Unveiling the role of histidine and tyrosine residues on the conformation of the avian prion hexarepeat domain

Adriana Pietropaolo[1], *Luca Muccioli*[1], *Claudio Zannoni*[1*], *Diego La Mendola*[2], *Giuseppe Maccarrone*[3], *Giuseppe Pappalardo*[2] *and Enrico Rizzarelli*[3]

[1] Dipartimento di Chimica Fisica e Inorganica and INSTM, Università di Bologna, v.le Risorgimento 4, 40136 Bologna (Italy)

[2] CNR-Istituto di Biostrutture e Bioimmagini Catania, v.le A. Doria 6, 95125 Catania (Italy)

[3] Dipartimento di Scienze Chimiche, Università di Catania, v.le A. Doria 6, 95125 Catania (Italy)

[*]corresponding author

Claudio.Zannoni@unibo.it

Abstract

The prion protein (PrP^{C}) is a glycoprotein that in mammals, differently from avians, can lead to prion diseases, by misfolding into a β sheet rich pathogenic isoform (PrP^{Sc}). Mammal and avian proteins show different N-terminal tandem repeats: PHGGGWGQ and PHNPGY, both containing histidine, while tyrosine is included only in the primary sequence of the avian protein. Here by means of potentiometric, circular dichroism (CD) and molecular dynamics (MD) studies at different pH, we have investigated the conformation of the avian tetra-hexarepeat $(PHNPGY)_4$ (TetraHexaPY) with both N- and C-termini blocked by acetylation and amidation respectively. We have found, also with the help of a recently proposed protein chirality indicator [A. Pietropaolo et al. *Proteins*, (2007) doi: 10.1002/prot.21578], a conformational dependence on the protonation states of histidine and tyrosine residues: the turn formation is pH driven and at physiological pH a pivotal role is played by the tyrosine OH groups which give rise to a very compact bent structure of backbone upon forming a hydrogen bond network.

keywords: MD, circular dichroism, chirality, prion proteins, pH, DSSP

Introduction

The prion protein is a highly conserved cell surface glycoprotein expressed in mammals as well as in several species of fish and birds [1-2]. The avian prion protein shares only about the 30% of identity in primary sequence with mammal prion proteins, although it possesses some essential features: multiple N-glicosylated sites; an amino-terminal signal sequence that is removed in the mature protein; an N-terminal domain featured by tandem amino acid repeats (PHNPGY in avian, PHGGGWGQ in mammals), followed by a highly conserved hydrophobic core [3] containing three α-helices, one short 3_{10}-helix and a short antiparallel β-sheet [4]. Despite these structural similarities, it should be noted that the normal isoform of mammalian prion protein is totally degraded by proteinase K, while avian prion protein is not, producing N-terminal domain peptide fragments stable to further proteolysis [5]. Besides, it is well known that prion diseases, neurodegenerative disorders in mammals, are associated with a misfolded and beta-sheet rich isoform (PrP^{Sc}) of prion protein, while the same disorders seem spared to non mammals [6]. Although the structural changes from PrP^{C} to PrP^{Sc} occur at the structured C-terminal domain, the N-terminal component is supposed to play a regulatory role in these processes [7,8]. This region can bind copper ions and the copper transport is considered one of the most likely biological functions carried out by the prion protein [9-12]. NMR data show that the N-terminal domain in mammals is flexible and unordered [13,14], even if it was also reported that the octa-repeat region can adopt an extended conformation like a polyproline II or a loop conformation and a beta-turn [15], depending on pH [16]. The avian hexarepeat sequences are also located in the flexible, disordered and highly solvent-accessible N-terminal region and analogously to the case of mammals, NMR experiments carried out at pH 4.5 on the whole avian prion protein failed out in the individual assignments of these residues [4]. However, CD data concerning the entire protein and different peptide fragments containing hexarepeat sequences clearly suggest the presence of more than one conformation and not simply of a random coil one [5, 17, 18]. In particular, this has been pointed out by previous studies

on bis-hexarepeats, different single hexarepeats and analogues with single residue mutation (i.e. a tyrosine replaced by phenylalanine), where the CD spectral shapes and thus the conformational equilibria have been found to be strongly dependent on pH. It has been also suggested that histidine and tyrosine residues play a role in stabilizing one of the conformers and that the location of the PXXP motif along the peptide backbone may also determine the conformational features [19].

Considering all these points, a more structured conformation adopted by chicken hexarepeat region with respect to the mammal analogue could explain the different behavior towards proteolysis [5]. To prove this speculation, we have recently carried out Molecular dynamics (MD) simulations and NMR measurements on the single Hexarepeat PHNPGY at different pH values, showing the presence of turn structures in the NPGY region and the dependence of the peptide shape on tyrosine deprotonation [20]. Here, in order of gain insight into the conformational preferences of a more extended region of the hexarepeats and to better clarify the role of the different protonation states of histidine and tyrosine residues, we report potentiometric, CD and MD studies of the Tetra-Hexarepeat Ac-$(PHNPGY)_4$-NH_2 (TetraHexaPY) carried out at different pH values.

Materials and Methods

Peptide synthesis and purification

The peptide Ac-$(PHNPGY)_4$-NH_2 was synthesized with N- and C-termini blocked on a Pioneer^{TM} Peptide Synthesizer. All residues were introduced according to HATU/DIEA activation method starting from an Fmoc chemistry on Fmoc-PAL-PEG-PS resin (substitution 0.25 mmol/g, 0.1 mmol scale synthesis, 0.4 g of resin). The synthesis was carried out under a four-fold excess of amino acid at every cycle. Removal of Fmoc protection during synthesis was achieved by means of 20% piperidine solution in DMF. N-terminal acetylation was performed by treating the fully assembled

and protected peptide resin (after removal of the N-terminal Fmoc group) with a solution containing acetic anhydride (6% v/v) and DIEA (5% v/v) in DMF. The peptide was cleaved off from the resin and deprotected by treatment with a mixture of water / triisopropylsilane/ trifluoroacetic acid (95/2.5/2.5 v/v) for 1.5 hours at room temperature. The solution containing the free peptide was filtered off from the resin and concentrated in vacuo at 30 °C. The peptide was precipitated with freshly distilled diethyl ether. The precipitate was then filtered, dried under vacuum, re-dissolved in water and lyophilised. The resulting crude peptide was purified by preparative reversed-phase high-performance liquid chromatography (Rp-HPLC).

Rp-HPLC was carried out by means of Varian PrepStar 200 model SD-1 chromatography system equipped with a Prostar photodiode array detector with detection at 222 nm. Purification was performed by eluting with solvent A (0.1% TFA in water) and B (0.1% TFA in acetonitrile) on a Vydac C_{18} 250x22 mm column (300 Å pore size, 10-15 μm particle size), at flow rate of 10 mL/min. The peptide TetraHexaPY was eluted using a linear gradient (0-20%) in solvent B. The elution profiles were monitored at 222 nm and 278 nm, and the peptide fractions were collected and lyophilised. Sample identity was confirmed by ESI-MS (Calculated mass TetraHexaPY $C_{126}H_{161}N_{37}O_{33}$ M=2720.21; found m/z $[M+2H]^{2+}$= 1361.10; [M+3H]3+= 907.73; [M+4H]4+=681.05).

