

# Expanded Turn Conformations: Characterization and Sequence-Structure Correspondence in $\alpha$ -Turns With Implications in Helix Folding

Bhaskar Dasgupta,<sup>1</sup> Lipika Pal,<sup>2</sup> Gautam Basu,<sup>3,4</sup> and Pinak Chakrabarti<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry, Bose Institute, Calcutta, India

<sup>2</sup>Bioinformatics Centre, Bose Institute, Calcutta, India

<sup>3</sup>Department of Biophysics, Bose Institute, Calcutta, India

<sup>4</sup>Graduate School of Information Science, Nara Institute of Science and Technology, Takeaway, Ika, Nara, Japan

**ABSTRACT** Like the  $\beta$ -turns, which are characterized by a limiting distance between residues two positions apart ( $i, i+3$ ), a distance criterion (involving residues at positions  $i$  and  $i+4$ ) is used here to identify  $\alpha$ -turns from a database of known protein structures. At least 15 classes of  $\alpha$ -turns have been enumerated based on the location in the  $\phi, \psi$  space of the three central residues ( $i+1$  to  $i+3$ )—one of the major being the class AAA, where the residues occupy the conventional helical backbone torsion angles. However, moving towards the C-terminal end of the turn, there is a shift in the  $\phi, \psi$  angles towards more negative  $\phi$ , such that the electrostatic repulsion between two consecutive carbonyl oxygen atoms is reduced. Except for the last position ( $i+4$ ), there is not much similarity in residue composition at different positions of hydrogen and non-hydrogen bonded AAA turns. The presence or absence of Pro at  $i+1$  position of  $\alpha$ - and  $\beta$ -turns has a bearing on whether the turn is hydrogen-bonded or without a hydrogen bond. In the tertiary structure,  $\alpha$ -turns are more likely to be found in  $\beta$ -hairpin loops. The residue composition at the beginning of the hydrogen bonded AAA  $\alpha$ -turn has similarity with type I  $\beta$ -turn and N-terminal positions of helices, but the last position matches with the C-terminal capping position of helices, suggesting that the existence of a “helix cap signal” at  $i+4$  position prevents  $\alpha$ -turns from growing into helices. Our results also provide new insights into  $\alpha$ -helix nucleation and folding. *Proteins* 2004;55:305–315.

© 2004 Wiley-Liss, Inc.

**Key words:**  $\alpha$ -turn;  $\beta$ -turn; conformation; secondary structure; protein folding; sequence preference

## INTRODUCTION

Of all the structural elements in proteins, loops pose the maximum problem for a systematic categorization,<sup>1–8</sup> though they are often parts of active sites<sup>9</sup> and an understanding of the determinants of their conformations is of intrinsic interest. The first loop conformations to be identified were three categories of four-residue  $\beta$ -turns.<sup>10</sup> This classification was subsequently extended and updated by

several workers.<sup>11–17</sup> Though it is possible to assign a class to a four-residue stretch of a polypeptide chain, the problem is rather complicated as the majority of the turns do not occur in isolation, but are in tandem with at least one residue in common between two consecutive turns.<sup>16</sup> There has not been any systematic study to delineate such multiple  $\beta$ -turns. Beyond  $\beta$ -turn, there is five-residue  $\alpha$ -turn (Fig. 1), but as in the case of earlier studies on  $\beta$ -turns,<sup>10</sup> the criterion used to identify them (which are not part of an  $\alpha$ -helix) is the existence of a hydrogen bond between the backbone carbonyl oxygen of residue  $i$  and the amide proton of residue  $i+4$ .<sup>18,19</sup> Subsequently, the definition of  $\beta$ -turn was enlarged to also include non-hydrogen-bonded turn, the main condition being a distance of less than 7 Å between the two terminal C $^{\alpha}$  atoms.<sup>12,15</sup> In a similar fashion, here we enlarge the definition of  $\alpha$ -turns by using a limiting distance (optimized from an analysis of hydrogen-bonded  $\alpha$ -turns) between the C $^{\alpha}$  atoms of residues  $i$  and  $i+4$ , and show that many of the multiple  $\beta$ -turns are indeed different classes of  $\alpha$ -turns.

Helices, sheets, and turns are three types of secondary structures with their characteristic pattern of backbone torsion angles and hydrogen bonding, and a large number of studies have been aimed at understanding them as distinct entities.<sup>20</sup> However, the occurrence of prion diseases has highlighted the importance of the interconversion between secondary structural elements.<sup>21</sup> Even to understand the structural changes that take place as the polypeptide chain moves along the folding pathway, it is important to find the common thread linking the sequence and structural patterns in various secondary structural elements. Along this line we have recently shown that the extension of different types of  $\beta$ -turns can lead to variants of  $3_{10}$ -helices<sup>22</sup> and the sequence in type I  $\beta$ -turns has considerable similarity to the N-terminal region of  $3_{10}$ - and  $\alpha$ -helices, indicating that the early nucleation event for the formation of an  $\alpha$ -helix can indeed be a  $\beta$ -turn,

BD and LP contributed equally to this work.

\*Correspondence to: Pinak Chakrabarti, Department of Biochemistry, Bose Institute, P-1/12 CIT Scheme VIIM, Calcutta 700 054, India. E-mail: pinak@boseinst.ernet.in

Received 13 August 2003; Accepted 27 October 2003

Published online 16 March 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20064

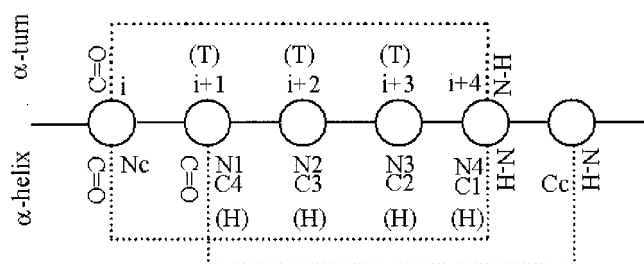


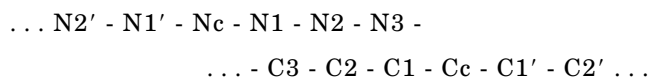
Fig. 1. Schematic representations and the convention to label each position in  $\alpha$ -turn and  $\alpha$ -helix, with the DSSP notation for these two secondary structures (T and H, respectively) given in parentheses. As the  $\alpha$ -helix is of length 4, each position is shown to have two different labels (for example, N1 and C4) counting from the two capping residues (Nc and Cc) at two ends; but for typical helices with length  $\geq 8$ , the labels are distinct. If the two consecutive  $5 \rightarrow 1$  hydrogen bonds are shortened to the  $4 \rightarrow 1$  type, the  $\alpha$ -helix is converted to a  $3_{10}$ -helix; a similar structural change will make a  $\beta$ -turn out of an  $\alpha$ -turn.

which on extension at the C-terminal end leads to a  $3_{10}$ -helix intermediate, followed by further extension and rearrangement to  $\alpha$ -helix.<sup>23</sup> In this context, we would also like to know if  $\alpha$ -turns bear any resemblance—in sequence and structure—to  $\beta$ -turns and the N-terminal region of helices. Moreover, the location of  $\alpha$ -turns in the tertiary structures would be analyzed.

## MATERIALS AND METHODS

Atomic coordinates were obtained from the Protein Data Bank (PDB) at the Research Collaboratory for Structural Bioinformatics (RCSB).<sup>24</sup> Five hundred and fifty-five (555) chains (in 531 files) were selected using PDB\_SELECT<sup>25</sup> from PDB files (as of April, 2002) with an R-factor  $\leq 20\%$ , and resolution of  $\leq 2.0$  Å and sequence identity less than 25%. Secondary structure assignments were made using the DSSP program.<sup>26</sup> Hydrogen-bonded  $\alpha$ -turns were identified as a stretch of five residues, such that the three central residues were not helical (G or H, for  $3_{10}$ - and  $\alpha$ -helices, respectively, according to DSSP notation). The residues in the “4-turn” (the hydrogen bond is between residues  $i$  and  $i+4$ ) are marked as “>444<,” where “>” and “<” indicate the positions that contribute the carbonyl and NH groups to the hydrogen bond, respectively (the symbols may be replaced by “X” if both the CO and NH groups of a given residue participate in separate hydrogen bonds).

