Volume 277, number 1,2, 185–188 FEBS 09192 December 1990

The influence of proline residues on α-helical structure

Derek N. Woolfson[1,2] and Dudley H. Williams[1]

[1] *University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, UK* and [2] *Department of Biochemistry, Tennis Court Road, Cambridge, CB2 1QW, UK*

Received 5 October 1990; revised version received 15 October 1990

Proline lacks an amide proton when found within proteins. This precludes hydrogen bonding between it and hydrogen bond acceptors, and thus often restricts the residue to the first four positions of an α-helix. Helices with proline after position four have a pronounced kink [(1988) J. Mol. Biol. 203, 601–619]. In these cases, we find that the proline residue almost always occurs on the solvent exposed face of each helix. This positioning facilitates the compensatory hydrogen bonding between solvent and residues P–3 and P–4 (relative to proline, P), through the formation of the kink. Further, it aids in the packing of long helical structures around globular protein structures.

Amphipathic helix; Helix packing; Membrane spanning helix; Proline; Protein design

1. INTRODUCTION

Statistical analysis of protein sequences and structures used to discern rules that determine protein structure has met with some success in determining the amino acid preferences for α-helical structure. Early work in this area included the Chou and Fasman rules for secondary structure prediction [1]. An increase in the size of the protein database has led, more recently, to further refinement of the earlier rules. For example, two groups noticed a preference for hydrogen bond acceptor and donor sidechains at the N- and C-termini of helices [2,3]. The stabilisation which can be produced by such sidechain/mainchain hydrogen bonding has been verified by experiments on the stability of mutants of barnase lacking interactions [4]. Our concern has been with the way in which sequence affects the ability of a helix to pack against the rest of the protein.

It has been known for some time that proline is rarely found after the fourth residue from the N-terminus in α-helices [5]. This is because it is an imino acid, and therefore lacks an amide proton when at any position other than the N-terminus of a protein. As a consequence, proline residues cannot take part in hydrogen bonding to proton acceptors such as backbone carbonyl oxygen atoms. In an α-helix, amide protons of the *i*th residue hydrogen bond to the carbonyl oxygen of the *i*–4th. Residues which lack this proton will disrupt this hydrogen bonding pattern and, presumably, destabilise the helix. As the protein database grows, more helices appear with proline residues after the fourth residue in their sequence. Further, many putative membrane spanning helices of ion channel peptides have a proline residue within their sequence. In contrast, similar helices in non-transport proteins seem to be devoid of such features in their primary structures [6]. For these reasons it appeared interesting to determine whether there is any common theme in the positioning and structural effects of proline residues within helices. This paper reports the results of such a study carried out on 15 helices in 13 proteins of known structure.

Correspondence address: D.H. Williams, University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, UK

2. MATERIALS AND METHODS

Protein structures were visualised on an Evans and Sutherland PS390 system with the aid of the Macromodel software package [7]. Atomic solvent accessibilities were calculated using a spherical probe of radius 1.4 Å, after Chothia [8]. Helical hydrophobic moments [9] were calculated by a program using one of two hydropathy scales. Either that of Janin [10], which is based on the relative occurrence of each amino acid at the surface or in the interior of proteins, or one derived from the scales of Roseman [11], which are based on transfer data of organic molecules, that are chemically similar to the amino acid sidechains, from aqueous to alcoholic solutions. The calculations assumed that all α-helices were regular structures with 3.6 residues per turn. Components of the hydrophobic moment from consecutive amino acids were, therefore, separated by 100°.

3. RESULTS AND DISCUSSION

Fig. 1A shows the structure of the A-helix from citrate synthase; the helix is highly kinked. The 11–15 *CO* to N*H* hydrogen bond is lost due to the presence of proline at residue 15 in the helix (numbers refer to residue positions in the overall sequence). The 12– 16 hydrogen bond is also lost, although residue 16 is lysine and not proline. In Fig. 1B the sequence of the same helix is plotted on a Schiffer and Edmundson helical wheel [12]. The helix is clearly amphipathic; that is, it

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

has a large number of its hydrophobic residues on one face and most of its hydrophilic residues on the other face of the structure. To emphasise this point the hydrophobic moment of the helix is also drawn on the diagram. The striking feature of this diagram is that the *hydrophobic* proline residue sits on the solvent exposed face of the helix (the hydrophobic moment does, in fact, point directly away from this amino acid). To discover whether or not this is a general feature of helices kinked by proline, we calculated helical hydrophobic moments for the 15 helices in the study. These results are plotted in Fig. 2A and B.