Potentiometric measurements

Potentiometric titrations were performed with a computer-controlled Metrohm digital pH meter (Model 654) and a Hamilton digital dispenser (mod Microlabm). The titration cell (2.5 ml) was thermostated at 25.0 ± 0.2 $^{\circ}$C and all solutions were kept under an atmosphere of argon, which was bubbled through another solution under the same conditions of ionic strength and temperature. A KOH solution was added through a Hamilton burette equipped with 0.25 or 0.50 cm^3 syringes. The combined microelectrode (ORION 9103SC) was calibrated on the pH= -log $[H^+]$ scale by titrating HNO_3 with CO_2 free base. The ionic strength of all solutions was adjusted to 0.10 mol dm^{-3} (KNO_3). The analytical concentrations of TetraHexaPY ranged from 2.5 x 10^{-3} to 5.0 x 10^{-3} mol dm^3. Stability

constants for proton complexes were calculated from three or four titrations carried out over the pH range 2.5–10.6. Calculations of the electrode system, E°, E_j and K_W values as well as ligand purity were determined by the least square ACBA computer program [21]. The protonation constants were calculated by means of the HYPERQUAD program [22].

CD measurements

CD spectra were recorded on a JASCO 810 spectropolarimeter at a scan rate of 50 nm/minute and 0.1 nm resolution. The pathlengths were 1 or 0.1 cm, in the 190-800 nm range. The spectra were recorded as an average of 10 or 20 scans. The CD instrument was calibrated with ammonium (+)-camphor-10-sulfonate. Peptide solutions were prepared in water in a concentration range of $1x10^{-5}$-$1x10^{-6}$ mol dm^{-3} and varying the pH by addition of a diluted solution of potassium hydroxide or hydrochloric acid.

Theoretical Calculations

Molecular Dynamics.

An extensive Molecular Dynamics (MD) study of the fragment Ac-$(PHNPGY)_4$-NH_2 (TetraHexaPY) was carried out in water. Taking into consideration the pK_a of histidine and tyrosine residues (see Table 1), by assuming acidic pH in the simulation, histidines are protonated and tyrosines are in the neutral state (labelled LH_8^{4+}), at neutral pH both histidines, (protonated at the δ nitrogen) and tyrosines are in the neutral state (labelled LH_4) and finally at basic pH histidines are in the neutral state and tyrosines are deprotonated (labelled L^{4-}). Four chloride ions and four sodium ions were added at acidic and basic pH conditions respectively, to ensure charge neutrality in the simulation box. All the simulations were run in water using GROMACS 3.3 [23] and the Amber94 force field [24], using the SPC model [25] for water. The ESP charges of deprotonated tyrosine, not available in the FF, were calculated as previously suggested [20]. The starting configuration of TetraHexaPY (Figure 1) was built by linking four times the most representative NMR structure of MonoHexaPY previously reported [20]; a cubic box containing one TetraHexaPY chain and 4338

water molecules with periodic boundary conditions (PBC) was used in the isothermal-isobaric ensemble (NPT, P=1 atm, T=300 K), with the temperature controlled using a Berendsen thermostat [26]. Long simulation runs, of about 150 nanoseconds for each of the three protonation states were performed. Preliminarily, to assess the effective equilibration of the peptide, we analyzed the hydrogen bonds formation, following L. J. Smith et al. [27], and the evolution of the end-to-end distances with time, reported respectively in Figure 2 [a] and [b]. The number of hydrogen bonds increases, while the end-to end distances (Figure 2 [b]) show a fast decrease in the first 40 ns, with a reorganization time of 10 ns (40 ns considering also the 30 ns of the equilibration) for the LH_4 state, and then remain substantially stable for each case. Accordingly, we considered the first 40 ns as equilibration runs and totally excluded them from the production analysis, to avoid a starting configuration bias. Moreover, after equilibration, we checked that volume, total energy, the number of H-bonds and end-to end distances fluctuated around their average value, without systematic drifts. The trajectory analysis was thus performed on 110 ns-long production runs, with configurations stored every 2.5 ps.

Results and Discussion

Protonation constants and Far UV-CD of TetraHexaPY

The protonation constants of the TetraHexaPY are reported in Table 1; it is clear that the peptide contains 8 protonation sites and their assignments are also shown in Table 1. The side chain phenolic-OH of tyrosine residues have the highest pK values and their deprotonation take place in overlapping processes between pH 9 and 11. Protonation of the four imidazole-N atoms takes place in the pH range 5-7 and shows as well overlapping processes; the average pK value is 6.1, which is the range of the imidazole pK values reported in literature.

Far UV-CD spectra of TetraHexaPY peptide, carried out at different pH values, are reported in Figure 3. At pH 4 the spectrum is broad with a minimum at 203 nm, a weak shoulder around 216 nm and a maximum at 230 nm. Generally, this shape indicates an equilibrium between different

conformations, suggesting the presence of other secondary structure elements besides the random coil. At this value of pH, all histidyl residues are protonated and the spectrum shape is similar to that found for shorter peptide fragments MonoHexaPY and BisHexaPY as previously reported [19]. For those fragments, we suggested the presence of both random coil and beta-turn structures [20], coherently with the primary sequence that encompasses a PXXP motif, generally supposed to favor beta-turn and/or polyproline II structure [28]. However, while an intensity enhancement of the minimum at 200 nm and the disappearing of the maximum at 230 nm was observed on increasing the pH for mono and bis-hexarepeats, a different trend is observed here (Figure 3) for the four-tandem repeat. In fact, up to pH 7 we do not observe any variation in the band centered at 230 nm, while there is a shift of the minimum towards 200 nm, together with a decrease of the band intensity. At basic pH values we notice a general broadening of the spectra and the significant decrease of the signal at 230 nm. In addition, at pH 10 a maximum at 250 nm is observed, that can be easily attributed to the deprotonation of tyrosine residues. The corresponding decrease and then disappearance of the maximum at 230 nm at basic pH allows to relate this latter band to a positive signal of the phenolic group of tyrosine residues, which fades out with their progressive deprotonation. Actually it is well known that aromatic side-chains can give rise to a contribute to the far UV-CD spectra of peptides and that this is red-shifted for the phenolate ion with respect to the phenol [29,30]. Besides the evidences of deprotonation of specific residues, secondary structure variations can also be identified as pH increases: the strong positive band at 190 nm (Figure 3) and the shoulder found at 216 nm are features typical of β turn like conformations [31, 32, 33], which appear to be predominant at neutral and basic pH. Concerning the mono and bis-hexarepeat, the folded states were destabilized at basic pH values, although, similarly, the adopted structure was mainly the type I β turn at physiological pH.

On the whole, the trend observed increasing the pH indicate that a β-turn conformation prevails at basic pH and that the conformational equilibria are strongly dependent on protonation steps of histidines and tyrosines.

Molecular Dynamics simulations

As above mentioned, NMR studies carried out on entire protein did not provide any particularly relevant data on the conformation of tandem hexarepeats. Therefore in order to better understand the conformational equilibria involved in solution, suggested from CD measurements, and to further rationalize the effects of pH, a fully atomistic molecular dynamics study was carried out in water.

