PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Assessing Backbone Solvation Effects in the Conformational Propensities of Amino Acid Residues in Unfolded Peptides

Niranjan Ilawe,[1] Alexandra E. Raeber,[1] Reinhard Schweitzer-Stenner[2*], Siobhan E. Toal,[2§] and Bryan M. Wong[1*]

[1]Department of Chemical & Environmental Engineering, and Materials Science & Engineering Program, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, and [2]Department of Chemistry, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104

AUTHOR INFORMATION

Corresponding Author

*Bryan M. Wong, Phone: 1-951-827-2163, E-mail: bryan.wong@ucr.edu, Web: http://www.bmwong-group.com and Reinhard Schweitzer-Stenner, Phone: 1-215-895-2268, E-mail: rschweitzer-stenner@drexel.edu, Web: http://www.schweitzer-stenner.com

§Present address: Department of Biophysics and Biochemistry, Yale University, New Haven, CT 06250

ABSTRACT. Conformational ensembles of individual amino acid residues within model GxG peptides (x representing different amino acid residues) are dominated by a mixture of polyproline II (pPII) and β-strand like conformations. We recently discovered rather substantial differences between the enthalpic and entropic contributions to this equilibrium for different amino acid residues. Isoleucine and valine exceed all other amino acid residues in terms of their rather large enthalpic stabilization and entropic destabilization of polyproline II. In order to shed light on these underlying physical mechanisms, we performed high-level DFT calculations to explore the energetics of four representative GxG peptides where x = alanine (A), leucine (L), valine (V), and isoleucine (I) in explicit water (10 H_2O molecules with a polarizable continuum water model) and in *vacuo*. We found that the large energetic contributions to the stabilization of pPII result, to a major extent, from peptide-water, water-water interactions, and changes of the solvent self-energy. Differences between the peptide-solvent interaction energies of hydration in pPII and β-strand peptides are particularly important for the $\text{pPII} \rightleftharpoons \beta$ equilibria of the more aliphatic peptides GIG and GLG. Furthermore, we performed a vibrational analysis of the four peptides in both conformations and discovered a rather substantial mixing between water motions and peptide vibrations below 700 cm^{-1}. We found that the respective vibrational entropies are substantially different for the considered conformations, and their contributions to the Gibbs/Helmholtz energy stabilize β-strand conformations. Taken together, our results underscore the notion of the solvent being the predominant determinant of peptide (and protein) conformations in the unfolded state.

Graphical Abstract sentence: Large energetic contributions to the stabilization of polyproline II result from peptide-water, water-water interactions, and changes of the solvent self-energy.

Introduction

Our current understanding of protein folding processes relies on a detailed knowledge of the conformations they can adopt in their folded and unfolded states. For some time it was commonly believed that the unfolded states of peptides and proteins were characterized by a random sampling of all possible backbone conformations.[1-4] This predominant view is largely based on Flory's classical independent site model, which allows the various protein residues to sample the entire allowed region of the Ramachandran space.[2] However, over the last 15 years, considerable evidence has accumulated suggesting that significantly more restricted conformational ensembles of individual amino acid residues are present in unfolded peptides[5-12] and restricted coil libraries.[13, 14] Specifically, amino acid residues of short peptides in solution as well as in truncated coil libraries were found to exhibit much more restricted conformational sampling and to differ in terms of their conformational propensities. A recent conformational analysis of GxG peptides in water based on vibrational spectroscopic and NMR data revealed that most residues *do* predominantly sample the upper left quadrant of the Ramachandran plot, which can be subdivided into two sub-conformational ensembles associated with the β-strand and less extended polyproline II (pPII) conformations.[12, 15, 16] Individual amino acid residues differ in terms of the equilibrium between these two conformations – while alanine shows a clear preference for pPII, β-strand like conformations are slightly more preferred for valine and isoleucine. The $\Delta G_{\mathrm{pPII}\text{-}\beta}$-values associated with these equilibria range between -3.5 and 2 kJ/mol, with most values lying in the $R{\cdot}T$ energy range.[17] However, these $\Delta G_{\mathrm{pPII}\text{-}\beta}$ values were found to

be composed of much larger ΔH and ΔS values, which also exhibit much larger variations with changing residues than ΔG. For two subsets of data, the analysis resulted in isoequilibria, e.g. temperatures at which the ΔG values of the different members of the subset were practically identical. This was interpreted as an indication that peptide-water interactions are the key determinant of the residue's propensities. In this context it was particular striking that valine and isoleucine were found to exhibit rather large ΔH and $T\Delta S$ values (in the 40-60 kJ/mol range at room temperature) that exceed those of other residues with aromatic or aliphatic side chains. The fact that the ΔH and $T\Delta S$ values of isoleucine exceed the corresponding values for the isomeric amino acid leucine by a factor of ~2.5 was particular astonishing and has thus far been rationalized, in general terms, as an intricate interplay between backbone and side chain solvation.[17] However, these results cannot be understood in terms of solvent accessible surface areas that are normally considered to be the decisive parameters in side chain–solvent interactions. Specifically, since isoleucine (I) and leucine (L) have practically the same accessible surface area, which exceeds that of valine (V),[18] this suggests a hierarchy with I, L > V, contrary to the observed I > V >> L hierarchy.

Multiple lines of evidence suggest that interactions between unfolded peptides/proteins and the solvent are pivotal for conformational preferences exhibited by individual amino acid residues. This notion is particularly applicable to alanine which exhibits the highest propensity for pPII, a conformation generally found in proline-rich proteins and peptides.[19-28] In non-aqueous solvents like DMSO and primary alcohols, the sampling of pPII is generally reduced or even absent.[25, 28] The role of water in stabilizing the pPII conformation of alanine residues was first proposed by Han et al. based on density functional theory (DFT) calculations of N-acetyl-L-alanine N'methylamide (alanine dipeptide) in complexes with four water molecules.[19] These

