximum deviation up to 1.7° for the angle $N_7{-}C_8{-}C_{13}$ in Pro–Thr) is more sensitive to the changes in surrounding environment than that around the C_2 atoms where the maximum deviation predicted is only up to 0.9° for the $N_1{-}C_2{-}C_6$ angle in Pro–Arg.

3.3. Intramolecular hydrogen bonds

In gas phase, the different conformers of a given molecule are stabilized by a delicate interplay between the stabilizing intramolecular H-bond interactions and destabilizing repulsive forces arising either from lone pair-lone pair electronic repulsion or steric effects. However, in solvent phase additional intermolecular interactions with the solvent molecules make the conformational behavior more complex and often it has been found that the gas phase stability order of a certain set of conformers is reversed in solvent phase [13]. The present DFT structural study conducted in both gas and aqueous phase provides us the opportunity to understand the combination of intramolecular H-bond interactions that may exist in the dipeptide structures and play crucial role in stabilizing the dipeptides in both the phases. Table 2 collects the geometric parameters for three types of intramolecular H-bonds, namely $O\cdots H{-}N$, $O\cdots H{-}C$ and $Se\cdots H{-}C_8$. A delicate balance among these intramolecular H-bonds seems to play a crucial role in imparting the observed deviations of the peptide planes from planarity as well as in determining the energetics of the dipeptides. The strength of the H-bonds is assessed by considering two geometric criteria, (a) the shorter the distance A—H···B, the stronger the H-bond and (b) the closer the angle A—H···B to 180° the stronger the H-bond; where A—H is H-bond donor and B is H-bond acceptor [50,51]. The A—H···B angle of all the intra- and intermolecular H-bond interactions reported in this study range from 85° to 180°, while their H···B distances span up to a maximum value of 2.75 Å [10,51]. In $Se\cdots H{-}C_8$ type of interactions, the H···B distances may range up to 3.035 Å because of the large atomic size of selenium atom compared to that of oxygen. It is clear from Table 2 that the two H-bonds viz. $O_{10}\cdots H{-}C_8$ and $O_{10}\cdots H{-}N_1$, whose B···H distances range from 2.247 to 2.750 Å in gas phase while the aqueous phase distances range from 2.234 to 2.588 Å, are predicted to occur regularly in all the seven systems. On the other hand, the gas and solvent phase data on the three H-bonds $O_{SC}\cdots H{-}C_8$, $O_{SC}\cdots H{-}N_7$ and $Se\cdots H{-}C_8$ clearly indicates the role of –SC group's identity in determining the number and type of H-bond interactions existing in the dipeptide structures. Thus, the presence of $O_{SC}\cdots H{-}N_7$ only in the case of Pro–Thr, $Se\cdots H{-}C_8$ only in Pro–Sec and $O_{SC}\cdots H{-}C_8$ only in Pro-Asp and Pro–Asn can be explained on the basis of the identity of the–SC groups belonging to Thr, Sec, Asp and Asn residues. The solvent effects do not alter the gas phase intramolecular H-bond combinations existing in most of the dipeptide structures studied here (with only exception in the case of Pro-Thr). It is worth mentioning that the $O_{14}\cdots H{-}N_7$ bond present in the gas phase structure of Pro–Thr (the B···H distance is equal to 2.604 Å while the value of the A—H···B angle is 87.3°) gets significantly weakened due to the solvation effects of the aqueous phase (the angle A—H···B is reduced to 66.8° and the B···H distance is increased up to 3.022 Å); and consequently the $O_{14}\cdots H{-}N_7$ bond is considered to be absent in the aqueous phase structure of the Pro–Thr dipeptide [10]. The absence of the $O_{14}\cdots H{-}N_7$ bond in the aqueous phase structure of Pro–Thr can also be justified in view of our discussion on the conformational propensity of the dipeptides in gas and aqueous phase where we have recognized that the aqueous phase structure of Pro–Thr is close to the β

type conformer that lacks a $O_{14}\cdots H{-}N_7$ type of H-bond interaction [12].