After verifying the effective equilibration of the peptide backbone as already described, we analyzed the different conformational states of TetraHexaPY on the production trajectories. To identify the contacts occurring inside the peptide structure, we calculated the average contact map concerning the C_α–C_α interatomic distances between two non consecutive aminoacids, using a cutoff window, to exclude trivial contacts, shorter than 8.5 Å [34]. Looking at the contact maps, which we report in Figure 4, end-to end contributes were found at acidic and mainly at neutral pH, thus attesting the presence of a bent structure in the LH_4 form and moreover, long range contacts are present in the left region of the map concerning LH_4 and LH_8^{4+} species.

To better characterize the different conformations adopted by TetraHexaPY emerging from the contact map, and to probe the role of specific aminoacids, we subsequently analyzed the local secondary structure during the simulation using the DSSP algorithm [35]. The resulting time evolution of TetraHexaPY conformers is shown in Figure 5. We found prevalently i-i+3 hydrogen bonds, as in the single repeat sequence [20], with few i-i+4 hydrogen bonds in the 16-20 and 11-15 regions (Figure 5) only for the LH_4 and L^{4-} state. In particular, at acidic pH two consecutive i-i+3 hydrogen bonds are present in the GYPHN sequence (5-8 and the 6-9, Figure 5) characteristic of a 3_{10} helix [35], which involves residues 6-8 (Figure 6). Two single i-i+3 hydrogen bonds were also found in the 9-12 (NPGY) and 21-24 (NPGY) regions (Figure 5), corresponding to a type I β turn structure for residues 10-11 and 22-23. Significantly, the number of residues involved in turn conformation increases with histidines deprotonation (see Tables 2, 3, 4). This is perfectly in good agreement with the formation of the shoulder at around 216 nm and the positive values approximately at 190 nm, observed in the CD spectra increasing the values of pH. In particular, at

neutral pH, the β turn conformation is basically driven by the interaction between the deprotonated imidazole of histidine 8 and the phenol hydrogen of tyrosine 18, which causes a strong tilting of the peptide backbone (Figure 6), disrupting the 3_{10} helix structure in the 6-8 YPH region, found at acidic pH.

Work is in progress, performing molecular dynamics studies of the whole chicken prion protein [36] to test if this interaction occurs also in the hexarepeat region bound to the protein globular core.

Such β turn conformation leads to the formation of a family of conformers in which a new 3_{10} helix is stabilized in the 17-20 GYPH region, carrying tyrosine 24 inside the backbone. Consequently, the phenolic hydroxyl groups of Tyr6, Tyr18 and Tyr24 get close, as shown in Figure 6. At basic pH, the deprotonation of tyrosine 24 determines another weak bond with the amide side chain hydrogens of asparagine 21 (Figure 6), previously found also in the MonoHexaPY at the same pH [20]. Consequently, glycine 17 and hystidine 20 stay much closer, thus making possible the formation of turns in the 17-20 GYPH regions, as underlined from Figure 5 for residues 18, 19 and 20 of L^{4-}. Here, the deprotonated Tyrosine 18 is stable in a turn conformation if compared to the LH_4 and LH_8^{4+} states, in which the protonated Tyrosines 18 spend the initial time in a bend, interconverting later to turn and a 3_{10} helix structure.

Chirality analysis.

As a complementary test of the secondary structure, we have also investigated the conformations of TetraHexaPY in terms of their chirality, using a methodology recently proposed by some of us [37]. The method consists in dividing up the entire backbone sequence in a number of fragments and assigning to each one a chirality index calculated from their geometry. This index, a scaled version [38] of the chiral index of Osipov et al. [39] suitably modified for probing local chirality along the backbone [37], is expressed as follows:

$$G^{a,Na} = 4!/(3N_a^4) \sum_{\substack{\text{all permutations} \\ \text{of i,j,k,l}}} [(\mathbf{r}_{ij} \times \mathbf{r}_{kl}).\mathbf{r}_{il}]\,(\mathbf{r}_{ij} \cdot \mathbf{r}_{jk})\,(\mathbf{r}_{jk} \cdot \mathbf{r}_{kl})/[(r_{ij}\, r_{jk}\, r_{kl})^2\, r_{il}] \quad \text{if } r_{ij}, r_{kl}, r_{il}, r_{jk} < r_c, \text{ and } a \le i,j,k,l \le N_a + a - 1$$

where N_a is the number of consecutive backbone atoms (N, C_α, C) involved in the chirality calculations, and a is the atom from which the calculation of the chiral index inside a fragment of N_a atoms starts; r_c is the cutoff radius (12 Å, see [37]), added to avoid the computation of unnecessary long-range terms. We have shown previously [37] that the chirality index allows assigning the motif type to the fragment, (see Table 5) complementing the DSSP classification. In Figure 7 we report the pattern of *G* along the TetraHexaPY (N, C_α, C) backbone atoms, with N_a equal to 15, corresponding to five residues at once involved in the calculation.

From the trend of the chirality index, averaged among the trajectories, the structures previous found using DSSP analysis can be recognized. For all the protonation states, the C-terminal region shows a pattern with chirality approaching zero, typical of coil structures. At acidic pH the broad negative peak centered at residue 7 (Pro), involving the 6-8 YPH region, confirms the presence of a 3_{10} helix structure [37] while the turn region centered approximately at residue 11 (Gly), namely the 9-12 NPGY region, is characterized by a negative sharp peak.

At neutral pH a turn region is found in the 4-6 PGY region, although in weak extent, as indicated by the wide standard deviations and the value of the negative peak, close to the higher values of the typical range for this motif [-0.06,-0.1]. Two other better characterized turn regions, signaled by negative peaks, are centered on residue 11 and 17 (Gly). At basic pH the negative peaks found at neutral pH, show lower values of the standard deviations, underlining an enhancement of turn regions inside the peptide. Moreover, looking at the chirality pattern along the backbone, we can notice that the regions which strongly differ in the chirality index values concerning the LH_8^{4+} state, on one hand, and LH_4, L^{4-}, on the other one, are the 4-8 PGYP and the C-terminal region, in which His8 and Tyr24 are mainly involved, thus singling out the pivotal region for the conformational change.

We have also examined the strong pH dependence of the local chirality evolution as a function of the simulation time. As an example, in Figure 8 we report the time evolution of *G* for Tyrosine18.

Here, the chirality index shows negative values, consistent with the presence of turn regions in the LH_4 and L^{4-} states and of a 3_{10} helix after 35 ns in LH_4. This is shown by lower values with respect to the L^{4-} form, in which the 3_{10} helix is less stable and thus the number of residues with the negative chirality, typical of 3_{10} helix, are less (see Table 5) than the LH_4 neutral state.

Looking at the LH_8^{4+} protonated state, we observe instead that, the chirality index is approaching zero, rather typical of a coil region, till approximately 75 ns, thus demonstrating the strong pH dependence of the conformational states of this molecular system.

In summary, the chirality analysis is consistent with the time evolution of the TetraHexaPY conformers reported in Figure 5, and highlights a significant contribution of turn structures progressively increasing with pH, expressed here by a shift towards more negative *G* values.

Conclusions

The presence of turns inside the hexarepeat peptide fragment was suggested by our previous work concerning the mono hexarepeat Ac-PHNPGY-NH_2 [20] and as well as in other different studies [40, 41].