authors attributed the stabilization of pPII to water bridges between CO and NH groups of the two peptide linkages of the molecule. Garcia, by means of molecular dynamics (MD) simulations with a modified AMBER force field, suggested that pPII conformations allowed for an optimal packing of water molecules in the hydration shell.[21] Fleming et al. proposed a correlation between pPII propensity and changes in the solvent accessibility area of side chains associated with transitions between pPII and β-strand like conformations.[18] However, this notion is at odds with results of MD studies reported by Mezei et al., who found instead that the pPII preference of alanine over β-strand results from stronger backbone-water interactions in the former.[24] Computational studies by Drozdov et al.[20] as well as Avbelj and Baldwin[29, 30] led these authors to suggest a more indirect role of solvation in the stabilization of pPII, namely the screening of electrostatic interactions between peptide groups, which would produce a preference for more extended β-strand conformations. Amino residues with side chains bulkier than alanine were predicted to perturb the hydration shell and thus destabilize pPII, a notion consistent with experimental results.[31] A recent DFT-based study of Lanza and Chiacchio on N-acetyl-L-alanine amide complexes with up to 13 water molecules revealed the existence of water clusters that were hydrogen bonded to the functional groups of the peptide backbone (i.e. CO and NH). They concluded that pPII conformations exhibit an energetic preference that leads to a greater reduction of water mobility,[32] further suggesting that pPII is energetically favored and entropically disfavored, in agreement with experimental data.[31] A more recent study by these authors confirmed this view in which the number of water molecules in the hydration shell was substantially increased up to 37.[33] A somewhat lower number of water molecules (22) was found to be necessary for obtaining stabilization energies of pPII relative to the β-strand that are comparable with experimentally obtained values.

The present study is aimed at shedding some light on the influence of side chains on the backbone hydration of GxG peptides. Toal et al. recently reported rather diverse enthalpy and entropy values for the $\text{pPII} \rightleftharpoons \beta\text{-strand}$ equilibria of fifteen GxG peptides in aqueous solution.[31] To this end, we performed several DFT calculations on a selected set of the earlier investigated GxG peptides surrounded by 10 explicit water molecules with a polarizable continuum water model. As representative peptides, we selected a set of aliphatic amino acid residues, namely, alanine (A), leucine (L), valine (V), and isoleucine (I) for our studies, as these residues showed the largest variance in thermodynamic values.[17] Alanine was also selected owing to its exceptional high pPII propensity (0.72 in GAG) for which multiple molecular dynamics (MD) and some quantum chemical studies suggest hydration as the main reason.[20-24, 29, 30, 32, 34-37] The other three aliphatic side residues were selected because the above thermodynamic analysis has yielded large and surprisingly different ΔH and $T\Delta S$ values for their respective $\text{pPII} \rightleftharpoons \beta$ equilibria. The experimental ΔH and ΔS values for these peptides as reported by Toal et al. are displayed in Figure 1.[31] In this study, we focus on determining the energetics of the $\text{pPII} \rightleftharpoons \beta$ equilibria for the investigated peptides rather than on a thorough exploration of the energy landscapes of these peptides. In view of the established preponderance of pPII and β-strand conformations in their Ramachandran plots, this restriction is justified.[12, 15, 16]

Computational Methods

All quantum chemical calculations in this study utilized density functional theory (DFT) calculations in conjunction with the dispersion-corrected, range-separated ωB97X-D functional as implemented in the Gaussian 09 package. Previous investigations by us[38-40] and others[41] have shown that these dispersion corrections in conjunction with nonlocal exchange are essential for

accurately calculating both the hydrogen-bonding interactions and thermochemistry of water-based complexes. All geometries were optimized with a large cc-pVTZ basis set in the presence of an aqueous polarizable continuum model (PCM). The specific PCM model used in this work is the implementation devised by Tomasi and co-workers[42-46] which creates a solute cavity via a set of overlapping spheres to calculate the solvent reaction field. The initial geometries of the water complexes were taken from Lanza and Chiacchio.[32] Geometry optimizations and harmonic frequencies at the same level of theory were calculated to verify that these stationary points were local minima.

Results and Discussion

Energetics of GxG $\mathrm{pPII} \rightleftharpoons \beta$ *equilibria in water.* Quantum calculations were performed for the selected cationic GxG peptides *in vacuo* and in a bath of 10 explicit H_2O molecules with the entire system embedded in a polarizable continuum solvent model, as described in the Computational Methods section. The number of water molecules is substantially less than what has recently been employed by Lanza and Chiacchio.[33] However, as shown in Figure 1 our approach allowed us to provide a minimal hydration of all functional groups of the peptide backbone, which includes 2 water molecules as hydrogen-bonded donors for the two carbonyl groups and one water molecule as a hydrogen-bond acceptor for the two amide protons. The remaining six water molecules were used to hydrate the terminal groups and to connect the different water molecules hydrogen bonded to the peptide. It is important to note that there exists many conformational isomers in each tripeptide, and we have only focused on a single conformer in each of the four GxG tripeptides. However, the goal of this study is the comparison of $\mathrm{pPII} \rightleftharpoons \beta$ equilibria rather than an accurate calculation of absolute values of energy

differences or a thorough sampling of the conformational space of water molecules in the hydration shell. The chosen setup of water molecules does not hydrate the side chains of the chosen x-residues, hence the contributions from side chain solvent interactions are neglected in our study (as they were in the studies of Lanza and Chiacchio[32, 33]). Furthermore, the previous studies of Lanza and Chiacchio investigated both AAA and AdP-type peptides, while our work focuses on the different GxG peptide geometries. While a comparison between our work and Lanza and Chiacchio's study is not entirely straightforward, we were able to confirm that the backbone geometry for our computed GAG structure has a similar topology with their previously published structure for AdP. This comparison, as well as our results described further below, demonstrate that our approach is capable of elucidating the influence of the side chains on backbone hydration. We selected the fully protonated form to directly correspond to the experiments this study is referring to.[12, 16, 31] Since we were solely interested in the $\text{pPII} \rightleftharpoons \beta$ equilibria of the investigated peptides, we carried out two fully unconstrained geometry optimizations per peptide, one starting in the pPII conformation and the other in the β-strand region. After a full optimization, a normal mode analysis was carried out. These calculations served three purposes: (1) determining the internal energy difference between β and pPII, (2) analyzing, in particular, the manifold of peptide modes that are vibrationally mixed with the hydration shell of 10 water molecules, and (3) checking for the appearance of any imaginary frequencies which would indicate that the observed conformation does not represent a true minimum in the energy landscape of the system.