The conventions for labeling residues when they are part of  $\alpha$ -turn and  $\alpha$ -helix are shown in Figure 1. While  $i$  through  $i+4$  represent the positions in an  $\alpha$ -turn, the nomenclature for helices and their flanking residues is as follows:



where N1 through C1 belong to helix proper, Nc and Cc represent helix capping positions, while the primed residues represents residues preceding and succeeding the helix. (As a 4-residue helix has been shown in Figure 1, the labels N1 and C4, N2 and C3, etc., indicate the same position). The propensity of a given residue to occur at a

particular location in a secondary structure was calculated as the ratio of the actual number of observations to the expected number of observations, as elaborated in Pal et al.<sup>22</sup> For each amino acid at a specific position in  $\alpha$ -turns and other secondary structural elements, a z-value was calculated. Tables of the number of occurrences of each residue at individual positions of different classes of turns, the propensity and z-values, etc., are provided as supplementary material, as are the correlation coefficients of the comparison of the residue composition between different secondary structural elements. A PDB file is mentioned in the text as the four letter PDB code (in small letter).

## RESULTS

### Identification of $\alpha$ -Turns

Five hundred and ninety-one (591) hydrogen-bonded  $\alpha$ -turns were identified by the program, DSSP, of which 439 were in the conformation AAA (discussed in the next section) (Table I). This is an improvement over the numbers (50<sup>18</sup> and 356<sup>19</sup>), observed in earlier studies. The average distance between the C $^{\alpha}$  positions of residues  $i$  and  $i+4$  for the  $\alpha$ -turns identified by DSSP was found to be 5.7 ( $\pm 0.7$ ) Å. As  $\beta$ -turns are normally identified using a distance criterion<sup>15</sup> and not by the presence of hydrogen bonds, we also used a distance less than 6.5 Å (at one-sigma level, found above) between the C $^{\alpha}$  positions of  $i/i+4$  pair, such that residues  $i$  through  $i+3$  are not part of any helix, to identify  $\alpha$ -turns. This search resulted in 3725  $\alpha$ -turns. Though there were examples of overlapping  $\alpha$ -turns, each was counted independently.

### Classification of $\alpha$ -Turns From the Distribution of $\phi, \psi$ Angles

The distributions of  $\phi, \psi$  angles for the three positions in the  $\alpha$ -turn are presented in Figure 2. To identify what combination of torsion angles constitute different types of  $\alpha$ -turn, it was necessary to divide the Ramachandran plot into distinct regions.<sup>27</sup> Taking the distribution into account the  $\phi, \psi$  space has been split into rectangular domains and their nomenclature is provided in Figure 3. This in general follows the work of Efimov<sup>28</sup> and Rooman et al.,<sup>29</sup> though for the ease of computation we have used rectangular shapes. If the distribution of points is sufficiently narrow and restricted to a smaller portion of the broad extended (E) and helical (A) regions, the latter were further divided into sub-regions (B and P, R and G, respectively). No sub-region could be used for the last ( $i+3$ ) position, indicating a more diffused nature of the distribution, which can also be inferred from the higher standard deviations on the average  $\phi, \psi$  values for this position as compared to those at the preceding two positions (Table I). It was not possible to split A into sub-regions in any of the three positions. Using a label corresponding to a  $\phi, \psi$  region at a given position, a three-letter code was used to designate the conformational features of various  $\alpha$ -turns. Fifteen major classes containing  $\approx 80\%$  of all  $\alpha$ -turns are shown in Table I.

While AAA is overwhelmingly the main class of hydrogen-bonded  $\alpha$ -turns identified by DSSP, PAA is almost

**TABLE I. The Occurrence of  $\alpha$ -Turns With Different Conformational Designations and the Average Backbone Torsion Angles ( $^\circ$ ) in the Turn Region<sup>†</sup>**

Turn designation	Identified using the distance criterion												Identified by DSSP, based on Hydrogen-bond between residues $i$ and $i+4$															
	$i+1$				$i+2$				$i+3$				$i+1$				$i+2$				$i+3$							
	Number	$\phi$	$\psi$		$\phi$	$\psi$			$\phi$	$\psi$			Number	$\phi$	$\psi$			$\phi$	$\psi$			Number	$\phi$	$\psi$				
AAA	596 <sup>b</sup>	-67 (15)	-30 (16)	-78 (16)	-33 (17)	-103 (21)	-17 (21)		-63 (9)	-28 (14)	-76 (14)	-32 (16)	439	-63 (9)	-28 (14)	-76 (14)	-32 (16)	-100 (20)	-21 (17)			25	-74 (13)	145 (9)	-93 (9)	2 (17)	-77 (18)	-30 (17)
PAA	506	-79 (13)	168 (14)	-63 (11)	-24 (14)	-88 (15)	-2 (17)																					
AAa	402	-65 (13)	-23 (13)	-93 (14)	5 (11)	83 (16)	11 (19)																					
AAE	239	-78 (20)	-29 (20)	-102 (20)	-25 (21)	-142 (21)	142 (28)																					
Bra	238	-136 (14)	126 (16)	53 (6)	41 (8)	77 (11)	6 (16)																					
BAA	226	-140 (16)	179 (13)	-63 (9)	-25 (14)	-96 (19)	-10 (20)																					
PPa	205	-76 (13)	161 (13)	-55 (5)	133 (7)	85 (14)	0 (30)																					
AAD	108	-71 (18)	-31 (14)	-89 (17)	-25 (20)	-134 (17)	72 (11)																					
rgE	90	53 (5)	43 (9)	84 (8)	-3 (9)	-115 (15)	145 (19)																					
BpA	75	-135 (15)	121 (25)	61 (7)	-125 (7)	-95 (12)	8 (14)																					
pAE	63	65 (11)	-125 (8)	-94 (13)	6 (15)	-108 (15)	142 (18)																					
BPp	60	-140 (18)	169 (14)	-55 (9)	130 (9)	85 (16)	-1 (18)																					
AAe	58	-62 (9)	-27 (13)	-98 (13)	4 (17)	128 (29)	180 (21)																					
PgA	59	-54 (7)	133 (9)	87 (12)	-11 (13)	-125 (16)	-16 (35)																					
pAA	51	79 (18)	-172 (30)	-73 (17)	-18 (18)	-97 (20)	-1 (16)																					

<sup>†</sup>The standard deviations are in parentheses.

<sup>a</sup>Though the total number of  $\alpha$ -turns is 3725, the numbers in the table add up to 2967 (80%) (591, 506, and 86%, respectively for turns identified by DSSP). Some other classes of turns (with numbers) are: PrA (36), PpA (36), rrE (34), rrA (22), ABa (21) and so on. Similarly, for those identified by DSSP: rrA (15), PrA (9), pAA (7), etc.

<sup>b</sup>On dividing into 390 hydrogen-bonded and 206 non-hydrogen bonded cases, the average values are: -63(9), -28(14), -77(14), -34(16); -101(20), -19(16) and -74(21), -34(20); -80(20), -31(20); -105(23), -13(28), respectively.