Here, we have shown that the longer hexarepeat fragment $(PHNPGY)_4$ is essentially turn rich and that the turn formation is driven by the increase of pH. This trend is consistently disclosed by the analysis of the pH dependent CD spectra and by MD simulations at different pH, where C_α contacts, H-bond patterns and local chirality indices were monitored. In our previous paper [19], we suggested that the difference between secondary structures of mammalian octarepeats and avian hexarepeats could be due to the different ratio between Gly and Pro residues. Actually the mammalian octarepeat peptides contain 50% of glycine and 12% of proline residues, respectively, while the chicken hexarepeats encompass 16% of glycine and 33% of proline residues. The greater flexibility conferred by glycine residues might explain why the mammalian prion tandem repeat region is unordered and, unlike the chicken one, does not form a stable protease resistant domain [5]. The high number of prolines of course stabilizes a turn structure, but here we show that a crucial role is played by the histidine residues, particularly histidine 8 which, if deprotonated, is able to stabilize

turn regions in the peptide backbone. The CD and MD results for TetraHexaPY at different pH values, also stress the essential role played by tyrosine OH group in the peptide secondary structure. It is worth noting that tyrosine residues play a regulatory role in the endocytosis process [41] and are present in avian tandem hexarepeats (16% like glycine residues), but not in mammalian octarepeats. Moreover, the chirality pattern of the tetra hexarepeat region, in particular for the LH_4 and L^{4-} states, possesses a periodic-like shape, thus reflecting the periodicity in the primary structure; it is therefore likely that a similar pattern could be adopted also by the full repeat region. Within this hypothesis, it is possible to foresee a different biological behavior of the N-terminal domain of avian and mammal prion protein: the presence of tyrosines in the avian protein allows forming a compact hydrogen bond network, which could be probably responsible for its high resistance to proteases.

Acknowledgements.

We thank MIUR (PRIN 2004032851, 2005035119, FIRB RBNEO3PX83_001) for financial support.

Figure Captions

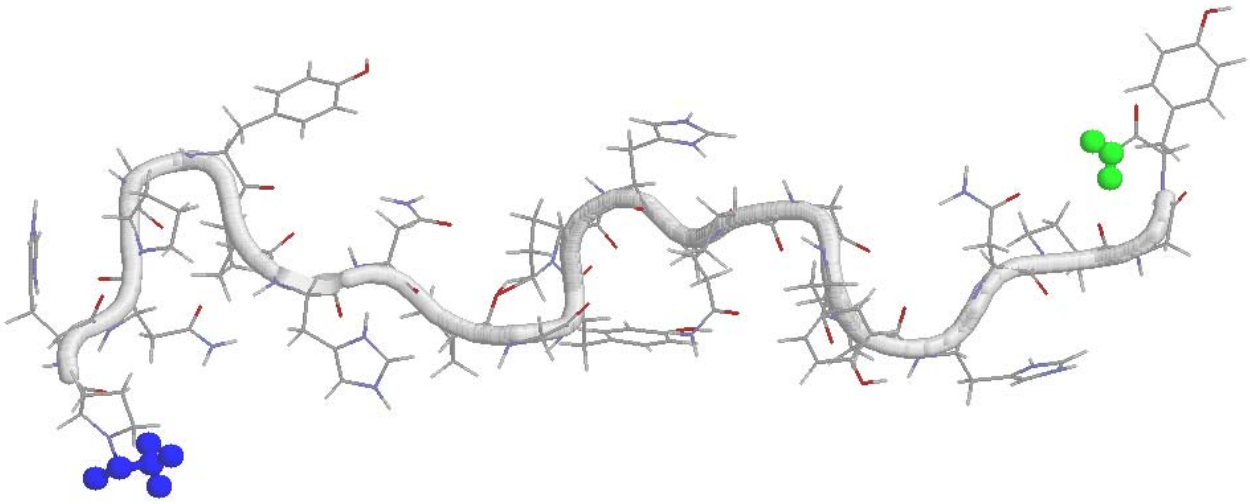
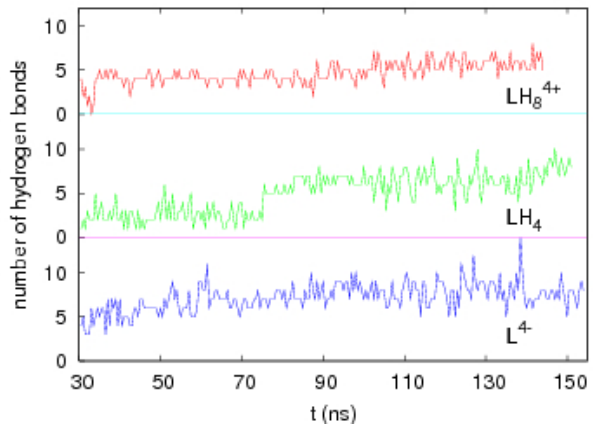
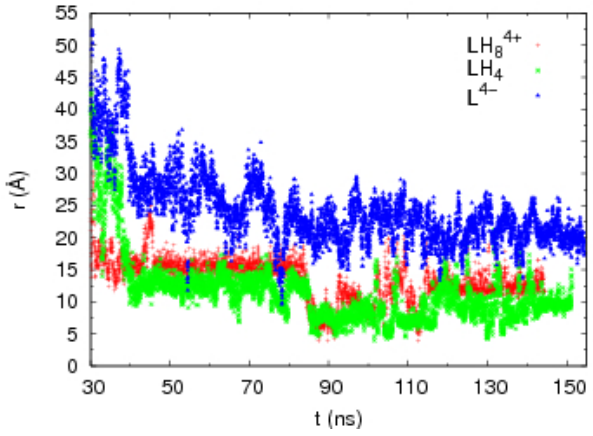


Figure 1: Starting configuration of TetraHexaPY, where the N-terminal Acetyl and the C-terminal amide are shown in blue and in green respectively.



[a]



[b]

Figure 2: Time evolution of hydrogen bonds number [a] and of the end-to-end distance relative to the C_α carbons [b]. Both of them are useful to check the equilibration of the starting configuration of the peptide, initially with an elongated structure and thus not having hydrogen bonds. This starting configuration is clearly lost during the run.

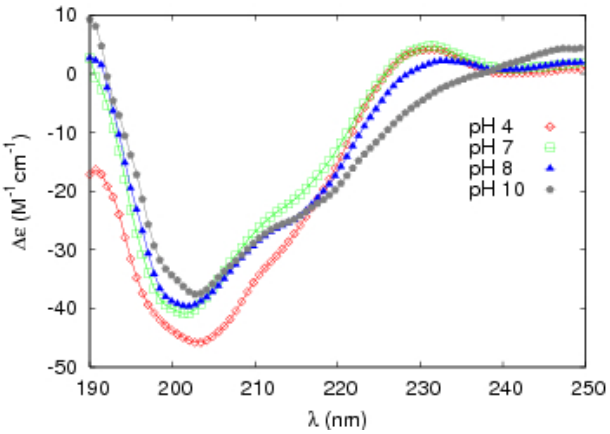


Figure 3: CD spectra of TetraHexaPY (4 x 10^{-6} mol dm^{-3}) as a function of pH. It is possible to observe the intense and broad negative peak at acidic pH, which instead becomes less intense increasing the pH. The band at around 190 nm, typical of type I β turn structures, turns from negative (acidic pH) to zero (neutral pH) to highly positive (basic pH) values. The shoulder at around 216 nm, typical of type I β turn region becomes evident with the increasing of pH.

