International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 2 Issue 6 ǁ June. 2013 ǁ PP.20-30

Effects of Local Interactions on the Structural Features of Dipeptides: A Theoretical Study

Gunajyoti Das, Shilpi Mandal
Department of Chemistry, North Eastern Hill University, Shillong-793022, India.

ABSTRACT: *Structural features of small amino acid sequences are known to determine the dynamic properties and functional specificity of proteins and polypeptides. In this study, the effects of solvation and identity of the varying C-terminal residue on the energetics, structural features of the peptide planes, values of the ψ and ϕ dihedrals, geometry around the α-carbon atoms and theoretically predicted vibrational spectra of a series of dipeptides are analyzed. The dipeptide geometries, constructed by considering Phe, Trp, Met, Gly, Cys or Tyr as varying C-terminal residue and pyrrolysine as the fixed N-terminal residue, are subjected to full geometry optimization and vibrational frequency calculations at B3LYP/6-31++G(d,p) level in gas and simulated aqueous phase using a polarizable continuum model (PCM). The identity of the varying C-terminal residue influences the values of ϕ, planarity of the peptide planes and geometry around the C_{17} α-carbon atoms while the solvation effects are evident on the values of bond lengths and bond angles of the amide planes. The interplay of intramolecular H-bonds influences the planarity of the peptide planes, geometry around the α-carbon atoms and determines the energetics of the dipeptides.*

KEYWORDS *Dipeptides, Peptide planes, Solvation effects, Vibrational frequencies.*

I. INTRODUCTION

Proteins are polymers of amino acids joined together by peptide bonds. The biological functions of proteins are dependent on their three-dimensional (3D) structure [1] and the 3D structure of protein is almost exclusively dependent on the primary structure or the linear sequence of amino acid residues [2]. Thus a clear knowledge about the conformations of the dipeptides can help us resolve the 3D structure of proteins and ultimately their biological activities. Although it is difficult to implement theoretical or computational approaches directly to the large systems such as polypeptides or proteins, model systems like the solitary amino acids or dipeptides can easily be studied by the modern day cutting-edge computational techniques. The importance of gas-phase structural studies on dipeptides lies in the fact that such studies can provide us the opportunity to understand their intrinsic properties free from the solvent or crystal phase effects. However, it is of fundamental importance to determine the conformational details of a biological molecule in aqueous solution since the vast majority of biochemical processes occur in an aqueous environment. The low-energy structures and their related properties derived from such computations have a meaningful relationship to their presence and functional activities performed in the macromolecular context of real life systems. It has now been realized that computational techniques are indispensable in elucidating atomic level structural information about biologically active molecules owing to certain limitations of experimental techniques as pointed out in literatures [3-5]. Computational studies on dipeptides arising from the genetically encoded amino acids [6-10] have been performed with a view toward understanding the structural features of small amino acid sequences and their possible roles in imparting the 3D structure to proteins. Structural studies [6, 7, 10] on a series of dipeptides have pointed out that in most of the dipeptides the amide planes are never perfectly planar; and these observations have been explained in terms of the cumulative effect of steric hindrance of –R group and H-bonding. Besides serving as model systems, dipeptides themselves have been shown to play numerous key biological roles [11-15].

The effects of solvation on the conformations and energies of dipeptides have been well documented in several literatures [10, 16-21]. In these studies the energetics and structural features of the dipeptides in gas and solvent phases are analyzed to understand the effects of the surrounding environment on the stabilities and conformational preferences of the dipeptides. 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) [22, 23] conferring markedly different conformations to the dipeptides in the aqueous phase. It has also been reported that solvation effects can enhance the planarity of the peptide planes [17].

The goal of the present theoretical study is to obtain full knowledge about the effects of solvation and identity 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, theoretically predicted vibrational spectra, dipole moments, rotational constants and types of intramolecular H-bonding interactions that may play crucial roles in determining the structure and stability of the pyrrolysine containing dipeptides. The dipeptides are constructed by keeping pyrrolysine as a fixed component in the N-terminus whereas the component in the C-terminus is varied with six different combinations. The six different amino acids chosen for the C-terminal position are phenylalanine (Phe), tryptophan (Trp), methionine (Met), glycine (Gly), cysteine (Cys) and tyrosine (Tyr). All these amino acid residues are taken as neutral (non-ionic) species. The standard three letter abbreviations are used to represent an amino acids while a particular dipeptide is named by listing the N-terminal residue first. Thus, Pyl-Phe dipeptide corresponds to a structure in which pyrrolysine is in the N-terminal position and phenylalanine is in the C-terminal position. Fig. 1 schematically represents the chemical structures of the six dipeptides studied here. The atom numbering of the pyrrolysine molecule is given in accordance with the schemes used earlier in the various literatures [24]. The C_{13}–N_{16} is the peptide bond of a given dipeptide structure while C_{12} and C_{17} are the α-carbon atoms of the N- and C-terminal residues respectively. To facilitate a clear representation of the intramolecular hydrogen bond interactions present in the pyrrolysine dipeptides some of the hydrogen atoms are named as H_a or H_b. This DFT study on dipeptides of pyrrolysine in the gas as well as in simulated aqueous phase is expected to provide the opportunity to know the structural features of the dipeptides at an atomic level which in turn may help us to understand the dynamics and functional specificity of proteins, to synthesize a new generation of pyrrolysine analogues, in discovering the biosynthetic pathway of pyrrolysine and in understanding the nature of the genetic code or amino acid code which is still evolving [25].

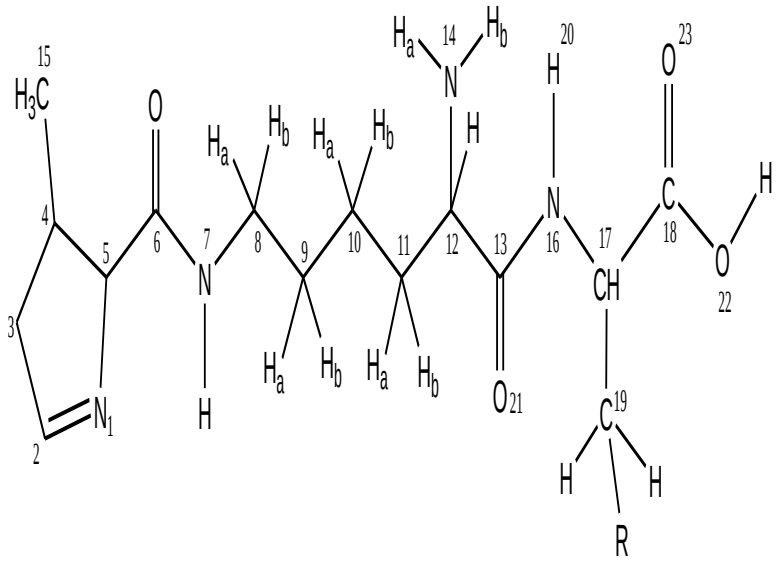


Fig. 1. Schematic representation of chemical structures of the dipeptides studied. -R= -C_6H_5, C_8NH_6, -CH_2SCH_3, –SH and -C_6H_4OH for Pyl-Phe, Pyl-Trp, Pyl-Met, Pyl-Cys and Pyl-Tyr respectively. Pyl-Gly does not have a side chain group.