Tables 1 and 2 list the dihedral angles of the obtained optimized structures, the energy differences between the pPII and β-strand geometries, and the permanent dipole moments for the various GxG systems *in vacuo* and in explicit water, respectively. Table 2 also lists the dihedral

angles of earlier reported maxima of pPII and β-strand sub-distributions obtained from amide I' profiles and various J-coupling constants.[12, 15, 16, 47] Most of the obtained values are in reasonable and some in excellent agreement with the reported experimentally-derived values. Exceptions are, to a minor extent, the ϕ-value of the β-strand conformation of GLG and, to a major extent, the ψ-value of the β-strand GVG. The experimentally-obtained values of pPII are generally very well reproduced by our calculations. Generally, our structure optimization with the larger cc-pVTZ basis set does a better job of reproducing experimentally obtained positions of energy minima (population maxima) than MD simulations, which generally yield rather extended structures for β-strand conformations.[27, 48-54]

Figure 1 depicts the conformations resulting from the geometry optimization of the investigated GAG + 10 H_2O configurations. The optimization process, to a major extent, has maintained the water-mediated hydrogen bonding network and the hydration of the backbone groups, consistent with the initial setup. These conformations resemble what Ben Naim has termed the HϕI interaction, which are hydrophilic in character.[55] A notable difference is observed for the arrangements of the water molecules above the peptide backbone. In the β-strand conformation, a water-mediated hydrogen bonding loop is maintained between the N- and the C-terminal which involves hydrogen bonding to and from the amide group of the C-terminal and the carbonyl group of the N-terminal peptide. In the pPII conformation, however, this loop is broken owing to the oxygen atom of the C-terminal's water molecule being now hydrogen bonded to the amide proton of the C-terminal peptide. On the contrary, the hydrogen bonding network below the peptide backbone is more extended in the pPII conformation. A distance analysis revealed that the C=O···H-O-H, NH···O, and H-O-H···OH_2 distances are generally shorter in the β-strand than in the corresponding pPII conformations, thus stabilizing the

former. Corresponding bending angles are also closer to their ideal 180° value in the β-strand conformation than in pPII. Such a stabilization of the β-strand conformation by hydrogen bonding was earlier reported by Mezei et al. for a polyalanine peptide in explicit water.[24] Lanza and Chiacchio reported a similar finding for their $AcANH_2 \cdot 13\ H_2O$ complex.[32] .

Based on the above analyses, one is tempted to suspect an overall energetic stabilization of β-strand over pPII. However, the corresponding electronic energy differences, $\Delta U_e(\text{pPII},\beta) = U_e(\text{pPII}) - U_e(\beta)$), as plotted in Figure 2 do not meet this expectation. These $\Delta U_e(\text{pPII},\beta)$ values can be directly compared to the experimental $\Delta H(\text{pPII},\beta)$ (= $H(\text{pPII}) - H(\beta)$) values reported by Toal et al.[31] Since no pressure and significant volume changes are involved, one can expect that the enthalpy ΔH equals the internal energy ΔU_{int}; i.e. the Gibbs free energy equals the Helmholtz energy. Therefore, in the following, we interpret the experimentally-obtained ΔH values reported by Toal et al. as internal energies that can generally be decomposed into electronic, vibrational, rotational, and translational contributions. For our system we can certainly ignore translation and rotation and we can further expect that the electronic part exceeds by far the vibrational contribution (*vide infra*). Hence, it is justified to compare the computational ΔU_e with the experimental ΔU_{int} values reported by Toal et al.[27] The calculated and experimental ΔU values are practically identical for GLG, whereas calculated values for GVG and GIG are lower than the experimental values by factors of 2 and 3, respectively. Surprisingly, our calculation leads to an overestimation of ΔU for GAG; the value we obtained more resembles the experimentally-obtained AAA and alanine dipeptide structures.[27] However, our calculated values reproduced the experimental hierarchy of I > V > L, which is one of the primary goals of this study. The corresponding ΔU-values for GxG *in vacuo* are also displayed in Figure 2. They are all positive, indicating a stabilization of the β-strand conformation and confirming that water is indeed

pivotal for pPII stabilization. Interestingly, GLG and GIG *in vacuo* now exhibit very similar energies, as one would expect from the fact that their side chains are chemical isomers. This suggests that the dissimilarity between pPII/β energy differences obtained (both experimentally and computationally) for hydrated leucine and isoleucine (Figure 1) is in part a direct result of differences in side-chain and backbone solvation.[31] Since the only difference between these two residues is the position of $-CH_3$ branching on the side-chain (i.e., β-branching in isoleucine and γ-branching is leucine), it is likely that the hydration shell about this group is the pivotal determinant. In general, these results strongly corroborate the notion that the unfolded state of peptides in water would be dominated by pPII-like conformations if the favorable hydration energy (Figure 1) was not compensated and sometimes overcompensated by entropic contributions. They further suggest that the $\text{pPII} \rightleftharpoons \beta$ equilibrium enthalpy/energy of GxG peptides is predominantly determined by peptide-solvent interactions.[28] Large differences between ΔU (and also between corresponding ΔS values) for different amino acid residues can give rise to large disparities between their respective $\text{pPII} \rightleftharpoons \beta$ equilibria at conditions significantly above and below their compensation temperatures. This further suggests a sequence dependence of the conformational entropies of unfolded states of peptides and proteins at high (thermal denaturation) and low temperatures (cold denaturation).[56]

The discrepancy between the calculated energetics and the β-strand stabilization suggested by the hydrogen bond network analysis is surprising, but not unprecedented. Fleming et al., from the MD-based analysis of a 12-residue polyalanine in explicit water, found no evidence for water bridges in pPII at all.[18] Nevertheless, they obtained a strong stabilization of pPII, attributable to peptide-water interactions, in agreement with our findings for all the peptides investigated. The DFT-based optimization of various $AcANH_2 \cdot nH_2O$ complexes by

Lanza and Chiacchio yielded water bridges for both conformations, but with more stable ones in β-strands. With regard to the total energy, however, hydration was found to stabilize pPII energetically.[32] The underlying reason for this overall stabilization of pPII remains elusive in all these studies.

