Chemical Physics Letters 563 (2013) 80–87

Contents lists available at SciVerse ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Helix compactness and stability: Electron structure calculations of conformer dependent thermodynamic functions

Imre Jákli [a], Imre G. Csizmadia [a,b,c], Szilard N. Fejer [b,d], Ödön Farkas [e], Bela Viskolcz [b], Svend J. Knak Jensen [f,*], Andras Perczel [a,g,*]

[a] Protein Modeling Group HAS-ELTE, Institute of Chemistry, Eötvös Loránd University, H-1538 Budapest, P.O. Box 32, Hungary
[b] Department of Chemical Informatics, Faculty of Education, University of Szeged, H-6701 Szeged, P.O. Box 396, Hungary
[c] Department of Chemistry, University of Toronto, 80 St George Street, Toronto, ON, Canada M5S 3H6
[d] Pro-Vitam Ltd., Str. Muncitorilor Nr. 16, Sfântu-Gheorghe 520032, Romania
[e] Department of Organic Chemistry, Institute of Chemistry, Eötvös Loránd University, H-1538, Budapest, P.O. Box 32, Hungary
[f] Department of Chemistry, 140 Langelandsgade, University of Aarhus, DK-8000 Aarhus C, Denmark
[g] Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University, Pázmány Péter Sétány 1/A, H-1117, Budapest, Hungary

ARTICLE INFO

Article history:
Received 19 October 2012
In final form 29 January 2013
Available online 8 February 2013

This Letter is dedicated to Professor Imre G. Csizmadia on the occasion of his 80th birthday by the co-authors and friends in appreciation of his contribution to the advancement of computational chemistry.

ABSTRACT

Structure, stability, cooperativity and molecular packing of two major backbone forms: 3_{10}-helix and β-strand are investigated. Long models HCO-(Xxx)$_n$-NH$_2$ Xxx = Gly and (L-)Ala, $n \leqslant 34$, are studied at two levels of theory including the effect of dispersion forces. Structure and folding preferences are established, the length modulated cooperativity and side-chain determined fold compactness is quantified. By monitoring $\Delta G^{\circ}_{\beta\rightarrow\alpha}$ rather than the electronic energy, $\Delta E_{\beta\rightarrow\alpha}$, it appears that Ala is a much better helix forming residue than Gly. The achiral Gly forms a more compact 3_{10}-helix than any chiral amino acid residue probed here for L-Ala.

1. Introduction

Polypeptides can adopt a number of different conformers, but helices and β-strands are the most abundant structural motifs of proteins. Among helices the α- or 3.6_{13}-helix (sometimes it is called 4_{13}), with the characteristic $i \leftarrow (i+4)$ type H-bonds is the most common [1]. Less frequently 3_{10}- and sporadically π-helical turns can also be assigned in proteins, all of right handedness. Regardless of their length α-helices often start and terminate by shorter 3_{10}-helical turn(s), where the intramolecular H-bonds are of $i \leftarrow (i+3)$ type. This results in a narrower fold typically found by computational methods. In vacuum, the latter helical structure is the more stable backbone fold [2]. Furthermore, the mixing and interconversion of 3_{10}- and α-helices is allowed and was successfully obtained by quantum mechanical (QM) calculations in various solvents [2]. Helices have noticeable dipole moments, as each of the component homo-conformers (α_L for short [3]) has their own dipole moment. The constructive summation is due to the similar orientation of the adjacent amide planes around and along the central axis of a helix. This results in a significant macro dipole moment, with a positive end at the N- and a negative end at the C-terminus.

Helices can be involved in a variety of motifs, such as all-α, βαβ, Rossman-fold and TIM barrel [4]. Even in *i*ntrinsically *d*ynamic *p*roteins (IDP), where the backbone of the protein presents an unexpectedly large structure fluctuation, residual helices could be assigned [5], as temporarily existing structural motifs [6]. Helix packing is an important issue, as there are only a limited number of modes how helices can self organize. In the most common four-helix bundle arrangement, the topology is 'up-and-down' allowing the side-chains formed ridges to ideally pack into each other's grooves. Different type of membrane proteins were already characterized; those equipped with a single α-helical tail anchoring the globular part to the membrane and those passing through the membrane several times (e.g., bacteriorhodopsin) resulting in seven transmembrane (7TM) helices. The longest helical regions are in keratine, the protein of hair, in myosine and tropomyosine [7,8], which form muscular fibers, winding around each other by forming the coiled-coil superstructure.

The length of an α-helix varies as function of several factors, but in globular proteins most helices comprise around 10–15 amino acid residues [9]. By analyzing over 150 globular proteins the average length was found to be ~18 ± 8, ranging between 10 and 50 [10]. In coiled-coils [11], the average length is ~17, based on a recent survey by using PDB select 2008 [12].

* Corresponding authors. Fax: +45 8619 6199 (S.J. Knak Jensen), +36 1 3722 620 (A. Perczel).
E-mail addresses: kemskj@chem.au.dk (S.J. Knak Jensen), perczel@chem.elte.hu (A. Perczel).

0009-2614/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cplett.2013.01.060

However, the helices forming coiled-coils can be rather long, comprising well over hundred residues (e.g., myosine). The analysis of a total of 160 transmembrane helices of 15 non-homologous proteins resulted in an average length of 17(18) ± 2(3) [10]. Finally, in a single α-helical motif, the length of this secondary structural element can be significantly longer, exceeding 60 helical residues [13]. Thus, elucidation of the backbone structural features and the inherent cooperativity of helices of various lengths are of significance. Long helices could have additional structural features, they can be curved or kinked [14].

