Influence of the Environment in the Conformation of α-Helices Studied by Protein Database Search and Molecular Dynamics Simulations

Mireia Olivella,* Xavier Deupi,* Cedric Govaerts,†‡ and Leonardo Pardo*

*Laboratori de Medicina Computacional, Unitat de Bioestadística, Facultat de Medicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain; †Institut de Recherche Interdisciplinaire en Biologie Humaine et Nucléaire, Université Libre de Bruxelles, Campus Erasme, B-1070 Bruxelles, Belgium; and ‡Service de Conformation des Macromolécules Biologiques, Université Libre de Bruxelles, 1050 Bruxelles, Belgium

ABSTRACT   The influence of the solvent on the main-chain conformation (ϕ and Ψ dihedral angles) of α-helices has been studied by complementary approaches. A first approach consisted in surveying crystal structures of both soluble and membrane proteins. The residues of analysis were further classified as exposed to either the water (polar solvent) or the lipid (apolar solvent) environment or buried to the core of the protein (intermediate polarity). The statistical results show that the more polar the environment, the lower the value of ϕ_i and the higher the value of Ψ_i are. The intrahelical hydrogen bond distance increases in water-exposed residues due to the additional hydrogen bond between the peptide carbonyl oxygen and the aqueous environment. A second approach involved nanosecond molecular dynamics simulations of poly-Ala α-helices in environments of different polarity: water to mimic hydrophilic environments that can form hydrogen bonds with the peptide carbonyl oxygen and methane to mimic hydrophobic environments without this hydrogen bond capabilities. These simulations reproduce similar effects in ϕ and Ψ angles and intrahelical hydrogen bond distance and angle as observed in the protein survey analysis. The magnitude of the intrahelical hydrogen bond in the methane environment is stronger than in the water environment, suggesting that α-helices in membrane-embedded proteins are less flexible than in soluble proteins. There is a remarkable coincidence between the ϕ and Ψ angles obtained in the analysis of residues exposed to the lipid in membrane proteins and the results from computer simulations in methane, which suggests that this simulation protocol properly mimic the lipidic cell membrane and reproduce several structural characteristics of membrane-embedded proteins. Finally, we have compared the ϕ and Ψ torsional angles of Pro kinks in membrane protein crystal structures and in computer simulations.

INTRODUCTION

α-Helices are major structural elements in both soluble and membrane proteins (Fasman, 1989; White and Wimley, 1999). The stability of α-helices is basically achieved by the hydrogen bonds between the N—H atoms of residue i to the carbonyl oxygen of residue $i - 4$ in the preceding turn of the helix. Importantly, in transmembrane proteins, the formation of this hydrogen bond network allows the polar polypeptide backbone to expand the hydrophobic lipid bilayer of the cell membrane. Thus, the helical bundle motif frequently builds the three-dimensional structure of membrane proteins along with the β-barrel motif also observed in membrane-spanning proteins (White and Wimley, 1999).

An early statistical analysis of the conformation of α-helices in crystal structures of mostly soluble proteins (Barlow and Thornton, 1988) showed average main-chain torsion ϕ and Ψ angles of $-62°$ and $-41°$, respectively. However, additional hydrogen bonds between the peptide carbonyl oxygen to a solvent molecule (Blundell et al., 1983) or to a protein side-chain (Ballesteros et al., 2000) produce a sig-

nificant change in ϕ and Ψ angles and in the curvature of the helix. Thus, it seems reasonable to assume that the conformation of α-helices located in hydrophilic environments, such as water, will differ from the conformation of α-helices located in hydrophobic environments, such as the cell membrane.

To assess the influence of the environment on the conformation of α-helices, complementary approaches were used in this study. A first approach consisted in surveying known protein structures. The results are presented for crystal structures of both soluble and membrane proteins. Despite the limited availability of membrane protein structures in the Brookhaven protein data bank (PDB), the significant increase in the number of deposited structures during the last years yields to an acceptable number of transmembrane helices for statistical analysis. Moreover, the residues of analysis are further classified as exposed, to either the water or the lipid environment, or buried to the core of the protein. A second approach involved nanosecond molecular dynamics simulations of poly-Ala α-helices in environments of different polarity: water and methane. The main-chain ϕ and Ψ torsional angles and intrahelical hydrogen bond parameters obtained in the analysis of protein crystal structures are compared with those obtained in computer simulations. Moreover, we have compared the ϕ and Ψ torsional angles of Pro kinks in membrane protein crystal structures and in computer simulations.