II. COMPUTATIONAL METHODOLOGY

The molecular geometries of all the selected pyrrolysine dipeptides are subjected to full geometry optimization and vibrational frequency calculations using the B3LYP/6-31++G(d,p) level of theory [26, 27] of Gaussian 03 package [28]. These computations are conducted in the gas as well as in aqueous phase using a polarizable continuum model (PCM) [29]. The accuracy of self-consistent reaction field (SCRF) model in predicting the structure and energetics of dipeptides has already been justified in literature [30]. Absence of imaginary frequency values in the vibrational frequency calculations proves that the optimized geometries are precise minima. Zero point energy (ZPE) corrections were applied to the total energies of all the dipeptides using a correction factor of 0.97 [31]. The vibrational frequencies below 1800 cm^{-1} are scaled with 0.977 and for those above 1800 cm^{-1} a correction factor 0.955 is used [32, 33]. Use of diffuse functions is important to take into account the relative diffuseness of lone pair of electrons when a molecule under investigation contains lone pair of electrons [34] while polarization functions are useful in studying the conformational aspects where stereoelectronic effects play an important role [35].

III. RESULTS AND DISCUSSION

Investigations of the numerous parameters involved in dipeptide structure prediction have now been regarded as a pivotal part of the computational studies concerning the structure of protein and energetics of protein folding [36]. The geometrical parameters that have been considered in this study are expected to give a clear account of the effects of solvation and identity 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 and theoretically predicted vibrational spectra. The Table 1 presents the gas and aqueous phase data on total energies, rotational constants and dipole moments of the dipeptides calculated at B3LYP/6-31++G(d,p) level of theory. Table 2 and 3 list the values of the bond lengths and bond angles of the amide planes of the dipeptides respectively (the gas phase values are given in brackets). The four dihedral angles considered to monitor the planarity of the peptide planes of the dipeptides, *viz.* C_{12}–C_{13}–N_{16}–C_{17}, O_{21}–C_{13}–N_{16}–H_{20}, C_{12}–C_{13}–N_{16}–H_{20} and O_{21}–C_{13}–N_{16}–C_{17}, are listed in Table 4. Table 4 also lists the two well known Ramachandran backbone dihedral angles ψ (N_{14}–C_{12}–C_{13}–N_{16}) and ϕ (C_{13}–N_{16}–C_{17}–C_{18}) which are useful in studying the effects of solvation on the dipeptide structures as well as in predicting the overall structure of proteins. Table 5 represents the gas and aqueous phase data of the geometrical parameters considered to examine the geometry around the α-carbon atoms. Table 6 lists some important 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 7 lists some of the characteristic frequency and intensity values (given in brackets) of the dipeptides calculated at the B3LYP/6-31++G(d,p) level of theory. Fig. 3 represents the theoretical IR spectra of the six dipeptides both in gas and aqueous phase (scaled with a correction factor 0.955).

3.1. Dipeptide structure

As listed in Table 1 all the six dipeptide geometries exhibit large values of total dipole moments, ranging from 1.565 to 6.735 D in gas phase and 2.481 to 8.991 D in aqueous phase, indicating that they have greater polar character and consequently possess greater affinity to polar solvents. Thus, the data on the total energies of dipeptides correctly predicts that the dipeptide geometries are thermodynamically more stable in a strong polar solvent such as water than in gas phase by an energy difference that may range from 17.6 to 21.69 kcal/mol. The accuracy of DFT method in predicting the rotational constants of conformers of some aliphatic amino acids has been discussed in literatures [37, 38]. In the absence of any experimental data on rotational constants and dipole moments, these theoretically predicted values may assist experimentalists in determining the other conformers of the six dipeptides studied here.

Table 1

Calculated total energies[a] (kcal/mol), rotational constants (GHZ) and dipole moments (Debye) of the pyrrolysine dipeptides in gas and aqueous phase using B3LYP/6-31++G(d,p) level of theory.

Dipeptides	Phases	Total Energies	Rotational Constants			Dipole moments
			A	B	C	
Pyl-Phe	Aqueous	-839545.51	0.32481	0.03644	0.03472	6.159
	Gas	-839527.08	0.29974	0.03798	0.03614	4.920
Pyl-Trp	Aqueous	-922097.45	0.18927	0.04454	0.03914	8.991
	Gas	-922075.76	0.23282	0.03290	0.03105	4.237
Pyl-Met	Aqueous	-993774.39	0.29274	0.04233	0.03919	2.481
	Gas	-993755.87	0.28997	0.04259	0.03946	1.565
Pyl-Gly	Aqueous	-669943.75	0.97850	0.05638	0.05575	6.536
	Gas	-669926.15	0.95460	0.05657	0.05594	5.268
Pyl-Cys	Aqueous	-944463.78	0.50389	0.04627	0.04516	8.145
	Gas	-944444.65	0.48071	0.04721	0.04616	6.735
Pyl-Tyr	Aqueous	-886750.94	0.26736	0.03445	0.03240	2.845
	Gas	-886730.68	0.26514	0.03467	0.03264	2.432

[a]ZPVE corrected; Scaled with a correction factor 0.97

Table 2

Calculated bond lengths (in angstrom) for the peptide planes of the pyrrolysine dipeptides; the gas phase values are given in brackets.

Dipeptides	C_{12}–C_{13}	C_{13}=O_{21}	C_{13}–N_{16}	N_{16}–H_{20}	N_{16}–C_{17}
Pyl-Phe	1.539 (1.541)	1.239 (1.231)	1.353 (1.360)	1.017 (1.016)	1.452 (1.455)
Pyl-Trp	1.538 (1.540)	1.240 (1.231)	1.350 (1.359)	1.016 (1.016)	1.446 (1.456)
Pyl-Met	1.538 (1.540)	1.240 (1.233)	1.350 (1.356)	1.017 (1.016)	1.450 (1.449)
Pyl-Gly	1.538 (1.539)	1.239 (1.230)	1.352 (1.361)	1.016 (1.015)	1.441 (1.440)
Pyl-Cys	1.539 (1.541)	1.237 (1.229)	1.356 (1.362)	1.018 (1.016)	1.453 (1.455)
Pyl-Tyr	1.538 (1.540)	1.240 (1.232)	1.350 (1.357)	1.017 (1.016)	1.451 (1.451)
Average	1.538 (1.540)	1.240 (1.231)	1.352 (1.359)	1.017 (1.016)	1.449 (1.451)
MD[a]	0.001 (0.001)	0.003 (0.002)	0.004 (0.003)	0.001 (0.001)	0.008 (0.011)