Today computational quantum chemistry (QM) is suitable to study even longer biopolymers at an acceptable level of accuracy. The great feature of a QM approach is that beside structural information (comparable to experiments), it can provide stability data on molecular conformers. Varieties of QM methods have been shown to give reliable information on the stability of foldamers and transition state structures. It has been shown previously [2], in the case of $(Ala)_8$, that the greater stability of the 4_{13}- or α-helix is due to solvation. Direct hydration with the inclusion of explicit water molecule(s) has also been reported [15]. Consequently the 3_{10}-helix must exhibit the intrinsic properties of the right-handed helical structure, since it is the most stable form when environmental effects are excluded. Several QM computations have been published on 3_{10}-helical structures [16–19]. However, these studies reported electronic energies rather than thermodynamic functions. Here we have used electron structure calculations to obtain thermodynamic properties of oligopeptides, namely $(Gly)_n$ and $(L\text{-}Ala)_n$ in their extended- and 3_{10}-helical conformers, where n ranges between 1 and 34. These data allow us to study the build-up of the internal H-bond network and the associated folding process as a function of the number of residues within the peptide. The thermodynamics of the folding process, particularly the associated entropy change, is expected to contain a great deal of useful information [20]. In helices, L-Ala is regarded as the most preferred (most stabilizing) residue, while Gly (beside Pro) is the one to destabilize helical motifs the most [21]. Although, the structural difference between the above two residues is extremely small (–H versus $–CH_3$), the propensity difference is very significant. The difference in helix stabilization is typically explained by hydrophobic and folding entropy effects. We will focus here on the entropy term with respect to the overall stabilization.

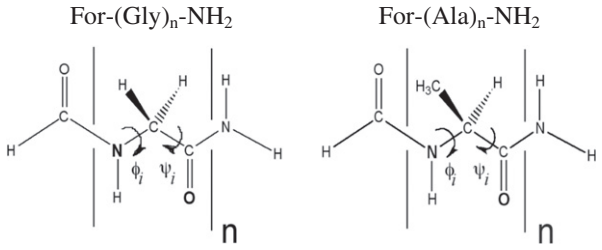
Our goal is to decipher the entropy and stability measures of a helix of various lengths. This Letter has to precede studies on more complex bundled helices (e.g., ion channels, 7TM systems) of great biochemical significance. In addition, we wanted to separate stability arising from backbone/backbone interaction, rather than from more specific backbone/side-chain interactions or from hydration.

Thus, the following specific questions were asked:

(i) Is a 3_{10}-helix more stable and compact than a single β-strand?
(ii) How do side chains and their chirality affect molecular packing and helix stability?
(iii) How accurately can the stability of a longer helix be predicted from its fragments?
(iv) What is the magnitude and location preference of subunit cooperativity in a helix?

2. Methods and computational details

The oligo- and poly(Gly) and Ala peptides used here have the following chemical structures:



where n stands for the number of residues.

Peptide geometries, both β-strand and 3_{10}-helical forms were determined by full optimization, using the GAUSSIAN 09 software [22] following the strict conformer selection rules published earlier [23,24]. The chosen levels of theory are *ab initio* Hartree–Fock and B3LYP implementation of the density functional theory. In both cases the chosen basis set was 6-31G(d). Selected structure calculations were also completed at the larger basis set 6-311+G(d,p), to assess the quality of the data. The optimized structures allow calculation of the harmonic vibration frequencies, which in turn allow determination of the thermodynamic functions ($H°$, $G°$, and $S°$). We note that the entropy is sensitive to low frequency modes which are found for long peptides. For a selected n the lowest frequency is found in the β-strand – both for L-Ala and for Gly.

For describing the 'folding process', thermodynamic functions and their normalized forms were calculated as follows:

$$\Delta H^\circ_{\beta\to\alpha} = H^\circ[(\alpha_L)_n] - H^\circ[(\beta_L)_n] \text{ and } \Delta H^\circ[n]_{\beta\to\alpha}/n \quad (1)$$

$$\Delta G^\circ_{\beta\to\alpha} = G^\circ[(\alpha_L)_n] - G^\circ[(\beta_L)_n] \text{ and } \Delta G^\circ[n]_{\beta\to\alpha}/n \quad (2)$$

$$\Delta S^\circ_{\beta\to\alpha} = S^o[(\alpha_L)_n] - S^\circ[(\beta_L)_n] \text{ and } \Delta S^\circ[n]_{\beta\to\alpha}/n \quad (3)$$

for $n \leqslant 34$. The subscript $\beta \to \alpha$ indicates conformational transition from the β-strand to the 3_{10}-helical conformer. We have completed geometry optimizations up to $n = 34$, where the overall length of the folded peptide is about 65 Å in its helical structure, similar to recent estimates [25]. For practical reasons the vibration frequencies were only determined for $n = 34$ in the case of Gly.

The calculations were done for the gas phase, as reliable methods for calculation of entropies in solution are computationally very demanding [26,27]. This restriction may not be too important if data are to mimic structure and stability of secondary structural elements shielded from the solvent (e.g., inside of a protein or a membrane).

In the world of the experimentalists the reference-state for the experimental observations favoring Ala as a helix maker is not a β-strand, which is a relatively uncommon secondary structural motif, but rather a random coil. In the context of the present calculations, this means that the reference-state should consists of a large number, x, of 'coil' conformations in thermal equilibrium, the large number itself adding to the entropy as $R \cdot ln\ x$. Since in the present case a single conformation, the β-strand, was used as reference-state the computed internal entropy cannot be compared quantitatively to the experimentalists' entropy, but qualitative comparisons can indicate trends.

3. Results

3.1. Molecular structure I: torsion angles and twisting

Molecular structures of both the β-stranded and helical structures of Ala_{10} (Figure 1) are very similar to those calculated