Can Ionic Effects Induce α-sheet Conformation of Peptides?

Nadiyah Alshammari and James A. Platts*

School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, UK

Phone: +44-2920-874950, Email: platts@cardiff.ac.uk

Abstract

We report coupled cluster, MP2 and DFT data on the relative energy and geometry of α-sheet and β-strand conformations of model peptides. We show that ionic effects have a strong effect on energy balance through formation of multiple cation-oxygen contacts in the α-sheet form. Such effects are markedly dependent on both sequence and ion: two peptides considered favour α-sheet in the presence of cations, whereas a third, non-polar one favours β-strand in the same conditions.

Keywords

Peptide; conformation; ab initio; DFT; ions

Introduction

Peptides are one of the main building blocks of life: their ability to fold swiftly and reliably into a wide range of three-dimensional structures is a key aspect of protein function. Backbone flexibility is largely due to rotation about N-Cα and Cα-C bonds, denoted as ϕ and ψ, respectively. Textbook definitions identify large regions of (ϕ, ψ) space corresponding to α-helix (α_R) and β-strand, and a smaller region for left-handed helix (α_L). More detailed analysis has identified further detail within this broad picture, including (ϕ, ψ) combinations for P_{II}-spirals and γ-turns as well as a "bridge" region that connects canonical helix and strand, while the specific structures of glycine and proline lead to quite different preferences.[1]

The α-sheet conformation is not as widely recognised as α-helix and β-strand. It was originally proposed by Pauling and Corey,[2] but recent interest has been sparked by the pioneering work of Daggett et al.[3] They identified transient α-sheets in molecular dynamics simulations of a range of proteins, many of which are implicated in amyloidosis, often prompted by low pH conditions. They

identify α-sheets as consisting of repeats of around four alternating α_L and α_R amino acids, resulting in extended chains that can be stabilised by bifurcated hydrogen bonding with neighbouring chains. This has led to the intriguing proposal that α-sheets may play a key role in the broad class of protein misfolding diseases.[4] Recent work has also proposed α-sheet conformations to be present in the non-amyloid-β component (NAC) region of α-synuclein, the peptide implicated in Parkinson's disease.[5] Careful analysis of Ramachandran propensity plots noted similarity in (ϕ, ψ) space for α-sheet and β-turn, denoted α_L and γ_L, and established preferences for different amino acids to adopt these.[6] Notably, glycine is found to strongly disfavour α_L, whereas polar residues such as asparagine favour this form.

As well as hydrogen bonding, the approximate alignment of C=O and N-H vectors on opposite faces of α-sheets gives rise to "nests" of cation and anion binding sites, seen for instance in a potassium channel structure (PDB entry 1BL8).[7] This alignment also makes α-sheets more polar than either α-helix or β-strand. Our interest was sparked by observation of resides in (ϕ, ψ) space close to α_L and γ_L in simulations of amyloid-β when bound to transition metal ions.[8] This prompted us to investigate in more detail the relative stability of α-sheets in some model peptides, and the effect of solvent and ions on this. In this context, it is notable that a recent study showed that human transthyretin, one of the main proteins used by Daggett et al to demonstrate the importance of α-sheets, is destabilised by the presence of Ca^{2+}.[9]

Methods

The model peptide Ala_4 was built manually in canonical α-helix, β-strand and α-sheet ($\alpha_L\alpha_R\alpha_L\alpha_R$) conformations in MOE,[10] with amide caps at N- and C-termini to avoid end effects of charged groups. This peptide was selected as a model as Armen et al show that all proposed α-sheet conformations in Protein Data Bank are no more than 4 amino acids in length. It should also be noted that this is used as a model peptide that avoids complications from side chains: we make no claim to its propensity for α-sheet conformations. For other peptides, proposed α-sheet conformations were extracted from the larger PDB entries for structures proposed to adopt this form by Armen et al.[3] The highlighted residues from that study were retained, including amide caps at termini. Canonical α-helix (ϕ = -65°, ψ = -39°) and β-strand (ϕ = -120°, ψ = +113°) forms of the same sequences were then manually constructed for comparison.