MATERIALS AND METHODS

Membrane protein structures

The atomic coordinates of bacteriorhodopsin (PDB access number 1c3w, 1.55-Å resolution), aa3 (1occ, 2.8 Å), and ba3 (1ehk, 2.4 Å) cytochrome c oxidases, photosyntethic reaction center (1prc, 2.3 Å), potassium channel (1bl8, 3.2 Å), mechanosensitive ion channel (1msl, 3.5 Å), rhodopsin (1f88, 2.8 Å), halorhodopsin (1el2, 1.8 Å), sensory rhodopsin (1h68, 2.1 Å), light harvesting complex (1lgh, 2.4 Å), photosystem I (1jbo, 2.5 Å), AQP1 (1hwo, 3.7 Å), and GlpF (1fx8, 2.2 Å) channels, P-type ATPasa (1eul, 2.6 Å), and fumerate reductase respiratory complex (1qla, 2.2 Å) were obtained from the Brookhaven PDB. The coordinates of the residues in the HELIX annotation of the PDB files, corresponding to transmembrane helices 1–7 of 1c3w; 2–3, 7, 9, 12, 14–15, 19–20, 23, 28–30, 32–35, 41, 54, 59–60, and 63–66 of 1occ; 1, 3–9, 13–14, 16, 18–19, and 22 of 1ehk; 6, 8–10, and 13–14 of 1prc; 1 and 3 of 1bl8; 2–4 of 1msl; 1–7 of 1f88; 1–6, 8–9, 13–14, and 16 of 1el2; 1–8 of 1h68; 2 and 5 of 1lgh; 4, 8, 10, 16, 20, 27, 34–36, 40, 44, 48, 53, 57, 59, 68, 71, 77, 80, 85, 94, 103, 105, 109, 113, and 115–116 of 1jbo; 1 of 1hwo; 1–6, 9, 11–12, and 15 of 1fx8; 2, 4–5, 10–12, 15–16, 20, 25, 28, 31, 36, 38, and 41–43 of 1eul; and 1, 13, 16–20, 22, 25, 28, 29, 32, 26, 38, 40–41, 43–44, 47–48, and 81 of 1qla, were extracted for analysis. This results in a total of 160 transmembrane helices. These helices were split into amino acid stretches of 12 residues long with 1) Ala (349 structures) or 2) Pro (27 structures) at the eighth position. Stretches with other Pro residues in the sequence were removed from the database.

Soluble protein structures

Iditis 3.1 (Oxford Molecular, Oxford, U.K.) was used for the selection of protein structures in the Brookhaven PDB. The chosen α-helices possess: 1) a resolution of 2.0 Å or better; 2) 12 residues length with Ala at the eighth position; and 3) no Pro residues in the sequence. If two α-helical segments have more than 80% sequence identity (if 10 or more than 10 residues of 12 are identical) only the structure with best resolution was considered.

Accessible surface

The accessible surface of the residues in the survey of protein crystal structures at the fourth ($i - 4$) and the eighth (i) positions, was obtained with the Naccess program (Hubbart and Thornton, 1993). The sum of the accessible surface of residues i and $i - 4$ was used to classify the helices as exposed (>60) or buried (<40). These cutoffs were chosen by visual inspection of the crystal structures. The structures between these values could not be visually assigned to either group and were not included in the analysis.

Molecular dynamics simulations

The model peptides Ace-Ala$_{25}$-Nme and Ace-Ala$_{12}$-Pro-Ala$_{12}$-Nme were built in the standard α-helical conformation (backbone dihedral angles ϕ and Ψ of -58 and $-47°$) using the SYBYL 6.5 program (Tripos Inc., St. Louis, MO). The Ace-Ala$_{25}$-Nme structure was placed in a rectangular box containing 808 water or 1532 methane molecules, and the Ace-Ala$_{12}$-Pro-Ala$_{12}$-Nme structure was placed in a rectangular box containing 1689 methane molecules. The sizes of the boxes were approximately 52 × 23 × 23 Å for the α-helix in water, and 60 × 36 × 36 Å for the α-helices in methane, resulting in a density of 1.0 g cm^{-3} and 0.5 g cm^{-3}, respectively. It is important to note that the density of the methane box is not the density observed in the hydrophobic core of the membrane (White and Wimley, 1999). This is due to the different equilibrium distance between carbons in the methane box and in the polycarbon chain of the lipid. The density of the
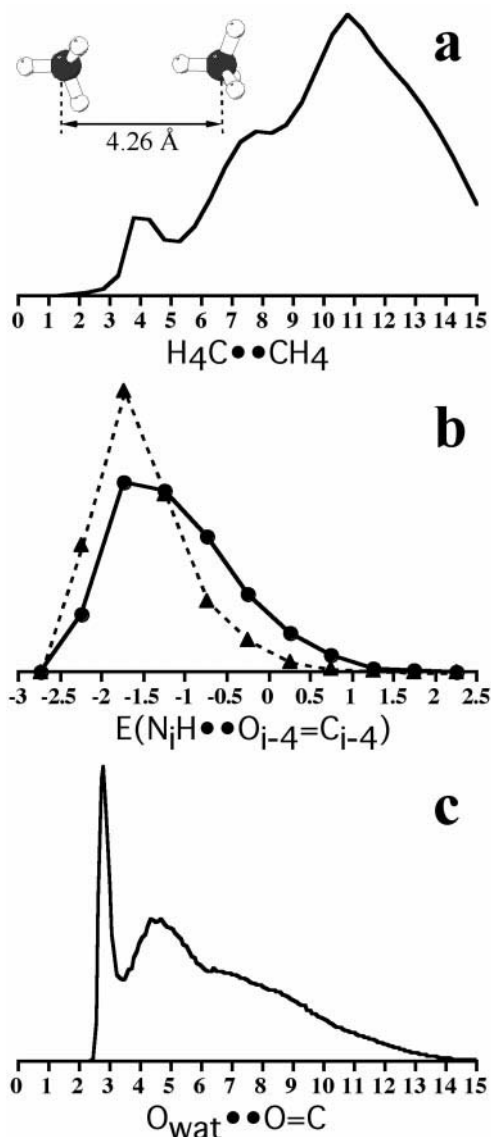

