PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

PCCP

RSCPublishing

PAPER

Interfacial Water at the Trialanine Hydrophilic Surface. A DFT Electronic Structure and Bottom-Up Investigation

Cite this: DOI: 10.1039/xoxxoooooox

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

Giuseppe Lanza* and Maria Assunta Chiacchio,

DFT-M062X quantum chemical computations on the $Ala_3H^+ \cdot nH_2O$ (n up to 37) complexes have been performed to model for hydration effects on the molecular properties of protonated trialanine. Following simple rules to arrange water molecules around the peptide, geometry optimization allows us to find four minima corresponding to the unfolded extended (β) and polyproline II (PPII) conformations. The peptide is incorporated into the network of hydrogen bonds of interfacial water molecules with hydration energy of about -85 kcal mol^{-1}. The progressive hydration of peptide shows a more efficient intermolecular hydrogen bonding in the PPII arrangement, and the following relative electronic energy stability β-β < β-PPII ≈ PPII-β < PPII-PPII has been found. The conformational entropy term proceeds in the reverse direction, thus these changes compensate in a way that leads to small changes in Gibbs free energy. These findings agree with experimental data which report an equilibrium between these conformers modulated by temperature.

Introduction

The interaction of water with biomolecules is at the forefront of biophysical research since water modulates molecular structures and hence all fundamental functions related to life.[1-4] Very simple di- and tripeptides dissolved in an aqueous environment have attracted great attention because, due to their chemical simplicity, it is presumably easy to understand molecular properties and water-peptide interactions.[5-20] Although there is no exhaustive experimental technique that directly probes water/peptide geometrical arrangement, neutron scattering,[5] NMR,[6,7] TeraHz,[8-10] dielectric relaxation,[11] and fluorescent[12] spectroscopies have provided invaluable information about dynamic processes occurring at the water/peptide interface. All these techniques have the time resolution in the picoseconds range, hence appropriate for investigating the network of hydrogen-bonds. It is commonly accepted that at the solute/water interface, molecules are more "connected" than in the water bulk and therefore the breaking/formation of H-bonds and the molecule reorientation occur more slowly than in bulk water. However, the slow dynamic at the interface is in part due to the strong peptide-water hydrogen-bonds but also to the heterogeneity of peptide surface that largely disrupts the water-water synergic reorientation mechanisms.[21]

Solvent molecules imposed large perturbations to the peptidic chain leading to important changes in the thermodynamic functions of various conformers and in particular on their relative stability. For example, short alanine based peptides adopt compact structures with strong intramolecular H-bonds in the gas-phase[22-25] or apolar solutions.[26] Conversely, in aqueous solutions intramolecular H-bonds are largely surmounted by more efficient intermolecular H-bonds with water molecules and unfolded conformations like extended (β) and polyproline II (PPII) become more stable.[13-20] For intrinsically-disordered peptides, the energy separation between stable secondary structures is rather small. There is a rapid interconversion of various conformers thus inferring high fluxionality of the peptidic chains.[27]

With the aim to get computational data that are as free as possible from empirical assumptions and hence able to provide complementary and independent information to the scientific community, there has recently been a flourishing of quantum chemical studies on the solvation effects upon peptide properties.[28-35] Among them is our procedure based on MP2 and DFT-M062X methods together with a bottom-up approach to account for the hydration effects on the molecular properties of N-acetyl-L-alanine amide, which turns out to be promising for an expanding application.[35] In this study, short range solute-solvent interactions are considered, including some water molecules explicitly in quantum chemical calculations, while bulk solvation is modelled using a reaction field method as the polarisable continuum model (PCM). The proposed bottom-up strategy reveals the formation of compact ring clusters of water molecules strongly bonded to peptidic polar groups through hydrogen bonds. The alternance of donor and acceptor groups along the extended and PPII conformers allows for synergy and extensive H-bonding.

The state-of-art quantum chemical methodology adopted allows for an unbiased quantitative evaluation of the strength of individual peptide-water molecule interactions that, currently, represent a difficult and intriguing matter. On the other side the incremental hydration in the bottom-up fashion allows us to

control and understand (step by step) variables that determine and maintain peptide conformations.[35]

At this point, the question that intrigues us is if it is possible to expand such a methodology of incremental water number to more complex peptides and, hence, to more interesting biophysical problems. To this purpose, protonated trialanine, Ala_3H^+ is an ideal candidate to be considered for several reasons. It is a simple model that well represents the class of intrinsically-disordered proteins and for this reason it has received a great deal of attention from experimentalists.[13-18] The equilibrium structure of this molecule has been investigated by several experimental techniques: at low temperature it populates mainly polyproline II but, upon heating, the fraction of extended conformation increases quickly and becomes predominant.[13,18] This indicates that the PPII structure is enthalpically favoured, whereas entropy stabilizes the β-strand conformation. Because of the presence of two couple ϕ/ψ angles, mainly four structures indicated PPII-PPII and β-β and mixed PPII-β and β-PPII have been identified and quantified in solution.

There are several molecular dynamics studies of the Ala_3H^+ peptide and many of them are reported together with experimental researches.[14-17] On the other hand, quantum chemical investigations are limited to the gas-phase[22-24,36] where compact structure with intramolecular H-bonds are reported as ground state while water effects have never been considered. Instead, there are earlier studies reporting on the protonated dialanine, Ala_2H^+, hydrogen bonded with some water molecules.[28]

In the present contribution, we report a detailed computational analysis at the DFT-M062X level[37] of the hydration phenomena occurring at the surface of protonated Ala_3H^+ peptide in various conformations. In particular, what is analyzed is the behaviour of equilibrium structures and the relative stability of the peptide on increasing the number of coordinating water units. It is shown that a simple "sheet of wrapping H-bonded connected water molecules" accounts for a large fraction of overall solvation effects and allows for the reproduction and rationalization of many experimental features.

Calculation methods

The choice of exchange-correlation functional type for the molecular properties of peptides computation is a critical decision for it also keeps in mind the importance of dispersive terms.[37] Recently, extensive benchmark MP2 and various DFT functionals showed that the M062X performs well.[35] Therefore, geometries were optimized at the DFT-M062X level employing the 6-31+G* basis set and including implicitly solvent effects (M062X/6-31+G*/PCM=WATER). Minima were characterized evaluating the hessian matrix and the associated harmonic vibrational frequencies. Implicit solvent effects were modelled using the polarized continuum method adopting a 78.36 dielectric constant for water as implemented in the G09 program.[38]

To improve energetics and to reduce intermolecular basis set superposition error, single point energy at the related optimized geometry has been performed using the more accurate aug-cc-pVTZ basis sets including implicit solvent effects. Furthermore, energetics reproducibility has been verified adopting other type of functionals specifically the pure TPSS-TPSS and the hybrid B3PW91 and B3LYP including Grimme-D3 dispersion terms.

To calculate the entropy, $S°_{298}$, the different contributions to the partition function were evaluated by using the standard expressions for an ideal gas in the canonical ensemble, the harmonic oscillator, and the rigid rotor approximations. For selected cases, the computed energies were corrected for zero-point vibrational and thermal contributions to obtain enthalpy changes at 298 K (ΔH_{298}).