Geometries were optimised at BP86-D3BJ/def2-SVP[11,12,13,14] level in CPCM[15] model of aqueous solvent, and vibrational frequencies calculated to ensure all structures were minima, and to extract zero-point and thermal corrections at 298 K. Benchmark *ab initio* data was then obtained from extrapolation of RI-MP2 correlation energy to the basis set limit from cc-pVTZ and cc-pVQZ data, corrected by difference in CCSD(T) and MP2 correlation energy at cc-pVDZ level. We also used the DLPNO-CCSD(T) method[16] with def2-TZVPP[14] basis set and default settings in Orca v5.0.0, incorporating self-consistent CPCM or SMD solvent correction where appropriate. *Ab initio* and DFT calculations were performed at the DFT-optimised geometry, making extensive use of the Rij and RijCosX approximations with appropriate auxiliary basis sets.[17] All *ab initio* and DFT calculations used the Orca package.[18]

Results and Discussion

Geometry optimisation at BP86-D3BJ/def2-SVP/CPCM level retains the structures of β-strand and α-sheet: values of (ϕ,ψ) angles in the central peptides of Ala_4 (Table S1) place the α-sheet firmly in the $\alpha_R\alpha_L$ region of the Ramachandran map, and the β-strand within the ranges expected for this form. α-helix values are not as close to expected values, presumably because a four-peptide sequence is too short to fully form the helix: we therefore do not consider this form in detail further. Dihedral angles outside the central region are also further from canonical values due to end effects. All conformations are confirmed as local minima on the potential energy surface by harmonic frequency calculation. We also find up to 20 cm^{-1} difference in wavenumber of carbonyl stretching bands (Table S2), which might form the basis of spectroscopic identification of α-sheets. Geometry optimisation with the larger def2-TZVPPD basis does not significantly alter these results, with heavy atom RMSD between smaller and larger basis of 0.35 and 0.17 Å for α-sheet and β-strand, respectively.

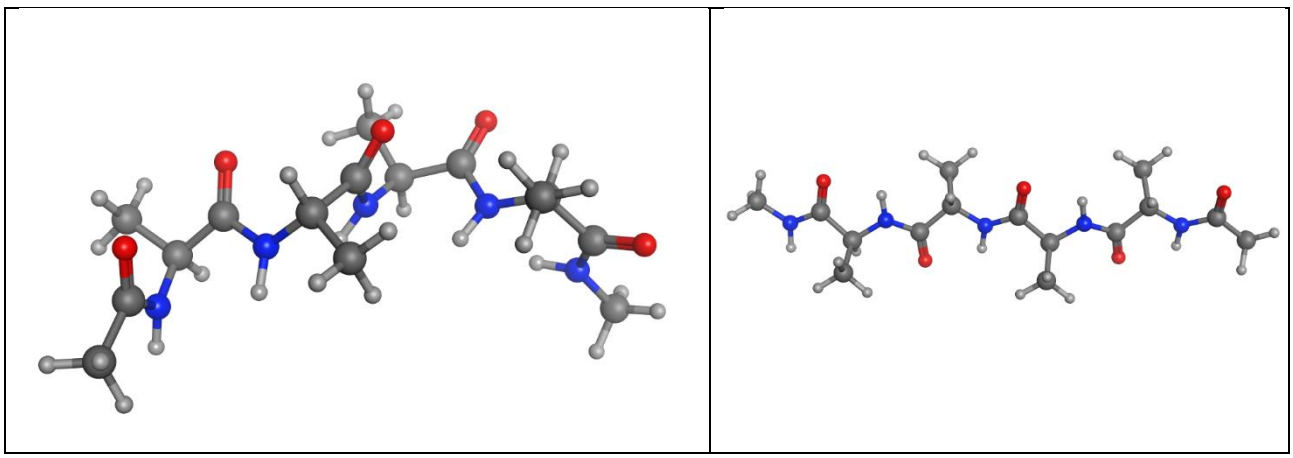

Figure 1: DFT optimised geometry of α-sheet and β-strand conformations of Ala_4.