FIGURE 1 (*a*) Radial distribution function for the H$_4$C··CH$_4$ distance (Å) obtained in molecular dynamics simulations of methane and the structure of H$_4$C··CH$_4$ obtained by full geometry optimization with ab initio quantum mechanical calculations at the MP2/6-31G** level of theory. (*b*) Distribution of the energy of interaction (kcal/mol) between the N—H atoms of residue i and the carbonyl group of residue $i - 4$ obtained from the molecular dynamics simulations of a poly-Ala α-helix in water (*circles*, *solid line*) and methane (*triangles*, *broken line*). (*c*) Radial distribution function for the distance (Å) between the peptide carbonyl oxygen and the oxygen of the water molecules obtained in the molecular dynamics simulations of a poly-Ala α-helix in water.

methane box was chosen to equal the first peak of the radial distribution function for the H$_4$C··CH$_4$ distance obtained in the molecular dynamics simulations with the interatomic distance between two methane molecules obtained by full geometry optimization with ab initio quantum mechanical calculations at the MP2/6-31G** level of theory (Fig. 1 *a*). An increase of the density of the methane box leads to short contacts between molecules and thus extreme behavior of the system.

Initially, the atoms of the model peptides were kept fixed, whereas the solvent molecules were energy minimized (500 steps), heated (from 0–300

TABLE 1 Means/standard deviations of the backbone torsion angles (ϕ_i and ψ_i, in degrees) of the residue at the eighth position (denoted as *i*) in the survey of α-helices containing Ala in protein crystal structures or at the 13th position (denoted as *i*) in the molecular dynamics simulations of poly-Ala α-helix, intrahelical hydrogen bond distance ($N_i \cdot \cdot O_{i-4}$, in Å), and angle ($N_i \cdot \cdot O_{i-4} = C_{i-4}$, in degrees) between the N atom of residue *i* to the carbonyl of residue *i* − 4, the energy of interaction between the N–H atoms of residue *i* and the carbonyl group of residue *i* − 4 ($E(N_i \cdot \cdot O_{i-4} = C_{i-4})$, in kcal/mol), and the intermolecular hydrogen bond distance ($O_{wat} \cdot \cdot O_{i-4}$, in Å) and angle ($O_{wat} \cdot \cdot O_{i-4} = C_{i-4}$, in degrees) between the peptide carbonyl oxygen and the oxygen of the water molecules obtained in the molecular dynamics simulations of the poly-Ala α-helix in water

	Protein database search			Molecular dynamics	
	$SOL_{hydrophilic}$	$SOL\text{-}MEM_{core}$	$MEM_{hydrophobic}$	Water	Methane
n	252	510	97	1000	1000
ϕ_i	−63.5/5.6	−62.9/5.3	−61.8/6.7	−65.9/10.0	−61.2/8.3
ψ_i	−40.9/5.4	−41.6/6.1	−43.1/7.0	−39.3/9.7	−44.1/8.5
$N_i \cdot \cdot O_{i-4}$	3.04/0.14	2.98/0.15	2.96/0.17	3.10/0.25	2.93/0.13
$N_i \cdot \cdot O_{i-4} = C_{i-4}$	151.5/6.0	153.3/7.1	153.5/7.5	148.9/10.5	154.4/8.3
$E(N_i \cdot \cdot O_{i-4}) = C_{i-4}$				−1.1/0.7	−1.5/0.6
$O_{wat} \cdot \cdot O_{i-4}$				2.94/0.2	
$O_{wat} \cdot \cdot O_{i-4} = C_{i-4}$				116.6/15.3	

K in 15 ps), and equilibrated (from 15–50 ps). Subsequently, the entire system was subjected to 500 iterations of energy minimization and then heated to 300 K in 15 ps. This was followed by an equilibration period (15–500 ps for Ace-Ala$_{25}$-Nme, and from 15–1000 ps for Ace-Ala$_{12}$-Pro-Ala$_{12}$-Nme) and a production run (from 500–1000 ps for Ace-Ala$_{25}$-Nme, and from 1000–1500 ps for Ace-Ala$_{12}$-Pro-Ala$_{12}$-Nme) at constant volume using the particle mesh Ewald method to evaluate electrostatic interactions (Darden et al., 1993). The equilibration time was chosen so that root mean square deviations relative to the first structure in the simulations remained constant (results not shown). The longer equilibration period of the Pro-containing structure is necessary to account for the flexibility of Pro kinks. Structures were collected for analysis every 0.5 ps during the last 500 ps of simulation (1000 structures). The energy of interaction between the N—H atoms of residue 13 and the carbonyl group of residue 9 was calculated with the Anal program of AMBER 5 (Case et al., 1997). The molecular dynamics simulations were run with the Sander module of AMBER 5, the all-atom force field (Cornell et al., 1995), SHAKE bond constraints in all bonds, a 2-fs integration time step, and constant temperature of 300 K coupled to a heat bath.