The plots are helical wheel representations of the helices with all 15 proline residues superimposed. The two plots shown (Fig. 2A,B) use hydrophobic moments calculated using the two different scales of hydrophobicity. In both plots the same general trend appears; the proline residues appear by far most fre-

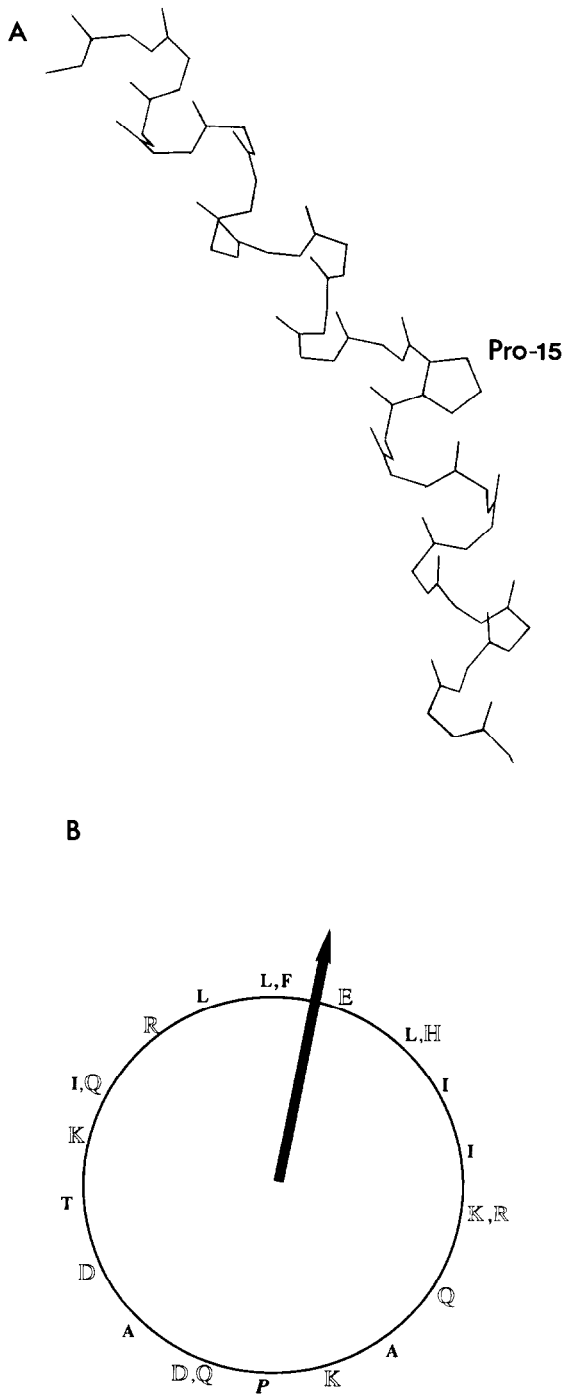

Fig. 1. (A) The backbone fold of the A-helix (residues 5 through 29) of citrate synthase. The helix is clearly kinked due to the presence of Pro-15 (position 11 in the helix). (B) The sequence of the helix plotted as a helical wheel. This representation shows the amphipathic nature of the helix which is emphasised by the hydrophobic moment of the helix also included in the plot. Note that the moment points directly away from the proline residue which lies on the hydrophilic side of the helix.