Taken together, our analysis suggests that it is difficult to pinpoint the reason for the enthalpic (energetic) stabilization of pPII in water. It is clear from our study and previous experimental results that the interaction between the peptide and the first hydration shell is pivotal in this regard. Furthermore, the present calculations confirm the notion of Toal et al.[31] that the hydration energy difference between pPII and β-strand is heavily side chain dependent. However, the net electronic energies that emerged from our calculations are likely to be a superposition of competing contributions, which might involve through-space electrostatic in addition to through-hydrogen-bond interactions. To quantify these various interactions, we carried out a detailed analysis of the configurations for all four peptides using the energy-decomposition analysis of Mirkin and Krimm.[57] Within this energy-decomposition scheme, one can rigorously define the total energy of a composite system as the sum of the energies of its individual components plus the interaction energy between them. For example, the energy $U_e(j_s w_n)$, of the conformation $j = P, \beta$ in explicit water is given by

$$U_e\left(j_s w_n\right) = U_e\left(j_s\right) + U_{e,j}\left(w_n\right) + U_{e,i}\left(j\right), \qquad (1)$$

where $U_e(j_s)$ is the energy of the peptide geometry adopting the state j in the solvated system, $U_{e,j}(w_n)$ is the self-energy of the n water molecules surrounding a peptide adopting the conformation j, and $U_{e,i}(j)$ represents the total intermolecular interaction energies in the composite state j (which includes water-water and water-peptide hydrogen bonds as well as all

electrostatic interaction effects). Thus, the difference between pPII and β-strand energies of the solvated systems is given by

$$\Delta U_e\left(\mathrm{P_s}\beta_s\mathrm{w_n}\right) = \Delta U_e\left(\mathrm{P_s}\beta_s\right) + \Delta U_{e,P\beta}\left(\mathrm{w_n}\right) + \Delta U_{e,i}\left(\mathrm{P}\beta\right), \qquad (2)$$

where $\Delta U_e(\mathrm{P_s}\beta_\mathrm{s}\mathrm{w_n}) = U_e(\mathrm{P_s}\mathrm{w_n}) - U_e(\beta_\mathrm{s}\mathrm{w_n})$, $\Delta U_e(\mathrm{P_s}\beta_\mathrm{s}) = U_e(\mathrm{P_s}) - U_e(\beta_\mathrm{s})$, $\Delta U_{e,\mathrm{P}\beta}(\mathrm{w_n}) = U_{e,\mathrm{P}}(\mathrm{w_n}) - U_{e,\beta}(\mathrm{w_n})$, and $\Delta U_e(\mathrm{P}\beta) = U_e(\mathrm{P}) - U_e(\beta)$. The energies $\Delta U_e(\mathrm{P_s}\beta_\mathrm{s}\mathrm{w_n})$, $\Delta U_e(\mathrm{P_s}\beta_\mathrm{s})$, and $\Delta U_{e,\mathrm{P}\beta}(\mathrm{w_n})$ in Eq. (2) can be directly obtained from DFT total energy calculations of the composite peptide-water system, the peptide geometry, and the water geometries, respectively. As the original energy-decomposition analysis by Mirkin and Krimm[57] does not account for basis set superposition errors, we neglect these contributions which is also justified by our fairly large cc-pVTZ basis set. With all of the energies in Eq. (2) determined, the intermolecular interaction energy $\Delta U_{e,,i}(\mathrm{P}\beta)$ can be obtained as well. To further understand how our results would change with other basis sets, we also carried out this analysis with the smaller cc-pVDZ basis (see Electronic Supplementary Information). Upon optimization of the various GxG peptide structures, we found that several of the geometries exhibited significant distortions for both the ϕ and ψ dihedral angles. As such, these findings highlight the importance of using fairly large basis sets, and we only discuss the ωB97X-D/cc-pVTZ results in the main text. Using this decomposition scheme, we calculated the various energy contributions to the obtained $\Delta U_e(\mathrm{P_s}\beta_s\mathrm{w_n})$ values for all 4 of the investigated peptides, which are summarized in Table 3. All of the $\Delta U_e(\mathrm{P_s}\beta_\mathrm{s})$ values are positive (cf. Figure 2), indicating that the β conformation in the geometry of the solvated system is more stable (the latter statement is also true for the β conformation *in vacuo*). This suggests that the consideration of only indirect solvation is insufficient to describe pPII preferences. The intermolecular interaction energy differences,

$\Delta U_{e,i}(P\beta)$ as well as the solvent self-energies ($\Delta U_{e,p\beta}(w_n)$) are all negative, and both contribute significantly to the stabilization of the pPII structure. The latter are nearly identical for GLG, GVG, and GIG, while the value for GAG is only slightly lower. Apparently, the obtained differences $\Delta U_e(P_s\beta_s w_n)$ between the total energies, particularly between those of GLG, GVG, and GIG result from the more residue-specific peptide-water and water-water interactions. With regard to GAG, our results are at variance with the findings of Mirkin and Krimm, who identified the solvent energy as the main contributor to the stabilization of the pPII conformation of an alanine dipeptide-like compound.[57]

Estimation of vibrational entropies and energies. Generally, DFT energy calculations do not provide any direct information about entropies and, therefore, do not enable one to gauge relative stabilities of conformations. However, we wondered to what extent the hydration shell could contribute to entropic differences between the considered conformations solely through peptide-water and water-water vibrational mixing. Such vibrational mixing is well-established for amide I and the water bending modes.[58, 59] However, similar interactions between low-frequency modes of the peptide and collective water modes have not been fully investigated on a quantum chemical level. Our vibrational analysis reveals that nearly all modes below 700 cm^{-1} are involved in peptide-water mixing. As an example, Figure 3 shows four mixed water-peptide modes of the GAG-water complex, two for pPII and two for β-strand.

The strong mixing between the water and peptide modes clearly indicates that peptide-water interactions constitute a dynamic entity with significant coherent vibrational dynamics. It is obvious that the degree of this mixing and, thus, the vibrational energy density should depend on the backbone conformation. To check this conjecture we calculated the difference between

the (harmonic) entropies of pPII and β at room temperature by utilizing the well-known relationship:

$$S = -R \cdot \sum_{i=1}^{N} \left\{ \ln\left(1 - e^{-hc\tilde{\nu}_i / k_B T}\right) + \frac{N_A hc\tilde{\nu}_i}{T} \cdot \frac{e^{-hc\tilde{\nu}_i / k_B T}}{1 - e^{-hc\tilde{\nu}_i / k_B T}} \right\}, \quad (3)$$

where N_A is the Avogadro constant, h the Planck constant, c the vacuum velocity of light, $\tilde{\nu}_i$ the wavenumber of the ith peptide-H_2O vibration, R the gas constant, k_B the Boltzmann constant, and T the absolute temperature. To avoid any arbitrary cutoffs, we included the entire vibrational manifold of N modes as obtained from the DFT based normal mode calculations. We are aware that this approach neglects the anharmonicity to be expected for low frequency modes and possible inaccuracies of our calculated frequencies. However, it is reasonable to expect that our calculations should provide us with reliable trends in the magnitude and sign of hydration-induced differences between the vibrational entropies of pPII and β-strand conformations. This is corroborated by recent comparisons between experimentally-obtained and computed entropy differences between pPII and β-strand of $AcANH_2$ in water.[32, 33]