Statistical analysis

One-way analysis of variance for independent samples plus a posteriori one-sided Tukey's test was used for contrasting the backbone torsion angles at position 8 (ϕ_i and Ψ_i) and intrahelical hydrogen bond distance ($N_i \cdot \cdot O_{i-4}$) and angle ($N_i \cdot \cdot O_{i-4} = C_{i-4}$) between residues in soluble proteins that are exposed to the hydrophilic aqueous solvent, in membrane proteins that are exposed to the hydrophobic lipid bilayer, and in both soluble and membrane proteins that are exposed to the core of the protein. Averages and standard deviations of ϕ_i, Ψ_i, $N_i \cdot \cdot O_{i-4}$, and $N_i \cdot \cdot O_{i-4} = C_{i-4}$ obtained in the molecular dynamics simulations were calculated from all the geometries in the production phase. The data obtained in molecular dynamics simulations are not independent, thus it is not possible to perform statistical tests as in the protein survey analysis. The statistical analysis was performed with the SPSS 10 program (SPSS Inc. Chicago, IL).

RESULTS AND DISCUSSION

Survey of helices in known protein structures

Table 1 summarizes the means and standard deviations for the backbone torsion angles of the residue at position 8 (ϕ_i and Ψ_i), populated by Ala, of α-helices (see Materials and

Methods) in soluble proteins that are exposed to the hydrophilic aqueous solvent ($SOL_{hydrophilic}$, 252 entries), in membrane proteins that are exposed to the hydrophobic lipid bilayer ($MEM_{hydrophobic}$, 97 entries), and in both soluble and membrane proteins that are exposed to the core of the protein ($SOL\text{-}MEM_{core}$, 510 entries). It has recently been proposed that, in contrast to previous hypothesis, the hydrophobicities of interior residues of both membrane and water-soluble proteins are comparable (Rees and Eisenberg, 2000; Stevens and Arkin, 1999). In consequence, the residues of α-helices pointing toward the core of soluble and membrane proteins have been grouped ($SOL\text{-}MEM_{core}$). Thus, the expected rank order of hydrophobicity, from hydrophobic to hydrophilic, of the environment to which the analyzed residues are exposed is: $MEM_{hydrophobic} > SOL\text{-}MEM_{core} > SOL_{hydrophilic}$. Besides, ϕ and Ψ angles vary depending on both side-chain type and side-chain conformation (Ballesteros et al., 2000; Chakrabarti and Pal, 1998). We limited the survey to alanine to avoid any direct or indirect effect of the side-chain in the conformation of the helix. In addition, Ala is the most helix-favoring residue in water (O'Neil and DeGrado, 1990), and it has one of the lowest turn propensities in transmembrane helices (Monne et al., 1999). Ala was favored over Gly because the lack of side chain in Gly provides additional flexibility (Kumar and Bansal, 1998). As shown in Table 1, the values of the backbone ϕ_i dihedral are found in the following rank order: $MEM_{hydrophobic}$ (−61.8°) > $SOL\text{-}MEM_{core}$ (−62.9°) > $SOL_{hydrophilic}$ (−63.5°). Thus, there is a positive correlation between hydrophobicity and ϕ_i: the more hydrophobic the environment, the higher the value of ϕ_i is. The values of the backbone Ψ_i dihedral are found in the following rank order: $MEM_{hydrophobic}$ (−43.1°) < $SOL\text{-}MEM_{core}$ (−41.6°) < $SOL_{hydrophilic}$ (−40.9°). Thus, in the case of Ψ_i the correlation is negative: the more hydrophobic the environment, the lower the value of Ψ_i is. It is important to remark that the difference between the conformation of an α-helix ex-

posed to either the hydrophilic aqueous solvent or the hydrophobic lipid bilayer is in average 1.7° for ϕ_i and 2.2° for Ψ_i. These differences in ϕ_i ($p = 0.016$) and Ψ_i ($p = 0.003$) are significant from a statistical point of view (see Materials and Methods). However, there are not statistical differences in ϕ_i and Ψ_i between SOL-MEM$_{core}$ and MEM$_{hydrophobic}$ or between SOL-MEM$_{core}$ and SOL$_{hydrophilic}$. Considering the small amplitudes of the difference, the influence of the lipidic or aqueous environment in the conformation of the α-helix will only be noticeable for long helices. The deviation between C-terminal positions of helices constructed with the ϕ_i and Ψ_i angles reported in Table 1 for SOL$_{hydrophilic}$ ($-63.5°$ and $-40.9°$) and MEM$_{hydrophobic}$ ($-61.8°$ and $-43.1°$), is 0.9 Å or 1.4 Å or 1.7 Å if helices 20 or 25 or 30 residues long are considered, respectively.