[a]Maximum deviation from average values

It is evident from the Table 2, which lists the gas and aqueous phase bond length values of the five bonds of the amide planes *i.e.* C_{12}–C_{13}, C_{13}=O_{21}, C_{13}–N_{16}, N_{16}–H_{20} and N_{16}–C_{17}, that very little variance in the bond length values of an amide plane results as the identity of the C-terminal residue of a given dipeptide changes. Maximum deviations of 0.011 Å in gas phase and 0.008 Å in aqueous phase from their respective average values indicate that the bond lengths are essentially fixed. However, due to solvation effects the aqueous phase bond length values of the above mentioned bonds deviate from their respective gas phase values. For example, in aqueous phase the exposed C_{13}=O_{21} bonds are elongated up to 0.009 Å while the buried C_{13}–N_{16} bonds are shortened by a range of 0.006 to 0.009 Å for all the systems. The N_{16}–H_{20} bonds are elongated up to 0.002 Å in five systems whereas for Pyl-Trp system the bond length value of the N_{16}–H_{20} bond remains unaltered. Table 3 lists the values of the six bond angles of the amide planes *i.e.* C_{12}–C_{13}–O_{21}, C_{12}–C_{13}–N_{16}, O_{21}–C_{13}–N_{16}, C_{13}–N_{16}–C_{17}, C_{13}–N_{16}–H_{20} and H_{20}–N_{16}–C_{17}; and the data in both the phases indicates very little changes in the bond angle values as the individuality of the C-terminal residue of the dipeptides changes. Maximum deviations of 1.2° in gas phase and 0.7° in aqueous phase indicate that the bond angles are also essentially fixed. The solvent effects on these bond angles are quite apparent when their aqueous phase data is compared with the corresponding gas phase values; a maximum deviation up to 1.9° is observed for the angle C_{13}–N_{16}–C_{17} in Pyl-Phe system.

Table 3

Calculated bond angles (in degrees) for the peptide planes of the pyrrolysine dipeptides; the gas phase values are given in brackets.

Dipeptides	C_{12}–C_{13}–O_{21}	C_{12}–C_{13}–N_{16}	O_{21}–C_{13}–N_{16}	C_{13}–N_{16}–C_{17}	C_{13}–N_{16}–H_{20}	H_{20}–N_{16}–C_{17}
Pyl-Phe	121.0 (120.9)	115.0 (115.2)	124.1 (123.9)	124.0 (122.1)	115.2 (116.4)	120.8 (121.1)
Pyl-Trp	121.1 (120.8)	115.4 (115.3)	123.5 (123.9)	123.3 (121.9)	115.5 (116.5)	121.1 (121.2)
Pyl-Met	121.0 (121.0)	115.4 (115.5)	123.6 (123.5)	123.4 (121.8)	116.5 (117.8)	120.1 (120.4)
Pyl-Gly	121.2 (121.1)	115.3 (115.0)	123.5 (123.9)	123.4 (122.3)	115.5 (115.5)	121.1 (121.7)
Pyl-Cys	121.2 (121.1)	114.8 (115.0)	124.0 (124.0)	123.8 (122.5)	115.1 (116.3)	121.1 (121.1)
Pyl-Tyr	121.0 (120.9)	115.3 (115.5)	123.7 (123.7)	123.5 (122.0)	116.4 (117.5)	120.1 (120.3)
Average	121.1 (121.0)	115.2 (115.3)	123.7 (123.8)	123.6 (122.1)	115.7 (116.7)	120.7 (121.0)
MD[a]	0.1 (0.2)	0.4 (0.3)	0.4 (0.3)	0.4 (0.4)	0.7 (1.2)	0.6 (0.7)

[a]Maximum deviation from average values

Table 4

Calculated dihedral angles (in degrees) for the peptide planes of the pyrrolysine dipeptides at B3LYP/6-31++G(d,p) level of theory; the gas phase values are given in brackets.

Systems	-SC Groups	C_{12}–C_{13}–N_{16}–C_{17}	O_{21}–C_{13}–N_{16}–H_{20}	C_{12}–C_{13}–N_{16}–H_{20}	O_{21}–C_{13}–N_{16}–C_{17}	ψ	ϕ
Pyl-Phe	$-CH_2C_6H_5$	-178.1 (-176.2)	-178.8 (175.9)	2.7 (-3.0)	0.4 (2.7)	-17.1 (-14.1)	-111.1 (-140.9)
Pyl-Trp	$-CH_2C_8N_1H_6$	-178.9 (-177.4)	-176.7 (175.6)	4.9 (-3.7)	-0.4 (1.9)	-19.2 (-13.0)	-84.1 (-143.6)
Pyl-Met	$-(CH_2)_2SCH_3$	-179.1 (-179.4)	-179.0 (-179.2)	2.1 (1.8)	-0.2 (-0.5)	-16.6 (-15.0)	-144.6 (-149.7)
Pyl-Gly	-H	-179.2 (177.8)	-177.3 (-176.7)	4.5 (5.2)	-0.9 (-4.0)	-18.3 (-18.0)	-93.4 (-104.4)
Pyl-Cys	$-CH_2SH$	-178.2 (-178.2)	-179.0 (177.5)	2.4 (-1.3)	0.4 (0.6)	-16.5 (-14.8)	-115.3 (-136.8)
Pyl-Tyr	$-CH_2C_6H_4OH$	-177.5 (-177.0)	179.8 (177.2)	0.9 (-1.9)	1.3 (2.1)	-16.9 (-13.9)	-145.1 (-151.0)

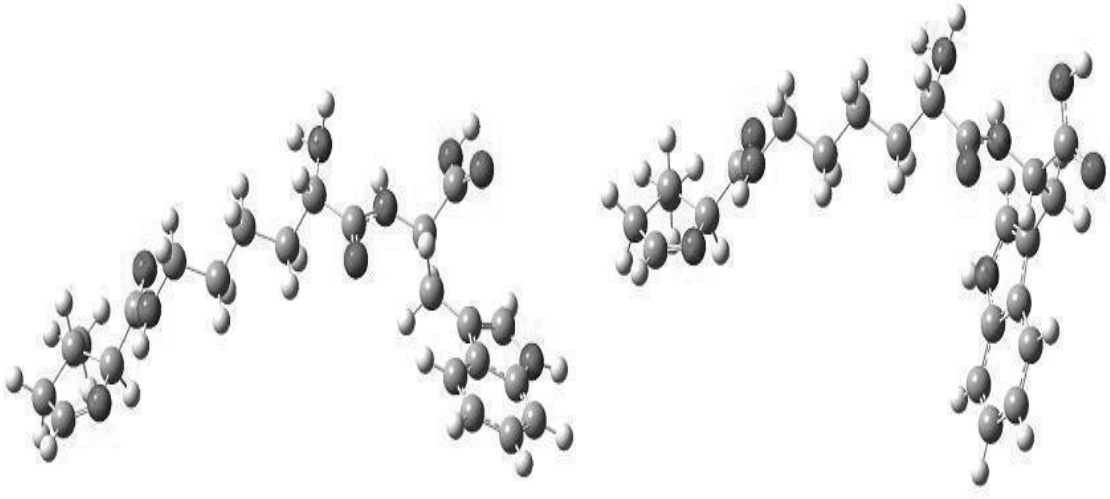
MD[a]	2.5 (3.8)	2.7 (4.4)	4.9 (5.2)	1.3 (4.0)

