Variants of 3_10-Helices in Proteins

Lipika Pal,1 Gautam Basu2* and Pinak Chakrabarti1*
1Department of Biochemistry, Bose Institute, Calcutta, India
2Department of Biophysics, Bose Institute, Calcutta, India

ABSTRACT An analysis of the shortest 3_10-helices, containing three helical residues and two flanking capping residues that participate in two consecutive i + 3 → i hydrogen bonds, shows that not all helices belong to the classic 3_10-helix, where the three central residues adopt the right-handed helical conformation (αR). Three variants identified are: 3_10-helix with all residues in the left-handed helical region (αL), 3_10-helix where the first residue is in the extended region followed by two residues in the αL conformation, and its mirror-image, the 3_10L-helix. In the context of these helices, as well as the equivalent variants of α-helices, the length dependence of the handedness of secondary structures in protein structure is discussed. There are considerable differences in the amino acid preferences at different positions in the various types of 3_10-helices. Each type of 3_10-helix can be thought to be made up of an extension of a particular type of β-turn (made up of residues i to i + 3) such that the (i + 3)th residue assumes the same conformation as the preceding residue. Distinct residue preferences at i and i + 3 positions seem to decide whether a particular stretch of four residues will be a β-turn or a 3_10-helix in the folded structure. Proteins 2002;48:571–579. © 2002 Wiley-Liss, Inc.

Key words: 3_10-helix; β-turn; conformation; handedness in protein structures; sequence preference

INTRODUCTION

The ideal 3_10-helix is comprised of three residues per helical turn, with the backbone carbonyl oxygen of residue i hydrogen bonded to the amide proton of residue i + 3, forming a 10-atom ring.1 3_10-Helical structures are relatively common in proteins comprising as much as 4% of all residues.2–5 Although typical protein 3_10-helices are only three or four residues long, the occurrence of longer 3_10-helices is not uncommon.6 The position-specific occurrence of different residues that can provide stability to these structures has been determined.2,7,8

As secondary structural elements, turns and helices are separate entities; however, it should be noted that occurring in isolation, the first three residues (i + 1 through i + 3) of an α-helix with a single 5 → 1 (i + 4 to i) hydrogen bond is basically an α-turn.9 Two such consecutive α-turns give rise to the shortest α-helix of four residues (i + 1 through i + 4). Similarly, two consecutive type-III β-turns, with two consecutive 4 → 1 hydrogen bonds, with the first β-turn defined by residues i to i + 3 and the second composed of residues i + 1 to i + 4, form the shortest 3_10-helix (i + 1 through i + 3), as shown in Figure 1. However, no helical structure can be constructed from the repetitive extension of other types of β-turns, primarily because the non-identical ϕ, ψ angles at the two central residues do not allow the extension of helical hydrogen bonding network. In the present work, we show that some types of β-turns (notably II and II’) can be part of variants of 3_10-helices. Analysis of the residues in these structures need not necessarily be stabilized by intra-turn hydrogen bonds.14 β-Turns have been classified into nine different types (I, I’, II, II’, V叙述, VIa1, VIa2, Vb, VIII, and IV) based on the values of the dihedral angles ϕ, ψ of the (i + 1)th and (i + 2)th positions in the turn.3,15 It is convenient to distinguish another category, viz., type III (and the main-chain mirror-image, type III’) occupying very similar regions of ϕ, ψ space as those of type-I β-turns.11 Overlapping type III β-turns give rise to the ideal 3_10-helix (ϕ = −60°, ψ = −30°). Besides the four-residue β-turn, one can also have short loops involving three or five residues (with a 3 → 1 or a 5 → 1 hydrogen bond between the carbonyl oxygen at position i and amide hydrogen at positions i + 2 or i + 4, respectively). The former constitutes a γ-turn16–18 and the latter an α-turn.19 Moreover, the conformational features of a π-turn involving a loop of six residues and usually following α-helices have been delineated.20

Grammar check or proofreading for this text is unnecessary as the content is already grammatically correct and professionally written. The text provides a comprehensive analysis of the shortest 3_10-helices in proteins, discussing their structures, characteristics, and implications for protein folding and secondary structure prediction.
shows that these have features different from classic 3₁₀-helices, as well as β-turns from which they are made. Distinct preferences are also observed for the closest secondary structures from the two ends. All of this information can be used for modeling and refining prediction of secondary structural elements in proteins.