The vibrational entropy contributions to the Helmholtz energies of the investigated GxG peptides as obtained by Eq. (3) are visualized in Figure 4. Apparently, all vibrational entropies stabilize the β-strand conformation in explicit water by contributing between 5 to 18 kJ/mol at room temperature. Interestingly, the obtained side chain dependence of the vibrational entropies correlate with the corresponding ΔU values. Even though they display the same I > V > L hierarchy, they underestimate the respective experimental values for these residues and overestimate the entropy of GAG. The underestimated entropy of GLG, GVG, and GIG can again be related to the inability of the considered 10 water molecules to fully hydrate these

peptides' side chains. However, our results clearly suggest that the vibrational dynamics of the combined peptide-water system is a major source of the entropic contribution to the Gibbs or Helmholtz energy differences between pPII and the β-strand, in agreement with Lanza and Chiacchio.[33]

In order to further corroborate the contributions of hydration to the vibrational entropy differences between pPII and β-strand conformations, we calculated the entropy differences between the respective vibrational manifolds of the investigated GxGs without solvent molecules. The respective contributions to the total free energy are also plotted in Figure 4. For GAG and GIG, these (positive) contributions are negligible. For GLG and GVG, the calculations actually yielded a room temperature stabilization of the pPII conformation by 4.3 and 1.6 kJ/mol, respectively. Together, all these calculations demonstrate that peptide solvation inverts the sign of the internal energy and entropy contribution to the Helmholtz energy governing the $\text{pPII} \rightleftharpoons \beta$ equilibrium.

Finally, for the sake of completeness, we also estimated the vibrational contribution to the internal energy for the above considered vibrational manifold of GxG-10·H_2O. This can be done by using the textbook equation:

$$U_{vib} = \sum_{i=1}^{N} \left(N_A \cdot h \cdot c \cdot \left(\frac{\tilde{\nu}_i e^{-hc\tilde{\nu}_i/k_B T}}{1 - e^{-hc\tilde{\nu}_i/k_B T}} + \right) + \frac{\tilde{\nu}_i}{2} \right) \qquad (4)$$

We thus found that the contribution of the vibrational energy to the total internal energy does not vary significantly with x, and the corresponding values lie between -1.4 and -1.7 kJ/mol. This justifies our focus on the electronic energies derived from our structure optimization procedure.

Conclusion

Taken together with our earlier reported experimental results, the complementary DFT calculations reveal the dominant role of backbone hydration in stabilizing the pPII conformation of amino acid residues of short peptides in water. Specifically, these interactions stabilize the pPII conformation energetically but destabilizing it entropically. We were able to reproduce an earlier observed I > V > L hierarchy of the experimentally-obtained $\Delta H/\Delta U$ and $T\Delta S$ values associated with the $\mathrm{pPII} \rightleftharpoons \beta$ equilibrium of GxG peptides in water on a qualitative level. We decomposed the obtained energies of the investigated $\mathrm{pPII} \rightleftharpoons \beta$ equilibria into intramolecular (peptide and water) and intermolecular contributions (water-peptide). The results of this procedure revealed that peptide-water interactions and solvent self-energies promote the energetic stabilization of pPII over water. Our results show that the hydration of the peptide backbone critically depends on the backbone structure and on the nature of the side chain. This notion seems to be consistent with the conditional solvation model of Ben Naim, but at variance with a model constructed from context free analogues of side chains.[55] It is further consistent with predictions from electrostatic theories.[30] Moreover, we provide evidence for the notion that the rather large entropic differences between pPII and β-strand are assignable to the vibrational dynamics of the peptide-water entity. Our calculations with only ten water molecules actually underestimate these contributions for residues with large and bulky side chains. Mezei et al. performed a molecular-dynamics-based analysis of a polyalanine 12-mer in explicit water and reported that the number of water molecules in the first hydration shell of both pPII and β is approximately 120.[24] Any further characterization of peptide hydration shells in terms of

hydrogen bonding and orientational distribution based on quantum chemical calculations will be the subject of future studies.

ACKNOWLEDGMENT

S. E. T. was in part supported by a research fellowship of the College of Arts and Sciences which is gratefully acknowledged. B.M.W. acknowledges the National Science Foundation for the use of supercomputing resources through the Extreme Science and Engineering Discovery Environment (XSEDE), Project No. TG-CHE150040.

Electronic supplementary information (ESI) available: Energy decomposition of the solvated peptides at the ωB97X-D/cc-pVDZ level of theory and reference Cartesian coordinates and energies for all 16 peptides.