3.4. Theoretical IR spectra

Table 3 lists some representative vibrational frequencies along with their intensity values (given in parentheses) which are sensitive to the structural changes corresponding to the changes in the identity of the varying C-terminal residues and solvent effects of implicit and explicit aqueous environments. The calculated quantum chemical harmonic vibrational frequencies are usually larger than their corresponding experimental values [52], and such discrepancies have been attributed to the neglect of anharmonicity effects in theoretical treatments, incomplete incorporation of electron correlation and the use of finite basis sets. Nevertheless, the theoretically predicted vibrational spectra (scaled with appropriate correction factors) of the seven dipeptides provide valuable information to understand the existence and nature of various types of intra- and intermolecular H-bonds in the dipeptides as well as their structural and molecular properties. From Table 3, we can see that some of the vibrational modes like the $\nu(N_7{-}H_{11})$, $\nu(N_1{-}H)$ stretch etc., remain more or less unchanged along with the structural changes in the dipeptide geometries while some are very sensitive to even small changes and consequently leave noticeable signature in the IR spectra. As it is mentioned in a preceding section of this paper that the geometry about the C_2 α-carbon atoms remains almost unaltered and those around C_8 α-carbon atoms are changed due to the changing identity of the −SC groups has been well reflected in the frequency values of $\nu(C_2{-}H)$ and $\nu(C_8{-}H)$ stretching modes. Thus, the variations in the $\nu(C_8{-}H)$ stretching values, ranging from 2937 to 3000 cm^{-1} in gas phase and 2969–3016 cm^{-1} in aqueous phase, can be attributed to the geometry changes around the C_8 α-carbon atoms caused by the changing identity of the −SC groups of the C-terminal residues. The aqueous phase frequency values of $\nu(C_6{=}O_{10})$ stretching are always lower than their corresponding gas phase values by a magnitude up to 47 cm^{-1} and such lowering in the frequency values of $\nu(C_6{=}O_{10})$ can be due to elongation of $C_6{=}O_{10}$ bonds in aqueous phase (elongated up to 0.008 Å). Similarly, shortening of $C_6{-}N_7$ bonds in aqueous phase (by a maximum magnitude of 0.009 Å) is also well reflected by the increased values of the $\nu(C_6{-}N_7)$ modes in aqueous phase. Table S6 of the SI presents a comparison of the gas phase IR assignments of the dipeptides considered in the present study with those obtained previously by various gas phase theoretical [33] and experimental [48] studies, which reveals that most of assignments are in good agreement with the earlier studies. On the other hand, in a previous experimental study of the vibrational spectral properties of the solid samples of aspartic acid and glutamic acid dipeptides [53], it was pointed out that the appearance of $\nu(C{=}O)$ and $\nu(N{-}H)$ vibrational modes of the CONH moieties around the range of 1727–1731 cm^{-1} and 3345–3346 cm^{-1} respectively could be used as evidence for the existence of intermolecular H-bonds in the CONH moieties. In view of the above mentioned observation, the effects of intermolecular H-bonds on the spectral properties of the CONH moieties of the dipeptides are well reflected by the DFT furnished results of this study. Because of the intermolecular H-bonding interactions established by the $C_6{=}O_{10}$ and $N_7{-}H_{11}$ bonds with the explicit water molecules, the $\nu(C_6{=}O_{10})$ and $\nu(N_7{-}H_{11})$ frequency values of the CONH moieties of Pro–Ala and Pro–Thr appearing at the range of 1697–1703 cm^{-1} and 3478–3491 cm^{-1} respectively in implicit aqueous phase are shifted toward the lower side of the frequency scale in the two dipeptide-water complexes Pro–Ala–W_3 and Pro–Thr–W_3 (appear at the range of 1665–1674 cm^{-1} and 3311–3291 cm^{-1}). Thus it can be concluded that B3LYP hybrid functional in combination with 6-311++G(d,p) basis set performs well in reproducing the experimental vibrational frequencies.