**RESULTS**

**Identification of Variants of 3₁₀-Helices From the Distribution of \( \phi, \psi \) Angles**

The cullpdb data set consists of 1085 protein chains of which 3.8% (10,086) residues were found to exhibit the 3₁₀-helical conformation. Of these, approximately 70% were part of three-length 3₁₀-helices. We have been interested in the pattern of residue distribution in shortest possible helices, viz., three-length 3₁₀-helices. The backbone \( \phi, \psi \) angles of residues in such structures identified by the commonly used program DSSP\(^{33} \) are shown in Figure 2(a). Although deviations from average \( \phi, \psi \) angles of \(-65^\circ, -20^\circ\) were noted earlier when 3₁₀-helices of various lengths were taken together,\(^{34} \) by restricting to only three-length helices the angles are found to cluster in four distinct regions, designated as: \( \alpha_R \) corresponding to the right-handed \( \alpha \)-helix; \( \alpha_L \), left-handed helix; \( E \), extended conformation and a region (allowed for Gly residues, but normally disallowed for non-Gly residues) \( E' \) related by a center of inversion. The average values of the angles in all the regions are given in Table I. On scrutiny, it was observed that the residues occur in these regions in specific order and this was used to classify these 3₁₀-helices (Table II). With all three residues in the \( \alpha_R \) conformation, one obtains the classic 3₁₀-helix, which we call \( 3_{10}^{\alpha_R} \), which is the major class. As pointed out in Table I, further analysis of the \( \alpha_R \) region showed that the average backbone \( \phi \) angle decreased \((-57^\circ, -67^\circ, \) and \(-95^\circ\) whereas the backbone \( \psi \) angle increased \((-32^\circ, -16^\circ, \) and \(0^\circ\) as the residue position is changed from 1 through 3 in the helix. The \( 3_{10}^{\alpha_R} \) helix is thus not the ideal 3₁₀-helix composed of two overlapping type-III \( \beta \)-turns \((\phi = -60^\circ, \psi = -30^\circ)\), rather it can be considered to be formed from overlapping type III/I \( \beta \)-turns. A small number of helices (designated as \( 3_{10}^{\alpha_R} \)-helices) are found with the opposite hand, where the residues have the \( \alpha_L \) conformation. Then there are two mirror-image related categories, \( 3_{10}^{EL} \) and \( 3_{10}^{ER} \), of nearly equal occupancy and with conformations, which can be designated as \( E_{\alpha_L, \alpha_L} \) and \( E'_{\alpha_R, \alpha_R} \), respectively.

It may be noted that the two central residues in type-II and -II’ \( \beta \)-turns have conformations \( E_{\alpha_L, \alpha_L} \) and \( E'_{\alpha_R, \alpha_R} \), respectively.\(^{15} \) An addition of an \( \alpha_L \) residue to II and \( \alpha_R \) to II’ \( \beta \)-turn leads to \( 3_{10}^{EL} \) and \( 3_{10}^{ER} \)-helices, respectively. It may be asked to what extent the addition of more residues with \( \alpha_L \) or \( \alpha_R \) conformation (as the case may be) can be sustained. Table II shows that there is a drastic reduction...
in the number of cases as the \(3_{10}\)-helix gets longer with contiguous \(3_1\) residues. A similar dependence on length can be seen when one compares the variant four-length \(3_1\)-helices [Fig. 2(b)] with those from the \(3_{10}\)-helix categories. \(\alpha\)-H helices are found in smaller numbers and the numbers get insignificant as the length increases for cases with contiguous \(3_1\) residues. To find out the secondary structure type exhibited by other sequences with \(E\), \(\alpha\)-helix, Gly is more prominent at position 3. Though some residues have reasonably large propensity values, no residue is over-represented (\(z > 1.96\)) are plotted in Figure 3 along with the values observed in \(3_{10}\)-helices. In conformity with what has been observed for \(3_{10}\)-helices in general, \(15\%\) was present as GGG, \(19\%\) was present as HHH, whereas \(25\%\) showed a stretch of two turn (T) conformations.