Table 1 also shows the means and standard deviations of the intrahelical hydrogen bond distance $N_i{\cdot}{\cdot}O_{i-4}$, and angle $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$. The $N_i{\cdot}{\cdot}O_{i-4}$ distance increases as the environment becomes more hydrophilic: MEM$_{hydrophobic}$ (2.96 Å) > SOL-MEM$_{core}$ (2.98 Å) > SOL$_{hydrophilic}$ (3.04 Å). There are statistical differences between SOL$_{hydrophilic}$ and both SOL-MEM$_{core}$ ($p < 0.0005$) and MEM$_{hydrophobic}$ ($p < 0.0005$). Clearly, the additional hydrogen bond between the peptide carbonyl oxygen to a solvent molecule, in water-exposed residues (SOL$_{hydrophilic}$), increases the intrahelical hydrogen bond distance. Correspondingly, the $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$ angle decreases in linearity in water exposed residues: MEM$_{hydrophobic}$ (153.5°) > SOL-MEM$_{core}$ (153.3°) > SOL$_{hydrophilic}$ (151.5°). Similarly to the $N_i{\cdot}{\cdot}O_{i-4}$ hydrogen bond distance, there are statistical differences between SOL$_{hydrophilic}$ and both SOL-MEM$_{core}$ ($p = 0.001$) and MEM$_{hydrophobic}$ ($p = 0.025$). Following the argument put forward by Blundell et al. (1983), the presence of a second hydrogen bond donor (i.e., a solvent molecule: O_{wat}) to the peptide carbonyl oxygen tends to bifurcate the $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$ and the $O_{wat}{\cdot}{\cdot}O_{i-4} = C_{i-4}$ angles toward 120° (see below).

Molecular dynamics simulations of poly-Ala α-helices

We have performed nanosecond molecular dynamics simulations of poly-Ala α-helices (see Materials and Methods) in two different environments: water to mimic hydrophilic environments that can form hydrogen bonds with the peptide carbonyl oxygen of the α-helix and methane to mimic hydrophobic environments without this hydrogen bond capabilities. Table 1 shows the obtained values of ϕ_i and Ψ_i and the intrahelical hydrogen bond parameters $N_i{\cdot}{\cdot}O_{i-4}$ and $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$ (in which i denotes residue number 13 in the poly-Ala α-helix). Notably, the effect of the environment observed in molecular dynamics simulations is the same in both magnitude and direction as the observed in the protein survey analysis. The polar environment formed by the water molecules tends to decrease ϕ_i ($-61.2°$ vs.

$-65.9°$), increase Ψ_i ($-44.1°$ vs. $-39.3°$), increase $N_i{\cdot}{\cdot}O_{i-4}$ (2.93 Å vs. 3.10 Å), and decrease $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$ (154.4° vs. 148.9°), relative to the apolar environment formed by the methane molecules. Thus, the presence or the absence of additional hydrogen bonds from the environment to the peptide carbonyl oxygen modifies the conformation of α-helices.

It is important to note that there is a remarkable coincidence between the values obtained in the analysis of exposed residues in membrane proteins (MEM$_{hydrophobic}$) and the results from computer simulations in the methane environment (ϕ_i: $-61.8°$ vs. $-61.2°$; Ψ_i: $-43.1°$ vs. $-44.1°$; $N_i{\cdot}{\cdot}O_{i-4}$: 2.96 Å vs. 2.93 Å; $N_i{\cdot}{\cdot}O_{i-4} = C_{i-4}$: 153.5° vs. 154.4°; see Table 1). Thus, we suggest, based on this analysis, that explicit methane molecules in molecular dynamics simulations properly mimic the lipidic cell membrane and reproduce several structural characteristics of membrane-embedded proteins.

The fact that the intrahelical hydrogen bond distance ($N_i{\cdot}{\cdot}O_{i-4}$) in water (3.10 Å) is longer than in methane (2.93 Å) suggests that this hydrogen bond in water is weaker than in methane. To corroborate this hypothesis we have calculated the mean and standard deviation (Table 1) and the distribution (Fig. 1 b) of the energy of interaction between the N—H atoms of residue i and the carbonyl group of residue $i - 4$ obtained from the molecular dynamics simulations in water (circles, solid line) and methane (triangles, broken line). The magnitude of the intrahelical hydrogen bond in water is smaller than in methane (-1.1 vs. -1.5 kcal/mol). The formation of a second hydrogen bond between the peptide carbonyl oxygen and the aqueous solvent enfeebles the intrahelical hydrogen bond that stabilize α-helices. This destabilization of the intrahelical hydrogen bond in water suggests that α-helices are more flexible in polar environments. The larger standard deviation (Table 1) of the dihedral angles that define the conformation of the helix, ϕ_i (10.0° vs. 8.3°) and Ψ_i (9.7° vs. 8.5°), in water than in methane reinforces this proposal. However, it is important to note that the standard deviations of ϕ_i and Ψ_i in the protein survey analysis of exposed soluble and membrane proteins do not follow this trend. We attribute this to the different number of structures in each category and the better resolution of soluble proteins compared with membrane proteins.