3.5. Effects of explicit water molecules

The intermolecular H-bond interactions formed by the water molecules with proteins play an important role in determining the three-dimensional structures adopted by proteins under physiological conditions [54]. For this reason, the interactions of water molecules with peptide moieties have been studied extensively, both experimentally [55,56] and theoretically [57–59]. In an effort to study the effects of explicit aqueous environment on the structural and molecular properties of dipeptides, the two systems Pro–Ala and Pro–Thr are complexed with three explicit water molecules via intermolecular H-bonds. The three explicit water molecules are placed in the vicinity of the peptide linkages of Pro–Ala and Pro–Thr, since one of the chief aims of this study is to investigate the effects of aqueous phase on the structural features of the amide planes. The intermolecular H-bonds present in Pro–Ala–W_3 and Pro–Thr–W_3 are depicted in Fig. 3 along with their intermolecular H-bond distances. The predicted total electronic energies (E) and Gibbs free energies (G) along with their corresponding ZPVE corrected values, E_{corr} and G_{corr} respectively, of Pro–Ala–W_3 and Pro–Thr–W_3 are listed in Table S1. Table 1 lists the results on the rotational constants, dipole moments and HOMO–LUMO energy gaps of Pro–Ala–W_3 and Pro–Thr–W_3. From these data, it is evident that the values of the total dipole moments for Pro–Ala–W_3 and Pro–Thr–W_3 are increased by 4.443 and 4.546 D respectively compared to those of Pro–Ala and Pro–Thr predicted in implicit aqueous phase. Similarly, the effects of the explicit aqueous environment are also evident on the HOMO–LUMO energy gaps of the two systems; the energy gap for Pro–Ala–W_3 is decreased while the same for Pro-Thr-W_3 is increased as compared to those of Pro–Ala and Pro–Thr respectively. The bond length values of the five bonds related to the amide planes of Pro–Ala–W_3 and Pro–Thr–W_3, listed in Table S2, suggest that the exposed polar bonds $C_6{=}O_{10}$ and $N_7{-}H_{11}$ are elongated up to 0.013 Å while the buried $C_6{-}N_7$ bonds are reduced by a maximum value of 0.011 Å because of the interaction of the explicit water molecules with the dipeptide structures. These deviations in the bond length values also cause significant variations in the values of the vibrational modes like $\nu(C_6{=}O_{10})$, $\nu(N_7{-}H_{11})$ and $\nu(C_6{-}N_7)$ of Pro–Ala–W_3 and Pro–Thr–W_3 systems (listed in Table 3). For example, the frequency value of $\nu(N_7{-}H_{11})$ mode is decreased up to 200 cm^{-1} in Pro–Thr–W_3 compared to that of Pro–Thr in aqueous phase. Among the six dihedral angles (listed in Table S4), considered to monitor the amide plane planarity and back-bone structural features of the dipeptides, the values of the ϕ and Ω dihedrals of Pro–Ala–W_3 are changed by 30.8° compared to that in Pro–Ala in aqueous phase. On the other hand, the values of the bond angles associated with the geometries of the amide planes and α-carbon atoms (listed in Table S3 and S5 respectively) remain more or less unchanged. The intramolecular H-bond combination (listed in Table 2) observed in Pro–Ala–W_3 and Pro–Thr–W_3 also remains unaltered.

4. Conclusions

This quantum mechanical study investigates the role of C-terminal residues in influencing the structural features of dipeptides containing proline at their N-terminal positions. In aqueous phase all the seven dipeptide geometries exhibit larger values of total dipole moments and greater HOMO–LUMO energy gaps than those in gas phase. The negative values of solvation enthalpies (E_{Sol}) and solvation free energies (G_{Sol}) indicate the enhanced thermodynamic stability of the dipeptides in aqueous phase. This study

predicts very little changes in the bond length and bond angle values of the amide planes as the C-terminal residue of the dipeptides changes. The gas and aqueous phase values on the four dihedral angles related to the amide plane geometries suggest that the planarity of a given amide plane mainly depends on two factors – (i) size and type of functional group/groups present in the side chain moiety of the C-terminal residue (—SC group) and (ii) intramolecular H-bond forming ability of the H- and O-atoms of the amide plane with their adjacent moieties belonging to the C- and N-terminal residues. The ϕ dihedral values are also sensitive to the identity of the —SC group of the C-terminal residues. The variations in the values of the $\nu(C_8—H)$ stretching frequencies of the dipeptides reflect the effects of the changing —SC groups on the geometry around the C_8 atoms. The $\nu(C_2—H)$ stretching values remain relatively unchanged since the geometry around the C_2 α-carbon atoms are not affected by the changing identity of the–SC groups. The gas phase intramolecular H-bond combinations of the dipeptides are more or less similar to those in the aqueous phase; however, the absence of $O_{SC}\cdots H—C_8$, $O_{SC..}H—N_7$ and $Se\cdots H—C_8$ bonds in the cases of Pro–Ala Pro–Leu and Pro–Arg clearly indicates the role of —SC group's identity in determining the number and type of H-bond interactions existing in the dipeptide structures. Intermolecular H-bond interactions with explicit water molecules greatly modify the structural features of the amide planes and other molecular properties like dipole moments, HOMO–LUMO energy gaps and vibrational spectra of the dipeptides.