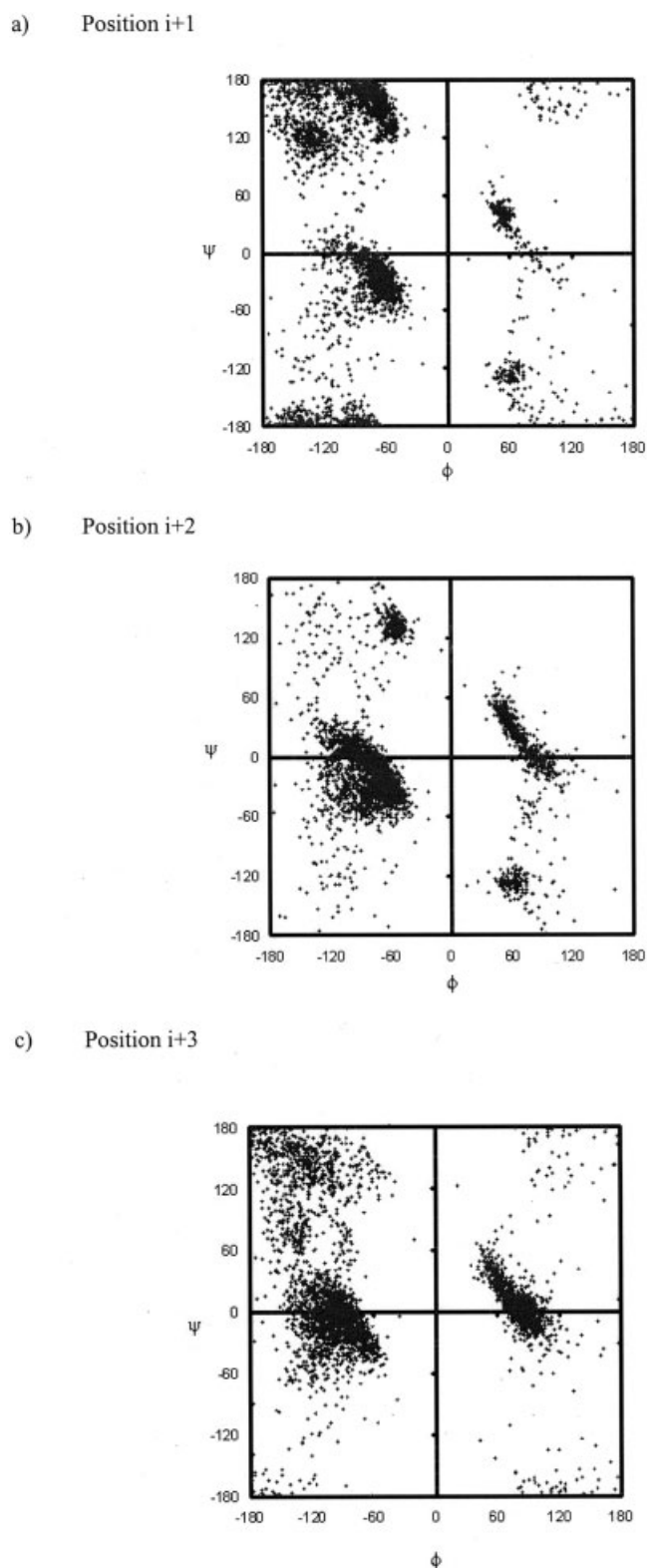


Fig. 2. The distribution of  $\phi, \psi$  angles (degrees) for residues ( $i+1, i+2, i+3$ ) in  $\alpha$ -turns.

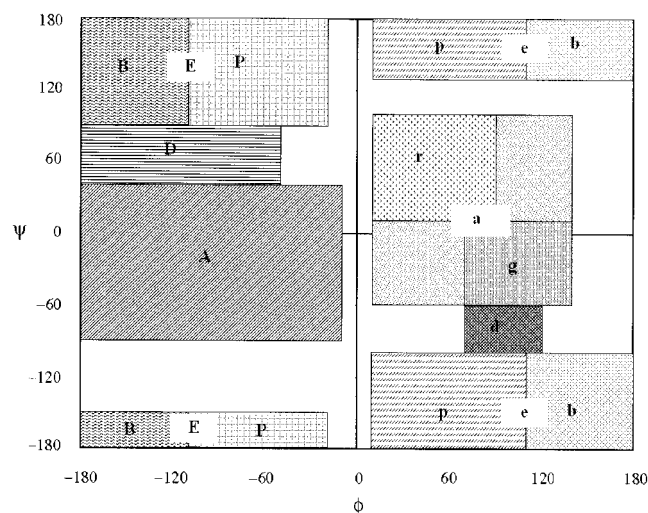


Fig. 3. The delineation of the  $\phi, \psi$  space into different regions (with their labels) encompassing the clusters of points in Figure 2. The labels are in capital letters when  $\phi$  is negative, while the centrosymmetrically-related regions are labeled in small letters. It may be noted that the symmetry-related regions are not necessarily of the same size, which has been adjusted to encompass the cluster of points. The regions are as follows: "E," extended conformation, which is further divided into "B" (largely observed in  $\beta$ -strands) and "P" (mainly associated with polyproline-like helices, but also observed in  $\beta$ -strands); "A," helical conformation, containing "R," right-handed  $\alpha$ -helical region, "G," where the C-terminal residues in  $3_{10}$ -helix are observed<sup>21,22</sup> and some outlying regions; and "D," the region bridging the above two. As the distribution of points did not warrant it, "A" has not been split into sub-regions, but "a" has been.

equally populated when turns are identified based on end-to-end distance. Hydrogen-bonded turns are quite rare (25 out of 506) in PAA, while these are found more (390, as opposed to 206, which are not hydrogen-bonded) in AAA. The number of hydrogen-bonded turns (in AAA) is slightly less than those identified by DSSP (439), as some of the  $\alpha$ -turns had end-to-end distance larger than 6.5 Å and were thus not picked up when distance was the criterion for selection. All calculations on hydrogen-bonded  $\alpha$ -turns are based on 390 cases identified above. As for hydrogen bonds in other classes, they are very rare, with no example being found in Bra, PPa, rgE, BpA, pAE, and AAe. For comparison, we also identified type I  $\beta$ -turns following the normal procedure and found that the numbers of hydrogen-bonded and non-hydrogen bonded cases are much more equitable (1946 and 1436, respectively, out of a total of 3382).

### Shift in Backbone Torsion Angles Along the Residues in $\alpha$ -Turns

There is a systematic variation of the backbone torsion angle  $\phi$  as one moves along the three central residues (from N- to C-termini) of class AAA, such that between the positions  $i+1$  and  $i+3$ ,  $\phi$  is changed by  $-36^\circ$  (Table I). A similar shift was also observed in the last three positions of  $3_{10}$ -helices.  $\psi$  also tends to change, but to a much smaller extent ( $+13^\circ$ ). If one constructs two models, one for an "ideal"  $\alpha$ -turn with classical helical  $\phi, \psi$  angles and another for the "real"  $\alpha$ -turn with the average  $\phi, \psi$  values observed,

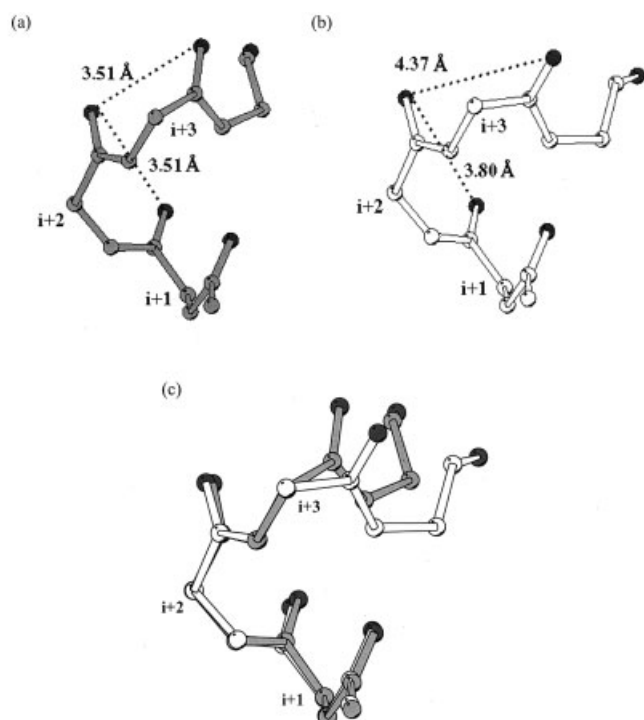


Fig. 4. Molecular diagrams of (a) an “ideal”  $\alpha$ -turn constructed with the ideal helical  $\phi, \psi$  angles ( $-65^\circ, -40^\circ$ ) at positions  $i+1$  to  $i+3$  and (b) a “real”  $\alpha$ -turn with the average  $\phi, \psi$  angles given in Table I (left half, category AAA). The two flanking residues have the  $\phi, \psi$  angles observed when the turn is part of a  $\beta$ -hairpin (Table III). (c) The superposition of (a) and (b) using C $\cdots$ O distances are shown in (a) and (b).

changing the ideal to the real structure has the effect of increasing O...O distances, especially the one involving  $i+2$  and  $i+3$  (Fig. 4), which would reduce the electrostatic repulsion between them. The  $\phi, \psi$  angles of the turn residues are very similar even when there is hydrogen bonding (Table I, footnote b), as the change in  $\phi, \psi$  angles at  $i+3$  does not significantly affect the hydrogen bond geometry (the average values in all the structures: the N...O distance, 3.15 Å and N-H...O angle,  $135^\circ$ ). If one were to consider an  $\alpha$ -helix passing through  $i+1$  and  $i+2$ , the  $\phi, \psi$  angles at  $i+3$  would make its carbonyl group splay out of the helix, like what is actually observed at the C-termini of  $3_{10}$ -helix. The reduction of the electrostatic repulsion between two consecutive carbonyl groups should be a general phenomenon, as in all the classes in Table I where there are two consecutive AA conformations the  $\phi, \psi$  angles of the second residue show the shift in the same direction relative to the first. Interestingly, when  $\phi$  is in the positive region for two consecutive residues (in classes Bra and rgE), the trend observed is just the opposite.