**Positional Potentials in Different Three-Length \(3_{10}\)-Helices**

We have calculated the propensity and z-value (see Materials and Methods) of 20 amino acid residues to occupy each position in \(3_{10}\)-helices (Table III). To facilitate comparison, the propensities of residues that are over-represented (\(z \geq 1.96\)) are plotted in Figure 3 along with the values observed in \(3_{10}\)-helices. In conformity with what has been observed for \(3_{10}\)-helices in general, \(15\%\) was present as GGG, \(19\%\) was present as HHH, whereas \(25\%\) showed a stretch of two turn (T) conformations.

**TABLE I. Average Backbone Torsion Angles (°) of the Shortest Protein Helices**

<table>
<thead>
<tr>
<th>Region (a)</th>
<th>(\alpha_E)</th>
<th>(\alpha_L)</th>
<th>(E)</th>
<th>(E')</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
<td></td>
</tr>
<tr>
<td>(3_{10})</td>
<td>(-73 (30))</td>
<td>(-15 (31))</td>
<td>(-70 (18))</td>
<td>(-32 (24))</td>
</tr>
<tr>
<td>(3)</td>
<td>(65 (15))</td>
<td>(22 (15))</td>
<td>(61 (13))</td>
<td>(35 (25))</td>
</tr>
<tr>
<td>(4)</td>
<td>(-61 (15))</td>
<td>(135 (15))</td>
<td>(-87 (33))</td>
<td>(136 (18))</td>
</tr>
<tr>
<td></td>
<td>(56 (13))</td>
<td>(-132 (14))</td>
<td>(49 (24))</td>
<td>(-137 (30))</td>
</tr>
</tbody>
</table>

\(a\) Regions are defined in Figure 2.
\(b\) Standard deviations are in parentheses.

*Position specific (from the N-termini) torsion angle values for the \(\alpha_E\) region are: 1: \(\phi = -57 (9), \psi = -32 (12)\); 2: \(\phi = -67 (9), \psi = -16 (12)\); 3: \(\phi = -95 (16), \psi = 0 (6)\).

TABLE II. Angle (°) of \(R\) and \(L\) Residues in Different Three-Length \(3_{10}\)-Helices

<table>
<thead>
<tr>
<th>Region (a)</th>
<th>(\alpha_E)</th>
<th>(\alpha_L)</th>
<th>(E)</th>
<th>(E')</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
<td></td>
</tr>
<tr>
<td>(3_{10})</td>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
</tr>
<tr>
<td>(3)</td>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
</tr>
<tr>
<td>(4)</td>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
</tr>
<tr>
<td></td>
<td>(\phi)</td>
<td>(\psi)</td>
<td>(\phi)</td>
<td>(\psi)</td>
</tr>
</tbody>
</table>

\(a\) Regions are defined in Figure 2.

The number of \(3_{10}\)-helices is very low (18 cases) for doing any statistical analysis. In 18 cases of \(3_{10}\)-helices, Ne is occupied by Asp, Asn, or Gly (eight cases), and Lys is found at Cc in four instances. In the total of 54 positions in the main body of the helices, Gly is found in 24 positions, followed by Asn (six), Phe (three), and Trp (three). Two extreme situations are found in the PDB file ImroB where all helical positions are occupied by Gly (residue range
106–110: Leu-Gly-Gly-Lys) and 2bbkH where no Gly is involved (39–43: Pro-Ala-His-Phe-Ala).

Comparison of Potentials for the Formation of 3\textsubscript{10}-Helix and \(\beta\)-Turn