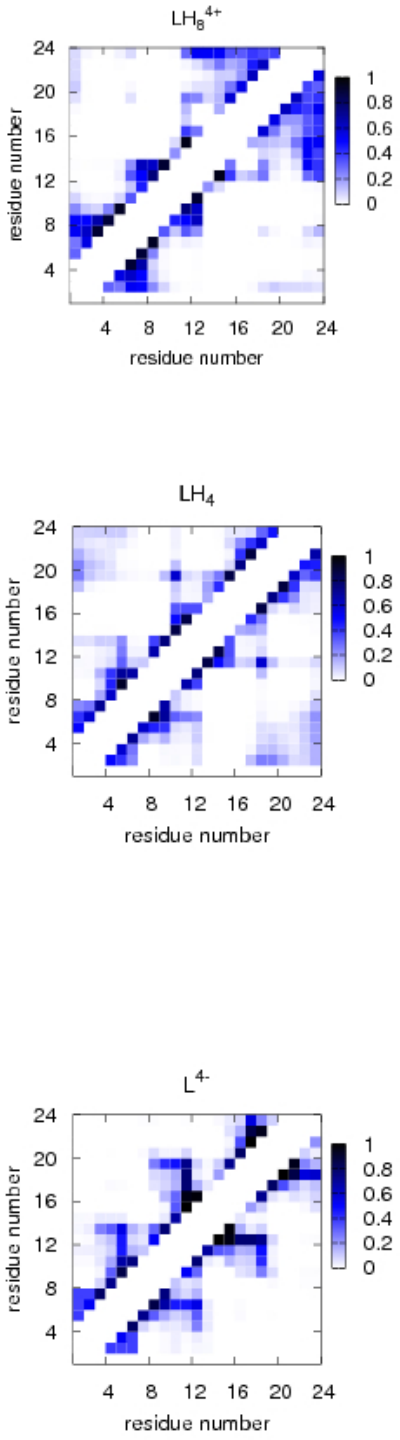


Figure 4: Contact maps of TetraHexaPY as a function of pH. It is worth to note a more compact structure at neutral pH (LH_4), underlined from the N- and C-terminal contacts. The L^{4-} state shows the most expanded structure with respect to the LH_4 and LH_8^{4+} states. The relative occurrence of the contacts is shown with a colour code ranging from white to black.

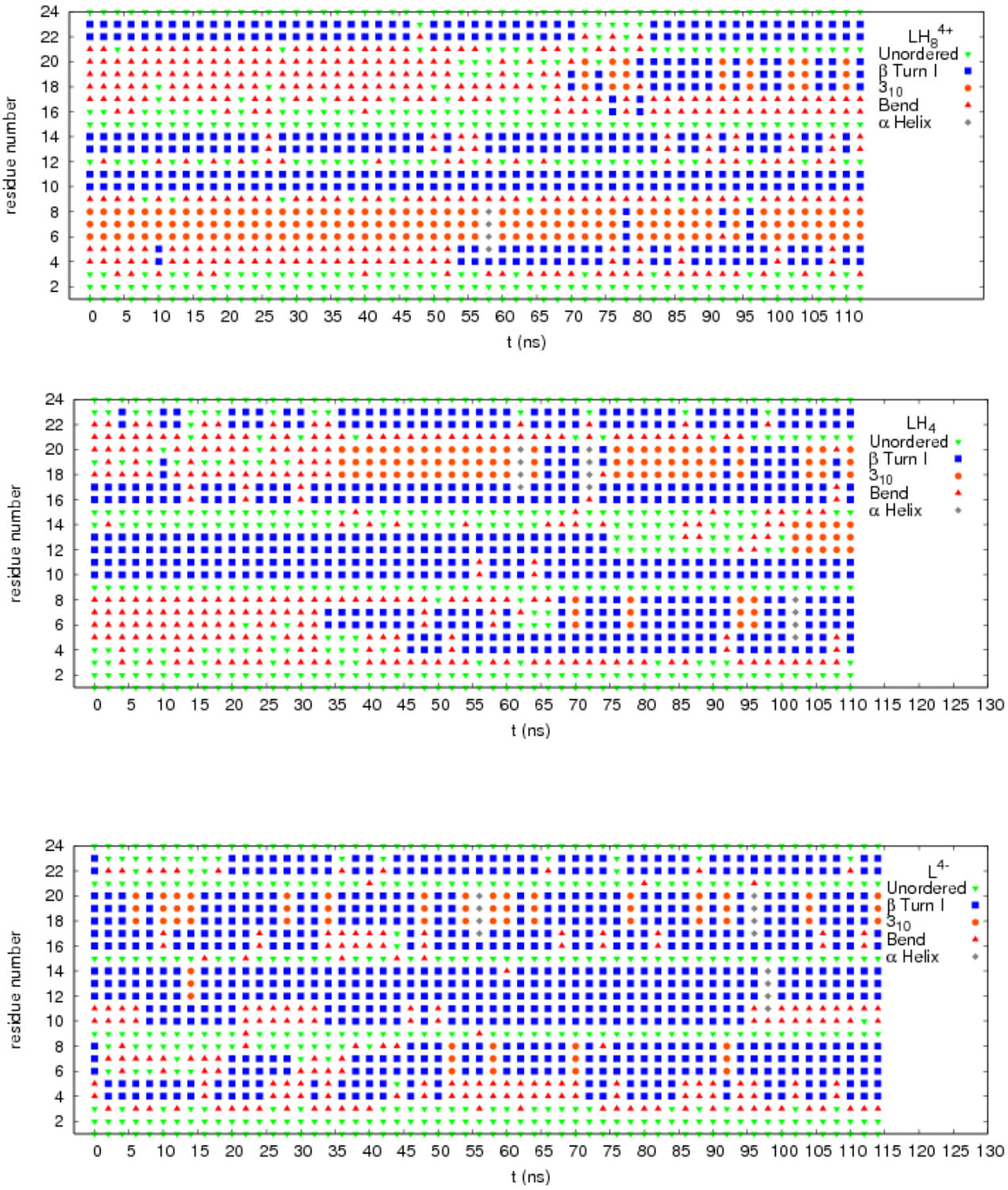


Figure 5: Time evolution of TetraHexaPY conformers in the three different protonation states obtained from MD simulations according with the DSSP criteria [35]. Turn regions increase with the histidine and tyrosine deprotonation, while the 3_{10} helix in the 6-8 region is present only in the LH_8^{4+} protonation state.