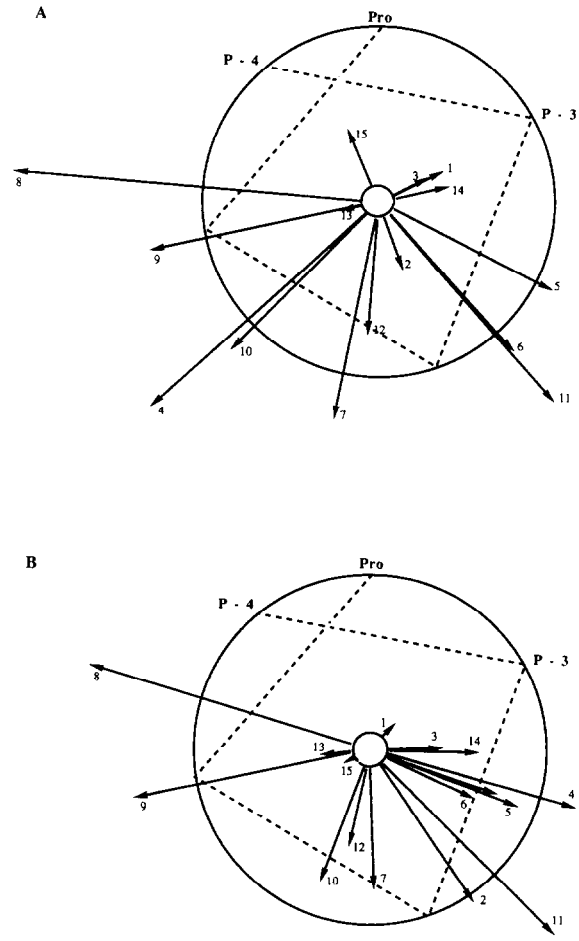

Fig. 2. The hydrophobic moments of 15 helices kinked by proline. The proline residues of each helix are superimposed. In (A) hydrophobic moments were calculated using a scale based on the data of Roseman [12], while those shown in (B) were calculated using the hydropathy scale of Janin [11]. The helices included are: 1, haemoglobin residues 13–30 (2LHB); 2, adenylate kinase 141–168 (3ADK); 3, oxidised glutathione reductase 55–80 (3GRS); 4, melittin 1–26 (1MLT); 5, glyceraldehyde-3-P-dehydrogenase 148–166 (3GPD); 6, cytochrome P450 192–214 (2CPP); 7,8,9, citrate synthase 5–29, 167–195, 344–365 (3CTS); 10, phosphoglycerate kinase 37–51 (3PGK); 11, cytochrome *c* peroxidase 84–99 (2CYP); 12, proteinase II 212–234 (3RP2); 13, subtilisin 220–238 (1CSE); 14, proteinase K 220–238 (2PRK); 15, catalase 347–366 (8CAT).

quently on the opposite side of the helix to which the hydrophobic moment is pointing. As a consequence of helix geometry residues P-4 and P-3 of the helix also lie on the solvent exposed face of the helix. This finding can be further tested by calculating the solvent accessibilities of the backbone atoms of the helix. The average results of calculations for the carbonyl oxygen atoms of the helices used in the study are shown in Fig. 3B. The formation of an α-helix usually offers good protection of these atoms from solvent ([13], Woolfson and Williams, unpublished data), as is shown in Fig. 3A. The increase in accessibility for the last four residues of the helix is due to them being the C-cap residues and so not being hydrogen bonded within the helix itself. For residues P-4, and to a lesser extent P-3,

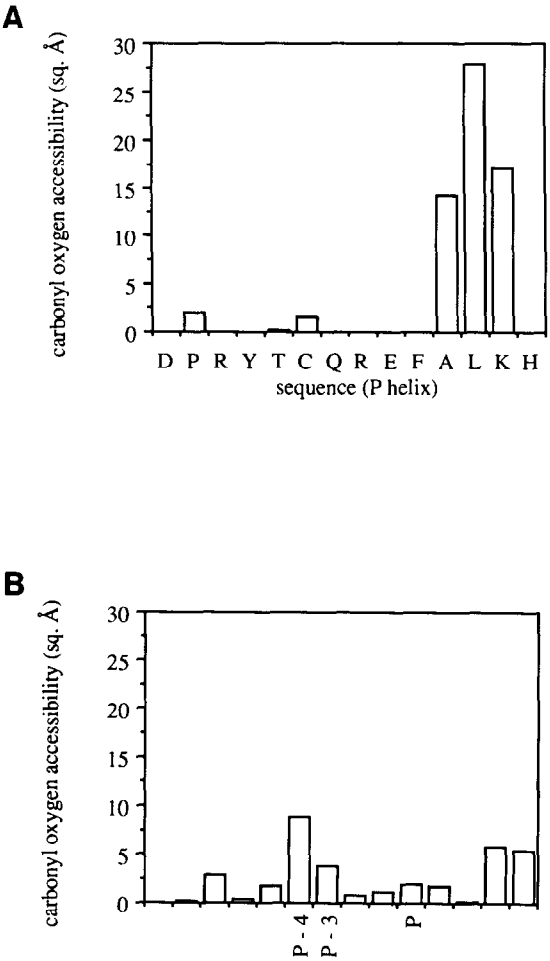

Fig. 3. Plots of solvent accessibility of the carbonyl oxygen atom against residue. (A) The plot for the P-helix (residues 327-341) of citrate synthase demonstrating the protection from solvent afforded by α-helical structure. (B) A similar averaged plot for the nine proline kinked helices (1,2,3,4,6,7,9,11,15), listed in the legend to Fig. 2. Again good protection from solvent is observed; however, carbonyl oxygen atoms of P-4 and P-3 have much increased accessibilities to solvent. High accessibilities of the carbonyl oxygen atoms at the C-termini of the helices arise because C-cap residues of helices are not hydrogen bonded within the helix, and are thus less protected from solvent.

relative to proline in a helix the accessibility to solvent is also high (Fig. 3B). Thus, good hydrogen bonding between solvent and these carbonyl groups is to be expected.