As shown in Figure 1, an additional hydrogen bond attributable to the placement of an extra residue in the right conformation extends a \(\beta\)-turn into a 3\textsubscript{10}-helix and it would be of interest to see if one can distinguish, based on the types of residues occurring at different positions, whether a given stretch of the polypeptide chain would form an isolated turn or get extended into a 3\textsubscript{10}-helix. Residues with high propensity values at different positions in various 3\textsubscript{10}-helices (Table III) are compared with the prominent residues at equivalent positions of \(\beta\)-turns, the positional residue potentials of which have been calculated by Hutchinson and Thornton\textsuperscript{15} (Table IV). The structures to be compared are 3\textsubscript{10}-helix and type-I \(\beta\)-turn (ideally one should use type-III \(\beta\)-turn, but because of paucity of data, potentials are not available and instead, we have used type-I \(\beta\)-turns, which are conformationally very similar; in addition we showed earlier that 3\textsubscript{10} helices are formed from overlapping type III/\(I\) \(\beta\)-turns), 3\textsubscript{10}-helix and type-II \(\beta\)-turn, and 3\textsuperscript{\textsubscript{ER}} helix and II’ \(\beta\)-turn. It can be seen that the residue preferences at position 3 of the helices are entirely different from the equivalent i + 3 positions in the turn, and the same is true even for the Nc/i position (except for the first pair). Thus, the two central residues at positions i + 1 and i + 2 determine the type of \(\beta\)-turn, but whether it will remain a \(\beta\)-turn in the final structure or extend to a 3\textsubscript{10}-helix is controlled by the two adjacent residues.

Secondary Structures on Either Side of 3\textsubscript{10}-Helix Variants

To see whether 3\textsubscript{10}-helix variants have any preference to be positional between common secondary structures (helix or \(\beta\)-strand), we looked at the structural features of the Nc and Cc positions in 91 3\textsubscript{EL} and 117 3\textsuperscript{ER} helices. For 34 cases (37\%) of 3\textsubscript{10} helices and 16 cases (14\%) of 3\textsuperscript{ER} helices, both the positions are present in C-conformation, as described in Materials and Methods. A count of the number of times a \(\beta\)-strand (E) or H (including G) occurs as the nearest secondary structure from the 3\textsubscript{10}-helix shows that for the 3\textsubscript{EL} helix, E occurs 62 times (36 times within three residues) before and 62 times (42 times within three residues) after the helix, whereas the corresponding numbers for H are 27 (five times within three residues) and 29 (six times within three residues), respectively. This shows that a 3\textsuperscript{ER}-helix is more likely to have a strand rather than a helix as a closest neighbor. For 3\textsuperscript{ER}-helices, the numbers of E and H are nearly equal on both the sides (E: 52(33), 47(31); H: 44(33), 57(36)).

When simultaneous occurrence of an E or H/G conformation on both sides of the helices were analyzed (Fig. 4), it was found that for 3\textsubscript{EL}-helices, there are 15 cases [Fig. 4(a)] with an E occurring before the Nc position and an E/H/G occurring after the Cc position (of these, 13 cases have an E following Cc). For five cases [Fig. 4(b)], a stretch of E ends at N1’ position and an E/G stretch starts after the C1’ position (of five cases, four cases have an E after C1’).

Sequence neighbors of 3\textsuperscript{ER}-helices showed a variety of secondary structure combinations: 19 cases were found [Fig. 4(c)] with an \(\alpha\)-helix starting at the Cc position whereas a stretch of E or H ends anywhere within four positions before the helix; eight cases were found [Fig. 4(d)] with an \(\alpha\)-helix leading directly into a 3\textsuperscript{ER}-helix, which, after a gap of two residues with a nonregular structure, leads into a E/H/G stretch (H stretch for six cases). When a stretch of E ends at the N1’ position of 3\textsuperscript{ER}-helices, two more prominent combinations are observed after the helix: one combination has six cases [Fig. 4(e)] with E/H stretch starting from the Cc position (H stretch for five cases), whereas the other combination has five cases [Fig. 4(f)] with an E stretch occurring after the Cc position as in 3\textsuperscript{ER}-helices [Fig. 4(a)]. Finally, four cases [Fig. 4(g)] consisted of an \(\alpha\)-helix ending at the Nc position and a \(\beta\)-strand starting from the Cc position.