### Multiple $\beta$ -Turns as $\alpha$ -Turns

To find out if two consecutive  $\beta$ -turns can comprise an  $\alpha$ -turn, we examined if the distances between  $i$  and  $i+3$ , and  $i+1$  and  $i+4$  positions in all  $\alpha$ -turns are simultaneously within 7 Å. Results presented in Table II indicate that 38% of major  $\alpha$ -turns are indeed made up of overlap-

TABLE II. Multiple  $\beta$ -Turns as Part of  $\alpha$ -Turns

$\alpha$ -Turn		Multiple $\beta$ -turns	
Designation	Number	Designation <sup>a</sup>	Number <sup>b</sup>
AAA	596	I-I	586 (3)
PAA	506	Vial-I	32 (29)
AAa	402		8
AAE	239	I-VIII	69 (3)
Bra	238	X-I'	220
BAA	226		5 (6)
PPa	205		0
AAD	108	I-VIII	106
rgE	90		0
BpA	75	X-II'	26
pAE	63		3
BPA	60		1 (1)
AAe	58		2
PgA	59	II-X	59
pAA	51	X-I	17

<sup>a</sup>The types of two overlapping  $\beta$ -turns are indicated (separated by a hyphen) when there are at least 10 cases. X indicates  $\beta$ -turns not categorized earlier.<sup>16</sup>

<sup>b</sup>Out of a total given in column 2. The number in parentheses corresponds to those with a *cis* peptide bond. In all the cases of PAA and five out of six cases of BAA the *cis* peptide is between residues  $i+1$  and  $i+2$ .

ping  $\beta$ -turns; however, the type of some of these turns do not match with earlier classifications.<sup>16</sup> Considering individual classes, all the 59 members of PgA represent multiple  $\beta$ -turns, while it is 98% for AAA, 92% for Bra, and 98% for AAD. On the other hand, none in PPa or rgE, and only a small fraction of PAA (6%) and AAa (2%) contain two  $\beta$ -turns within them.

### Comparison of Residue Composition in Hydrogen-Bonded and Non-Hydrogen-Bonded $\alpha$ -Turns

As AAA  $\alpha$ -turns can be both hydrogen-bonded or non-hydrogen-bonded we examined if there is any similarity in the types of residues that are found across them by finding out the correlation coefficients of the percentage composition of residues at each position of the two categories of  $\alpha$ -turns (Fig. 5). The correlation is very strong (0.88) at the  $i+4$  position, fairly high (range, 0.70–0.80) at  $i+2$  and  $i+3$  positions and rather weak at the first two positions. This difference prompted us to investigate how similar is the residue composition in hydrogen-bonded and non-hydrogen-bonded  $\beta$ -turns of type I (which is the most highly populated  $\beta$ -turn and has backbone conformation quite akin to AAA  $\alpha$ -turn). The resemblance between the two categories is more striking except at the  $i+1$  position, for which the correlation coefficient is 0.56.

The residues in the two major classes of  $\alpha$ -turns, AAA and PAA, were also compared. As the conformational angles are in the same region, one might assume that the composition of residues at  $i+2$  and  $i+3$  positions would be similar. Contrary to this, Figure 5 indicates that the only equivalent position where the correlation coefficient is high (0.90) is  $i+4$ . There is no obvious trend in other positions. Thus except for the last position, there may not

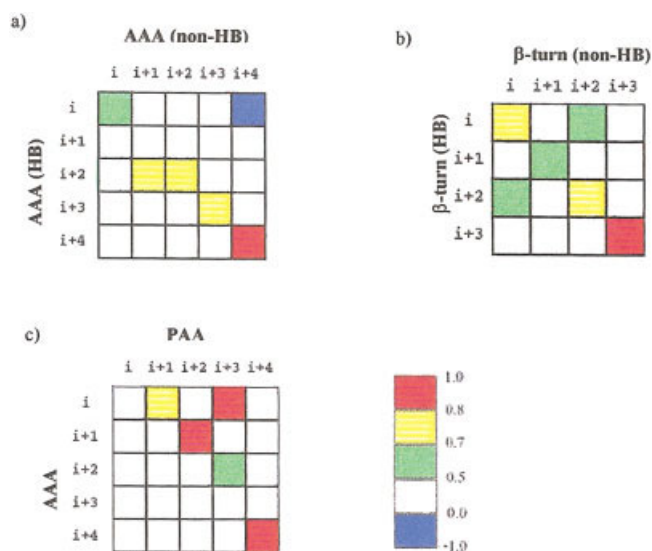


Fig. 5. Color-coded correlation coefficients of percentage composition of amino acids at different positions of (a) hydrogen-bonded and non-hydrogen-bonded  $\alpha$ -turns in class AAA; (b) hydrogen-bonded and non-hydrogen-bonded type I  $\beta$ -turns; (c)  $\alpha$ -turns of classes AAA and PAA (both inclusive of hydrogen-bonded, as well as non-hydrogen-bonded cases). The color code is shown at the end.

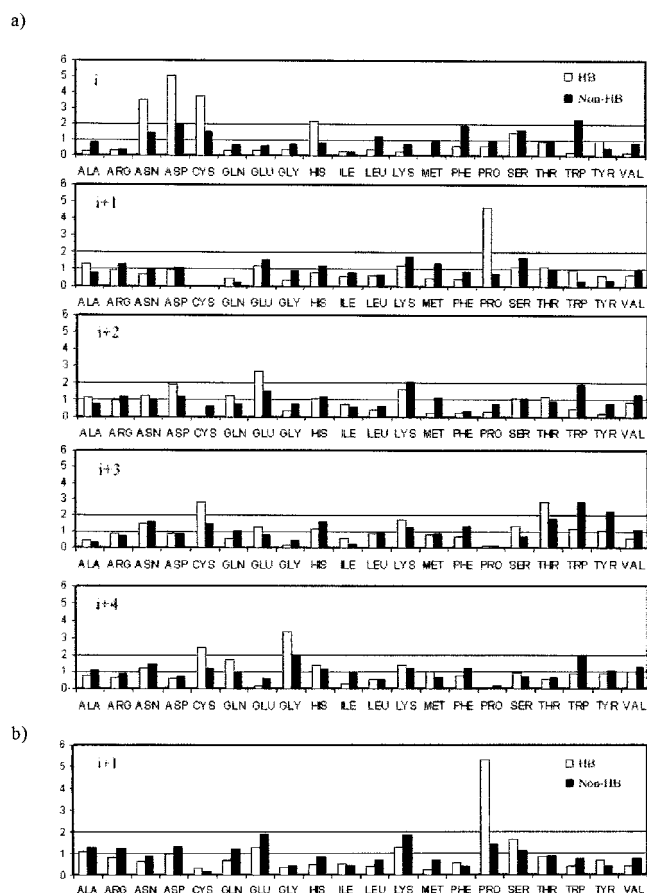


Fig. 6. Histogram of residue propensities at different positions in hydrogen-bonded and non-hydrogen-bonded (a)  $\alpha$ -turns in class AAA, and (b) type I  $\beta$ -turns, position  $i+1$ .