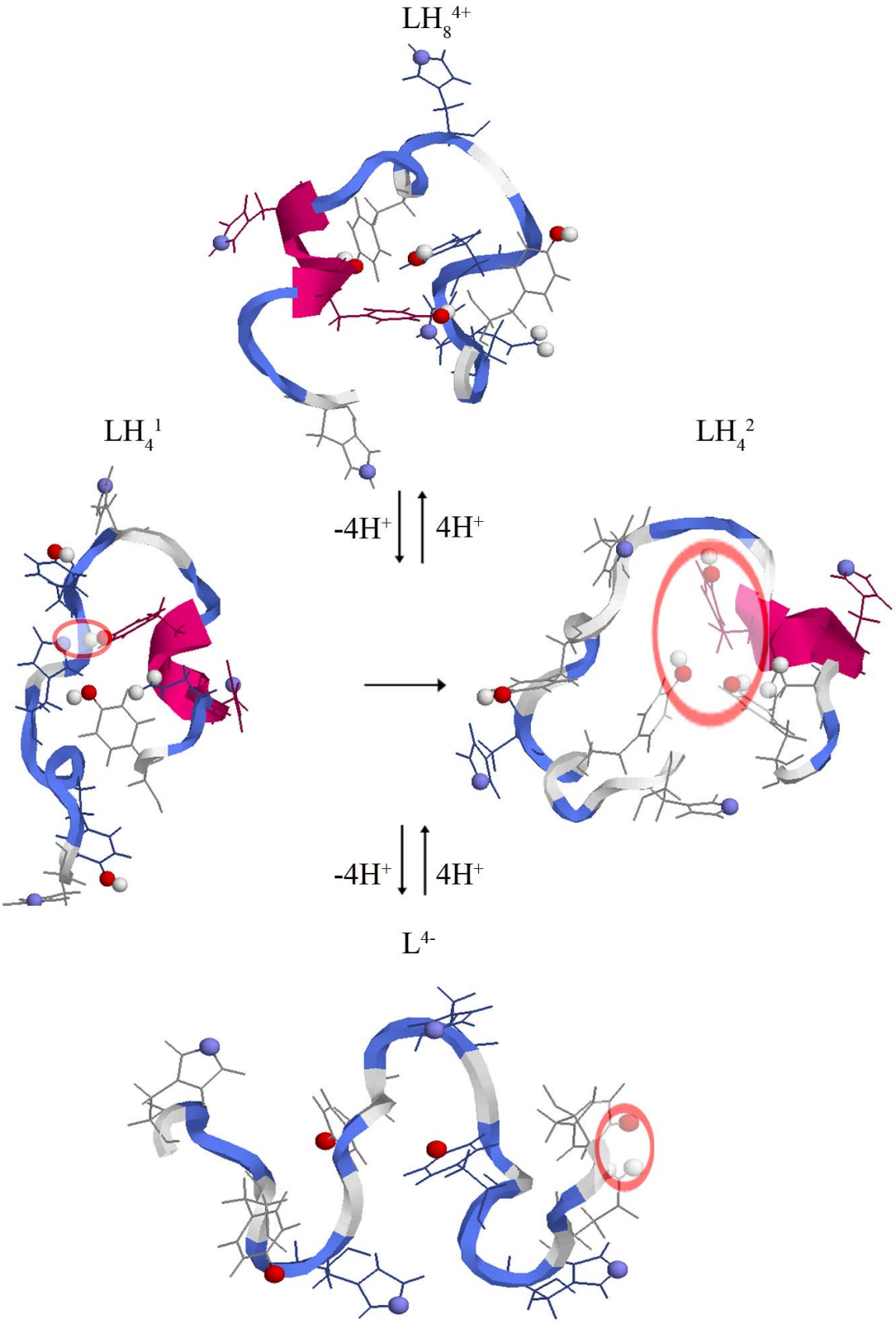


Figure 6: TetraHexaPY typical conformations in the LH_8^{4+}, in the LH_4 and in the L^{4-} form obtained from MD simulations. It is worth to note the 3_{10} helix, in the 6-8 YPH region, featured by two consecutive hydrogen bonds (5-8, 6-9, up LH_8^{4+} state). Upon the deprotonation of the four histidines, the 6-9 hydrogen bond, inside the 3_{10} helix, is disrupted

because of the interaction between the imidazole nitrogen of His8 and the phenol hydrogen of tyrosine18 (LH_4^1). Such as conformation leads to the formation of another one in which the 3_{10} helix in the 17-20 region, get involved tyrosine 6, 18 and 24 in a hydrogen bond network, as shown in LH_4^2. At basic pH, when the four tyrosines are deprotonated (L^{4-}, bottom), a new interaction between the phenolate oxygen of tyrosine24 and the side chain amide hydrogens of asparagine21, causes a bending which stabilizes a turn structure in the 17-20 GYPH region and at the same time it provokes a tilting of the backbone. The different structures are shown according with a colour code: Blue for turn, violet for 3_{10} helix and gray for coil regions. The phenolate hydrogen and oxygen of tyrosines 6, 12, 18 and 24 are shown respectively in white and red, the imidazole nitrogens of histidines 2, 8,14, 20 are shown in silver blue and the amide hydrogens of asparagine 21 are shown in white. The side chains hydrogen bonds are circled in red. N and C termini are shown respectively from left to right.

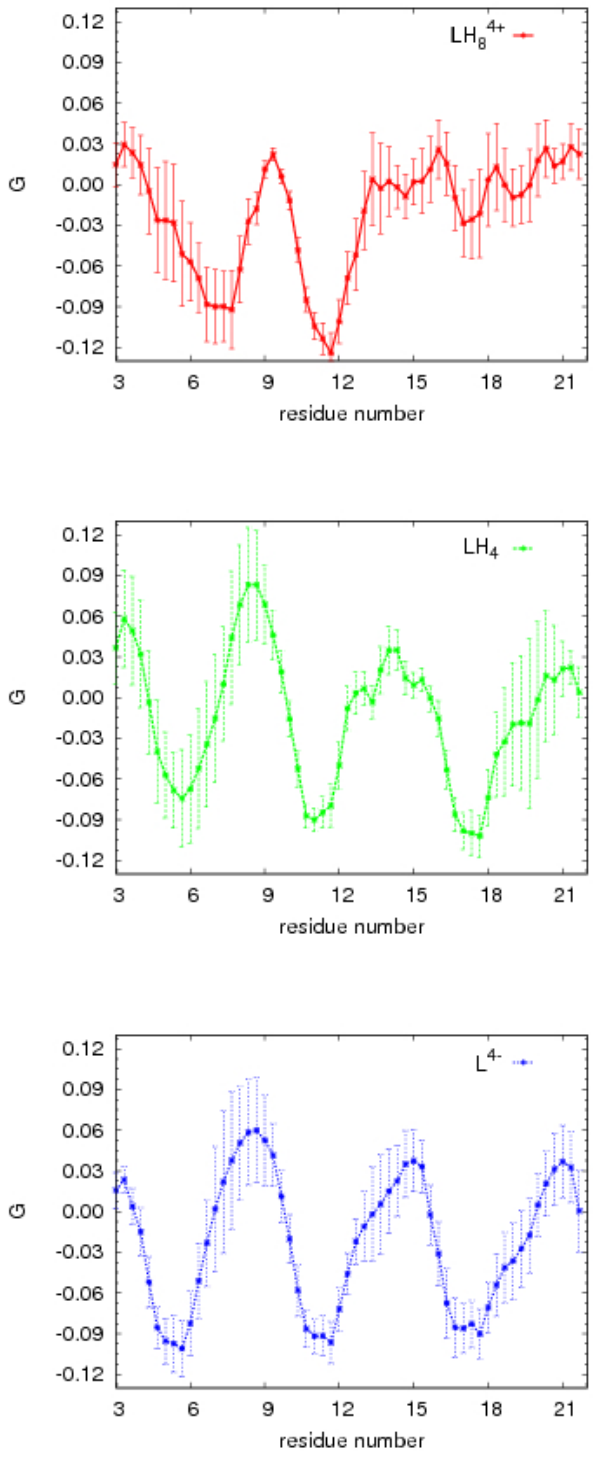


Figure 7: Chirality index, G, averaged over the trajectories of TetraHexaPY as a function of pH. The more negative peaks (-0.1< G <-0.06) underline the presence of turn regions, while the positive peak centered at residue Asn9, underlines a small amount of polyproline structure in the 7-9 PHN region. The index shows a marked difference concerning the number of turn regions in the LH_8^{4+}state with respect to the LH_4 and the L^{4-} ones. Error bars are reported for each of the *G* values as standard deviations on the ensemble of trajectories.