Fig. 1 c shows the radial distribution function for the distance between the peptide carbonyl oxygen and the oxygen of the water molecules obtained in the molecular dynamics simulations of a poly-Ala α-helix in water. The first peak in the distribution occurs at distances up to 3.3 Å, which implies an explicit hydrogen bond between the carbonyl oxygen of the α-helix and water. To characterize the geometric parameters of this hydrogen bond ($O_{wat}{\cdot}{\cdot}O_{i-4}$ and $O_{wat}{\cdot}{\cdot}O_{i-4} = C_{i-4}$), we selected the
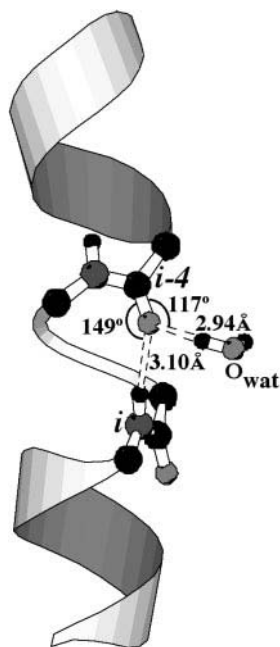
FIGURE 2   Representative structure obtained in the molecular dynamics simulations of the poly-Ala α-helices in water (see Materials and Methods). Distances (Å) and angles (degrees) are shown relative to the heavy atoms.

bound water molecules (O_{wat}··O_{i-4} < 3.3 Å) to the carbonyl oxygen from the 1000 structures computed during the last 500 ps of simulation (see Materials and Methods) for statistical analysis. Fig. 2 shows a representative structure of the interaction between the water molecule and the carbonyl group that occurs at a O_{wat}··O_{i-4} distance of 2.94 Å and at a O_{wat}··O_{i-4} = C_{i-4} angle of 116.6° (see Table 1). The electronic nature of the carbonyl oxygen allows the formation of a hydrogen bond with both the N—H group of the residue in the following turn of the helix and a water molecule.

Structural analysis of Pro-containing α-helices in hydrophobic environments

Pro induce distorsion in α-helices as their cyclic side-chains introduce a local break, denoted Pro kink, to avoid a steric clash between the pyrrolidine ring and the carbonyl oxygen of residue $i - 4$ (Barlow and Thornton, 1988; Milner-White et al., 1992; Sankararamakrishnan and Vishveshwara, 1992; Von Heijne, 1991). Pro kinks impart backbone flexibility, due to the absence of the hydrogen bond with the carbonyl oxygen in the preceding turn of the helix. This structural flexibility is an important functional element in membrane proteins that transduce extracellular signals across the membrane through conformational changes in the transmembrane α-helices (Gether et al., 1997; Govaerts et al., 2001a; Ri et al., 1999; Sansom and Weinstein, 2000). We have

studied the main-chain ϕ and Ψ torsional angles of Pro kinks in membrane protein crystal structures and in computer simulations. Pro kinks alter the conformation of a complete turn of the helix, from the Pro residue i to $i - 4$. Thus, the ϕ and Ψ angles of all these residues must be taken into account in the conformational analysis. In the protein survey analysis some of these residues forming the Pro kink will be exposed to the lipidic membrane and others to the core of the protein. In contrast, in the molecular dynamics simulation all these residues will be exposed to the hydrophobic environment made of methane molecules. Moreover, we have searched for Pro kinks with the xxxxP sequence in the crystal structures, where x is any residue except Pro, whereas we have run the AAAAP sequence in the molecular dynamics simulation (see Materials and Methods). Therefore, some divergences between crystal structures and computer simulations are expected due to the effect of the environment and the different residues forming the Pro kink. However, the effect of the environment (see above and Table 1) and the type of residue (Ballesteros et al., 2000; Chakrabarti and Pal, 1998) in the ϕ and Ψ torsional angles are much lower than the influence of the Pro residue in the conformation of the helix (Fig. 3). Fig. 3 shows the evolution of ϕ (squares) and Ψ (circles) torsional angles along the α-helix as observed during the molecular dynamics simulations (black line) and in the crystal structures of membrane proteins (broken line). The helical distortion induced by the Pro residue is clearly seen at the level of the dihedral angles up to residue four positions upstream. Clearly the simulation in the methane environment reproduces the dihedral angles profile of the Pro kink observed in the analysis of crystal structures (see Table 2), indicating that the methane box can reliably reproduce the conformational behavior of helical deformations as well.
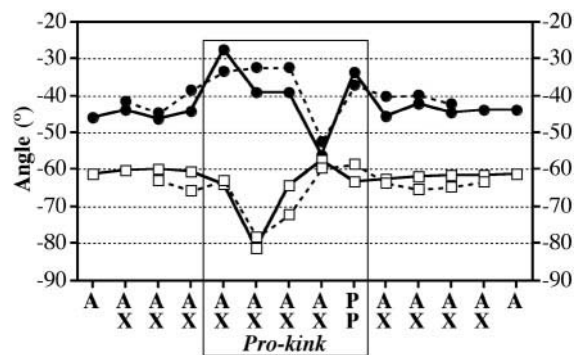