Acknowledgments

Financial assistance from the Special Assistance Program of the University Grants Commission to the Department of Chemistry, NEHU, is gratefully acknowledged. GD is thankful to Council of Scientific and Industrial Research, New Delhi, India, for generous allocation of computational facilities through the Research Project No. 37(1481)/11/EMR-II. SM is also grateful to the University Grants Commission, Government of India, New Delhi, for financial assistance through a research fellowship.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2013.12.009.

References

[1] L.M. Gierasch, J. King (Eds.), Protein Folding, Deciphering the Second Half of the Genetic Code, American Association for the Advancement of Science, Washington, 1990.
[2] O. Antohi, F. Naider, A.M. Sapse, Ab initio calculations on Pro-Ala and Pro-Gly dipeptides, J. Mol. Struct. (Theochem.) 360 (1996) 99–108.
[3] C. Kapota, G. Ohanessian, The low energy tautomers and conformers of the dipeptides HisGly and GlyHis and of their sodium ion complexes in the gas phase, Phys. Chem. Chem. Phys. 7 (2005) 3744–3755.
[4] T. Kolev, B.B. Koleva, M. Spiteller, Spectroscopic, theoretical and structural characterization of hydrogensquarates of L-threonyl-Ll-serine and L-serine, Amino Acids 33 (2007) 719–725.
[5] B.B. Koleva, T.M. Kolev, S. Todorov, Structural and spectroscopic analysis of dipeptide L-methionyl-glycine and its hydrochloride, Chem. Pap. 61 (2007) 490–496.
[6] C. Clavaguera, F. Piuzzi, J.P. Dognon, Electronic spectrum of Tryptophan-Phenylalanine. A correlated *Ab initio* and time-dependent density functional theory study, J. Phys. Chem. B 113 (2009) 16443–16448.
[7] S. Ghosh, S. Mondal, A. Misra, S. Dalai, Investigation on the structure of dipeptides: a DFT study, J. Mol. Struct. (Theochem.) 805 (2007) 133–141.
[8] C.D. Keefe, J.K. Pearson, Ab initio investigations of dipeptide structures, J. Mol. Struct. (Theochem.) 679 (2004) 65–72.
[9] I. Saada, J.K. Pearson, A theoretical study of the structure and electron density of the peptide bond, Comp. Theor. Chem. 969 (2011) 76–82.
[10] R. Vargas, J. Garza, B.P. Hay, D.A. Dixon, Conformational study of the alanine dipeptide at the MP2 and DFT levels, J. Phys. Chem. A 106 (2002) 3213–3218.
[11] D.J. Tobiast, C.L. Brooks III, Conformational equilibrium in the alanine dipeptide in the gas phase and aqueous solution: a comparison of theoretical results, J. Phys. Chem. 96 (1992) 3864–3870.
[12] Z.X. Wang, Y. Duan, Solvation effects on alanine dipeptide: a MP2/cc-pVTZ//MP2/6-31G** study of (Phi, Psi) energy maps and conformers in the gas phase, ether, and water, J. Comp. Chem. 25 (2004) 1699–1716.
[13] F.F. García-Prieto, I.F. Galvan, M.A. Aguilar, M.E. Martin, Study on the conformational equilibrium of the alanine dipeptide in water solution by using the averaged solvent electrostatic potential from molecular dynamics methodology, J. Chem. Phys. 135 (2011) 194502–194509.
[14] I.R. Gould, W.D. Cornell, I.H. Hillier, A quantum mechanical investigation of the conformational energetics of the alanine and glycine dipeptides in the gas phase and in aqueous solution, J. Am. Chem. Soc. 116 (1994) 9250–9256.
[15] T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III, J. Pople, Theoretical study of blocked glycine and alanine peptide analogs, J. Am. Chem. Soc. 113 (1991) 5989–5997.
[16] C. Adamo, V. Dillet, V. Barone, Solvent effects on the conformational behavior of model peptides. A comparison between different continuum models, Chem. Phys. Lett. 263 (1996) 113–118.
[17] G. Das, Investigations of dipeptide structures containing pyrrolysine as N-terminal residues: a DFT study in gas and aqueous phase, J. Mol. Model. 19 (2013) 1901–1911.
[18] S. Mandal, G. Das, Structure of dipeptides having N-terminal selenocysteine residues: a DFT study in gas and aqueous phase, J. Mol. Model. 19 (2013) 2613–2623.