REFERENCES

1. D. A. Brant and P. J. J. Flory, *J. Am. Chem. Soc.*, 1965, 87, 2791-2800.
2. P. J. Flory, *Statistical Mechanics of Chain Molecules*, Wiley & Sons, New York, 1969.
3. G. N. Ramachandran, C. Ramachandran and V. Sasisekharan, *J. Mol. Biol.*, 1963, 7, 95-99.
4. C. Tanford, *Adv. Protein Chem.*, 1968, 23, 121-282.
5. Z. Shi, K. Chen, Z. Liu, A. Ng, W. C. Bracken and N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA*, 2005, 102, 17964 - 17968.
6. Z. Shi, C. A. Olson, G. D. Rose, R. L. Baldwin and N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA*, 2002, 99, 9190-9195.
7. Z. Shi, K. Shen, Z. Liu and N. R. Kallenbach, *Chem. Rev.* , 2006, 106, 1877-1897.
8. S. Woutersen and P. Hamm, *J. Phys. Chem. B*, 2000, 104, 11316-11320.
9. S. Woutersen and P. Hamm, *J. Chem. Phys.*, 2001, 114, 2727-2737.
10. J. Graf, P. H. Nguyen, G. Stock and H. Schwalbe, *J Am Chem Soc*, 2007, 129, 1179-1189.
11. F. Eker, X. Cao, L. Nafie and R. Schweitzer-Stenner, *J Am Chem Soc*, 2002, 124, 14330-14341.
12. A. Hagarman, T. J. Measey, D. Mathieu, H. Schwalbe and R. Schweitzer-Stenner, *J Am Chem Soc*, 2010, 132, 540-551.
13. A. K. Jha, A. Colubri, K. F. Freed and T. R. Sosnick, *Proc Natl Acad Sci U S A*, 2005, 102, 13099-13104.
14. J. DeBartolo, A. Jha, K. Freed, F. and T. R. Sosnick, in *Proteins and Peptides. Folding, Misfolding and Unfolding.*, ed. R. Schweitzer-Stenner, Wiley & Sons, Chichester, 2012, pp. 79-98.
15. A. Hagarman, D. Mathieu, S. Toal, T. J. Measey, H. Schwalbe and R. Schweitzer-Stenner, *Chemistry*, 2011, 17, 6789-6797.
16. R. Schweitzer-Stenner, A. Hagarman, S. Toal, D. Mathieu and H. Schwalbe, *Proteins*, 2013, 81, 955-967.
17. S. E. Toal and R. Schweitzer-Stenner, *Biomolecules*, 2014, 4, 725-773.
18. P. J. Fleming, N. C. Fitzkee, M. Mezei, R. Srinivasan and G. D. Rose, *Prot. Sci.*, 2005, 14, 111-118.
19. W.-G. Han, K. J. Jakanen, M. Elstner and S. Suhai, *J. Phys. Chem. B*, 1998, 102, 2587-2602.
20. A. N. Drozdov, A. Grossfield and R. V. Pappu, *J Am Chem Soc*, 2004, 126, 2574-2581.
21. A. E. Garcia, *Polymer*, 2004, 120, 885-890.
22. S. Gnanakaran and A. E. Garcia, *J. Phys. Chem. B*, 2003, 107, 12555-12557.
23. A. Kentsis, M. Mezei, T. Gindin and O. R., *Proteins: Structure, Function and Genetics*, 2004, 55, 493-501.
24. M. Mezei, P. J. Fleming, R. Srinivasan and G. D. Rose, *Proteins: Structure, Function and Genetics*, 2004, 55, 502-507.
25. F. Eker, X. Cao, L. Nafie, Q. Huang, K. Griebenow and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2003, 107, 358-365.
26. Z. Liu, K. Chen, A. Ng, Z. Shi, R. W. Woody and N. R. Kallenbach, *J Am Chem Soc*, 2004, 126, 15141-15150.

27. S. Toal, D. Meral, D. Verbaro, B. Urbanc and R. Schweitzer-Stenner, *J Phys Chem B*, 2013, 117, 3689-3706.
28. S. Toal, A. Omidi and R. Schweitzer-Stenner, *J. Am. Chem. Soc.*, 2011, 133, 12728.
29. F. Avbelj, in *Protein and Peptide Folding, Misfolding and Non-Folding*, ed. R. Schweitzer-Stenner, John Wiley & Sons,, Hobocken, 2012, pp. 131-158.
30. F. Avbelj and R. L. Baldwin, *Proc Natl Acad Sci U S A*, 2003, 100, 5742-5747.
31. S. E. Toal, D. J. Verbaro and R. Schweitzer-Stenner, *J Phys Chem B*, 2014, 118, 1309-1318.
32. G. Lanza and M. A. Chiacchio, *ChemPhysChem*, 2014, 15, 2785–2793.
33. G. Lanza and M. A. Chiacchio, *Phys. Chem. Chem, Phys.*, 2015, in press.
34. F. Avbelj and R. L. Baldwin, *Proc Natl Acad Sci U S A*, 2004, 101, 10967-10972.
35. F. Avbelj, S. G. Grdadolnik, J. Grdadolnik and R. L. Baldwin, *Proc Natl Acad Sci U S A*, 2006, 103, 1272-1277.
36. W. Li, M. Qin, Z. Tie and W. Wang, *Phys. Rev. E*, 2011, 84, 041933.
37. C. D. Poon, E. T. Samulsi, C. F. Weise and J. C. Weisshaar, *J. Am. Chem. Soc.*, 2000, 122, 5612-5613.
38. M. A. Hossain, M. A. Saeed, A. Pramanik, B. M. Wong, S. A. Haque and D. R. Powell, *J Am Chem Soc*, 2012, 134, 11892-11895.
39. M. A. Saeed, A. Pramanik, B. M. Wong, S. A. Haque, D. R. Powell, D. K. Chand and M. A. Hossain, *Chem Commun*, 2012, 48, 8631-8633.
40. M. A. Saeed, B. M. Wong, F. R. Fronczek, R. Venkatraman and M. A. Hossain, *Cryst Growth Des*, 2010, 10, 1486-1488.
41. E. E. Dahlke, R. M. Olson, H. R. Leverentz and D. G. Truhlar, *J Phys Chem A*, 2008, 112, 3976-3984.
42. A. Fortunelli and J. Tomasi, *Chem Phys Lett*, 1994, 231, 34-39.
43. J. Tomasi, R. Bonaccorsi, R. Cammi and F. J. O. Delvalle, *Theochem-J Mol Struc*, 1991, 80, 401-424.
44. F. J. O. Delvalle, R. Bonaccorsi, R. Cammi and J. Tomasi, *Theochem-J Mol Struc*, 1991, 76, 295-312.
45. F. J. O. Delvalle and J. Tomasi, *Chem Phys*, 1991, 150, 139-150.
46. M. A. Aguilar, F. J. O. Delvalle and J. Tomasi, *Chem Phys*, 1991, 150, 151-161.
47. K. Rybka, S. E. Toal, D. J. Verbaro, D. Mathieu, H. Schwalbe and R. Schweitzer-Stenner, *Proteins*, 2013, 81, 968-983.
48. K. Kwac, K. K. Lee, J. B. Han, K. I. Oh and M. Cho, *J Chem Phys*, 2008, 128, 105106.
49. B. Zagrovic, J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, S. Doniach and V. S. Pande, *Proc Natl Acad Sci U S A*, 2005, 102, 11698-11703.
50. S. Gnanakaran and A. E. Garcia, *Proteins*, 2005, 59, 773-782.
51. M. Feig, *J. Chem. Theo. Comput.* , 2008, 4, 1555-1564.
52. S. Hovmoller, T. Zhou and T. Ohlson, *Acta Crystallogr D Biol Crystallogr*, 2002, 58, 768-776.
53. Y. Mu, D. S. Kosov and G. Stock, *J. Phys. Chem. B*, 2003, 107, 5064-5073.
54. Y. Mu and G. Stock, *J. Phys. Chem. B*, 2002, 106, 5294-5301.
55. A. Ben Naim, *Molecular Theory of Water and Aquesous Solutions-Part I: Understanding Water*, World Scientific, Singapur, 2009.
56. R. Schweitzer-Stenner and S. E. Toal, *Phys. Chem. Chem, Phys.*, 2014, 16, 22527-22536.
57. N. G. Mirkin and S. Krimm, *Biopolymers*, 2012, 97, 789-794.