FIGURE 3   Mean values of the ϕ (squares) and Ψ (circles) torsional angles (degrees) along the α-helix containing the Pro kink as observed during the molecular dynamics simulations (black line) and in the crystal structures of membrane proteins (broken line). The x axis shows the sequence of the helix in the molecular dynamics simulations (top) and in the membrane protein survey (bottom). A, Ala; P, Pro; X, any residue except Pro.

TABLE 2 Means/standard deviations of the backbone torsion angles (ϕ and ψ, in degrees) of α-helices containing Pro in membrane crystal structures (27 entries) or in molecular dynamics simulations of Pro-containing α-helix in a methane environment (1000 entries)

Position	Membrane proteins		Molecular dynamics	
	ϕ_i	ψ_i	ϕ_i	ψ_i
$i-8$			−61.1/8.2	−46.0/8.3
$i-7$		−41.5/6.6	−60.3/8.2	−43.7/8.2
$i-6$	−62.9/5.9	−44.5/9.5	−59.8/8.0	−46.2/8.3
$i-5$	−65.7/7.7	−38.4/10.7	−60.6/7.9	−44.2/7.5
$i-4$	−62.9/6.7	−33.4/9.1	−63.9/8.2	−27.5/12.0
$i-3$	−78.3/13.1	−32.4/11.3	−81.2/12.6	−38.9/9.5
$i-2$	−72.1/14.3	−32.1/9.2	−64.3/9.2	−39.2/9.9
$i-1$	−59.3/9.2	−52.4/8.7	−57.5/8.7	−56.2/7.3
Pro	−58.5/7.7	−37.0/11.5	−63.1/9.5	−33.7/9.6
$i+1$	−63.7/7.3	−40.2/8.7	−62.4/7.8	−45.6/8.4
$i+2$	−65.3/6.3	−39.8/7.3	−62.0/8.2	−42.0/8.0
$i+3$	−64.7/6.5	−42.2/9.5	−61.4/7.5	−44.4/8.1
$i+4$	−63.1/7.3		−61.6/8.0	−43.8/8.2
$i+5$			−61.2/7.9	−44.0/7.8

The residues that constitute the Pro kink are highlighted.

CONCLUSIONS

The influence of the environment in the conformation of α-helices has been studied by surveying crystal structures of both soluble and membrane proteins and by molecular dynamics simulations of poly-Ala α-helices in water and methane. The results of both approaches show that polar environments tend to decrease ϕ_i and increase Ψ_i, relative to hydrophobic environments. Thus, there is a significant change in the conformation of the α-helix depending whether the peptide bond is exposed to bulk water or to the lipidic membrane. This effect is produced by an additional hydrogen bond between the peptide carbonyl oxygen to a water molecule (Blundell et al., 1983), which is not possible in membrane-embedded α-helices. Moreover, the participation of the carbonyl oxygen in the hydrogen bond with both the N—H group of the residue in the following turn of the helix and the water molecule increases the intramolecular N_i··O_{i-4} hydrogen bond distance and decreases the N_i··O_{i-4} = C_{i-4} angle. The fact that the intrahelical hydrogen bond in apolar environments is stronger suggests that α-helices in membrane-embedded proteins are more rigid than in soluble proteins. However, conformational changes in the transmembrane α-helices are necessary to transduce extracellular signals across the membrane (Sansom and Weinstein, 2000). Thus, membrane proteins incorporate in the sequence of their transmembrane helices specific residues like Pro, Gly, Ser, and Thr (Senes et al., 2000), which add flexibility and assist in the conformational change (Ballesteros et al., 2000; Gether et al., 1997; Govaerts et al., 2001a; Palczewski et al., 2000; Ri et al., 1999). Notably, in soluble proteins, these residues are mostly located in loop regions and acts as helix breaker (O'Neil and DeGrado, 1990).

Membrane proteins are particularly difficult to crystallize, yielding to only a few available structures (White and Wimley, 1999). Thus, molecular dynamics simulations are becoming a powerful tool to study the structure and dynamics of membrane proteins (Forrest and Sansom, 2000). We have observed a remarkable coincidence between the ϕ and Ψ angles obtained in the analysis of residues exposed to the lipid in membrane proteins and the results from computer simulations in methane. Thus, the simulation technique described here, where the membrane environment is replaced by explicit methane molecules, is a fast and reliable method that appears to reproduce several important characteristics of membrane-embedded proteins. Similar procedure has been recently used to mimic the membrane in molecular dynamics simulations of the potassium channel (Åqvist and Luzhkov, 2000). This approach is therefore well suited to study, in a reasonable amount of time, conformational arrangements and dynamic behavior of membrane proteins, and study the structural effects of specific mutations in their transmembrane domain (Govaerts et al., 2001b).