[a]Maximum deviation from expected values

Investigating the four dihedral angles of the dipeptides *viz.* C_{12}–C_{13}–N_{16}–C_{17}, O_{21}–C_{13}–N_{16}–H_{20}, C_{12}–C_{13}–N_{16}–H_{20} and O_{21}–C_{13}–N_{16}–C_{17}, listed in Table 4, can provide valuable information regarding the planarity of the peptide planes. The values of the two dihedral angles C_{12}–C_{13}–N_{16}–C_{17} and O_{21}–C_{13}–N_{16}–H_{20} should be close to 180° and those for the other two *i.e.* C_{12}–C_{13}–N_{16}–H_{20} and O_{21}–C_{13}–N_{16}–C_{17} should be close to 0° if indeed the amide plane is planar. The data presented in Table 4 shows that in aqueous phase the values of the four dihedral angles deviate up to a maximum value of 4.9° from the expected values whereas in the gas phase the maximum deviation observed is 5.2°. Thus, these dihedral angles do not deviate dramatically from their expected values in both the phases, however, the extent of deviations observed in the values of the four dihedral angles obviously suggests that the geometry of the amide planes are not perfectly planar regardless of whether the systems are in gas phase or in strong polar solvents like water. A previous observation that solvation effects can enhance the planarity of the peptide planes [17] is evident in the cases of Pyl-Phe, Pyl-Gly and Pyl-Tyr out of the six systems considered in this paper. However, in the other three systems Pyl-Trp, Pyl-Met and Pyl-Cys the solvent effects could not enhance the planarity of the amide planes. We expect that the conformations of the six dipeptides predicted at B3LYP/6-31++G(d,p) level are reliable since it has been pointed out that full geometry optimization of gaseous tryptophan conformers at B3LYP/6-311G(d) and MP2/6-311++G(d,p) levels do not produce any noticeable structural changes, only the conformer energies change by small amounts [39]. Therefore, it is reasonable to assume that solvation effects cannot drastically improve the planarity of the amide planes and the extent of the deviations from planarity primarily depends on two factors – (a) steric interactions of the side chain moieties of the C-terminal residues (–SC group) and (b) intramolecular H-bond formation by the H- and O-atoms of the amide planes with their adjacent moieties belonging to the C- and N-terminal residues. The intramolecular H-bond interactions that play crucial roles in deviating the amide planes from planarity and in imparting the observed conformations to the dipeptides in gas and aqueous phase are listed in Table 6 and a discussion on these interactions is also offered in a succeeding section of this paper.

Table 4 also lists the –SC groups of the C-terminal residues of the dipeptides as well as the gas and aqueous phase values of the ψ and ϕ angles. A thorough analysis of the dipeptide structures reveals that both size as well as the type of functional groups present in a –SC group may influence the planarity of a given amide plane. A large sized –SC group may compete for its physical space requirements to accommodate itself in between the amide plane and carboxylic group of the C-terminal residue of a given dipeptide and consequently influence the planarity of the amide plane.

The gas phase values of the ϕ angles reveal that the value of ϕ increases as the size of a given –SC group increases. This point has been well discussed in various literatures [6, 7]. Among the six dipeptides studied in this paper the ϕ value of Pyl-Gly in gas phase is -104.4° while in the other five systems, which are having much bigger sized –SC groups compared to that of Pyl-Gly, the ϕ values are above -136.8°. On the other hand, the –SC groups, depending on the type of functional groups present in them, may exert electrostatic repulsive or electrostatic attractive forces on their neighboring atoms belonging to the peptide planes and the carboxylic group of the C-terminal residues of the dipeptides which may also influence the values of the ϕ as well as planarity of the amide planes. The aqueous phase values of the ϕ angles of the dipeptides reveal that in solvent phase the type of functional groups present in the –SC groups is more important in influencing the ϕ values of the dipeptides than the size of the -SC groups. For example, in aqueous phase the ϕ value of Pyl-Trp is smaller than that of Pyl-Gly even though the –SC group of Trp is much bigger in size than Gly. As shown in Fig. 2, in aqueous phase the $-CH_2C_8N_1H_6$ group of Trp residue adopts a different orientation from that in gas phase allocating more physical space to the carboxylic group of the Trp residue. It has been suggested that polar solvents remarkably influence the conformational properties of dipeptides, by weakening the intra-residue hydrogen bonds and leading to the appearance of new energy minima [19-21]. Thus, as a result of this new orientation of the $-CH_2C_8N_1H_6$ group of Trp residue adopted in aqueous phase the ϕ value is reduced to -84.1° (-143.6° in gas phase).



Pyl-Trp (Gas Phase) Pyl-Trp (Aqueous Phase)

Fig. 2. Optimized structures of Pyl-Trp system in gas and aqueous phase

3.2. α-Carbon geometry

Since the protein structures usually contain thousands of amino acid residues, the geometries about the α-carbon atoms of the individual residues play important role in deciding the overall structure of the proteins. The three bond angles considered to monitor the geometry around the C_{12} α-carbon atoms of the dipeptides are N_{14}–C_{12}–C_{11}, N_{14}–C_{12}–C_{13} and C_{11}–C_{12}–C_{13} for C_{12} while N_{16}–C_{17}–C_{18}, N_{16}–C_{17}–C_{19} and C_{19}–C_{17}–C_{18} are the same for the C_{17} atoms. The α-carbon atoms of the amino acids are sp^3 hybridized and therefore the ideal bond angle should be 109.5°, however, this is not expected due to their stereogenic character. By monitoring the above mentioned bond angles around each α-carbon atom of the dipeptides one can get an idea about how the change in identity of the C-terminal residue can affect the geometries about these α-carbon atoms. This DFT study also provides us the opportunity to probe the effects of solvation on the geometries of the α-carbon atoms. Table 5 lists the gas and aqueous phase data on the bond angles about the α-carbon atoms. Maximum deviations of 0.1° in aqueous and 0.2° in gas phase from their respective average values suggest that the geometries about the C_{12} atoms do not change much with the change in the identity of the C-terminal residues. On the other hand, with maximum deviations up to 5.2° in aqueous and 6.1° in gas phase from their respective average values, the bond angles around the C_{17} change appreciably with the change in identity of the C-terminal residue of the dipeptides. These observations can be justified 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. The stereoelectronic effects of the varying –SC groups on the geometry of the C_{12} atoms are very little as they reside at a distance of four bonds away from these α-carbon atoms. On the contrary, since the varying –SC groups are situated adjacent to the C_{17} atoms the geometry around them are affected by the changing identity of the –SC groups. The solvation effects are more prominent on the geometry of the C_{17} atoms, a maximum deviation up to 4.2° is observed for the angle N_{16}–C_{17}–C_{18} in Pyl-Trp system, than that on the C_{12} atoms where the maximum deviation predicted is 0.5° for the N_{14}–C_{12}–C_{13} angle of Pyl-Trp.

Table 5

Calculated bond angles (in degrees) for the α-carbon atoms of the pyrrolysine dipeptides; the gas phase values are given in brackets.