Results and discussion

Chemical theory

From a computational point of view the modelling of the aqueous environment around a peptide is a challenging task due to the difficulties in describing solute-solvent and solvent-solvent interactions and dynamical/configurational complexity. Furthermore, peptidic surfaces are heterogeneous in composition, topographically complex and water molecules are not uniformly distributed. They are expected to be more densely arranged around polar and charged groups while they are far away from hydrophobic groups.[35]

It is not clear how to determine, *a priori*, the number and the orientation of solvent molecules in direct contact with the peptide. Therefore, the starting peptide-water cluster orientation is a critical step in the present quantum chemical bottom-up analysis. The most used approach to generate the initial structures for quantum chemical minimum search is to perform preliminary molecular dynamics simulations of the biomolecule using a molecular mechanics force-field.[39,40] In alternative to this, recently we proposed a "chemical intuition" method to construct the initial arrangement of water molecules around the peptide using some well established structural features of H-bonding.[35,41]

1) A water molecule can form four H-bonds with adjacent molecules, two as H-donors and two as H-acceptors, and maintains a local pseudo-tetrahedral symmetry. There is a synergy in the formation of the H-bonds and, in general, the best energy match is reached when a network of alternation acceptor and donor H-bonds take place.

2) Water molecules surround the peptide maximizing the number of H-bonds and their strength to realize the highest density packing.

3) H-bond is highly directional with the D-H···A bond angle close to the linearity. This poses some structural constraints in the construction of the water molecules network around the peptide and the proper alignment can be used as a criterion to determine the quality of the final structure in geometry optimization.

4) The >NH groups of the peptide can donate a proton to a solvent molecule while the >CO group can accept two H-bonds because of the presence of two oxygen lone pairs.

5) To avoid artificial distortions of the peptide conformation, the size of wires and clusters of water molecules around the peptide should be neither too short nor too long compared to the distance that separates hydrophilic groups to be connected. In other words, peptide polar groups and water assembly should have a good molecular match. To connect adjacent >CO and >NH groups of the peptide often various water molecules are necessary. For this reasons progressive hydration occurs with step of 2, 3, 4, 5 or 6 water molecules.

6) To make a direct comparison of energies for the various conformations, models must be built with the same number of water molecules.

In addition to H-bond structural features above discussed, the formation electronic energy (1) and the peptide-water electronic interaction energy (2) are two important energetic parameters to

be considered with great care in our molecular model development.

$Ala_3H^+ + nH_2O \rightarrow Ala_3H^+ \cdot nH_2O \quad \Delta E_{form} \quad (1)$

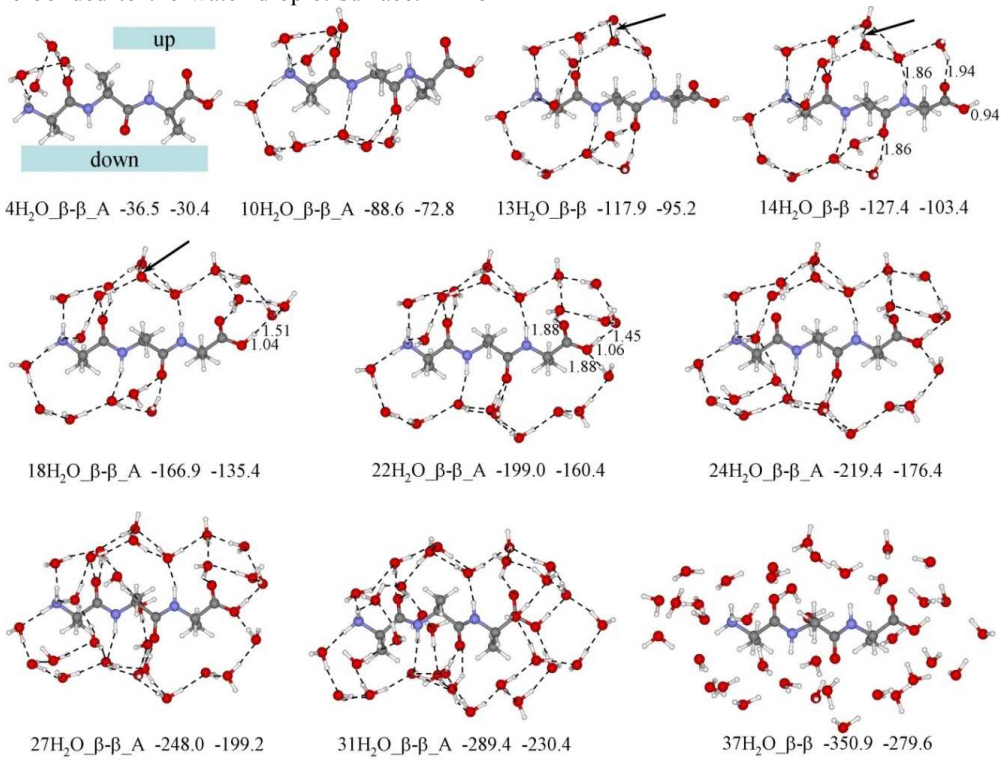
$Ala_3H^+ + (H_2O)_n \rightarrow Ala_3H^+ \cdot nH_2O \quad \Delta E_{int} \quad (2)$

The ΔE_{form} of various structures is computed as the difference of electronic energy of reagents and product at their minima of energy. For the bare Ala_3H^+ peptide, the fully-extended conformation has been taken as a reference. For a given model, the lowest value, the more exothermic, corresponds to the most stable structure that is obtainable from global energy minimization. The peptide-water interaction energy, ΔE_{int}, is estimated as the difference between the total electronic energies of isolated water cluster and bare peptide and their assembly complex at the peptide-water optimized geometry. This parameter accounts for the solute-solvent forces, and for a given model, the more exothermic corresponds to the more hydrated structure.

At this point two questions arise: do ΔE_{form} and ΔE_{int} follow the same trend in our hydration model? And if they do not, which should we follow? For peptide-water complexes with few solvent molecules, the ΔE_{form} and ΔE_{int} follow the same trend, the more stable structure is also the more hydrated. As the number of solvent molecules increase (n>7), the global minimum is obtained arranging water units over the already formed water cluster in a region in which there is no contact with the peptide. The homogeneity of solvent molecules allows for the best packing, hence the highest number of H-bonds, thus the trialanine bonds to a "water droplet" surface. A similar situation has already been reported for a systematic global minimum search of the simple alanine amino acid. It has been shown that up to 42 water molecules the global minimum corresponds to a partially hydrated amino acid with the nonionized alanine bonded to the water droplet surface.[44] The full hydration global minimum occurs after 46 water molecules. It should be clear that for the present trialanine, hundreds of water molecules would be necessary to get the global minimum with a full hydrated peptide. Presently, it is an intractable problem and we devote our efforts to find the number of minimal water molecules to cover all polar groups of the peptide (maximum hydration, the ΔE_{int} criteria) and see what effects they produce on the conformations. The formation energy, ΔE_{form}, is extremely useful in comparing the thermodynamic properties of a model with the same number of water molecules and a different conformation of the peptidic chain.

Taking these general features into account, the large and detailed search of the minima for the $Ala_3H^+ \cdot nH_2O$ model systems in an incremental way is reported from n=2 to 37.