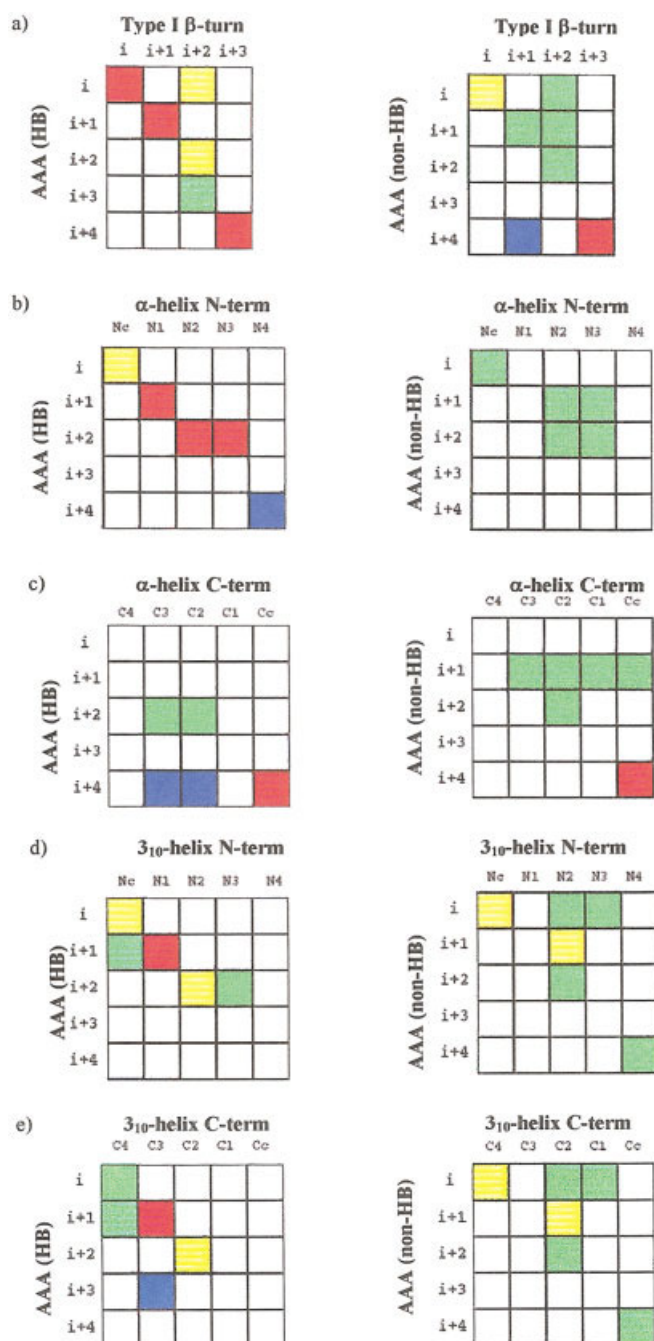


Fig. 7. Correlation coefficient of percentage composition of amino acids at different positions of hydrogen-bonded or non-hydrogen-bonded  $\alpha$ -turn in class AAA and those in (a) type I  $\beta$ -turn, (b, c) N- and C-termini of  $\alpha$ -helix and (d, e) N- and C-termini of  $3_{10}$ -helix. The color code is explained in Figure 5.  $\beta$ -turns include both hydrogen-bonded, as well as non-hydrogen-bonded cases.

be easily comprehensible trend in residue composition at other positions between different classes of  $\alpha$ -turns.

#### Propensities of Residues to be in $\alpha$ -Turns

As the compositions of residues in hydrogen-bonded and non-hydrogen-bonded AAA  $\alpha$ -turns are quite different at

positions  $i$  and  $i+1$  (Fig. 5), it is of interest to find out what residues have different propensities to occur in these two categories of turns [Fig. 6(a)]. Asn, Asp, Cys and His are highly favored at positions  $i$ , Pro at  $i+1$  in hydrogen-bonded turns. Non-hydrogen-bonded turns have more of hydrophobic residues (Phe and Trp) at positions  $i$  and  $i+4$  and these may provide hydrophobic interactions across the turn (as observed in the shortest helices<sup>23</sup>), stabilizing it. Trp also has rather high propensities at  $i+2$  and  $i+3$ , and for the latter position there also exists more hydrophobic interactions across the turn.

### Correlation of Residue Composition in $\alpha$ -Turns With Those in Other Secondary Structures

AAA  $\alpha$ -turn, type I  $\beta$ -turn and terminal turns in  $3_{10}$ - and  $\alpha$ -helices all have the constituent residues occupy the same region in the  $\phi, \psi$  space (Fig. 3) and it is of interest to see if the residue compositions at different positions are similar across these secondary structural elements. For  $\alpha$ -turns, the hydrogen-bonded and non-hydrogen-bonded categories were considered separately (left and right panels in Fig. 7), but there is no resemblance except when the  $i+4$  position is compared with the  $i+3$  position of  $\beta$ -turn [Fig. 7(a)] or Cc of  $\alpha$ -helix [Fig. 7(c)]. The similarity at the last position is not surprising as the two categories of AAA turns show the strongest similarity at this position [Fig. 5(a)]. Considering the hydrogen-bonded AAA  $\alpha$ -turns, there is almost sequential match with type I  $\beta$ -turn, such that  $\alpha$ -turns may be deemed to be  $\beta$ -turns with an extra residue inserted after the  $i+2$  position. Residue compositions at three consecutive positions in  $\alpha$ -turn also match with the N-terminal region of  $\alpha$ -helix, the last position matching with the Cc position at the other end of the helix [Fig. 7(b, c)]— $i+3$  is the only position which does not match with the helical positions considered here. Though slightly weaker as compared to  $\alpha$ -helix, the N-terminal region of  $3_{10}$ -helix shows a very similar trend. Though a bit confusing at the first glance, the C-terminal end of  $3_{10}$ -helices essentially reproduces what has been found at the N-terminal end [Fig. 7(d,e)]; as most of the  $3_{10}$ -helices are three residues long,<sup>23</sup> N1 and C3, N2 and C2, and N3 and C1 indicate the same positions.

The positions  $i+4$  of  $\alpha$ -turn,  $i+3$  of  $\beta$ -turn and Cc of  $\alpha$ -helix has very similar composition. The propensities of residues to occur at these positions are plotted together in Figure 8. The distribution is quite similar and characterized by very high occurrence of Gly and absence of Pro. Residues such as Asn, His, and Lys have high propensities to occur, especially in  $\alpha$ -turn and  $\alpha$ -helix.

### Secondary Structures Around $\alpha$ -Turns

The patterns of secondary structural elements on either side of the  $\alpha$ -turns were investigated. The most consistent feature across all the classes is the occurrence of  $\alpha$ -turns in the loop region<sup>28,30</sup> between two antiparallel  $\beta$ -strands, or  $\beta$ -hairpins. Though various types of  $\beta$ -hairpins are possible the most consistent patterns in each class are given in Table III. While only 13% of conventional  $\alpha$ -turns (in AAA) follow this pattern, the percentage increases with

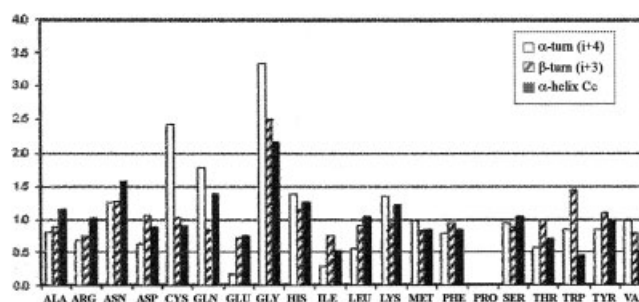


Fig. 8. Histogram of residue propensities at the last position of AAA  $\alpha$ -turn (hydrogen-bonded), type I  $\beta$ -turn (all structures) and the Cc position of  $\alpha$ -helix.

the classes (such as Bra, rgE, and pAE), which are populated to a lesser extent. However, with many of these either the position  $i+1$  or  $i+3$  becomes a part of the  $\beta$ -strand. In Figure 9 all the structures in some representative classes are superimposed. Though only two residues on each side were used for the least-square fit, the structural match for the central residues is quite remarkable, suggesting that all the structures have the same  $\beta$ -hairpin motif. The motifs have also been classified following the convention of Sibanda et al.<sup>30</sup> (Table III) and some individual structures, along with the hydrogen bond pattern used in the classification, are shown in Figure 9.