[19] G.N. Ramachandran, Need for nonplanar peptide units in polypeptide chains, Biopolymers 6 (1963) 1494–1496.
[20] G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, Stereochemistry of polypeptide chain configurations, J. Mol. Biol. 7 (1963) 95–99.
[21] D.Q. Mc Donald, W.C. Still, Molecular mechanics parameters and conformational free energies of proline-containing peptides, J. Org. Chem. 61 (1996) 1385–1391.
[22] M. MacArthur, J.M. Thornton, Influence of proline residues on protein conformation, J. Mol. Biol. 218 (1991) 397–412.
[23] J.S. Richardson, D.C. Richardson, Amino acid preferences for specific locations at the ends of alpha helices, Science 240 (1988) 1648–1652.
[24] E.J. Milner-White, L.H. Bell, P.H. Maccallum, Pyrrolidine ring puckering in *cis* and *trans*-proline residues in proteins and polypeptides: different puckers are favoured in certain situations, J. Mol. Biol. 228 (1992) 725–734.
[25] M. Ramek, A.-M. Kelterer, S. Nikolic, Ab initio and molecular mechanics conformational analysis of neutral L-proline, Int. J. Quant. Chem. 65 (1997) 1033–1045.
[26] S.G. Stepanian, I.D. Reva, E.D. Radchenko, L. Adamowicz, Conformers of nonionized proline. Matrix-isolation infrared and post-hartree-fock *ab initio* study, J. Phys. Chem. A 105 (2001) 10664–10672.
[27] Y.K. Kang, Ab initio and DFT conformational study of proline dipeptide, J. Mol. Struct. (Theochem.) 675 (2004) 37–45.
[28] B. Maigret, D. Perahia, B. Pulman, Molecular orbital calculations on the conformation of polypeptides and proteins. IV. The conformation of the prolyl and hydroxyprolyl residues, J. Theor. Biol. 29 (1970) 275–291.
[29] H.A. Baldoni, A.M. Rodrigues, M.A. Zamora, G.N. Zamarbide, R.D. Enriz, O. Farkas, P. Csaszar, L.L. Torday, C.P. Sosa, I. Jakli, A. Perzel, J. Gy Papp, M. Hollosi, I.G. Csizmadia, Peptide models – XXIV: an *ab initio* study on N-formyl-L-prolinamide with trans peptide bond. The existence or non-existence of alpha(L) and epsilon(L) conformations, J. Mol. Struct. (Theochem.) 465 (1999) 79–91.
[30] I. Hudaky, H.A. Baldoni, A. Perczel, Peptide Models XXXVIII. Proline conformers from X-ray crystallographic database and from *ab initio* computations, J. Mol. Struct. (Theochem.) 582 (2002) 233–249.
[31] E. Czinki, A.G. Csaszar, Conformers of gaseous proline, Chem. Eur. J. 9 (2003) 1008–1019.
[32] W.D. Allen, E. Czinki, A.G. Csaszar, Molecular structure of proline, Chem. Eur. J. 10 (2004) 4512–4517.
[33] S. Kecel, A.E. Ozel, S. Akyuz, S. Celik, G. Agaeva, Conformational analysis and vibrational spectroscopic investigation of L-proline–tyrosine (L-Pro–Tyr) dipeptide, J. Mol. Struct. 993 (2011) 349–356.
[34] S. Miertus, E. Scrocco, J. Tomasi, Electrostatic interaction of a solute with a continuum. A direct utilizaion of *ab initio* molecular potentials for the prevision of solvent effects, Chem. Phys. 55 (1981) 117–129.
[35] E.S.J. Arner, Selenoproteins – what unique properties can arise with selenocysteine in place of cysteine? Exp. Cell. Res. 316 (2010) 1296–1303.
[36] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[37] C. Lee, W. Yang, R.G. Parr, Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B 37 (1988) 785–789.
[38] M.J. Frisch, et al., Gaussian 03, Revision B.04, Gaussian, Pittsburgh, PA, 2003.
[39] M.J. Frisch, et al., Gaussian 09, Revision D02, Gaussian, Wallingford, CT, 2009.
[40] Z.A. Tehrani, E. Tavasoli, A. Fattahi, Conformational behavior and potential energy profile of gaseous histidine, J. Mol. Struct. (Theochem.) 960 (2010) 73–85.
[41] I.R. Gould, I.H. Hillier, Solvation of alanine dipeptide: a quantum mechanical treatment, J. Chem. Soc., Chem. Commun. (1993) 951–952.
[42] M.P. Andersson, P. Uvdal, New scale factors for harmonic vibrational frequencies using the B3LYP density functional method with the triple-(basis set 6-311+G(d,p), J. Phys. Chem. A 109 (2005) 2937–2941.