Geometry Optimizations of Bare Peptide. Minima searches performed on the M062X/6-31+G*/PCM=WATER Born-Oppenhimer surface reveals the presence of four minima for the unfolded structure corresponding to "pure" and "mixed" extended and polyproline II arrangements in a narrow energy range (<0.5 kcal mol^{-1}, Fig. S1). The 3_{10}- or α-helices are not energetically favored because the peptide is too short to allow for the formation of an intramolecular H-bond in the C10 or C13 enclosures. Instead, two compact structures involving H-bonding of the terminally charged -NH_3^+ group and terminal -COOH with a reversed C11 enclosure (Fig. S1) are found ≈1.5 kcal mol^{-1} more stable than unfolded ones. These structures resemble a hairpin with two antiparallel H-bonds. However, torsional angles of these arrangements do not match optimal values, while important repulsion between neighboring groups of the chain occurs. Once a water molecule is placed in the



$4H_2O$_β-β_A -36.5 -30.4 $10H_2O$_β-β_A -88.6 -72.8 $13H_2O$_β-β -117.9 -95.2 $14H_2O$_β-β -127.4 -103.4

$18H_2O$_β-β_A -166.9 -135.4 $22H_2O$_β-β_A -199.0 -160.4 $24H_2O$_β-β_A -219.4 -176.4

$27H_2O$_β-β_A -248.0 -199.2 $31H_2O$_β-β_A -289.4 -230.4 $37H_2O$_β-β -350.9 -279.6

Fig. 1 Optimized molecular structures of $Ala_3H^+ \cdot nH_2O$ (n=4-37) complexes with the peptide in the β-β conformation. Formation energies (kcal mol^{-1}) at the 6-31+G* (value on the left) and aug-cc-pVTZ (value on the right) have been computed relative to the Ala_3H^+ (in the β-β conformation) and "n" isolated water molecules. Other structures are reported in Fig. S2.

proximity of the intramolecular H-bond, it spontaneously inserts itself between -NH_3 and -COOH groups, two more efficient H-bonds with water molecule are formed and some peptidic chain strain is released. The C11 structures cannot exist in aqueous solutions but they can play some role when the peptide is in the gas-phase.[22,23] On the basis of the above considerations, it should be clear that unfolded β and PPII structural elements are energetically accessible in aqueous solutions of the Ala_3H^+ peptide, and hereafter our discussion relates only to these conformations.

The classical Ramachandran plot distribution of amino acid residues shows a very large basin in the upper left quadrant where the β and PPII minima are located. It is not easy to define the borderline between β and PPII minima. In reference to the $AcAlaNHCH_3$ dipeptide structure, Cho et al.[42] define the β-conformation, $-180° \le \phi \le -120°$ and $60 \le \psi \le 180°$ while for the PPII, the borderline values are: $-120° < \phi \le -30°$ and $60 \le \psi \le 180°$. More recently, a larger definition of β-strand region ($-180° \le \phi \le -100°$) has been reported at PPII's expense ($-100° < \phi \le -40°$).[43] Hereafter, we will adopt the latter nomenclature to define the peptidic chain structure. Nevertheless, often the ϕ=-134°, ψ=145° and the ϕ=-75°, ψ=145° are indicated as β and PPII canonical conformations, respectively.[32]

Hydrating the β-β conformer. To disperse the net positive charge on the peptide, the H-bonding donation to water molecules involving terminal -NH_3 group assumes a prominent role; therefore, in our bottom-up approach it is appropriate to add, initially, water molecules to this highly polar group. One or two couples of water molecules can be added to connect the NH_3 group and the adjacent >CO in the "up" region that satisfy the alternance and the linearity of H-bonding criteria (**$2H_2O$_β-β** and **$4H_2O$_β-β** structures in Figs. 1 and S2). For the **$4H_2O$_β-β** model, four isoenergetic structures are found changing the orientation of the two water molecules H-bonded to the carbonyl. These structures involve the formation of four H-bonds between peptide and water molecules and two water-water H-bonds with a noticeable electronic energy of formation, -30.4 kcal mol^{-1}. As far as the electronic structure, the four water molecules play a dual role. On one hand they allow for an extensive electronic density transfer from water molecules to the positively charged peptide (0.14 e.u.) and, on the other, they also allow for a complementary tool for electron delocalization among various groups of the peptide.

The successive water molecules coordination involves the bare N-H bond of the terminal protonated amine in the "down" region. Close to this bond there is the >NH group of the central residue and after that there is the >CO H-bonds acceptor. A long wire of water molecules is necessary to get a reliable synergic H-bonding donor/acceptor. The size of the water wire is not easy to be evaluated *a priori*, so two model structures have been considered: the **$9H_2O$_β-β** and the **$10H_2O$_β-β** (Figs. 1 and S2). Both structures are minima with sizeable formation energy and with a similar structural arrangement for the >CO-$(H_2O)_2$ H-bonding at the central residue. The –NH_3---OH_2 and >N-H···OH_2 H-bond distances are shorter in the **$10H_2O$_β-β** model while related N-H···OH_2 angles in **$10H_2O$_β-β** structure are closer to the linearity than those in the **$9H_2O$_β-β** one. Both data suggest the presence of some strains in the **$9H_2O$_β-β** because of the non-optimal geometrical match thus, the **$10H_2O$_β-β** structure better models hydrating phenomena in the "down" region. Both models have the same number of peptide-water H-bonds thus, the peptide-water interaction energy -49.4 and -50.8 kcal mol^{-1} for **$9H_2O$_β-β** and **$10H_2O$_β-β** respectively, is in agreement with the better structural match found in the latter case.

Analogously to the **$4H_2O$_β-β** model, the **$10H_2O$_β-β** can exist in various configuratons depending on the relative orientation of the two water molecules H-bonded to the >CO of the N-terminal residue in the "up" region (Figs. 1 and S2). These structures are very close in energy and suggest that these two water molecules can be involved as both donor or acceptor of H-bonds with oncoming water molecules on adjacent hydrophilic groups.

Several attempts have been made to place water molecules in the "up" region over the >NH group of the C-terminal residue; however, the only arrangements that do not destroy the β-β peptidic conformation are the **$13H_2O$_β-β** and **$14H_2O$_β-β** models (Fig. 1). A common feature of these models is the presence of an H-bond between the two wires of water molecules as indicated by the arrow in Fig. 1. Starting from these models, peptide micro-hydration can further go on and two groups of structures can be formed. The **A** group is obtained by further hydrating the **$14H_2O$_β-β** (Fig. 1) while the **B** group structures (Fig. S2) derive from the **$13H_2O$_β-β**.

For the **$14H_2O$_β-β** model, the >CO---HOH bond distance for carboxylic group is significantly larger than those observed for carbonyl amidic (1.94 vs. 1.86 Å), thus indicating that there is a lower availability of lone pairs of the carboxylic group to be involved in H-bonding acceptance. This is due to the presence of the second electronic withdrawing oxygen atom and suggests the possibility that the >CO of the carboxylic group can be involved in only one H-bond at variance to the amidic carbonyl. The next hydrated structure is obtained forming a cluster of five water molecules over the carboxylic group in the "up" region, the **$18H_2O$_β-β_A** models. Because of its acidic nature, the COOH---OH_2 H-bond is very strong with a very short bond distance (1.51 Å) and the -COO-H distance lengthens significantly (from 0.98 to 1.04 Å).

Minimal hydration of all polar groups is reached placing a wire of four water molecules in the "down" region connecting the >CO of the central residue and the -OH of the terminal carboxylic functionality (**$22H_2O$_β-β_A** model). This structural feature endorses the COOH group to be involved in H-bonding donation, the COO-H distance further increases while the COOH---OH_2 distance decreases with respect to those found for the **$18H_2O$_β-β_A** model (Fig. 1). To further improve our hydration model, water molecules have been added to link dangling H-bonds of solvent molecules placed on the "up" and on the "down" regions of the minimal hydrated **$22H_2O$_β-β_A** structure. For both **$24H_2O$_β-β_A** and **$27H_2O$_β-β_A** there are no significant structural changes, and actually, the peptidic dihedral angles are maintained (Table S1). Conversely, some changes on the peptidic structure are obtained when four water molecules have been added to improve hydration on the carboxylic group (**$31H_2O$_β-β_A**) and six water molecules on the terminal charged amide (**$37H_2O$_β-β_A**). In both cases dihedral angles (Table S1) become close to those expected for the β-strand (ϕ=-134°; ψ=145°).