### Residue Preferences in $\alpha$ -Turns Forming $\beta$ -Hairpin Loops

As different combinations of conformational angles are involved in the  $\alpha$ -turns that are part of  $\beta$ -hairpins, it is of interest to see if there exists any pattern in the most common residues observed at different positions within these turns. When the turn is defined by two or more consecutive residues in the A conformation (Fig. 3) the first residue with this conformation (i.e.,  $i+1$  in AAA and AAa,  $i+2$  in PAA) has a preponderance of Pro, which is usually absent in the other positions (Table IV). When there are two consecutive residues in the right half of the  $\phi, \psi$  plot (as in the classes Bra and rgE) (Fig. 3), the first residue ( $i+2$  and  $i+1$ , respectively, in the two classes) is usually Asp or Asn, with Gly occupying the second position. Interestingly, concomitant with this change in the residue type there is a change in the  $\phi, \psi$  angles of the two residues (as mentioned in the section on the shift in backbone torsion angles).

## DISCUSSION

### Identification of $\alpha$ -Turns and Their Conformational Features

$\alpha$ -turns are conventionally defined as five consecutive residues ( $i$ , to  $i+4$ ) which are not part of an  $\alpha$ -helix, but still have a hydrogen bond connecting positions  $i$  and  $i+4$  (Fig. 1).<sup>18,19</sup> Based on the average distance observed in these turns, here we use a limiting distance of 6.5 Å between the C $^{\alpha}$  coordinates of the terminal residues to identify 3725 of  $\alpha$ -turns in a database of known protein structures. Based on the  $\phi, \psi$  angles at the three central positions (Fig. 2), these have been grouped in 15 major

**TABLE III. Most Common Secondary Structure Involving Different Types of  $\alpha$ -Turns and Their Immediate Neighbors<sup>†</sup>**

Turn <sup>a</sup>	Secondary structures at positions <sup>b</sup>							Number	Percentage <sup>c</sup>
	<i>i</i> - 1	<i>i</i>	<i>i</i> + 1	<i>i</i> + 2	<i>i</i> + 3	<i>i</i> + 4	<i>i</i> + 5		
AAA (4:4)	E	E	T	T	T	T	E	75	13
PAA <sup>d</sup> (7:7)	E	E	C	T	T	S	C	99	20
AAa (7:7)	E	C	T	T	S	C	E	96	24
AAE (4:6)	E	C	S	S	C	E	E	19	8
Bra (2:2)	E	E	E	T	T	E	E	168	71
BAA (2:2)	E	E	E	T	T	E	E	39	17
PPa	E	E	C	T	T	C	E	24	12
rgE (2:2)	E	E	T	T	E	E	E	57	65
	E	E	T	T	E	E	C	12	13
BpA (2:2)	E	E	E	T	T	E	E	49	65
pAE (2:2)	E	E	T	T	E	E	E	50	79
BPa (2:2)	E	E	E	T	T	E	E	10	17
AAe (2:2)	E	E	T	T	E	E	E	20	35

<sup>†</sup>The combination of secondary structural elements with the highest member of occurrence (if >5) is shown. If there are other combinations constituting >10% of cases in a given turn type, these are also presented. There is no entry for the three types of  $\alpha$ -turns, AAD, PgA and pAA.

<sup>a</sup>Except PPa, the other  $\alpha$ -turns are part of a  $\beta$ -hairpin structure and the nomenclature<sup>30</sup> is provided in parentheses.

<sup>b</sup>As depicted by DSSP: E,  $\beta$ -strands; T, hydrogen bonded turn; S, bend; C, no regular structure.

<sup>c</sup>Based on the number observed with the given combination of secondary structures (given in the previous column) relative the total number in the class.

<sup>d</sup>Residues at *i* + 6 and *i* + 7 have the structure E.

classes (Table I). The population of various classes varies depending on the presence or absence of hydrogen bonds between the main-chain atoms of terminal positions (*i*, *i*+4). While AAA (where the individual letters correspond to the region of the  $\phi$ ,  $\psi$  space occupied by residues, *i*+1 to *i*+3 (Fig. 3)—all in the  $\alpha$ -helical region in the present case) is the most conspicuous group when there is hydrogen bonding, PAA is as populated when non-hydrogen-bonded turns are also considered.

There is a systematic shift in the  $\phi$ ,  $\psi$  angles of residues as one moves from position *i*+1 to *i*+3, with  $\phi$  changing by as much as  $-36^\circ$  (Table I), the reason for which is likely to be electrostatics, as the O...O distance gets progressively longer, especially between the last two positions (Fig. 4). In fact, the shift in the torsion angles seem to be the general trend in all the classes when there are two consecutive residues with the AA conformation. When  $\phi$  is in the positive region for two consecutive residues (classes Bra and rgE), the shift is the opposite (such as a more positive  $\phi$ ), but interestingly, along with the shift in torsion angles one also observes the presence of different types of residues at the two positions—the first residue is usually Asp or Asn and the second, Gly (Table IV). This is a unique example, where two residues in the sequence can bring about a systematic change in the local backbone geometry.

$\beta$ -Turns usually occur in tandem and it is rather difficult to characterize the overlapping turns systematically.<sup>16</sup> However, when defined from the distance criterion, many multiple  $\beta$ -turns can be considered  $\alpha$ -turns. Table II indicates that 38% of  $\alpha$ -turns (and almost all in classes AAA, Bra and PgA) encompass overlapping  $\beta$ -turns.

### $\alpha$ -Turns in the Context of Protein Tertiary Structure

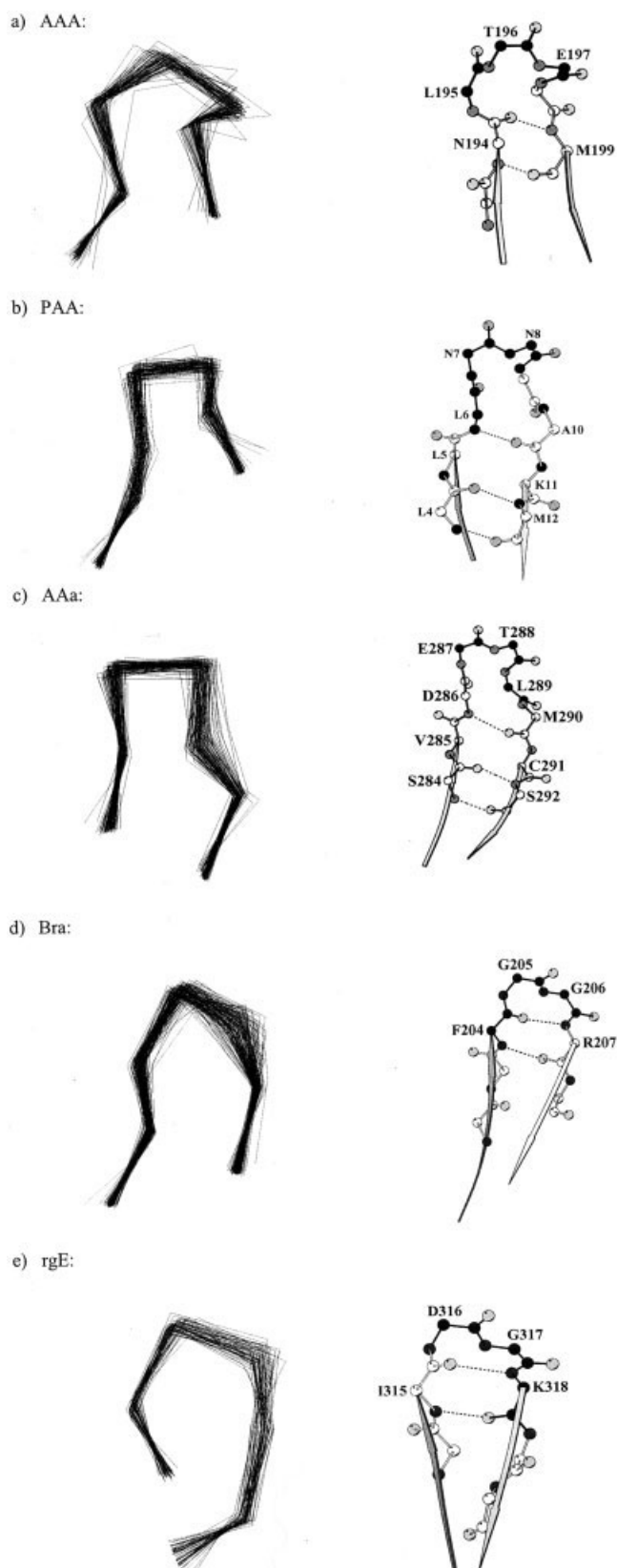
The most common structural motif where  $\alpha$ -turns are found is  $\beta$ -hairpin (Table III, Fig. 9). Such structures have

been understood from an analysis of the loop region spanning the strands.<sup>30</sup> As only  $\beta$ -turns were used to characterize the conformation, loop length beyond two residues could not be properly included. By using  $\alpha$ -turns as a basic structural unit, a larger number of  $\beta$ -hairpin loops become amenable to conformational characterization.