Dipeptides	α-carbon atoms C_{12}			α-carbon atoms C_{17}		
	N_{14}–C_{12}–C_{11}	N_{14}–C_{12}–C_{13}	C_{11}–C_{12}–C_{13}	N_{16}–C_{17}–C_{18}	N_{16}–C_{17}–C_{19}	C_{19}–C_{17}–C_{18}
Pyl-Phe	111.3 (111.2)	110.9 (111.2)	108.9 (108.7)	110.7 (109.4)	110.8 (111.5)	111.6 (112.1)
Pyl-Trp	111.1 (111.2)	110.8 (111.3)	109.0 (108.9)	113.9 (109.7)	111.3 (111.7)	109.0 (111.6)
Pyl-Met	111.2 (111.1)	111.0 (111.2)	108.9 (108.8)	108.2 (108.0)	111.0 (111.4)	111.5 (111.8)
Pyl-Gly	111.1 (111.2)	110.8 (111.0)	109.0 (109.0)	116.4 (116.2)	110.0 (111.4)	106.8 (106.7)
Pyl-Cys	111.3 (111.2)	110.9 (111.2)	108.8 (108.8)	109.5 (109.0)	110.2 (110.5)	112.9 (112.7)
Pyl-Tyr	111.1 (111.2)	111.0 (111.2)	108.8 (108.7)	108.2 (108.1)	111.0 (111.1)	110.9 (110.8)
Average	111.2 (111.2)	110.9 (111.2)	108.9 (108.9)	111.2 (110.1)	110.7 (111.3)	110.5 (111.0)
MD[a]	0.1 (0.1)	0.1 (0.2)	0.1 (0.2)	5.2 (6.1)	0.7 (0.8)	3.7 (4.3)

[a]Maximum deviation from average values; For Pyl-Gly the -SC Group is a H-atom

3.3. Intramolecular hydrogen bonds

Intramolecular hydrogen bonds (H-bonds), the strongest non-covalent interactions, play an important role in stabilizing the different conformations of a dipeptide molecule [7]. The strength of these H-bonds depends on two factors, (a) shorter is the distance A–H….B than the sum of their van der waals radii and (b) closer the angle A–H….B to 180° [40], where A–H is H-bond donor and B is H-bond acceptor. Table 6 lists three types of intramolecular H-bonds, namely O...H–N, N...H–N and O…H–C, whose interplay is very crucial in imparting the observed deviations of the peptide planes from planarity as well as in determining the energetics of the pyrrolysine containing dipeptides. The gas phase intramolecular H-bond combinations of the dipeptides are similar to those in the aqueous phase. In gas phase the B….H distances of the four H-bonds $N_{14}...H_{20}$-N_{16}, $O_{21}...H_b$-C_{11}, $O_{21}...H$-C_{12} and $O_{21}...H$-C_{17} range from 2.150 to 2.703 Å while in aqueous phase they range from 2.125 to 2.739 Å. On the other hand, the gas and solvent phase data on the three H-bonds $O_{23}...H_{20}$-N_{16}, $O_{22}...H$-C_{17} and $O_{23}...H$-C_{17} clearly indicates the effects of size and the type of functional groups present in the –SC groups on the conformation of the dipeptides as well as on the number and type of H-bond interactions existing in the dipeptide molecules. For example, the absence of $O_{23}...H$-C_{17} and presence of $O_{23}...H_{20}$-N_{16} and $O_{22}...H$-C_{17} H-bonds only in the cases of Pyl-Met and Pyl-Tyr systems can be explained on the basis of identify of the –SC groups of the C-terminal residues.

Table 6

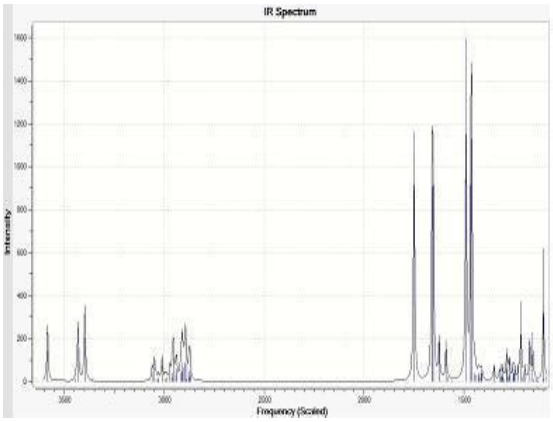
H-bond distances[a] (in angstrom) of the intramolecular H-bond interactions detected in the pyrrolysine dipeptides in gas and aqueous phases.

Dipeptides	Phases	$N_{14}...H_{20}$-N_{16}	$O_{21}...H_b$-C_{11}	$O_{21}...H$-C_{12}	$O_{21}...H$-C_{17}	$O_{23}...H_{20}$-N_{16}	$O_{22}...H$-C_{17}	$O_{23}...H$-C_{17}
Pyl-Phe	Aqueous	2.133	2.707	2.580	2.411	*abs*	*abs*	2.538
	Gas	2.173	2.658	2.585	2.440	*abs*	*abs*	2.554
Pyl-Trp	Aqueous	2.155	2.739	2.573	2.527	*abs*	*abs*	2.595
	Gas	2.176	2.648	2.586	2.456	*abs*	*abs*	2.559
Pyl-Met	Aqueous	2.165	2.708	2.580	2.492	2.450	2.494	*abs*
	Gas	2.193	2.669	2.579	2.488	2.411	2.530	*abs*
Pyl-Gly	Aqueous	2.148	2.725	2.579	2.504	*abs*	*abs*	2.690
	Gas	2.150	2.703	2.569	2.385	*abs*	*abs*	2.759
Pyl-Cys	Aqueous	2.125	2.704	2.586	2.396	*abs*	*abs*	2.536
	Gas	2.162	2.668	2.583	2.433	*abs*	*abs*	2.556
Pyl-Tyr	Aqueous	2.166	2.707	2.579	2.495	2.500	2.464	*abs*
	Gas	2.193	2.655	2.585	2.496	2.502	2.463	*abs*

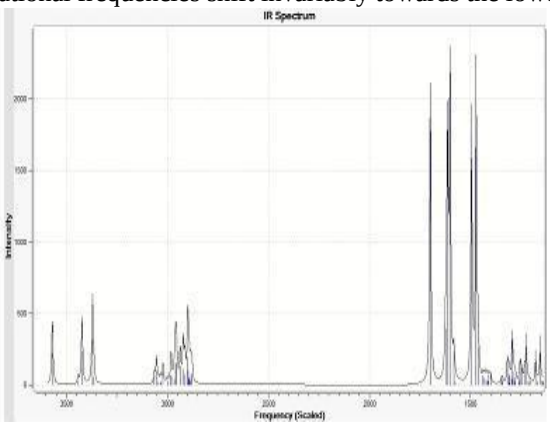
[a]Only the (B….H) distances are listed where B is H-bond acceptor; *abs*=absent