The group **B** models construction, starting from the **$13H_2O$_β-β**, have been developed in a similar way to group **A** with hydration of the carboxylic group in the "up" region (**$18H_2O$_β-β_B**, Fig. S2) and after in the "down" region. Analogously, twenty-two water molecules are necessary for the minimal hydrophilic group hydration (**$22H_2O$_β-β_B**, Fig. S2). During various attempts, carried out to get the above mentioned structures, another type of water arrangement was also found (**$18H_2O$_β-β_C** and **$22H_2O$_β-β_C**, Fig. S2). The water

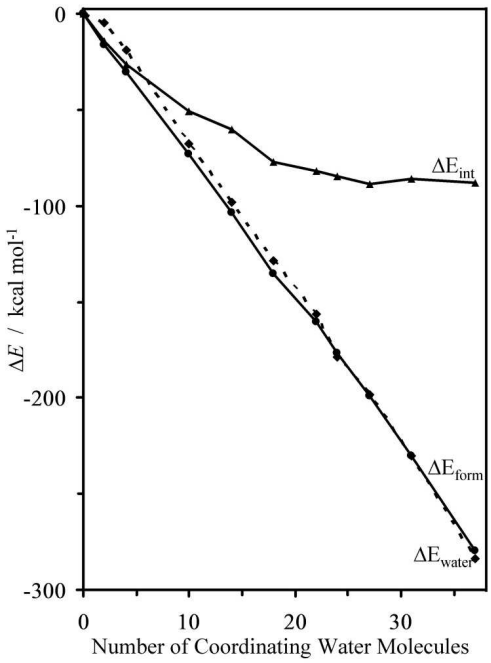


Fig. 2 Calculated M062X/aug-cc-pVTZ/PCM=water formation and interaction energies (ΔE_{form} and ΔE_{int}, respectively) for the $Ala_3H^+ \cdot nH_2O$ (n=2-37) complexes with the peptide in the β-β conformation. The energy of formation of $(H_2O)_n$ water clusters, ΔE_{water}, at the optimized geometry.

molecules cage differs from the **18H$_2$O_β-β_B** and **22H$_2$O_β-β_B** mainly for the absence of the weak H-bond of the two water-wires in the "up" region.

All these features are in accordance with many pieces of experimental evidence and many molecular dynamics studies, which suggested a small number of water molecules coordinated to acceptor and donor groups of peptides.[5-12,21] The formation of strong H-bonds, which disrupts water structure and avoids synergic water molecule reorientation in the immediate vicinity of peptide, is responsible for the slowdown of water motion.

The electronic energy of formation, ΔE_{form}, for the group **A** structures becomes more exothermic almost linearly as the number of hydrating molecules increases because of the increase in number of peptide-water and water-water formed H-bonds (Fig. 2). At the very beginning of the curve, the hydration energy, ΔE_{int}, falls quite quickly (almost overlaps with the ΔE_{form} line) because of the big interactions between water molecules and charged amine group. After that, the ΔE_{int} curve gets less steep, separates from ΔE_{form} line and the water-water interactions become predominant in the overall energy of formation. The ΔE_{int} data distributed approximately on a rectangular hyperbola with an asymptotic value of about -85 kcal mol^{-1}.

The dotted line in Fig. 2 represents the energy of formation of water clusters:

$$n(H_2O) \rightarrow (H_2O)_n + \Delta E_{water} \qquad n \text{ up to } 37 \qquad (3)$$

The first part of this curve (n<24) lies slightly above the ΔE_{form} line, while for the second part, the formation of simple water clusters is energetically favored over the peptide-water complexes. This is due to the higher number of H-bonds that are formed in clusters of pure water with respect to those that are formed on the surface of the heterogeneous peptide. For example, there are 38 H-bonds in the optimal geometry of the $(H_2O)_{22}$ cluster while, 25 are formed in the **22H$_2$O_β-β_A** model. From a computational point of view, this means that a global minimization approach would invariably produce the formation of a water droplet with the peptide coordinated on the surface as more stable configurations for hydrated trialanine. Therefore, our bottom-up method based on the intuitive geometrical arrangement of solvent over peptide and the hydration energy maximization is the only practicable way to get information from quantum chemical methods at a reasonable cost.

Hydrating the β-PPII conformer. Because of the geometrical analogies with the β-β conformer, the solvation around the protonated amine should occur in a similar way with four water molecules on the "up" region and six on the "down" side. To obtain a useful and stable **10H$_2$O_β-PPII** model, Fig. 3, a chain of water molecules connecting the H_3N- and -COOH termini is necessary. Instead of two water molecules bonded to the carbonyl group of the central residue in the "bottom" side as in the **10H$_2$O_β-β** case, for the **10H$_2$O_β-PPII** complex, a water molecule has been moved to link the terminal carboxylic group. The next three water molecules connect >NH of the C-terminal residue with the already present $(H_2O)_4$ cluster in the "up" region and one on the "down" side lead to the formation of the **14H$_2$O_β-PPII** complex. Two arrangements with the same energy are obtained (**A**-**B**). They differ mainly for the water molecule coordinated on the >NH group of the C-terminal residue. For the **14H$_2$O_β-PPII_A** complex, the water molecule acts as a donor of one H-bond and an acceptor of two H-bonds, whereas the reverse situation occurs for the **14H$_2$O_β-PPII_B** complex. Both systems present an H-bond between the two wires of water molecules in the up region as indicated by the arrow in the Fig. 3 in analogy to what happens for the **13H$_2$O_β-β** and **14H$_2$O_β-β** systems.

The two **14H$_2$O_β-PPII_A** and **14H$_2$O_β-PPII_B** structures can be further hydrated capping the carboxylic group with four water molecules (**18H$_2$O_β-PPII_A**, **18H$_2$O_β-PPII_A'** and **18H$_2$O_β-PPII_B**, respectively). Two other structures, **18H$_2$O_β-PPII_C** and **18H$_2$O_β-PPII_D**, has been obtained removing the H-bond between the two wires of water molecules. All the structures **18H$_2$O_β-PPII** have almost the same energy ($\Delta E \leq 0.7$ kcal mol^{-1}) and show a substantial gain in hydration energy because of the involvement of the carboxylic acid hydrogen in analogy to that observed for the **18H$_2$O_β-β** model.

Four additional water molecules placed in the region between the >NH and >CO of C-terminal residue connect the "top" and "down" water layers realize the minimal hydration of the peptide. The **22H$_2$O_β-PPII_A**, **22H$_2$O_β-PPII_B**, and **22H$_2$O_β-PPII_D** have a similar energy (within 1 kcal mol^{-1}) because they have a similar water covering. Instead, the **22H$_2$O_β-PPII_C** water molecules have a more compact H-bonding network with a greater number of H-bonds thus this structure is significant more stable (Fig. 3). In other words, the greater stability of the **22H$_2$O_β-PPII_C** structure is not due to an effective increase in hydration energy rather to an increase in water-water interaction energy. This different water molecules packing induces a large deviation in the ϕ_N angle (Table S1).

The dihedral angles of the part with extended β conformation undergoes to large variations, ϕ_N=-130±30° and ψ_N=150±10°, with average values comparable to those generally expected for β-strand (ϕ=-134° and ψ=145°, Table S1). For the PPII part, the angles spread in narrow values range (ϕ_C=-60±3°; and ψ_C=140±10°) and averaged values are close to those experimentally expected (ϕ_C=-75° and ψ_C=145°).