### Difference in Residue Composition in Hydrogen-Bonded and Non-Hydrogen-Bonded $\alpha$ - and $\beta$ -turns

$\beta$ -turns are identified based on close distance between *i* and *i*+3 positions, irrespective of whether there is any hydrogen-bond linking them, and propensities of residues to occur at various positions of the turn have been calculated.<sup>16</sup> It is however pertinent to ask if residues in the turn have no influence on the occurrence of the hydrogen-bond. Though there are strong similarities in the residue composition at all other positions, results in Fig. 5 indicate that at *i*+1 it is quite different depending on the presence or absence of hydrogen bonding. The propensity of Pro to occur at this position for hydrogen-bonded turn is very high (5.29) [Fig. 6(b)]. As the  $\alpha$ -turn is longer by one residue, the effect that a particular residue may exert on the existence of a hydrogen bond linking the terminal residues should be lesser. As it turns out, the correlation coefficients are rather poor at *i* and *i*+1 positions between the hydrogen-bonded and non-hydrogen-bonded cases, but even here Pro seems to be the favored residue at the *i*+1 position in the former category [Fig. 6(a)]. As to why Pro should have such an effect, it not only imparts conformational restriction on its own position, but also on the preceding position, which has a marked preference to occur in the  $\beta$  region.<sup>31</sup> Additionally, depending on the presence of an aromatic residue on a neighboring location, there could also be stacking interaction with the proline residue,<sup>32</sup> which would also reduce the structural fluctua-



**TABLE IV. Residue Preference in the  $\alpha$ -Turn Region of  $\beta$ -Hairpin Motifs Shown in Figure 9.**

Turn	Residues <sup>a</sup> at positions		
	$i + 1$	$i + 2$	$i + 3$
AAA	P (20), A (15)	E (21), [P]	T (23), K (20), [P]
PAA	D (31), N (12), [P]	P (19), A (15)	D (39), N (17), [P]
AAa	P (21), A (14)	D (43), N (16), [P]	G (85)
Bra	V (17), I (12), [P]	N (24), D (23)	G (73), D (9)
rgE	D (28), N (19)	G (95)	K (32), [P]

<sup>a</sup>The number of occurrence of each residue was noted; if the two highest values were  $> 10$ , these residues are reported (with the percentage occurrence in parentheses). If Pro is absent in a position (with a negative  $\phi$  value) it is shown in square bracket.

tion of the peptide segment, leaving the terminal residues to explore various conformational possibilities and arrive at the one amenable to the formation of the hydrogen bond.

### Resemblance of $\alpha$ -Turn With That of $\beta$ -Turn and Different Positions of Helices, and Implications in Helix Folding

A comparison of the amino acid compositions at consecutive positions between two secondary structural elements, such  $\beta$ -turns,  $3_{10}$ - and  $\alpha$ -helices provided important insight into helix folding,<sup>23</sup> where the absence of a large enough database ruled out the inclusion of  $\alpha$ -turns. The identification of a much larger number of  $\alpha$ -turns in the present study offered us an opportunity and these, separated into hydrogen- and non-hydrogen-bonded categories could be compared (Fig. 7). An interesting trend was observed for the former. Based on residue usage one can suggest that  $\beta$ -turn, on incorporation of a residue after the  $i+2$  position, becomes an  $\alpha$ -turn. The similarity is quite strong between the first three positions of  $\alpha$ -turn and positions Nc to N2 of  $\alpha$ -helix, while the last position matches with Cc, indicating that the presence of a "helix C-capping signal" at the  $i+4$  position may prevent the  $\alpha$ -turn from growing into an  $\alpha$ -helix. The similarity extends to  $3_{10}$ -helix also, though it gets weaker, as due to the shortness of  $3_{10}$ -helices<sup>23</sup> the N- and C-terminal ends

Fig. 9. Pair of diagrams showing major types of  $\alpha$ -turns when they constitute  $\beta$ -hairpin motif. On left are superpositions of all the structures under a given class as mentioned in Table III. A model was first constructed with the average  $\phi, \psi$  angles for all the structures and the individual structures were then superimposed onto this using the C $\alpha$ -positions of the four terminal residues ( $i-1, i, i+4$ , and  $i+5$ ). On right, one typical example is shown: the strands are in cartoon representation, the backbone atoms in the loop are in ball-and-stick (with those in  $\alpha$ -turn in black), N atoms are darker than the O atoms, the hydrogen bonds indicative of the type of  $\beta$ -hairpin loop are in dashed line. The residues in the  $\alpha$ -turn and the two residues in  $\beta$ -strands involved in the first set of hydrogen bonding are labeled (one letter residue name, followed by residue number); in (b) and (c) residues in the loop region involved in a hydrogen bond and the two terminal residues (not hydrogen-bonded) in the strands are also labeled. The PDB codes and protein names are: (a) 1ayl, phosphoenolpyruvate carboxykinase; (b) 1ads, aldose reductase; (c) 1jk7, serine/threonine protein phosphatase PP1-gamma; (d) 1b65, aminopeptidase, and (e) 1b6a, methionine aminopeptidase. The superimposition was done using the "Biopolymer" module of InsightII (Accelrys Inc., San Diego) and the molecular diagrams were generated using MOLSCRIPT.<sup>41</sup>

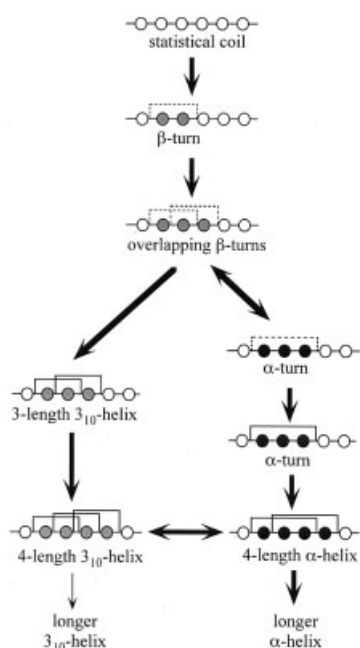


Fig. 10. A schematic description of helix nucleation and propagation in proteins. Broken lines indicate potential hydrogen bonds, but whose presence is not essential; solid lines represent hydrogen bonds (required). Shaded residues indicate that the backbone angles are restricted; as long  $3_{10}$ -helices form rarely, the arrow between 4-length  $3_{10}$ - and longer  $3_{10}$ -helices is thinner to show that the path is less probable.

overlap and these therefore leave a less definite impression on the residue composition. Interestingly, the last positions of  $\alpha$ -helix,  $\alpha$ - and  $\beta$ -turns have similar residue composition. This can also be seen in the distributions of propensities of residues to occur at these positions (Fig. 8).