REFERENCES

Åqvist, J., and V. Luzhkov. 2000. Ion permeation mechanism of the potassium channel. *Nature.* 404:881–884.

Ballesteros, J. A., X. Deupi, M. Olivella, E. E. J. Haaksma, and L. Pardo. 2000. Serine and threonine residues bend α-helices in the $\chi1$ = g-conformation. *Biophys. J.* 79:2754–2760.

Barlow, D. J., and J. M. Thornton. 1988. Helix geometry in proteins. *J. Mol. Biol.* 201:601–619.

Blundell, T., D. Barlow, N. Borkakoti, and J. Thornton. 1983. Solvent-induced distortions and the curvature of alpha-helices. *Nature.* 306:281–283.

Case, D. A., D. A. Pearlman, J. W. Caldwell, T. E. Cheatham, III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Ferguson, R. J. Radmer, G. L. Seibel, U. C. Sing, P. K. Weiner, and P. A. Kollman. 1997. AMBER 5. University of California, San Francisco, CA.

Chakrabarti, P., and D. Pal. 1998. Main-chain conformational features at different conformations of the side-chains in proteins. *Prot. Eng.* 11:631–647.

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second-generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Darden, T. A., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

Fasman, G. D., editor. 1989. Prediction of Protein Structure and the Principles of Protein Conformation. Plenum Press, New York.

Forrest, L. R., and M. S. P. Sansom. 2000. Membrane simulations: bigger and better? *Curr. Opin. Struct. Biol.* 10:174–181.

Gether, U., S. Lin, P. Ghanouni, J. A. Ballesteros, H. Weinstein, and B. K. Kobilka. 1997. Agonists induce conformational changes in transmembrane domains III and VI of the β2 adrenergic receptor. *EMBO J.* 16:6737–6747.

Govaerts, C., C. Blanpain, X. Deupi, S. Ballet, J. A. Ballesteros, S. J. Wodak, G. Vassart, L. Pardo, and M. Parmentier. 2001a. The TxP motif in the second transmembrane helix of CCR5: a structural determinant in chemokine-induced activation. *J. Biol. Chem.* 276:13217–13225.

Govaerts, C., A. Lefort, S. Costagliola, S. Wodak, J. A. Ballesteros, L. Pardo, and G. Vassart. 2001b. A conserved Asn in transmembrane helix 7 is an on/off switch in the activation of the thyrotropin receptor. *J. Biol. Chem.* 276:22991–22999.

Hubbart, S. J., and J. M. Thornton. 1993. NACCESS. Department of Biochemistry and Molecular Biology, University College London.

Kumar, S., and M. Bansal. 1998. Geometrical and Sequence Characteristics of α-helices in globular proteins. *Biophys. J.* 75:1935–1944.

Milner-White, E. J., L. H. Bell, and P. H. Maccallum. 1992. Pyrrolidine ring puckering in *cis* and *trans*-proline residues in proteins and polypeptides. *J. Mol. Biol.* 228:725–724.

Monne, M., M. Hermansson, and G. von Heijne. 1999. A turn propensity scale for transmembrane helices. *J. Mol. Biol.* 288:141–145.

O'Neil, K. T., and W. F. DeGrado. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science.* 250:646–651.

Palczewski, K., T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. L. Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, and M. Miyano. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science.* 289:739–745.

Rees, D. C., and D. Eisenberg. 2000. Turning a reference inside-out: commentary on an article by Stenvens and Arkin entitled "Are membrane proteins 'inside-out' proteins?" *Proteins Struct. Funct. Genet.* 38:121–122.

Ri, Y., J. A. Ballesteros, C. K. Abrams, S. Oh, V. K. Verselis, H. Weinstein, and T. A. Bargiello. 1999. The role of a conserved proline residue in mediating conformational changes associated with voltage gating of Cx32 gap junctions. *Biophys. J.* 76:2887–2898.

Sankararamakrishnan, R., and S. Vishveshwara. 1992. Geometry of proline-containing alpha-helices in proteins. *Int. J. Pept. Protein Res.* 39:356–363.

Sansom, M. S. P., and H. Weinstein. 2000. Hinges, swivels and switches: the role of prolines in signalling via transmembrane α-helices. *Trends Pharmacol. Sci.* 21:445–451.

Senes, A., M. Gerstein, and D. M. Engelman. 2000. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β-branched residues at neighboring positions. *J. Mol. Biol.* 296:921–936.

Stevens, T. J., and I. T. Arkin. 1999. Are membrane proteins "inside-out" proteins? *Proteins Struct. Funct. Genet.* 36:135–143.

Von Heijne, G. 1991. Proline kinks in transmembrane alpha-helices. *J. Mol. Biol.* 218:499–503.

White, S. H., and W. C. Wimley. 1999. Membrane protein folding and stability: physical principles. *Ann. Rev. Biophys. Biomol. Struct.* 28:319–365.