3.4. Vibrational spectra

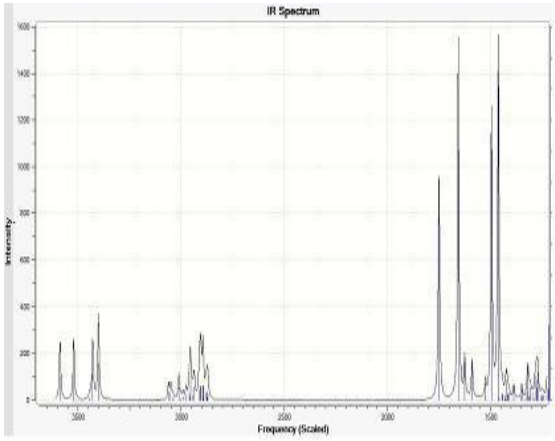
The theoretically predicted vibrational spectra of the six pyrrolysine dipeptides in both the phases provide valuable information to understand the existence and nature of various types of intramolecular H-bonds in the dipeptides. The Table 7 lists the characteristic frequency and intensity (given in brackets) values of only those vibrational modes which are sensitive to the structural changes caused by the varying C-terminal residues and solvent effects. It is evident from Table 7 that the vibrational frequencies shift invariably towards the lower



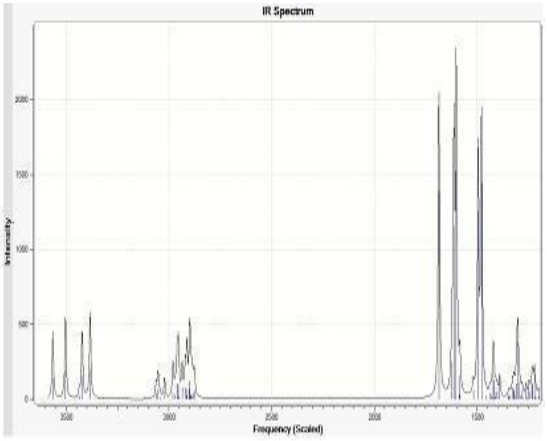
Pyl-Phe (Gas Phase)



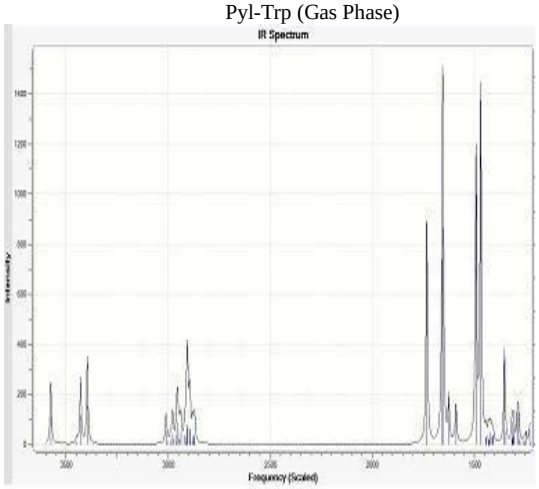
Pyl-Phe (Aqueous Phase)



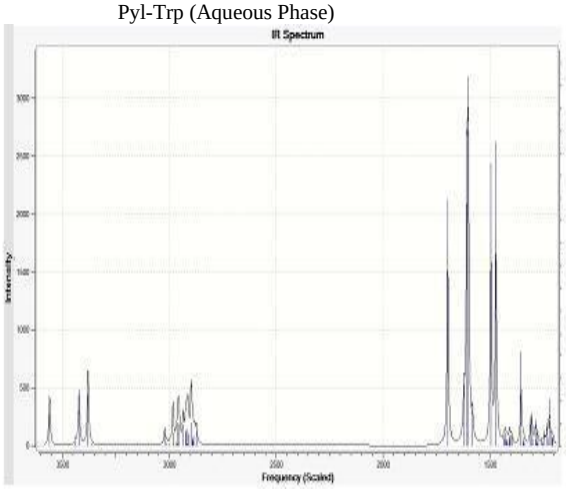
Pyl-Trp (Gas Phase)



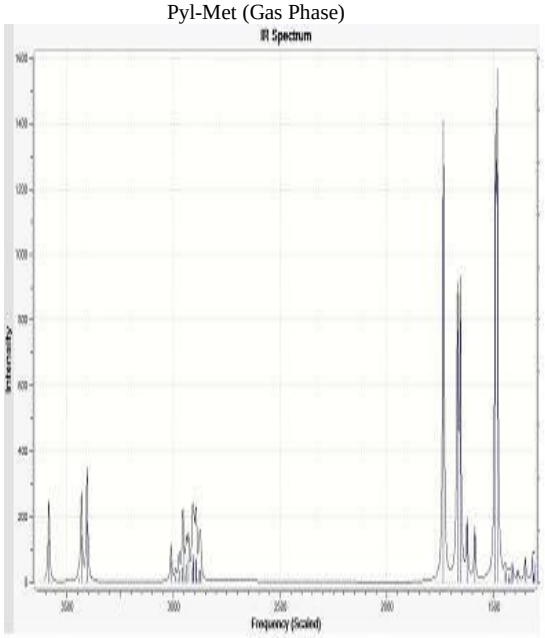
Pyl-Trp (Aqueous Phase)



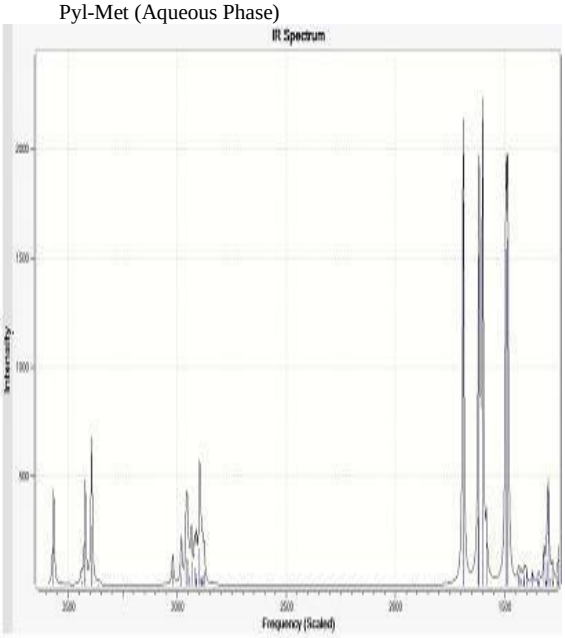
Pyl-Met (Gas Phase)