Capping Hydrogen Bond Interaction

We investigated whether the NH groups at Nc, 1 and 2 positions and the CO groups at 2, 3, and Cc positions are involved in any local (within five sequence contiguous
residues) hydrogen bonding. Only the groups at Nc and Cc were found participating in substantial numbers (about 44 and 30%, respectively) in both the 3_{10}^{EL} and 3_{10}^{ER}-helices. In 3_{10}^{EL}-helices (91 cases), the largest group has the NH of Nc and CO of Cc hydrogen bonded to each other (19 cases). Of these, 11 cases correspond to Fig. 4(a) with a stretch of E ending at N1’ and another starting from C1’. Among the rest, cases in which H or E stretch is present within four residues on both sides of the helix have: 1. E at Nc with E at Cc, 2. E at Nc with E at C1’, 3. E at N1’ with E at Cc, 4. E at N1’ with E at C3’, and 5. E at Nc with H at C3’. In the majority of these cases, the 3_{10}^{ER}-helices are part of a β-hairpin turn, a representative example of which is shown in Figure 5. Five-residue β-hairpin turns with a single hydrogen bond between the anchoring β-strand residues (in our case Nc and Cc) have been classified as 3.5 β-hairpin turns. 3^{ER} From an analysis of 106 β-hairpin turns in proteins, Sibanda et al. 3^{ER} had identified G1 β-bulge as a prominent subclass of 3.5 β-hairpin turns. The 3_{10}^{ER}-helix containing 3.5 β-hairpin turn, as identified here, can be considered to be a new subclass of 3.5 β-hairpin turn (conformation of the three central residues in G1 β-bulge is: α_{2}Ω_{2}Ω_{1}Ω_{3}, whereas in 3_{10}^{ER}-helix containing 3.5 β-hairpin turns it is Eα_{2}Ω_{1}). Among four cases of 3_{10}^{ER}-helices [Fig. 4(b)], when a β-strand ends at N1’ position and the other β-strand starts from C2’ position, three cases have hydro-

**TABLE III. Propensities and z-Values of Residues at Different Positions of Three-Length 3_{10}^{EL}- and 3_{10}^{ER}-Helices**

<table>
<thead>
<tr>
<th>Residues</th>
<th>Nc</th>
<th>1</th>
<th>Cc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>7</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Cys</td>
<td>6</td>
<td>4.6</td>
<td>4.2</td>
</tr>
<tr>
<td>Asp</td>
<td>5</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Glu</td>
<td>2</td>
<td>0.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Phe</td>
<td>2</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Gly</td>
<td>21</td>
<td>3.0</td>
<td>5.5</td>
</tr>
<tr>
<td>His</td>
<td>3</td>
<td>1.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Ile</td>
<td>3</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Lys</td>
<td>2</td>
<td>0.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Leu</td>
<td>4</td>
<td>0.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Met</td>
<td>1</td>
<td>0.5</td>
<td>0.7</td>
</tr>
<tr>
<td>Asn</td>
<td>7</td>
<td>1.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Pro</td>
<td>2</td>
<td>0.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Gln</td>
<td>2</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Arg</td>
<td>1</td>
<td>0.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Ser</td>
<td>8</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Thr</td>
<td>9</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Val</td>
<td>4</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Trp</td>
<td>2</td>
<td>1.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Tyr</td>
<td>0</td>
<td>0.0</td>
<td>1.8</td>
</tr>
</tbody>
</table>

*3_{10}^{ER} and 3_{10}^{ER} helices are defined in Table II. P stands for propensity and Z for z-values of amino acid residues. P > 1.3 and Z > 1.96 are given in bold.*
gen bonding between NH of Nc and CO of C1'. This type belongs to the class of 4:6 β-hairpin in accordance to the convention of Sibanda et al. In the case of 3\textsuperscript{ER}-helices, when a β-strand ends at N1' position and another β-strand starts from C1' position [Fig. 4(d)], three of five cases have a 3.5 β-hairpin-like hydrogen bonding pattern.