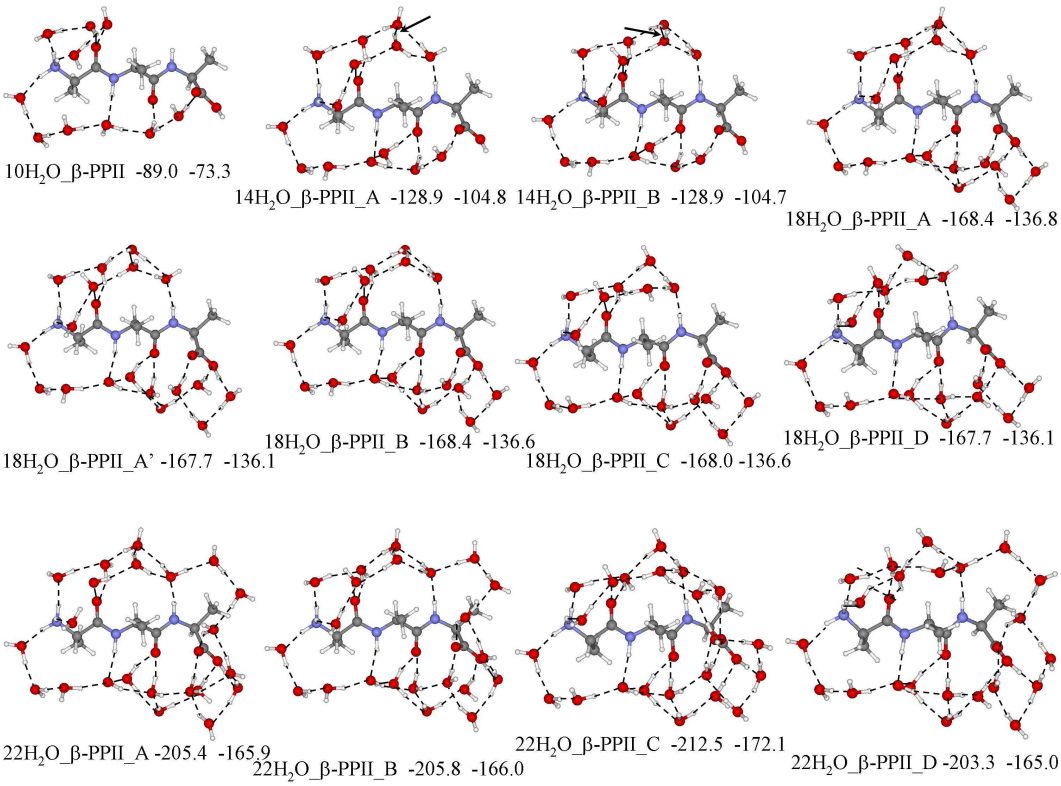


Fig. 3 Optimized molecular structures of $Ala_3H^+ \cdot nH_2O$ (n=10-22) complexes with the peptide in the β-PPII conformation. Formation energies (kcal mol^{-1}) at the 6-31+G* (value on the left) and aug-cc-pVTZ (value on the right) have been computed relative to the Ala_3H^+ (in the β-β conformation) and "n" isolated water molecules.

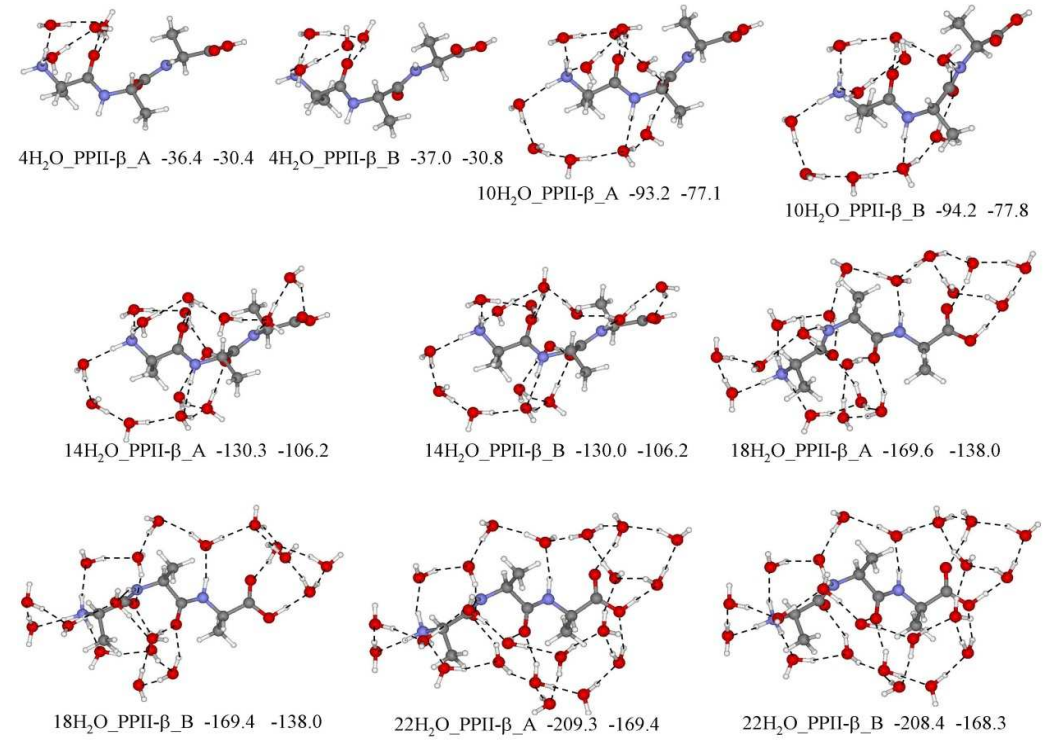


Fig. 4 Optimized molecular structures of $Ala_3H^+ \cdot nH_2O$ (n=4-22) complexes with the peptide in the PPII–β conformation. Formation energies (kcal mol^{-1}) at the 6-31+G* (value on the left) and aug-cc-pVTZ (value on the right) have been computed relative to the Ala_3H^+ (in the β-β conformation) and "n" isolated water molecules.

Hydrating the PPII-β conformer. For the **4H_2O_PPII-β** model, only two structures maintain a water coordination similar to those already reported for the β-β and β-PPII conformers (Fig. 4). The **10H_2O_PPII-β** structures are obtained by adding a water molecule that lengthens a pre-existing two-wire water and a new wire of five units in the distal region.Both lengthening and new wires connect the terminal charged -NH_3 and the central >CO groups. In analogy to other fourteen water molecule systems, depending on the orientation of the water molecule bonded to the >NH group of C-terminal residue and related water molecules over >CO group of N-terminal residue, two structures close in energy are found (**14H_2O_PPII-β_A** and **14H_2O_PPII-β_B**, Fig. 4). These structures expand with additional four water molecules to form a $(H_2O)_5$ cluster over the carboxylic group, the **18H_2O_PPII-β_A** and **18H_2O_PPII-β_B**, respectively. The microsolvation of hydrophilic groups is completed with four more water molecules on the other side of the carboxylic group to form the **22H_2O_PPII-β_A** and **22H_2O_PPII-β_B** models.

Hydrating the PPII-PPII conformer. Initial microsolvation around the PPII-PPII conformer occurs in a similar way to that found for the PPII-β one with the formation of the **4H_2O_PPII-PPII** model (Fig. 5). The next six water molecules bond the peptide in two alternative ways. In the **10H_2O_PPII-PPII_A** there is a complete solvation of polar groups of the N-terminal and central residues, while for the **10H_2O_PPII-PPII_B** there is a partial coordination of the carboxylic group. The greater stability of the **10H_2O_PPII-PPII_A** complex again emphasizes the greater H-bond acceptor tendency of >CO on amide with respect to >CO of carboxylic group. Depending on the orientation of the water molecule bonded to the >N-H of the C-terminal residue, two structures close in energy are obtained for the fourteen water molecules model (Fig. 5). With the formation of a water cluster over the carboxylic group, in the **18H_2O_PPII-PPII** models, all hydrophilic groups of peptide are hydrated, thus for compact structures like PPII, a lower number of water molecules is necessary to reach the minimal hydration. To get the **22H_2O_PPII-PPII** models, a chain of three water molecules is added between the water molecules bonded to >NH of N-terminal and >CO of C-terminal residues and the fourth molecule enlarges the water cluster over the carboxylic group.