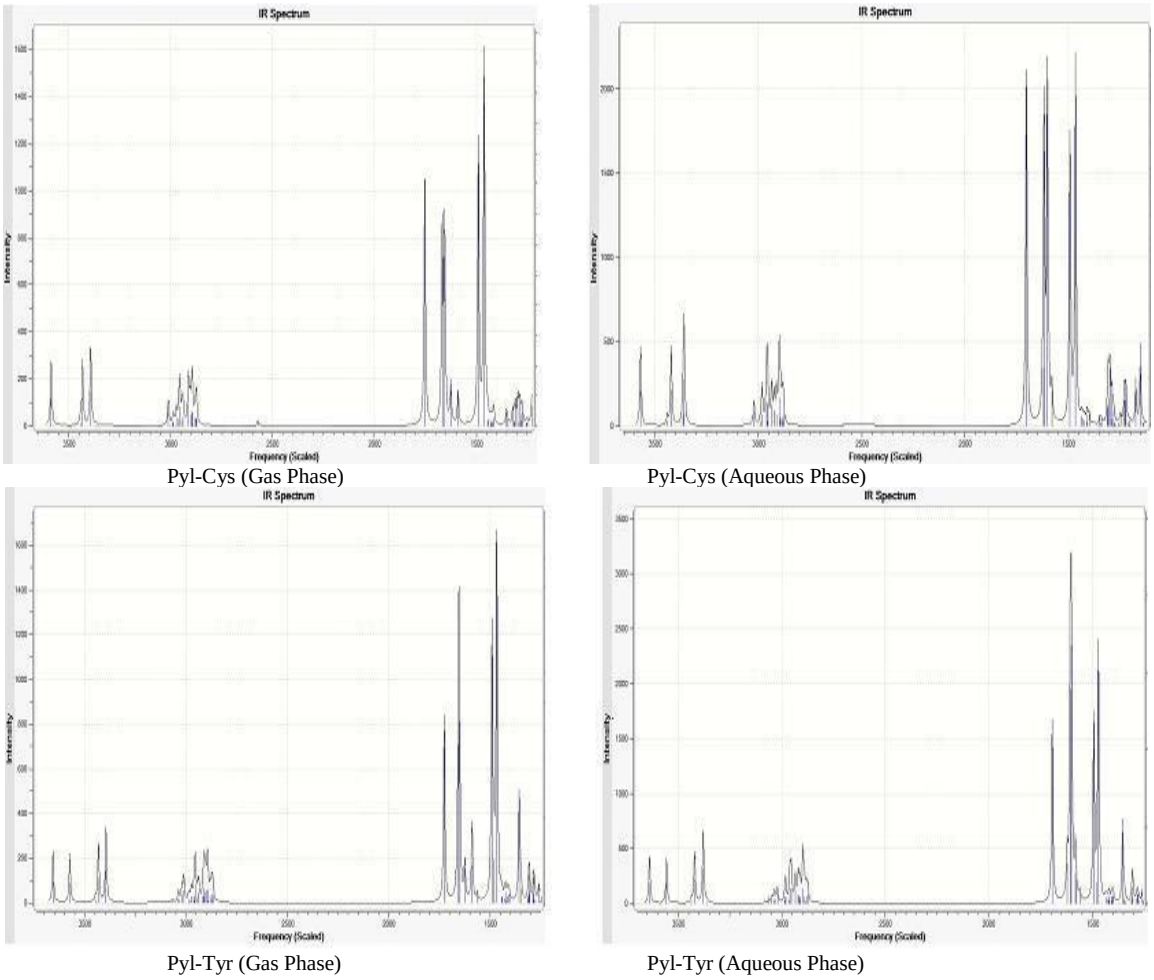
Pyl-Met (Aqueous Phase)



Pyl-Gly (Gas Phase)



Pyl-Gly (Aqueous Phase)



Pyl-Cys (Gas Phase) Pyl-Cys (Aqueous Phase)

Pyl-Tyr (Gas Phase) Pyl-Tyr (Aqueous Phase)

Fig. 3. Vibrational spectra of the dipeptides in gas and aqueous phase

side of frequency scale corresponding to the presence of intramolecular H-bond interactions. The shortening of $N_{14}...H_{20}$-N_{16} bonds in aqueous phase structures is well reflected by the lowering in the frequency values of the $\nu(N_{16}$-$H_{20})$ stretching by a range of 13 to 30 cm^{-1} than those in the gas phase. Solvent effects also lower the frequency values of the $\nu(C_{13}=O_{21})$, $\nu_s(N_{14}$-$H)$, $\nu_{as}(N_{14}$-$H)$ and $Sis(N_{14}$-$H)$ by a magnitude up to 62 cm^{-1} in the aqueous phase which can be due to elongation in the bond length values in solvent phase (the $C_{13}=O_{21}$ bonds are elongated up to 0.009 Å in the aqueous phase). The variations in the $\nu(C_{17}$-$H)$ stretching values can be attributed to the effects of the changing identity of the –SC groups of the C-terminal residues.

Table 7

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

Systems	Phases	$\nu(C_{13}=O_{21})$	$\nu(N_{16}$-$H_{20})$	$\nu(C_{13}$-$N_{16})$	$\nu_s(N_{14}$-$H)$	$\nu_{as}(N_{14}$-$H)$	$Sis(N_{14}$-$H)$	$\nu(C_{17}$-$H)$	$\nu(C_{12}$-$H)$
Pyl-Phe	Aqueous	1689 (496)	3528 (177)	1540 (571)	3520 (29)	3602 (16)	1656 (60)	3141 (11)	3013 (39)
	Gas	1738 (221)	3558 (101)	1534 (197)	3521 (1)	3606 (7)	1665 (45)	3138 (6)	3006 (14)
Pyl-Trp	Aqueous	1688 (491)	3543 (171)	1547 (560)	3516 (4)	3598 (14)	1658 (65)	3109 (21)	3010 (44)
	Gas	1735 (103)	3561 (106)	1531 (447)	3521 (1)	3606 (7)	1664 (45)	3141 (4)	3005 (15)
Pyl-Met	Aqueous	1643 (529)	3378 (200)	1505 (521)	3359 (6)	3438 (15)	1617 (66)	2969 (11)	2871 (45)
	Gas	1691 (280)	3395 (102)	1501 (419)	3364 (1)	3447 (7)	1625 (46)	2948 (12)	2867 (18)
Pyl-Gly	Aqueous	1692 (511)	3552 (194)	1555 (494)	3519 (4)	3601 (15)	1657 (64)	3084 (28)	3011 (40)
	Gas	1748 (251)	3565 (100)	1553 (380)	3520 (1)	3605 (7)	1662 (44)	3067 (33)	3011 (17)
Pyl-Cys	Aqueous	1643 (592)	3259 (200)	1505 (521)	3359 (6)	3438 (15)	1617 (66)	2969 (11)	2871 (45)
	Gas	1705 (243)	3278 (102)	1494 (465)	3363 (2)	3444 (8)	1625 (43)	2985 (7)	2874 (16)
Pyl-Tyr	Aqueous	1682 (587)	3537 (200)	1539 (662)	3517 (6)	3598 (14)	1655 (63)	3126 (12)	3009 (43)
	Gas	1730 (273)	3555 (99)	1536 (492)	3522 (1)	3609 (7)	1664 (49)	3122 (7)	3002 (19)

[a] The frequencies below 1800 cm^{-1} are scaled with 0.977 and for those above 1800 cm^{-1} a correction factor 0.955 is used
[b]Vibrational modes: ν=stretching; Sis= scissoring; s=symmetric; as=asymmetric