It is important to check requirements for suitable prediction of relative energy of β-strand and α-sheet conformations of Ala_4, using *ab initio* data at the DFT-optimised geometry as a benchmark (Table 1, full data in Table S3). In gas phase, this and all other methods tested indicate strong preference for β-strand, by 40 to 50 kJ mol^{-1}. Energy difference is not strongly dependent on basis set, and the performance of GGA, hybrid and range separated DFT functionals is very similar.

Table 1: Relative energy of α-sheet relative to β-strand Ala_4 (kJ mol^{-1})

Method	Basis	Solvent	Rel E
DLPNO-CCSD(T)	def2-TZVPP	None	+39.92
RI-MP2	def2-TZVPP	None	+36.87
BP86-D3BJ	def2-SVP	None	+38.63
wB97x-D3	def2-TZVP	None	+37.51
CCSD(T)	CBS	CPCM	-6.17
DLPNO-CCSD(T)	def2-TZVPP	CPCM	-5.55
RI-MP2	def2-TZVPP	CPCM	-7.23
BP86-D3BJ	def2-SVP	CPCM	-8.57
wB97x-D3	def2-TZVP	CPCM	+3.40

Incorporating solvent effects using implicit models changes the energy balance markedly: *ab initio* methods predict small preference for α-sheet, as does BP86-D3/def2-SVP, while DFT methods with larger basis sets indicate slight preference for β-strand. This strong solvent effect is ascribed to the greater dipole moment in α-sheet (13.5 *vs.* 10.5 Debye from gas phase DFT data) that results from alignment of C=O and N-H bonds. Table S3 indicates a change of less than 0.5 kJ mol^{-1} when using the SMD solvent model, such that specifics of solvent model employed do not change the overall picture. The excellent performance of BP86-D3/def2-SVP compared to *ab initio* benchmark is notable: that this gets worse as basis set is improved indicates fortuitous cancellation of errors. Nevertheless, it represents a useful compromise for estimation of energies for relatively little computational expense that we will exploit for larger peptides. Table S3 also demonstrates that more approximate methods, including the semi-empirical PM7 approach and the AMBER (ff14SB) molecular mechanics exhibit at least qualitative agreement with *ab initio* data. However, the semi-empirical GFN2-xTB method overestimates the stability of β-strand in this case.

Harmonic frequency data at this level indicates an overall difference in zero-point vibrational energy of 3.68 kJ mol^{-1} in favour of β-strand, and a difference in thermal energy at 298 K of 1.55

kJ mol^{-1}. Entropy differences, as T*S for T = 298 K, between conformations is 7.23 kJ mol^{-1}, stemming largely from the vibrational partition function. Incorporating these corrections into the *ab initio* electronic energy differences suggests that α-sheet and β-strand conformations of this model peptide are almost equal in ZPE-corrected, thermal or Gibbs free energy. The latter, at 1.68 kJ mol^{-1} in favour of β-strand, is almost certainly within the error of these calculations.

We then used similar calculations to examine whether the presence of ions can affect the predicted equilibrium between these conformations by manually placing Na^+, Mg^{2+}, K^+ or Ca^{2+} ion in the proximity of C=O group(s) in the central portion of each conformation.[1] Geometry optimisation places cations close to two central carbonyls in α-sheet, whereas the conformation of β-strand means that only one such contact is possible. As a result, the α-sheet is significantly stabilised: for a single Na^+ ion this is a small effect (less than 2 kJ mol^{-1} at *ab initio* level), but Mg^{2+} induces a change of over 30 kJ mol^{-1}. K^+ and Ca^{2+} ions are similar to Mg^{2+} in inducing much larger changes in relative energy, indicating that size as well as charge of ion is important in determining its effect (Table 2).