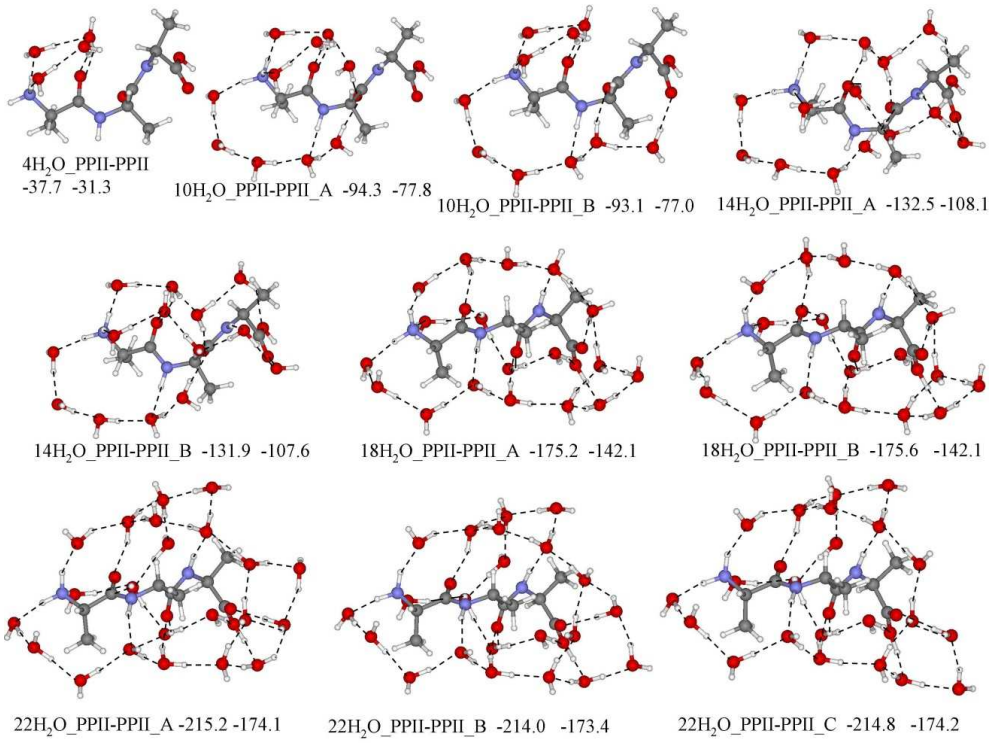


Fig. 5 Optimized molecular structures of $Ala_3H^+ \cdot nH_2O$ (n=4-22) complexes with the peptide in the PPII–PPII conformation. Formation energies (kcal mol^{-1}) at the 6-31+G* (value on the left) and aug-cc-pVTZ (value on the right) have been computed relative to the Ala_3H^+ (in the β-β conformation) and "n" isolated water molecules.

Energetics

All models have been built up with the same number of water molecules and generally various structures have about the same number of peptide-water and water-water H-bonds (12-13 and 22-24 for 22H_2O models, respectively) thus a direct comparison of thermodynamic parameters of the most stable configuration of each model becomes relevant. Fig. 6 reports relative electronic energy and relative entropy of various structures with respect to the fully-extended one, i.e. the β-β = β-PPII, β-β = PPII-β and β-β = PPII-PPII reactions energy and entropy, as the number of coordinating solvent molecules increase up to twenty-two.

For the bare peptide and models up to four water molecules, the structures have almost the same electronic energy thus indicating that at this stage, explicit water coordination acts in a similar way on all the structures. For models with ten water molecules there is the complete hydration of the N-terminal residue and the partial hydration of the central residue. The

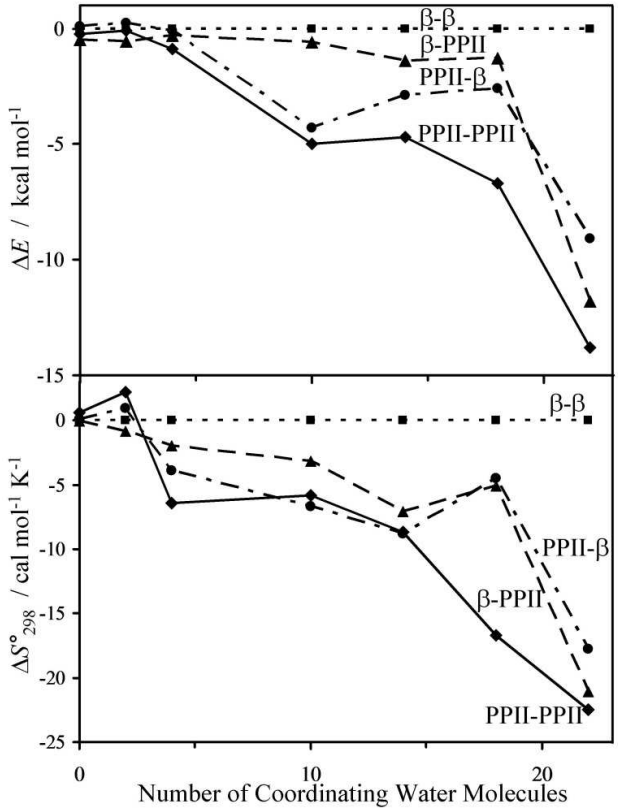


Fig. 6 Calculated relative electronic energies (ΔE) and relative entropies at 298 K (ΔS_{298}) for the $Ala_3H^+ \cdot nH_2O$ (n=2–22) complexes with the peptide chain in unfolded conformations. The fully extended β-β conformation is taken as a reference.

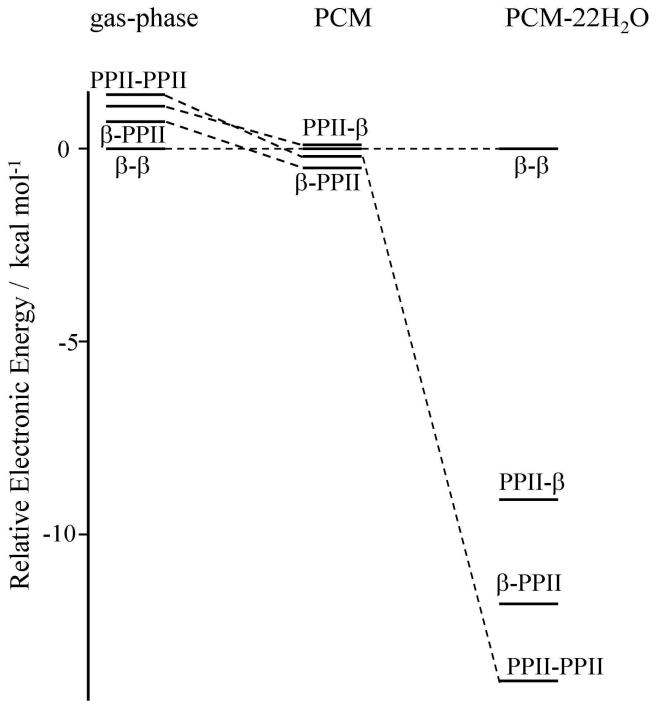


Fig. 7 Relative energy evolution of the four peptide conformations in the gas-phase, in the presence of the dielectric medium and plus twenty-two coordinated water molecules.