**DISCUSSION**

**Classification of 3\textsubscript{10}-Helices**

Under IUPAC convention rule 6.3,\textsuperscript{37} a residue is considered to be the part of an α-helix (or β-sheet) if either its NH or CO groups are involved in the appropriate hydrogen bond. It does not imply that the φ, ψ angles for this residue should be in the appropriate region for α-helix (or β-sheet). Based on the occurrence of two consecutive i + 3 to i hydrogen bonds (Fig. 1) we have identified all cases of three-length 3\textsubscript{10}-helices from the PDB. The φ, ψ distribution (Fig. 2) shows that not all constituent residues have their torsion angles correspond to the canonical conformation in the α\textsubscript{R} region of the Ramachandran plot. Based on the φ, ψ angles, three variants of the classic 3\textsubscript{10}-helices can be identified (Table II and Fig. 6). In 3\textsubscript{10}-helices, all the residues have α\textsubscript{L} conformation. The 3\textsuperscript{EL}-helices have the first residue in the extended conformation, followed by two residues with α\textsubscript{L}-conformation. The last category is the mirror-image related 3\textsubscript{10}-helix, where the conformation of the three residues can be specified as E'α\textsubscript{R}α\textsubscript{L}. These helices can also be considered as the extension of two-residue β-turns (corresponding to a stretch of four residues, i to i + 3, where the φ, ψ angles of the two central residues specify the turn type), such that the (i + 3) residue has the same conformation as the (i + 2) residue. Thus, the type-III β-turn leads to the classic 3\textsubscript{10}-helix (the more commonly observed 3\textsubscript{10}-helix is initiated by a type-III β-turn followed by a type-I turn), type II to 3\textsuperscript{ER}-helices (Table IV).

**Handedness of Secondary Structures**

Most structural elements in protein structures, for which handedness can be defined, prefer one enantiomer over the other. Handedness manifests itself in the right-handedness of α-helix, the twist of β-sheet, and the topology of connectivity in parallel β-sheet.\textsuperscript{3} Nevertheless, mirror images are observed, although at much lower proportions in short fragments, such as I and I', II and II' β-turns. The ratios of the numbers observed in the two categories are 8:1 and 5:5:1, respectively.\textsuperscript{31} The 3\textsubscript{10}- helices can be grouped into two pairs of structures of opposite hands: 3\textsuperscript{ER} and 3\textsuperscript{EL}, 3\textsubscript{10}\textsuperscript{L} and 3\textsubscript{10}\textsuperscript{R}. For three-length helices, the ratios of the numbers observed are 119:1 and 0.78:1, respectively (Table II). Thus, all pairs of structures of opposite handedness, 3\textsuperscript{ER} and 3\textsubscript{10}\textsuperscript{L}, and 3\textsubscript{ER} and 3\textsubscript{10}\textsuperscript{R} helices seem to be of comparable occurrence. However, as the length is increased, the number of helices containing α\textsubscript{L} conformation is drastically reduced, indicating that it is difficult to have a structure with more than two residues with α\textsubscript{L}.
conformation. The same conclusion can be arrived at considering α-helices also. Whereas there has been just one reported example of a left-handed α-helix,38 our survey revealed three cases of four-length α-helices, with an additional nine having the conformation E_{L}, E_{L}, E_{L}, L. However, as the length is increased beyond four, we are left with just one example of E_{L}, E_{L}, E_{L}, E_{L}.

Disallowed Conformation as Part of Regular Secondary Structure

The occurrence of a non-Gly residue in the disallowed region of the Ramachandran plot may be a source of concern while validating a newly determined structure.39 The conformation E′ (Table II, Fig. 2) found in the 3_{10}^{E/R}-helix has been identified as the region II in the disallowed region of the Ramachandran plot.40 Although the occurrence of such conformations is normally indicative of the presence of local strains, it is likely that when they form a part of a larger secondary structure identified here, such conformations can be accepted with a higher degree of confidence.

Potentials for 3_{10}-Helix Variants and Their Differences From β-Turns

Different types of 3_{10}-helices are also quite distinct from one another on the basis of amino acid usage, and many of the residues are selected at a given position because they are endowed with the ϕ, ψ angles required at the position. Thus, positions 2 and 3 with α_{L} conformation in the 3_{10}^{EL}-helix are predominantly occupied by Gly, Asn, and Asp, the preference of Gly being particularly strong at 3 (Fig. 3, Table III). Gly is also the highly favored residue at 1 in 3_{10}^{E/R}-helices, where an E′ conformation is normally not allowed for a non-Gly residue. Pro has a high propensity value for this position in 3_{10}^{R} and 3_{10}^{EL}-helices. Two residues with the highest propensity values at the Nc position are Asp and Pro in the 3_{10}^{R}-helix, Cys and Gly in the 3_{10}^{EL}-helix, and Thr and Leu in the 3_{10}^{E/R}-helix. In general, this position is usually hydrophobic for 3_{10}^{R}-helices and hydrophilic for 3_{10}^{E/R}-helices. The Cc position is generally hydrophobic for 3_{10}^{R} and 3_{10}^{EL}-helices (the latter also has Gly), whereas the propensities are quite high for Ala, Ser, and Gly to be at this position in 3_{10}^{E/R}-helices.