Table 2: Relative electronic energy of α-sheet to β-strand form of Ala_4-ion complexes in CPCM water (kJ mol^{-1})

Method	Basis	Na^+	Mg^{2+}	K^+	Ca^{2+}	Na^+/Cl^-
DLPNO-CCSD(T)	def2-TZVPP	-7.15	-37.63	-63.08	-106.88	-16.45
BP86-D3BJ	def2-SVP	-36.40	-81.75	-65.65	-113.51	-2.03

To explore this further, ion-oxygen distances in optimised geometry were examined (Table S4). This shows that the α-sheet conformation can form multiple contacts to all ions considered, whereas β-strand forms only one such contact. However, the single contacts to β-strand are uniformly shorter, indicating that the bifurcated arrangement of C=O groups in α-sheet conformation cannot accommodate all ions with ideal contact distances. This is especially true of Na^+, which forms only one contact that is close to ideal distance. Mg^{2+} and Ca^{2+} form two close contacts, while K^+ also forms two close contacts and a third, slightly longer contact. Thus, it seems that the size of the pocket, or "nest" formed by the α-sheet matches the ionic radius of K^+

[1] Multiple starting positions were tested for Na^+ by manually placing this ion in the proximity of one or more oxygen nuclei, the lowest final energy of which is reported. This position was then used for remaining ions. A similar approach was used for adding chloride in proximity to N-H.

particularly well (Figure 2) whereas Na^+ is too small to fit well here. Adding a chloride counterion proximal to N-H groups along with Na^+ enhanced the effect of the cation alone by formation of two Cl…H contacts in addition to two Na…O ones.

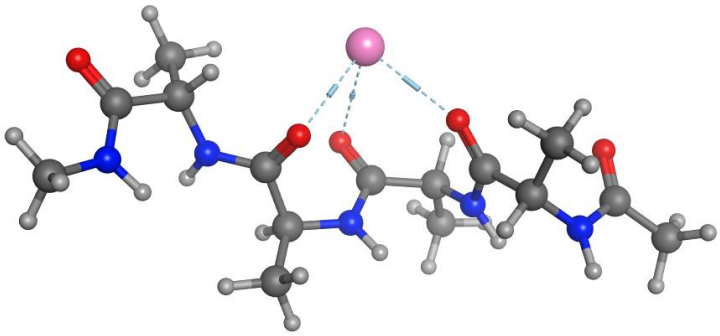

Figure 2: DFT optimised geometry of Ala_4-K^+ in α-sheet conformation, with K…O contacts shown

While Ala_4 is a useful model to test methods, it is not a likely candidate to form α-sheets as Ala-rich peptides are known to prefer helical structures. Instead, we took two peptides from the search for alternating α_L and α_R amino acids of a nonredundant version of the Protein Data Bank reported by Armen et al.[3] The first, taken from PDB entry 1A05, consists of Ala31-His-Leu-Gly34 (AHLG); the siting of this sequence between a helix and a β-strand and the presence of glycine, suggest this might be better characterised as a turn. The second, taken from PDB entry 1FEP, consists of Ser379-Asn-Thr-Gln382 (SNTQ). This shows more promise as a genuine α-sheet as it consists of polar residues, and is found at the C-terminus of a protein chain that makes up the larger β-barrel, bringing it close to other chains that could stabilise this form. In both cases, the experimental conformation of the peptide was extracted from the larger protein structure, retaining COMe and NHMe from neighbouring residues to avoid artificial end effects. Equivalent structures were built manually in canonical α-helix and β-strand forms, and all structures optimised at BP86-D3/def2-SVP/CPCM level.