58. X. G. Chen, R. Schweitzer-Stenner, S. A. Asher, N. G. Mirkin and S. Krimm, *J. Phys. Chem.* , 1995, 99, 3074-3083.
59. G. Sieler and R. Schweitzer-Stenner, *J. Am. Chem. Soc.* , 1997, 119, 1720.

Table 1: Dihedral angles ϕ and ψ, electronic energies, and total permanent dipole moment of the geometry-optimized pPII and β-strand type conformations of GxG peptide+reaction sphere. ΔU_e is the electronic energy difference calculated as $\Delta U_e(P_s\beta_s w)$ described in the main text.

	$\phi[^0]$	$\psi[^0]$	ΔU_e(kJ/mol)	$\lvert\mu\rvert$ (D)
GAG-pPII	-65.86	148.09	7.60	18.67
GAG-β	-159.49	161.17		14.69
GLG-pPII	-72.94	153.38	4.89	18.31
GLG-β	-115.17	153.68		14.06
GVG-pPII	-83.33	161.85	6.83	17.71
GVG-β	-134.33	138.28		13.00
GIG-pPII	-80.58	163.63	5.10	17.55
GIG-β	-132.13	131.66		12.85

Table 2: Dihedral angles ϕ and ψ, electronic energies, and total permanent dipole moment of the geometry-optimized pPII and β-strand type conformations of GxG peptides + 10 H_2O + polarizable continuum. ΔU_e is the energy difference calculated as $\Delta U_e(P_s\beta_s w)$ described in the main text. Numbers shown in parenthesis reflect the maxima positions of sub-populations observed from experimental data.[12, 16]

	$\phi[^0]$	**$\psi[^0]$**	**ΔU_e (kJ/mol)**	**$\|\mu\|$ (D)**
GAG-pPII	-71.93 (-74)	163.59 (152)	-21.24	3.89
GAG-β	-137.84 (-115)	131.55 (120)		18.93
GLG-pPII	-70.42 (-76)	160.81 (145)	-16.42	3.89
GLG-β	-121.24 (-98)	132.98 (145)		20.13
GVG-pPII	-77.81 (-80)	169.07 (170)	-18.25	3.70
GVG-β	-130.26 (-120)	118.29 (170)		18.69
GIG-pPII	-78.84 (-77)	169.58 (170)	-20.34	3.96
GIG-β	-129.81 (-118)	119.37 (115)		18.39

Table 3: Energy decomposition (Eqs. 1-2) of the solvated peptides at the ωB97X-D/cc-pVTZ level of theory.

	$\Delta U_e(P_s\beta_s w_n)$[a] (kJ/mol)	$\Delta U_e(P_s\beta_s)$[b] (kJ/mol)	$\Delta U_{e,P\beta}(w_n)$[c] (kJ/mol)	$\Delta U_{e,i}(P\beta)$[d] (kJ/mol)
GAG	-21.24	10.88	-11.28	-20.84
GLG	-16.42	7.49	-13.26	-10.65
GVG	-18.25	8.61	-13.70	-13.15
GIG	-20.34	7.85	-13.13	-15.06

[a] Electronic Energy difference between pPII and β solvated systems consisting of the peptide and 10 water molecules
[b] Electronic Energy difference between individual pPII and β structures in the geometry of the solvated system
[c] Electronic Energy difference between 10 water molecules associated with the pPII and β structures
[d] Electronic Interaction energy difference between pPII and β systems

Figure legends

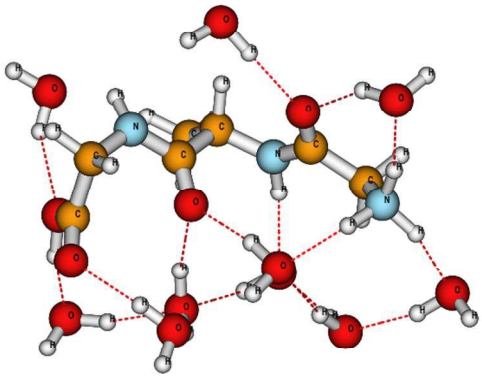
Figure 1: pPII and β-strand conformation of a cationic GAG + 10 H_2O complex obtained from geometry optimizations as described in the text.

Figure 2: Comparison of experimental enthalpies (black bars)[28] and calculated energies (light grey: explicit water; grey: in *vacuo*) for the $\text{pPII} \rightleftharpoons \beta\text{-strand}$ equilibrium of the indicated amino acid residues in cationic GxG peptides in H_2O.

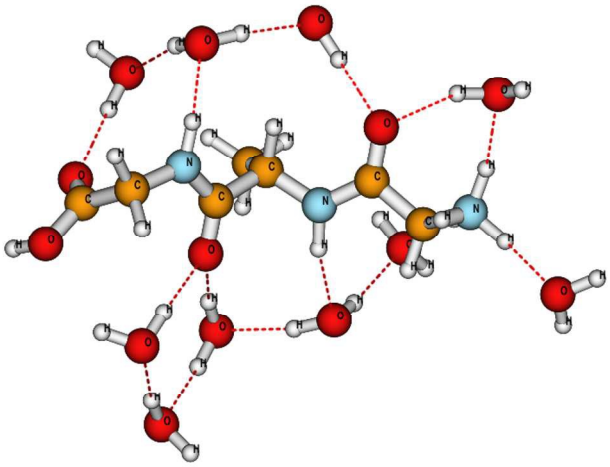
Figure 3: Representative mixed water-peptide modes of the GAG-water complex. For both the pPII and β-strand conformations, there is a strong mixing between water and peptide modes.

Figure 4: Difference between the Gibbs/Helmholtz energy contributions of the vibrational entropy of the indicated amino acid residue in cationic GxG peptides in explicit water (grey) and *in vacuo* (light grey) in the pPII and β-strand conformation. Experimentally obtained $T\Delta S$ values of the corresponding the experimentally obtained $pPII \rightleftharpoons \beta$-strand equilibrium are plotted in black. *In vacuo* values for L and I are so small that they are barely visible on the chosen scale.