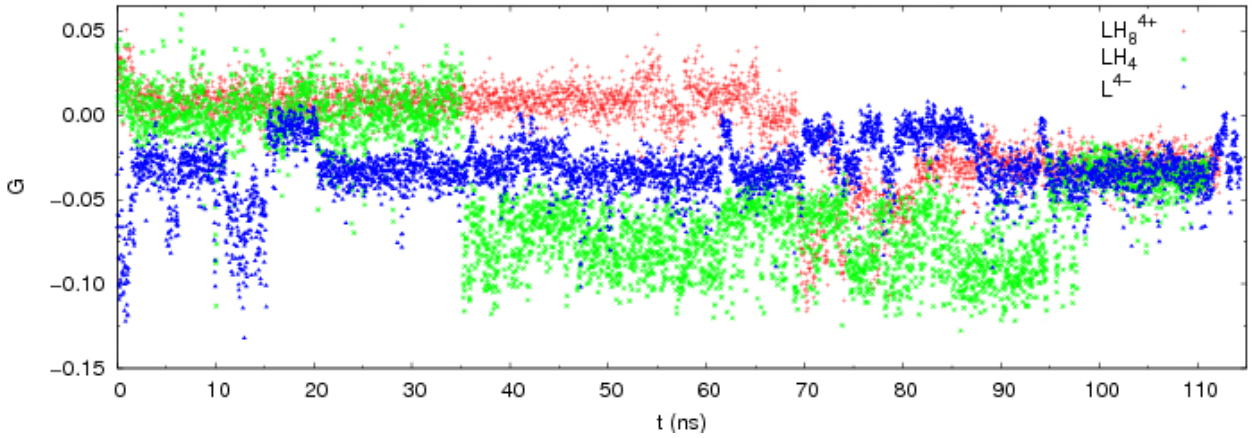


Figure 8: Time evolution of the chirality index *G* for Tyrosine 18. It should be noted the strong dependence on pH of the local chirality for tyrosine 18, pointed out by the variation of the chirality index in the three different protonation states, LH_8^{4+}, LH_4, L^{4-}.

Table 1. Protonation constants (log β) and pK values of TetraHexaPY (T= 298.0 K, I=0.1 mol dm^{-3} KNO_3).

* 3σ *10^{-2} values are shown in parentheses

Species	Log β*	pK	Site of protonation
$[HL]^{3-}$	10.65 (3)	10.65	OH group of Tyr6, Tyr12, Tyr18 and Tyr24
$[H_2L]^{2-}$	20.64 (3)	9.99	
$[H_3L]^{-}$	30.40 (6)	9.76	
$[H_4L]$	39.45 (6)	9.05	
$[H_5L]^{+}$	46.26 (9)	6.81	Imidazole of His2, His8, His14 and His20
$[H_6L]^{2+}$	52.58 (9)	6.32	
$[H_7L]^{3+}$	58.59 (9)	6.01	
$[H_8L]^{4+}$	63.72 (9)	5.13	

Table 2: Percentage of the different conformations (Turn, 3_{10} helix, α Helix, Bend, Coil) adopted by the TetraHexaPY at acidic pH.

Residues	**Turn**	**3_{10} Helix**	**α Helix**	**Bend**	**Coil**
Pro1	-	-	-	-	100
His2	-	-	-	-	100
Asn3	-	-	-	47.4	52.6
Pro4	45.6	-	-	54.4	-
Gly5	43.9	-	1.7	54.4	-
Tyr6	3.5	93	1.7	1.8	-
Pro7	5.3	93	1.7	-	-
His8	5.3	93	1.7	-	-
Asn9	-	-	-	86.0	14.0
Pro10	100	-	-	-	-
Gly11	100	-	-	-	-
Tyr12	-	-	-	38.6	61.4
Pro13	82.5	-	-	17.5	-
His14	82.5	-	-	17.5	-
Asn15	-	-	-	-	100
Pro16	3.5	-	-	45.6	50.9
Gly17	3.5	-	-	79.0	17.5
Tyr18	22.8	14.0	-	56.2	7.0
Pro19	24.6	14.0	-	52.6	8.8
His20	21.0	14.0	-	52.6	12.4
Asn21	-	-	-	54.4	45.6
Pro22	89.5	-	-	7.0	3.5
Gly23	89.5	-	-	-	10.5
Tyr24	-	-	-	-	100
Total	30.2	13.4	-	27.7	28.7

Table 3: Percentage of the different conformations (Turn, 3_{10} helix, α Helix, Bend, Coil) adopted by the TetraHexaPY at neutral pH.

Residues	Turn	3_{10} Helix	α Helix	Bend	Coil
Pro1	-	-	-	-	100
His2	-	-	-	-	100
Asn3	-	-	-	73.2	26.8
Pro4	53.6	-	-	19.6	26.8
Gly5	51.8	-	1.8	39.3	7.1
Tyr6	51.8	7.1	1.8	30.4	8.9
Pro7	51.8	7.1	1.8	35.7	3.6
His8	30.4	7.1	1.8	57.1	3.6
Asn9	-	-	-	-	100
Pro10	96.4	-	-	3.6	-
Gly11	96.4	-	-	3.6	-
Tyr12	67.9	8.9	-	3.6	19.6
Pro13	67.9	8.9	-	7.1	16.1
His14	-	8.9	-	23.2	67.9
Asn15	-	-	-	12.5	87.5
Pro16	91.1	-	-	8.9	-
Gly17	89.3	-	3.6	7.1	-
Tyr18	17.9	46.4	3.6	32.1	-
Pro19	19.6	46.4	3.6	14.3	16.1
His20	16.1	46.4	3.6	30.3	3.6
Asn21	-	-	-	73.2	26.8
Pro22	75.0	-	-	23.2	1.8
Gly23	75.0	-	-	-	25
Tyr24	-	-	-	-	100
Total	39.7	7.8	0.9	20.7	30.9

Table 4: Percentage of the different conformations (Turn, 3_{10} helix, α Helix, Bend, Coil) adopted by the TetraHexaPY at basic pH.

Residues	**Turn**	**3_{10} Helix**	**α Helix**	**Bend**	**Coil**
Pro1	-	-	-	-	100
His2	-	-	-	-	100
Asn3	-	-	-	55.2	44.8
Pro4	55.2	-	-	44.8	-
Gly5	55.2	-	-	43.1	1.7
Tyr6	70.7	6.9	-	13.8	8.6
Pro7	70.7	6.9	-	17.2	5.2
His8	53.5	6.9	-	10.3	29.3
Asn9	-	-	-	3.4	96.6
Pro10	62.1	-	-	36.2	1.7
Gly11	62.1	-	1.7	36.2	-
Tyr12	96.6	1.7	1.7	-	-
Pro13	96.6	1.7	1.7	-	-
His14	94.9	1.7	1.7	1.7	-
Asn15	-	-	-	8,6	91.4
Pro16	75.9	-	-	22.4	1.7
Gly17	72.5	-	3.4	22.4	1.7
Tyr18	69.0	26.6	3.4	-	-
Pro19	69.0	26.6	3.4	-	-
His20	69.0	26.6	3.4	-	-
Asn21	-	-	-	5.2	94.8
Pro22	74.1	-	-	19.0	6.9
Gly23	74.1	-	-	-	25.9
Tyr24	-	-	-	-	100
Total	50.9	4.5	0.9	14.1	29.6

Table 5: Average values and standard deviations of the Chirality index, *G*, for the different secondary structures according with reference [37].