Table 3 shows that AHLG is actually more stable in the α-sheet form than in β-strand: in fact, it is more stable by 12 kJ mol^{-1} than the structure that results from DFT optimisation of the α-helical form of this peptide (data not shown). Again though, this does not retain proper helical form so is not considered further. However, introduction of ions changes this order of stability, favouring β-

strand for all ions considered, in contrast with the situation with Ala_4. Closer inspection shows that side-chain effects may be responsible for this: in the α-sheet form, ions cannot form the same number of contacts to O as in the simpler Ala_4 model, perhaps due to the steric bulk of Leu. In the β-strand form, however, additional contact with Nδ of His complements the single contact with O to further stabilise this form. This is in agreement with Hayward and Milner-White's conclusion that non-polar amino acids are significantly disfavoured.[6]

Table 3: Relative electronic energy of α-sheet to β-strand for AHLG and SNTQ peptides and their ion complexes in CPCM water (kJ mol^{-1})

Method	Free	Na^+	Mg^{2+}	K^+	Ca^{2+}
AHLG	-25.70	+14.85	+13.31	+2.34	+43.05
SNTQ	+10.36	-10.88	-102.08	+3.06	-25.49

The situation in the more polar SNTQ is reversed: here, α-sheet is less stable than β-strand in the absence of ions, although the energy difference is relatively small. The α-helix peptide is more stable again by 4.5 kJ mol^{-1} (data not shown), but is again distorted from the classical helical structure. Ions change the energy balance markedly, this time stabilising the α-sheet relative to β-strand. Ionic radius is key to this: Na^+ leads to a change in order of stability, whereas K^+ does not although the energy difference does narrow. This seems to be related to the details of contacts formed: both Na^+ and K^+ have four ion…O close contacts, but the distances (Table S4) are further from ideal for K+. Divalent cations have a larger effect, with Mg^{2+} particularly notable in its stabilisation of the α-sheet. In this conformation, the ion forms 5 close contacts with oxygen, including three backbone carbonyl as well as side chain of serine and glutamine to form a square-based pyramidal coordination (Figure 3). Ca^{2+} forms the same number of contacts, with distances increased to reflect the larger ionic radius. Comparing these distances to those Table S4 shows that the smaller ions Na^+ and Mg^{2+} form contacts closer to the ideal length than the larger ions, such that the match in size of cation with the "nest" of potential cation binding groups is key to the effect of ions. The involvement of side chains may also offer an explanation for Hayward's observation that polar amino acids favour α-sheet form. In β-strand, ion…O contacts also form, but these are either fewer and/or longer than those found in α-sheet (Table S4). This is particularly so for Mg^{2+}, for which only three contacts are formed.

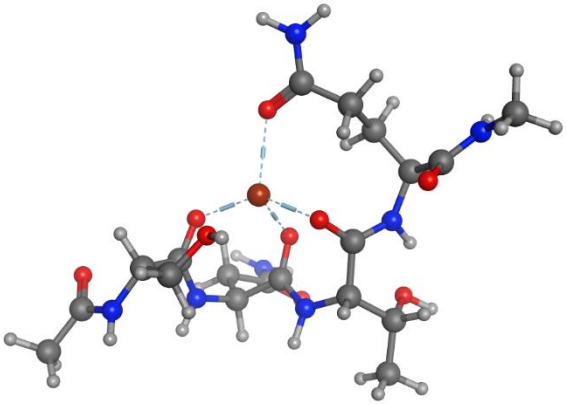

Figure 3 DFT optimised geometry of SNTQ-Mg^{2+} in α-sheet conformation, with Mg…O contacts shown

Conclusions

We have used *ab initio* and DFT calculations on model peptides to assess the relative stability of α-sheet and β-strand conformations. In gas phase the β-strand form of Ac-Ala_4-NMe is distinctly more stable, but in implicit solvent and especially in the presence of monovalent or divalent ions this changes markedly such that in most cases α-sheet is preferred. The size of this effect depends on ionic charge and radius: Na^+ has very little effect on relative stability, whereas K^+ has a large effect, as do Mg^{2+} and Ca^{2+}. This is traced to the number of short ion…O contacts that are able to form, for which K^+ seems especially well-suited to the "nest" of donor sites in the α-sheet.