The three-length 3_{10}-helices are a single-residue extension of a particular type of β-turn (specified by the ϕ, ψ angles of the two central residues in a segment of i to i + 3 residues). For instance, type-I (more precisely type-III, of

---

**Fig. 4.** Schematic diagram showing nearest neighbor residues in helical (G/H) or β-strand (E) conformation at both ends of 3_{10}^{EL} and 3_{10}^{E/R}-helices (see Materials and Methods for the designation of all secondary structural states). Comments: *Of which 13 cases have only E after the helix; in two cases, Nc position is S, and in one case the Cc position is B (instead of C). Only E in four cases after the helix; Nc position is S in one case, and in two cases, the Cc position is S (instead of C). The nearest E or H residue is anywhere within four positions before the helix; E occurs five times at N1′ position. Two cases have C at Cc position. Five cases have only H after the helix. Two cases have only T and one case has B at the Cc position.

**Fig. 5.** MOLSCRIPT stereo representation of a β-hairpin structure (residues 147–161 in labrB) containing a 3_{10}^{EL}-helix (residues 152–154). The 3_{10}^{EL}-helix is represented by ball-and-stick model. The usual 3_{10}^{EL}-type hydrogen bonds are in thinner dashed line and the Nc NH to Cc CO hydrogen bond is shown in thicker dashed line. In the ball-and-stick region, only those amide hydrogens are shown that are involved in backbone hydrogen bonding. Carbon atoms are white, oxygen atoms are black, and the amide nitrogen atoms are gray in color.
which there are not many well-characterized examples) β-turn leads to the classic $3_{10}$-helix, type II the $3_{10}^{EL}$-helix and type-II' to the $3_{10}^{ER}$-helix. When we compared the equivalent position within a pair of $3_{10}$-helix and β-turn (Table IV), we found that the two types of structures can be distinguished based on residue types that occur at position i and especially at $i+3$. This shows that the signature of whether a four-residue segment will be a β-turn or a $3_{10}$-helix lies in the two terminal residues.

Secondary Structural Motifs Containing $3_{10}$-Helix Variants

Like the super-secondary structural motifs containing the classic $3_{10}$-helices, the variants can also be embedded in between different structural elements. The most prominent location of a $3_{10}^{EL}$-helix is in β-hairpins (Figs. 4 and 5), which are usually of the class 3:5 according to the notation of Sibanda et al. For $3_{10}^{ER}$-helices, a considerable number of structures have an α-helix at their C-terminal end [Figs. 4(c) and 6(b)].

Finally, for modeling of protein structures, it is important to identify local structural motifs (sequence of specific conformational angles) and the residue patterns that give rise to them. With increasing size of the available database, it is possible to refine the conformational parameters and the sequences involved in the motifs. In this study, based on $\phi$, $\psi$ angles, we have classified $3_{10}$-helices into four types and identified the residue characteristics that determine whether a four-residue segment will form a β-turn or a $3_{10}$-helix in the folded structure. However, because of the limited size of the data set used in this work, any attempt to quantitatively investigate the predictive power of the sequence rules is prone to errors, and therefore was not attempted. Recent studies indicate that

Fig. 6. Cartoon representations of (a) $3_{10}^{EL}$- (residues 40–42 in 1bkjA), (b) $3_{10}^{ER}$- (residues 61–63 in 3pte_), and (c) $3_{10}'$- (residues 191–193 in 2bgu_) helices (all in darker shade), shown in stereo and generated by MOLSCRIPT. The secondary structural elements in the immediate environment are also included in the figure.
3₁₀-helices may act as intermediates in α-helix formation. The results presented herein suggest that β-turn and 3₁₀-helices may share similar folding pathways. Although overwhelmingly helices are conformationally repetitive structures, we have identified a number of helices, which have the first residue in a nonstandard conformation. In addition, we have also shown that the population of mirror images of such variants are found to be almost equally populated in proteins, especially for three-length 3₁₀-helices ($3₁₀^{CR}$ and $3₁₀^{CL}$).

ACKNOWLEDGMENTS
We thank the Bioinformatics Centre for computational facilities.

REFERENCES