Structure	$\langle G \rangle$	σ_G	Number of residues
α helix	-0.04	0.01	>3
3_{10} helix	-0.08	0.02	>3
β Turn I	-0.08	0.02	2,3
β Sheets	0.000	0.003	≥2
PPII	0.11	0.01	>3
π helix	-0.00	0.01	>3

References

1. Wopfner, F.; Weidenhöfer, G.; Schneider, R.; von Brunn, A.; Gilch, S.; Schwarz, T. F.; Werner, T. ; Schätzl, H. M. *J. Mol. Biol.*, **1999,** 289, 1163.

2. Suzuki, T.; Kurokawa, T.; Hashimoto, H.; Sugiyama, M. *Biochem. Biophys. Res. Comm.*, **2002**, 294, 912.

3. Harris, D. A.; Falls, D. L.; Johnson, F.A.; Fischbach, G. D. *Proc. Natl. Acad. Sci. U.S.A.*, **1991**, 88, 7664.

4. Calzolai, L.; Lysek, D. A.; Pérez, D. R.; Güntert, P.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.*, **2005**, 102, 651.

5. Marcotte, E. M.; Eisenberg, D. *Biochemistry*, **1999**, 38, 667.

6. Prusiner, S. B. *Proc. Natl. Acad. Sci. U.S.A.*, **1998**, 95, 13363.

7. Nunziante, M.; Gilch, S.; Schätzl, H. M. *J. Biol. Chem.*, **2003**, 278, 3726.

8. Frankenfield, K. N.; Powers, E. T.; Kelly, J. W. *Protein Sci.*, **2005**, 14, 2154.

9. Brown, D. R.; Qin, K. F.; Herms, J. W.; Madlung, A.; Manson, J.; Strome, R.; Fraser, P. E.; Kruck, T.; von Bohlen, A.; Shulz-Schaeffer, W.; Giese, A.; Westaway, D.; Kretzschmar, H. A. *Nature*, **1997**, 390, 684.

10. Brown, D. R.; Clive, C.; Haswell, S. J. *J. Neurochem.*, **2001**, 76, 69.

11. Lehmann, S. *Curr. Opin. Chem. Biol.*, **2002**, 6, 187.

12. Taylor, D. R.; Watt, N. T.; Perera, W. S.; Hooper, N. M. *J. Cell Sci.*, **2005**, 118, 5141.

13. Zahn, R.; Liu, A.; Luhrs, T.; Riek, R.; von Schroetter, C.; Garcia, F. L.; Billeter, M.; Calzolai, L.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.*, **2000**, 97, 145.

14. Garcia, L. F.; Zahn, R.; Riek, R.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.*, **2000**, 97, 8334.

15. Pushie, M. J.; Vogel, H.J. *Biophys. J.*, **2007** online

16. Zahn, R. *J. Mol. Biol.*, **2003**, 334, 477.

17. Smith, C. J.; Drake, A. F.; Banfield, B. A.; Bloomberg, G. B.; Palmer, M. S.; Clarke, A. R.; Collinge, J. *FEBS Lett.*, **1997**, 405, 378.

18. Hornshaw, M. P.; McDermott, J. R.; Candy, J. M.; Lakey, J.H. *Biochem. Biophys. Res. Comm.*, **1995**, 214, 993.

19. La Mendola, D.; Bonomo, R. P.; Impellizzeri, G.; Maccarrone, G.; Pappalardo, G.; Pietropaolo, A.; Rizzarelli, E.; Zito, V. *J. Biol. Inorg. Chem.*, **2005**, 10, 463.

20. Pietropaolo, A.; Raiola, L.; Muccioli, L.; Tiberio, G.; Zannoni, C.; Fattorusso, R.; Isernia, C.; La Mendola, D.; Pappalardo, G.; Rizzarelli, E. *Chem. Phys. Lett.*, **2007**, 442, 110.

21. Arena, G.; Rizzarelli, E.; Sammartano, S.; Riganò, C. *Talanta*, **1979**, 26, 1.

22. Gans, P.; Sabatini, A.; Vacca, A. *Talanta*, **1996**, 43, 1739.

23. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, J. C. *J. Comput. Chem.*, **1995**, 26, 1701.

24. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.*, **1995**, 117, 5179.

25. Berendsen, H. J. C.; Postma, J. P. M; van Gunsteren, W. F.; Hermans, J. *In Intermolecular Forces*, B. Pullman, Ed.; Reidel, Dordrecht, 1981; pp. 331-342.

26. Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A; Haak, J. R. *J. Chem. Phys.*, **1984**, 81, 3684.

27. Smith, L. J.; Daura, X.; van Gusteren, W. F. *Proteins*, **2002**, 48, 487.

28. Feller, S. M.; Ren, R.; Hanafusa, H.; Baltimore, D. *Trends Biochem. Sci.*, **1994**, 19, 453.

29. Krittanai, C.; Johnson, W. C. *Anal. Biochem.*, **1997**, 253, 57.

30. Chakrabartty, A.; Kortemme, T.; Padmanabhan, S.; Baldwin, R. L. *Biochemistry*, **1993**, 32, 5560.

31. Woody, R. W. In *Circular Dichroism and the conformational analysis of biomolecules*, Fasman, G.D., Ed.; Plenum Press, New York, 1996; pp 25-67.

32. Brahms, S.; Brahms, J. *J. Mol. Biol.*, **1980**, 138, 149.

33. Bandekar, J.; Evans, D. J; Krimm, S.; Leach, S. J.; Lee, S.; McQuie, J. R.; Minasian, E.; Némethy, G.; Pottle, M. S.; Scheraga, H. A.; Stimson, E. R.; Woody, R. W. *Int. J. Pept. Protein Res.*, **1982**, 19, 187.

34. Vendruscolo, M.; Najmanovich, R.; Domany, E. *Phys. Rev. E*, **1999**, 82, 656.

35. Kabsch, W.; Sander, C. *Biopolymers*, **1983**, 22, 2577.

36. A.Pietropaolo, L. Muccioli, C. Zannoni, E. Rizzarelli, *work in progress.*

37. Pietropaolo, A.; Muccioli, L.; Berardi, R.; Zannoni C. *Proteins*, **2007** doi: 10.1002/prot.21578.

38. Solymosi, M.; Low, R. J.; Grayson, M.; Neal. M. P. *J. Chem. Phys.*, **2002**, 116, 9875.

39. Osipov, M. A.; Pickup, B. T.; Dunmur, D. A. *Mol. Phys.*, **1995**, 84, 1193.

40. Bazan, J. F.; Fletterick, R. J.; McKinley, M. P.; Prusiner, S. B. *Protein Eng.*, **1987**, 1, 125.

41. Paccaud, J. P.; Reith, W.; Johansson, B.; Magnusson, K. E.; Mach, B.; Carpenter, J. L. *J. Biol. Chem.*, **1993**, 268, 23191.