structures in which the hydrated part adopts the polyproline arrangement (PPII-PPII and PPII-β) undergo to a significant stabilization (≈4 kcal mol^{-1}) compared to the structures with the hydrated extended conformation (β-β and β-PPII). There is an extra energy gain due to the better coordination of the water molecules in the residue with the PPII conformation. The relative electronic energy sequence is maintained for the $14H_2O$ and $18H_2O$ models even though small fluctuations are noted. A significant increase in the stability of structures that have one or two residues with a polyproline conformation is obtained for the model $22H_2O$. Again, there is a differential effect of explicit water coordination on residues with extended or polyproline conformations and finally, the following scale of energy stability can be drawn:

β-β < PPII-β ≈ β-PPII < PPII-PPII

Although the data reported in Fig. 6 clearly show the importance of explicit coordination of water molecules on the stability of PPII conformation,[45] it is appropriate to remark that, in the PPII helix, intramolecular steric-repulsive interactions are minimized due to the staggered conformation of all substituents on the peptidic chain.[46] Actually, for the present peptide, data in the gas-phase indicate that the PPII conformers are all minima, but they are less stable than extended β conformations (Fig. 7).

Less evident, but nevertheless very important for the PPII stability, are the implicit solvent effects, i.e. the mutual polarization of peptidic dipole and solvent. The dielectric medium significantly stabilizes the residue with the PPII conformation because the coupling of dipoles is much more pronounced with respect to the β extended conformation.[27,34] Fig. 7 shows the energy sequence of bare peptide in the gas phase, bare peptide in the dielectric medium, and the peptide with the twenty-two coordinated water molecules. It is interesting to follow the energy evolution once implicit and explicit solvent contributions are considered.

In regard to the present proposed PCM/bottom-up approach, it is important to note that the dielectric medium significantly shields explicit peptide-water and water-water interactions as already discussed for the N-acetyl-L-alanine amide.[35] This is particularly true for the present case where the positively charged peptide produces ion-dipole interactions that are much stronger than dipole-dipole ones, present in neutral protected peptides. For these reasons many of the above described structures do not exist once the geometry optimization is carried out in vacuum and generally more thickened structures are obtained. The small number of water molecules explicitly treated in the present study hampers the long-range solvent effect inclusion and an appropriate methodology, as the PCM, need to be considered in our bottom-up model.[35]

Entropy is another important parameter to be considered to describe extended – polyproline transformations with "hydrated" structures. The hydrogen bond potential is largely anharmonic due to extensive intermolecular coupling and long range cooperative motion, thus it is difficult to obtain reliable absolute entropy for presently analyzed systems.[47] For example, the standard thermochemical procedure applied to the formation of $(H_2O)_m$ clusters (m up to 10) overestimates ΔS by about 10-20 %. This shortcoming undermines the effectiveness of free energy estimation and in particular the $-T\Delta S$ term. Nevertheless, we can assume that the rigid rotor, harmonic oscillator and ideal gas approximations act on a similar way on entropy computation, hence some trends can be derived at the qualitative level. Fig. 6 reports the entropy variation for the β-β

= β-PPII, β-β = PPII-β and β-β = PPII-PPII transformations obtained for the most stable and homologous **nH_2O_β-β**, **nH_2O_β-PPII**, **nH_2O_PPII-β** and **nH_2O_PPII-PPII** complexes. Even though there are fluctuations, some trends are evident and allow us to draw the following scale for entropy content:

β-β > PPII- β ≈ β-PPII > PPII-PPII

The β-β conformation has the highest entropy content but the lowest energy stability, while the PPII-PPII has the lowest entropy and highest energy stability. For the mixed β-PPII and PPII-β conformations, an intermediate situation occurs. This is in agreement with experimental evidence that at low temperature shows this molecule mainly adopting the PPII helix, while on increasing the temperature, the fraction of β-strand increases quickly and becomes predominant.[13-18]

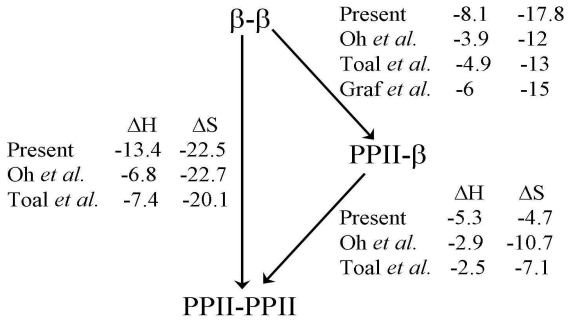


Fig. 8 Experimental and computed thermodynamic parameters (ΔH in kcal mol^{-1}; ΔS in cal mol^{-1} K^{-1}) related to β=PPII conformational transitions for Ala_3H^+ at 298K.[14-16]

Computed values of $\Delta H°_{298}$ and $\Delta S°_{298}$ for the complexes with the highest hydration degree, the $22H_2O$ models, are reported in Fig. 8 together with related experimental data derived from NMR, UVCD, VCD and Raman optical activity spectroscopies.[14-16] Considering the experimental uncertainty, the experimental/computation comparison for ΔS is satisfactory.

For the enthalpy, experimental/computation comparison is useful at the qualitative level because calculated values are considerably more exothermic, about twice the experimental values. The excess in hydration effect indicates a huge interaction of water molecules with the peptide. In other words, the water molecules having the dangling H-bonds unsatisfied have a high freedom and interact more significantly with the peptide than how they would do when immersed in a bath of water. This increased interaction stabilizes artificially the residues with PPII conformation.

Table 1. Relative electronic energy (ΔE, in kcal mol^{-1}) of the four $22H_2O$ models computed with various functionals and the aug-cc-pVTZ basis set.

	M06-2X	TPSS-D3	B3PW91-D3	B3LYP-D3
β-β	0.0	0.0	0.0	0.0
β-PPII	-11.8	-14.1	-14.5	-13.6
PPII-β	-9.1	-10.6	-11.3	-10.1
PPII-PPII	-13.8	-12.6	-14.0	-12.5

The reliability of M06-2X has been previously verified in a comparative MP2 study on the microhydration of the $AcAlaNH_2$ peptide and conformational stability of the $AcAla_4NH_2$ peptide.[27,35] Nevertheless, to test the robustness of present results selected calculations have been performed using other type of DFT, specifically the pure TPSS-TPSS and the hybrid B3LYP and B3PW91 including Grimme-D3 dispersion terms (Table 1). The energetic trend is similar to that found for the M06-2X and the only significant difference is found for the strong stabilization of the β-PPII conformer. Nevertheless, the abnormal stability of the **$22H_2O$_β-PPII_C** structure has already been discussed (*vide supra*) and it is due to a more compact water cluster that form an extra water-water H-bond. These results indicate that the M06-2X method can be safely used to investigate hydration phenomena of peptides.

The validity of present models has been also tested changing the degree of hydration and structures with five, six, eight, eleven, thirteen, seventeen and nineteen water molecules have been optimized. Results reported in Fig. S3 and Table S1 show that the dihedral angles of backbone undergo to minor variations with respect to those already discussed and, that the models with the peptide in the PPII-PPII conformation results always more stable.