[43] J.B. Foresman, A. Frisch, Exploring Chemistry With Electronic Structure Methods, 2nd ed., Gaussian, Inc., Pittsburgh, PA, 1996.
[44] F. Freeman, K.T. Le, A Computational study of conformations and conformers of 1,3-dithiane (1,3-dithiacyclohexane), J. Phys. Chem. A 107 (2003) 2908–2918.
[45] J. Casanova, D.R. Kent IV., W.A. Goddard III, J.D. Roberts, Quantum-mechanical calculations of the stabilities of fluxional isomers of $C_4H_7^+$ in solution, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 15–19.
[46] S.G. Stepanian, I.D. Reva, E.D. Radchenko, M.T.S. Rosado, M.L.T.S. Duarte, R. Fausto, L. Adamowicz, Matrix-isolation infrared and theoretical studies of the glycine conformers, J. Phys. Chem. A 102 (1998) 1041–1054.
[47] M.T. Baei, S.Z. Sayyed-Alangi, Effects of zinc binding on the structure and stability of glycylglycine dipeptide: a computational study, E – J Chem. 9 (2012) 1244–1250.
[48] Y. Grenie, M. Avignon, C. Garrigoulagrang, Molecular structure study of dipeptides isolated in an argon matrix by infrared spectroscopy, J. Mol. Struct. 24 (1975) 293–307.
[49] C. Cabezas, M. Varela, V. Cortijo, A.I. Jimenez, I. Pena, A.M. Daly, J.C. Lopez, C. Cativielab, J.L. Alonso, The alanine model dipeptide Ac-Ala-NH_2 exists as a mixture of C_7^{eq} and C_5 conformers, Phys. Chem. Chem. Phys. 15 (2013) 2580–2585.
[50] E. Arunan, G.R. Desiraju, R.A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D.C. Clary, R.H. Crabtree, J.J. Dannenberg, P. Hobza, H.G. Kjaergaard, A.C. Legon, B. Mennucci, D.J. Nesbitt, Definition of the hydrogen bond (IUPAC Recommendations 2011), Pure Appl. Chem. 83 (8) (2011) 1637–1641.
[51] D. Kaur, P. Sharma, P.V. Bharatam, M. Kaur, Understanding selenocysteine through conformational analysis, proton affinities, acidities and bond dissociation energies, Int. J. Quant. Chem. 108 (2008) 983–991.
[52] W.J. Hehre, L. Radom, P.v.R. Schleyer, J.A. Pople, Ab Initio Molecular Orbital Theory, Wiley, New York, 1986.
[53] J.T.L. Navarrete, V. Hernandez, F.J. Ramirez, Vibrational study of aspartic acid and glutamic acid dipeptides, J. Mol. Struct. 348 (1995) 249–252.
[54] P. Selvarengan, P.G. Kolandaivel, Molecular modeling of dipeptide and its analogous systems with water, J. Mol. Model. 10 (2004) 198–203.
[55] G. Sieler, R. Schweitzer-Stenner, The amide I mode of peptides in aqueous solution involves vibrational coupling between the peptide group and water molecules of the hydration shell, J. Am. Chem. Soc. 119 (1997) 1720–1726.
[56] T. Kameda, N. Takeda, S. Ando, I. Ando, D. Hashizume, Y. Ohashi, Structure of GlyGly peptide in the crystalline state as studied by X-ray diffraction and solid State ^{13}C NMR methods, Biopolymers 45 (1998) 333–339.
[57] S.G. Kalko, E. Guardia, J.A. Padro, Molecular dynamics simulation of the hydration of the alanine dipeptide, J. Phys. Chem. B 103 (1999) 3935–3941.
[58] K. Tazaki, K. Shimizu, Molecular dynamics simulations in aqueous solution: application to free energy calculation of oligopeptides, J. Phys. Chem. B 102 (1998) 6419–6424.
[59] N.A. Besley, J.D. Hirst, Ab initio study of the electronic spectrum of formamide with explicit solvent, J. Am. Chem. Soc. 121 (1999) 8559–8566.