CONCLUSIONS

This DFT study on dipeptides containing pyrrolysine as a fixed component at their N-terminal positions predicts large values of total dipole moments for the dipeptides, 1.565 to 6.735 D in gas phase and 2.481 to 8.991D in aqueous phase, and as a consequence the aqueous phase structures show more thermodynamic stabilities by a range of 17.6 to 21.69 kcal/mol than those in the gas phase. The identity of the varying C-terminal residue influences the values of ϕ, planarity of the peptide planes and geometry around the C_{17} α-carbon atoms while the solvation effects are evident on the values of bond lengths and bond angles of the amide planes. The geometry of the amide planes are not perfectly planar regardless of whether the systems are in gas or in strong polar solvents like water and the deviations from planarity primarily depends on two factors – (a) steric interactions of the side chain moieties of the C-terminal residues and (b) intramolecular H-bond formation by the H- and O-atoms of the amide planes with their adjacent atoms belonging to the C- and N-terminal residues. In gas phase the ϕ values depend on the size of a given –SC group which is evident from the fact that the ϕ value of Pyl-Gly in gas phase is -104.4° while in the other five systems, which are having much bigger sized –SC groups compared to that of Pyl-Gly, the ϕ values are above -136.8°. However, in solvent phase the type of functional group present in the –SC groups is more important in influencing the ϕ values of the dipeptides than the size of the–SC groups which is evident from the fact that the ϕ value of Pyl-Trp is the smaller than that of Pyl-Gly even though the –SC group of Trp is much bigger in size than Gly. The geometry around the C_{17} atoms are affected by the changes in the identity of the –SC groups. The presence or absence of three types of intramolecular H-bonds, namely O...H–N, N...H–N and O…H–C that leave noticeable signatures in the IR spectra, play crucial roles in influencing the geometry of the peptide planes and in determining the energetics of the pyrrolysine dipeptides.

ACKNOWLEDGMENTS

Financial assistance from the Special Assistance Program of the University Grants Commission to the Department of Chemistry, NEHU, is gratefully acknowledged. SM is also grateful to the University Grants Commission, Government of India, New Delhi, for financial assistance through a research fellowship.

REFERENCES

[1] N.A. Campbell, (1996) In Biology, 4th ed., Cummings Publishing Company, New York.
[2] R.H. Garrett, (1999) In Biochemistry, Saunders College Publishing, New York, Chapter 6.
[3] M.R. Wormald, A.J. Petrescu, Y.L. Pao, A. Glithero, T. Elliott, R.A. Dwek, Chem. Rev. 102 (2002) 371.
[4] N. Foloppe, B. Hartmann, L. Nilsson, A.D. MacKerell, Jr. Biophysical Journal 82 (2002) 1554.
[5] J. Sponer, M. Zgarbova, P. Jurecka, K.E. Riley, J.E. Sponer, P. Hobza, J. Chem. Theory Comput. 5 (2009) 1166.
[6] S. Ghosh, S. Mondal, A. Misra, S. Dalai, J. Mol. Struct. (Theochem) 805 (2007) 133.
[7] C.D. Keefe, J.K. Pearson, J. Mol. Struct. (Theochem) 679 (2004) 65.
[8] I. Saada, J.K. Pearson, Computational and Theoretical Chemistry 969 (2011) 76.
[9] R. Vargas, J. Garza, B.P. Hay, D.A. Dixon, J. Phys. Chem. A 106 (2002) 3213.
[10] G. Das, J. Mol. Model. (2013) DOI 10.1007/s00894-013-1754-7.
[11] O. Antohi, F. Naider, A.M. Sapse, J. Mol. Struct. (Theochem) 360 (1996) 99.
[12] C. Kapota, G. Ohanessian, Phys. Chem. Chem. Phys. 7 (2005) 3744.
[13] T. Kolev, B.B. Koleva, M. Spiteller, Amino Acids 33 (2007) 719.
[14] B.B. Koleva, T.M. Kolev, S. Todorov, Chem. Papers 61 (2007) 490.
[15] C. Clavaguera, F. Piuzzi, J.P. Dognon, J. Phys. Chem. B 113 (2009) 16443.
[16] D.J. Tobiast, C.L. Brooks III, J. Phys. Chem. 96 (1992) 3864.
[17] Z.X. Wang, Y. Duan, J. Comput. Chem. 25 (2004) 1699.
[18] F.F. García-Prieto, I.F. Galvan, M.A. Aguilar, M.E. Martin, J. Chem. Phys. 135 (2011) 194502.
[19] I.R. Gould, W.D. Cornell, I.H. Hillier, J. Am. Chem. Soc. 116 (1994) 9250.
[20] T. Head-Gordon, M. Head-Gordon, M.J. Frisch, C.L. Brooks III, J. Pople, J. Am. Chem. Soc. 113 (1991) 5989.
[21] C. Adamo, V. Dillet, V. Barone, Chem. Phys. Lett. 263 (1996) 113.
[22] G.N. Ramachandran, Biopolymers 6 (1963) 1494.
[23] G.N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, J. Mol. Biol. 7 (1963) 95.
[24] B. Hao, G. Zhao, P.T. Kang, J.A. Soares, T.M. Ferguson, J. Gallucci, J.A. Krzycki, M.K. Chan, Chemistry & Biology 11 (2004) 1317.
[25] S. Osawa, T.H. Jukes, K. Watanabe, A. Muto, Microbiological reviews 56 (1992) 229.
[26] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
[27] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
[28] M.J. Frisch et al., Gaussian 03, Revision A.7, Gaussian Inc., Pittsburg, PA, 2003.
[29] S. Miertus, E. Scrocco, J. Tomasi, Chem. Phys. 55 (1981) 117.
[30] I.R. Gould, I.H. Hillier, J. Chem. Soc., Chem. Comm. (1993) doi:10.1039/C39930000951
[31] G. Das, S. Mandal, J. Mol. Model. (2013) DOI 10.1007/s00894-012-1740-5
[32] S. Kecel, A.E. Ozel, S. Akyuz, S. Celik, G. Agaeva, J. Mol. Struct. 993 (2011) 349.
[33] M.P. Andersson, P. Uvdal, J. Phys. Chem. A 109 (2005) 2937.
[34] J.B. Foresman, A. Frisch, Exploring Chemistry with Electronic Structure Methods, 2nd edition, Gaussian, Inc. Pittsburgh, PA 1996.

[35] F. Freeman, K.T. Le, J. Phys. Chem. A 107 (2003) 2908.
[36] G.A. Chasse, A.M. Rodriguez, M.L. Mak, E. Deretey, A. Perczel, C.P. Sosa, R.D. Enriz, I.G. Csizmadia, J. Mol. Struct. (Theochem) 537 (2001) 319.
[37] S.G. Stepanian, I.D. Reva, E.D. Radchenko, M.T.S. Rosado, M.L.T.S. Duarte, R. Fausto, L. Adamowicz, J. Phys. Chem. A 102 (1998) 1041.
[38] S.G. Stepanian, I.D. Reva, E.D. Radchenko, L. Adamowicz, J. Phys. Chem. A 102 (1998) 4623.
[39] Z. Huang, Z. Lin, J. Phys. Chem. A 109 (2005) 2656.
[40] D. Kaur, P. Sharma, P.V. Bharatam, M. Kaur, Int. J. Quant. Chem. 108 (2008) 983.