Applying the same approach to two sequences proposed as possible α-sheet from a review of the Protein Data Bank shows that ion effects may also be sequence-dependent. One example, the polar SNTQ taken from PDB entry 1FEP, shows strong preference for α-sheet in the presence of most ions considered, with particularly significant effects from divalent ions that form interactions with both backbone and sidechain oxygen. In contrast, the non-polar AHLG (from PDB entry 1A05) is predicted to favour β-strand in the presence of ions, due to steric hindrance from the large Leu residue as well as favourable ion-His interactions in the strand form.

These static calculations therefore give some evidence for ionic effects in conformational preference, but do not include the dynamic processes that are vital for peptide structure and function. Use of accurate *ab initio* methods as benchmarks also limit the size of peptide that can

be considered. Moreover, implicit solvent models are inevitably limited in description of specific peptide-solvent and ion-solvent interactions. We hope to report explicit solvent molecular dynamics simulation of relevant and realistic peptides and proteins in a future publication.

CRediT authorship contribution statement

N. Alshammari: Visualization, Validation, Writing- Original draft preparation. J. A. Platts: Conceptualization, Investigation, Validation, Writing- Reviewing and Editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Advanced Research Computing at Cardiff (ARCCA) for computational resources.

Supplementary material

Additional structural, spectroscopic and energetic data available as Supplementary Material.

References

1. Hollingsworth, S. A. & Karplus, P. A. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol. Concepts* **1**, 271–283 (2010).
2. Pauling, L. & Corey, R. B. The Pleated Sheet, A New Layer Configuration of Polypeptide Chains. *Proc. Natl. Acad. Sci.* **37**, 251–256 (1951).
3. Armen, R. S., Alonso, D. O. V. & Daggett, V. Anatomy of an Amyloidogenic Intermediate. *Structure* **12**, 1847–1863 (2004).
4. Daggett, V. α-Sheet: The Toxic Conformer in Amyloid Diseases? *Acc. Chem. Res.* **39**, 594–602 (2006).
5. Balupuri, A., Choi, K.-E. & Kang, N. S. Computational insights into the role of α-strand/sheet in aggregation of α-synuclein. *Sci. Rep.* **9**, 59 (2019).
6. Hayward, S. & Milner-White, E. J. Determination of amino acids that favour the αL region using Ramachandran propensity plots. Implications for α-sheet as the possible amyloid intermediate. *J. Struct. Biol.* **213**, 107738 (2021).
7. Doyle, D. A. The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **280**, 69–77 (1998).
8. Al-Shammari, N., Savva, L., Kennedy-Britten, O. & Platts, J. A. Forcefield evaluation and accelerated molecular dynamics simulation of Zn(II) binding to N-terminus of amyloid-β. *Comput. Biol. Chem.* **93**, 107540 (2021).
9. Wieczorek, E. *et al.* Destabilisation of the structure of transthyretin is driven by Ca2+. *Int. J. Biol. Macromol.* **166**, 409–423 (2021).
10. *Molecular Operating Environment (MOE)*. (Chemical Computing Group Inc., 2013).
11. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).

12. Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822–8824 (1986).

13. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).

14. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).

15. Tomasi, J., Mennucci, B. & Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **105**, 2999–3094 (2005).

16. Liakos, D. G., Guo, Y. & Neese, F. Comprehensive Benchmark Results for the Domain Based Local Pair Natural Orbital Coupled Cluster Method (DLPNO-CCSD(T)) for Closed- and Open-Shell Systems. *J. Phys. Chem. A* **124**, 90–100 (2020).

17. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057 (2006).

18. Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **8**, (2018).