Conclusions

To understand the detailed interactions of the Ala_3H^+ peptide with surrounding water molecules and to develop a simple, intuitive and promising methodology for the hydration of small sized peptides, quantum chemical investigation was carried out using the DFT-M062X approach to generate putative poses of water molecules around the peptide, $Ala_3H^+ \cdot nH_2O$ (n up to 37). The four "unfolded" conformers almost isoenergetic and isoentropic in the bare peptide, progressively separate as the number of coordinate water molecules directly bound to the hydrophilic groups of peptide increase. The variations on energy and entropy among various structures reflect experimental results at the semiquantitative level. These results have been obtained without introducing any external parameters intended to quantify empirical or semiempirical potential-energy functions thus the proposed methodology is a promising way to understand and predict hydration phenomena at the peptide/water interface.

Given the many variables available in solvation modelization (i.e., water molecules number, H-bond anharmonicity, etc.), and given the peptidic surface heterogeneity and its topographical complexity, additional work is under way to optimize methodology and exploit its applicability to more elaborate peptides.

Acknowledgements

This work was supported by the Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN 2010-2011, 20109Z2XRJ_003).

Notes and references

Dipartimento di Scienze del Farmaco, Università di Catania
Viale A. Doria 6, 95125 Catania (Italy).
E-mail: glanza@unict.it.

Dedicated to the 70th birthday of Prof. Antonino Corsaro.

Electronic Supplementary Information (ESI) available: Figs. S1, S2 and S3 and a complete list of dihedral angles and Cartesian coordinates of all structures presently analyzed. See DOI: 10.1039/b000000x/

1 R.G Bryant, M.A. Johnson and P.J. Rossky, *Acc. Chem. Res.*, 2012, **45**, 1, and following articles.
2 G. Hummer and A. Tokmakoff, *J. Chem. Phys.*, 2014, **141**, 22D101, and following articles.
3 S.K. Kim, T. Ha and J.-P. Schermam, *Phys. Chem. Chem. Phys.*, 2010, **12**, 10145, and following articles.
4 R. Winter and R. Ludwig, *ChemPhysChem*, 2008, **9**, 2635, and following articles.
5 C. Malardier-Jugroot, M. E. Johnson, R. K. Murarka and T. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4903.
6 J. Qvist, E. Persson, C. Mattea and B. Halle, *Faraday Discuss.*, 2009, **141**, 131.
7 P. G. Takis, K. D. Papavasileiou, L. D. Peristeras, V. S. Melissas and A. N. Troganis, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7354.
8 K. Mazur, I. A. Heisler and S. R. Meech, *J. Phys. Chem. B*, 2010, **114**, 10684.
9 L. Comez, S. Perticaroli, M. Paolantoni, P. Sassi, S. Corezzi, A. Morresi and D. Fioretto, *Phys. Chem. Chem. Phys.*, 2014, **16**, 12433.
10 B. Born, H. Weingärtner, E. Bründermann and M. Havenith, *J. Am. Chem. Soc.*, 2009, **131**, 3752.
11 P. Sasisanker and H. Weingartner, *ChemPhysChem*, 2008, **9**, 2802.
12 K. Bhattacharyya, *Chem. Commun.*, 2008, 2848.
13 S. Woutersaen and P. Hamm, *J. Phys. Chem.*, 2000, **104**, 11316.
14 J. Graf, P. H. Nguyen, G. Stock and H. Schwalbe, *J. Am. Chem. Soc.*, 2007, **129**, 1179.
15 S. Toal, D. Meral, D. Verbaro, B. Urbanc and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2013, **117**, 3689.
16 K.-I. Oh, K.-K. Lee, E.-K. Park, D.-G. Yoo, G.-S. Hwang and M. Cho, *Chirality*, 2010, **22**, E186.
17 B. Sharma and S. A. Asher, *J. Phys. Chem. B*, 2010, **114**, 6661.
18 X. Xiao, N. Kallenbach and Y. Zhang, *J. Chem. Theory Compt.*, 2014, *10*, 4152.
19 J. Kubelka, R. Huang and T. A. Keiderling, *J. Phys. Chem. B*, 2005, **109**, 8231.
20 J. Grdadolnika, V. Mohacek-Grosevb, R. L. Baldwin and F. Avbelj, *Proc. Natl. Acad. Sci USA*, 2011, **108**, 1794.
21 A. Kuffel and J. Zielkiewicz, *J. Phys. Chem. B*, 2012, **116**, 12113.
22 R. Wu and T. B. McMahon, *J. Am. Chem. Soc.*, 2007, **129**, 11312.
23 T. D. Vaden, T. S. J. A. De Boer, J. P. Simons, L. C. Snoek, S. Suhai, and B. Paizs, *J. Phys. Chem. B*, 2008, **112**, 4608.
24 D. Liu, T. Wyttenbach, P. E. Barran and M. T. Bowers, *J. Am. Chem. Soc.*, **2003**, *125*, 8458.
25 R. J. Plowright, E. Gloaguen and M. Mons, *ChemPhyChem,* 2011, **12**, 1889.
26 a) S. Gnanakaran, and R. M. Hochstrasser, *J. Am. Chem. Soc.*, 2001, **123**, 12886; b) G. Pohl, A. Perczel, E. Vass, G. Magyarfalvi and G. Tarczay, *Phys. Chem. Chem. Phys.*, 2007, **9**, 4698.
27 G. Lanza and M. A. Chiacchio, *ChemPhysChem*, 2013, **14**, 3284.
28 M. Knapp-Mohammady, K.J. Jalkanen, F. Nardi, R.C. Wade and S. Suhai, *Chemical Physics*, 1999, **240**, 63.
29 J. Sebek, J. Kapitan, J. Sebestik, V. Baumruk, and P. Bour, *J. Phys. Chem. A*, 2009, **113**, 7760.
30 J. Ireta, *Int. J. Quantum Chem.,* 2012, **112**, 3612.
31 S. Boopathi and P. Kolandaivel, *J. Biomol. Struct. Dyn.,* 2013, **31**, 158.
32 N. G. Mirkin and S. Krimm, *Biopolymers,* 2009, **91**, 791.
33 M. Marianski and J. J. Dannenberg, *J. Phys. Chem. B,* 2012, **116**, 1437.
34 G. Lanza, U. Chiacchio, S. Motta, S. Pellegrino and G. Broggini, *ChemPhysChem*, 2011, **12**, 2724.
35 G. Lanza and M. A. Chiacchio, *ChemPhysChem*, 2014, **15**, 2785.
36 A. Cimas, T. D. Vaden, T. De Boer, L. C. Snoek and M. P. Gaigeot, *J. Chem. Theory Comput.*, 2009, **5**, 1068.
37 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.
38 Gaussian 09, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.
39 C. Grebner, J. Kästner, W. Thiel, B. Engels, *J. Chem. Theory Comput.*, 2013, **9**, 814.
40 M. Mobli, A. Almond, *Org. Biomol. Chem.*, 2007, **5**, 2243.
41 G. Lanza, M. A. Chiacchio, S. V. Giofrè, R. Romeo, P. Merino, *Eur. J. Org. Chem.*, 2013, 95.
42 K. Kwac, K.-K. Lee, J. B. Han, K.-I. Oh, M. Cho, *J. Chem. Phys.*, 2008, **128**, 105106.
43 R. Schweitzer-Stenner, *Mol. BioSyst.*, 2012, **8**, 122.
44 J. M. Mullin and M. S. Gordon, *J. Phys. Chem. B*, 2009, **113**, 14413.
45 C. F. Weise and J. C. Weisshaar, *J. Phys. Chem. B*, 2003, **107**, 3265.
46 A. N. Drozdov, A. Grossfield and R. V. Pappu, *J. Am. Chem. Soc.*, 2004, **126**, 2574.
47 B. Temelso and G. C. Shields, *J. Chem. Theory Comput.*, 2011, **7**, 2804.z