Figure 1



pPII GAG



β-strand GAG

Figure 2

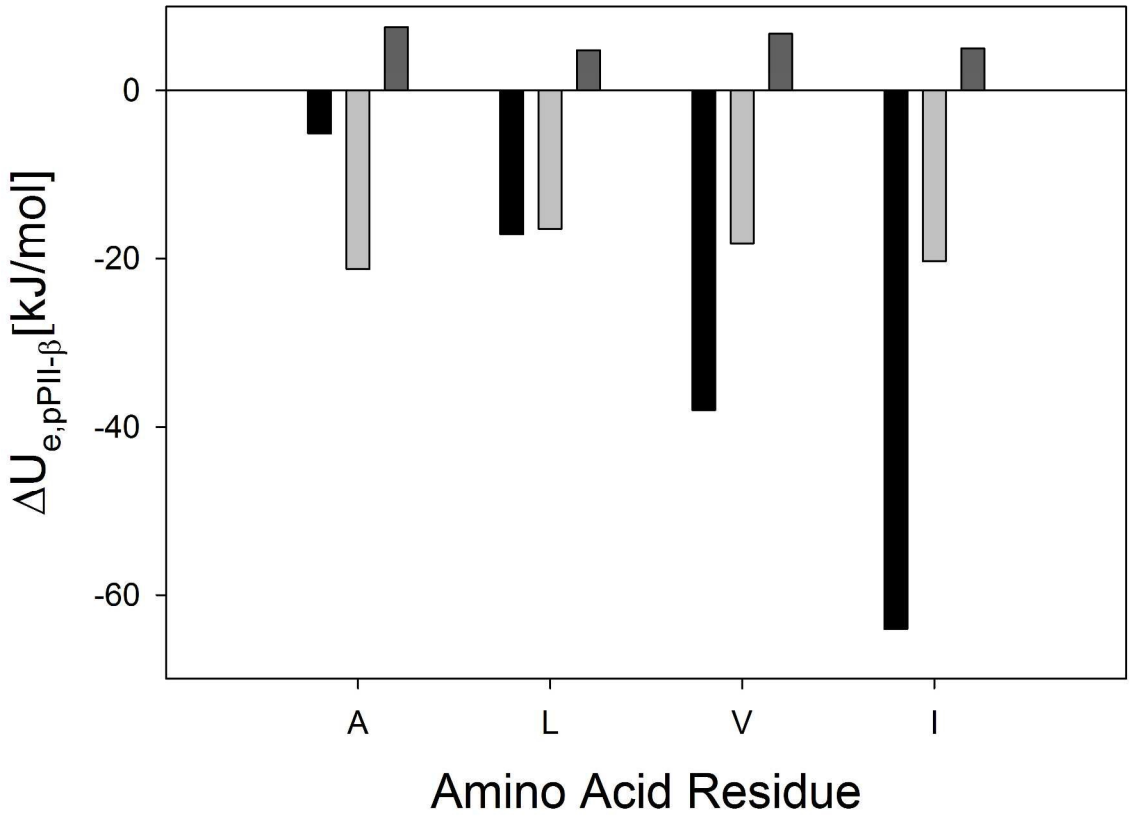
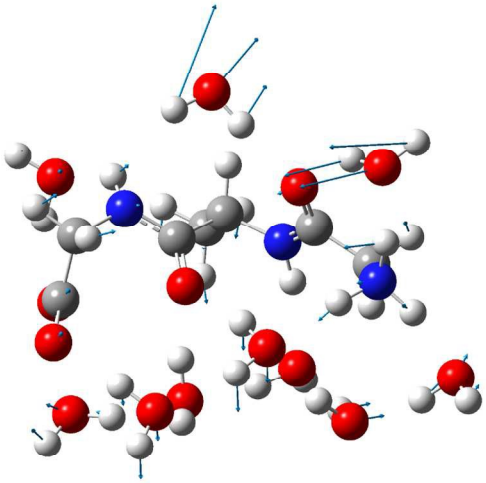
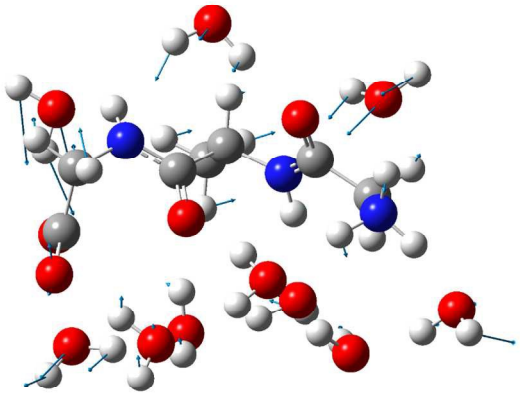


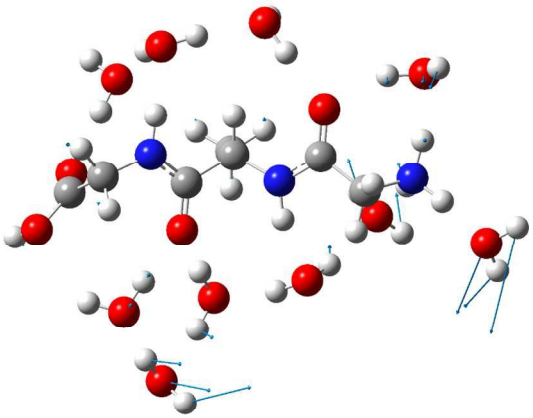
Figure 3



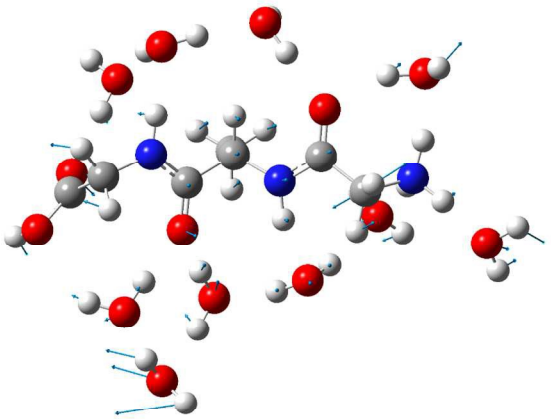
pPII GAG
$\nu_1 = 26.5\ cm^{-1}$



pPII GAG
$\nu_2 = 32.3\ cm^{-1}$



β-strand GAG
$\nu_1 = 16.4\ cm^{-1}$



β-strand GAG
$\nu_2 = 19.3\ cm^{-1}$

Figure 4