In light of these findings, it is interesting to speculate as to the role of proline residues found in the sequences of some membrane spanning helices. Such sequences of transport proteins and peptide ion channels often contain proline residues; those of non-transport proteins are often devoid of this residue. Our findings agree with the hypotheses of Vogel and Jahnig, Bazzo *et al.* and Fox and Richards concerning the structures of the gated ion-channels of melittin and alamethicin [14-16]. In these models, the hydrophilic sides of helices come together to form a pore. The proline residue, being on this side of each helix, helps kink the helices and form a funnel-shaped pore, i.e. a pore with a constriction along its length.

In conclusion, proline occurring after the fourth residue of an amphiphilic α-helix is almost always positioned on the solvent-exposed face of the structure. We see the reasons for this as 2-fold. (i) Loss of the P to P-4 N*H* to C*O* hydrogen bond needs to be compensated for energetically. This is achieved by kinking and allowing in water or hydrogen bond acceptor sidechains (in the case of buried helices), to fulfil the hydrogen bonding role. (ii) As a consequences of this, and as indicated in Fig. 1, kinking is always away from the side with the proline residue. Thus, surface helices always kink *around* the core of a protein, i.e. a proline residue within a helix appears to be used to accentuate the curvature often observed in helices lying on the surface of a protein [17,18]. Proline kinking of long helices may have this additional function to ensure good packing of the structures with the protein.

Loss of the hydrogen bond from the C*O* of residue P-3 to the N*H* of residue P + 1, in some of the helices studied, cannot straightforwardly be accounted for in physicochemical terms. It appears likely that it is a knock effect: (i) allowing better access of solvent to the carbonyl group of residue P-4; and/or (ii) allowing better packing of the helix against the core of the protein molecule.

The observation reported here may by of use to those designing α-helical structure into proteins, to the development of structure prediction algorithms and in developing an understanding of gating mechanisms of peptide ion channels.

Acknowledgements: We thank P.A. Evans, J.M. Thornton and R.J. Mortishire-Smith for helpful discussions. We are grateful to R.J.M.-S. for much help in writing a program to calculate hydrophobic moments, M.W. MacArthur and J.M.T. for providing a recent update of a database search on proline containing helices and M. Swindell and J.M.T. for providing the solvent accessibility files for the proteins used in Fig. 3. We acknowledge financial support from Pfizer Central Research (UK), and from the SERC (UK). D.W.W. is a member of the Cambridge Centre for Molecular Recognition.

REFERENCES

[1] Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 211-222.
[2] Presta, L.G. and Rose, G.D. (1988) Science 240, 1632-1641.
[3] Richardson, J.S. and Richardson, D.C. (1988) Science, 1648-1652.
[4] Serrano, L. and Fersht, A.R. (1989) Nature 342, 296-299.
[5] Madison, V. (1977) Biopolymers 16, 2671-2692.
[6] Brandl, C.J. and Deber, C.M. (1986) Proc. Natl. Acad. Sci. USA 83, 917-921.
[7] Mohamadi, F., Richards, N.G.J., Guida, W.C., Liskamp, R., Lipton, M., Caufields, C., Chang, G., Hendrickson, T. and Still, W.C. (1990) J. Comp. Chem. 11, 440-467.
[8] Chothia, C.H. (1975) Nature 254, 304-308.
[9] Eisenberg, D., Weiss, R.M. and Terwilliger, T.C. (1982) Nature 299, 371-374.
[10] Janin, J. (1979) Nature 277, 491-492.
[11] Roseman, M.A. (1988) J. Mol. Biol. 200, 513- 522.
[12] Schiffer, M. and Edmundson, A.B. (1967) Biophys. J. 7, 121-135.
[13] Lee, B.K. and Richards, F.M. (1971) J. Mol. Biol. 55, 379-400.
[14] Vogel, H. and Jahnig, F. (1986) Biophys. J. 50, 573-582.
[15] Bazzo, R., Tappin, M.J., Pastore, A., Harvey, T.S., Carver, J.A. and Campbell, I.D. (1988) Eur. J. Biochem. 173, 139-146.
[16] Fox, R.O. and Richards, F.M. (1982) Nature 300, 325-330.
[17] Blundell, T.L., Barlow, D.J., Borkakoti, N. and Thornton, J.M. (1983) Nature 306, 281-283.
[18] Barlow, D.J. and Thornton, J.M. (1988) J. Mol. Biol. 203, 601-619.