Based on the above results we may propose an alternative hypothesis on helix-nucleation.<sup>33,34</sup> We earlier proposed that  $\beta$ -turn, 3- and 4-length  $3_{10}$ -helices are intermediates during the formation of  $\alpha$ -helix, while  $\alpha$ -turns (as the literature did not indicate the existence of many of these in protein structures) are off the pathway.<sup>23</sup> Now that we find that  $\alpha$ -turns exist in reasonable numbers (especially if we remove the condition of hydrogen bonding to define them), we suggest that there may be an alternative (or parallel) pathway with an  $\alpha$ -turn as an intermediate. As proposed earlier,<sup>23</sup> the N-terminal end of the helix first starts as a  $\beta$ -turn, one residue extension of which leads to two overlapping  $\beta$ -turns. As most of the  $\alpha$ -turns (class AAA) are also simultaneously overlapping  $\beta$ -turns also (Table II), the above structure can rearrange to an  $\alpha$ -turn. As the composition similarity between  $\alpha$ -turn and  $3_{10}$ -helix is substantial, the rearrangement can also pass through a  $3_{10}$ -helix.

The proposed scheme, shown in Figure 10, is consistent with long time molecular simulation results on polyalanine. Huo and Straub<sup>35</sup> carefully computed and analyzed several coil-helix transition pathways and showed that 1) localized  $i, i+3$  hydrogen-bonded turns appear during the early stages of folding (nucleation), and, 2) a  $3_{10}$ -helix intermediate (three or more consecutive  $i, i+3$  hydrogen bonds) is not a necessary requirement for subsequent helix

propagation (not observed for eight out of eighteen trajectories), instead, often short  $\alpha$ -helical stretches appear without a  $3_{10}$ -helical intermediate or through mixed ( $3_{10}$ -/ $\alpha$ -) intermediates. Other molecular dynamics,<sup>36</sup> as well as Monte Carlo simulations<sup>37</sup> also indicate that the nucleation of  $\alpha$ -helical segments can occur by several different mechanisms, the development of an initial  $\beta$ -turn into an  $\alpha$ -turn being one of them. These observations are consistent with Figure 10 with respect to the helix nucleation ( $\beta$ -turn) and subsequent parallel pathways for helix propagation. An experimental observation involving the crystal structure and packing of terminally blocked Val-Ser dipeptide<sup>38</sup> is also pertinent here. The structure of t-Boc-Val-Ser-NHMe shows an ideal type I  $\beta$ -turn backbone but lacks the typical intra-molecular  $i, i+3$  hydrogen bond. Instead, it participates in inter-molecular hydrogen-bond network more typical of  $\alpha$ -helix, demonstrating the ease with which a stretch of backbone, defined by characteristic  $\beta$ -turn dihedral angles, can participate in  $\alpha$ -helical hydrogen-bonded network. Essentially this suggests the possibility of  $\alpha$ -helix formation directly from  $\beta$ -turns, without any need of  $3_{10}$ -helix intermediates. We also arrived at similar conclusions from similarity of residue composition, as proposed in Figure 10.

In conclusion, in this paper we identify and use backbone torsion angles to characterize various types of  $\alpha$ -turns. The presence of Pro (at position  $i+1$ ) seems to contribute to the existence of a hydrogen bond linking the two terminal positions of both  $\alpha$ - and  $\beta$ -turns. There have been attempts to identify sequence motifs that are predictive for the type of secondary structures they are located in.<sup>39</sup> Here we take a rather contra perspective to show the similarity of the residue composition across secondary structures, viz. in consecutive positions of  $\alpha$ -turns with those in  $\beta$ -turns and helices, an observation that should be useful in our understanding of the protein folding problem, especially the nucleation of  $\alpha$ -helix. The sequence and conformational features of  $\alpha$ -turns should be useful in the prediction of such structures in proteins.<sup>40</sup>

## ACKNOWLEDGMENT

We thank Mr. A. Gupta (who was supported from a CSIR-NMITLI grant) for help in computation. The Council of Scientific and Industrial Research and the Department of Biotechnology are thanked for fellowships to BD and LP, respectively. GB acknowledges support from special coordination fund promoting science and technology, MEXT, Japan.

## REFERENCES

1. Ring CS, Kneller DG, Langridge R, Cohen FE. Taxonomy and conformational analysis of loops in proteins. *J Mol Biol* 1992;224: 685–699.
2. Kwasigroch JM, Chomilier J, Moron JP. A global taxonomy of loops in globular proteins. *J Mol Biol* 1996;259:855–872.
3. Donate LE, Rufino SD, Canard LHJ, Blundell TL. Conformational analysis and clustering of short and medium size loops connecting regular secondary structures: a database for modeling and prediction. *Protein Sci* 1996;5:2600–2616.
4. Oliva B, Bates PA, Querol E, Aviles FX, Sternberg MJ. An automated classification of structure of the protein loops. *J Mol Biol* 1997;266:814–830.

5. Wojcik J, Mornon JP, Chomilier J. New efficient statistical sequence-dependent structure prediction of short to medium-sized protein loops based on an exhaustive loop classification. *J Mol Biol* 1999;289:1469–1490.
6. Burke DF, Deane CM, Blundell TL. Browsing the S-Loop database of structurally classified protein loops connecting elements of protein secondary structure. *Bioinformatics* 2000;16:513–519.
7. Burke DF, Deane CM. Improved loop prediction from sequence alone. *Protein Eng* 2001;14:473–478.
8. Pal M, Dasgupta S. The nature of the turn in omega loops of proteins. *Proteins* 2003;51:591–606.
9. Tramontano A, Chothia C, Lesk AM. Structural determinants of the conformations of medium-sized loops in proteins. *Proteins* 1989;6:382–394.
10. Venkatachalam CM. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* 1968;6:1425–1436.
11. Lewis PN, Momany FA, Scheraga HA. Chain reversals in proteins. *Biochim Biophys Acta* 1973;303:211–229.
12. Nemethy G, Scheraga HA. Stereochemical requirements for the existence of hydrogen bonds in beta-bends. *Biochem Biophys Res Commun* 1980;95:320–327.
13. Richardson JS. The anatomy and taxonomy of protein structure. *Adv Protein Chem* 1981;34:167–339.
14. Rose GD, Gierasch LM, Smith JA. Turns in peptides and proteins. *Adv Protein Chem* 1985;37:1–109.
15. Wilmot CM, Thornton JM. Analysis and prediction of the different types of  $\beta$ -turn in proteins. *J Mol Biol* 1988;203:221–232.
16. Hutchinson EG, Thornton JM. A revised set of potentials for  $\beta$ -turn formation in proteins. *Protein Sci* 1994;3:2207–2216.
17. Pal D, Chakrabarti P. Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J Mol Biol* 1999;294:271–288.
18. Nataraj DV, Srinivasan N, Sowdhamini R, Ramakrishnan C.  $\alpha$ -Turns in protein structures. *Current Sci* 1995;69:434–447.
19. Pavone V, Geata G, Lombardi A, Natri F, Maglio O, Isernia C, Saviano M. Discovering protein secondary structures: classification and description of isolated alpha-turns. *Biopolymers* 1996;38:705–721.
20. Creighton TE. *Proteins: structures and molecular properties*. New York: WH Freeman and Co.; 1993.
21. Soto C. Protein misfolding and disease; protein refolding and therapy. *FEBS letters* 2001;498:204–207.
22. Pal L, Basu G, Chakrabarti P. Variants of  $3_{10}$ -helices in proteins. *Proteins* 2002;48:571–579.
23. Pal L, Chakrabarti P, Basu G. Sequence and structure patterns in proteins from an analysis of the shortest helices: implications for helix nucleation. *J Mol Biol* 2003;326:273–291.
24. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–242.
25. Hobohm U, Scharf M, Schneider R. Selection of representative protein data sets. *Protein Sci* 1993;1:409–417.
26. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983;22:2577–2637.
27. Chakrabarti P, Pal D. The interrelationships of side-chain and main-chain conformations in proteins. *Prog Biophys Mol Biol* 2001;76:1–102.
28. Efimov AV. Standard structures in proteins. *Prog Biophys Mol Biol* 1993;60:201–239.
29. Rooman MJ, Kocher JA, Wodak SJ. Prediction of protein backbone conformation based on seven structure assignments. Influence of local interactions. *J Mol Biol* 1991;221:961–979.
30. Sibanda BL, Blundell TL, Thornton JM. Conformation of  $\beta$ -hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* 1989;206:759–777.
31. MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. *J Mol Biol* 1991;218:397–412.
32. Bhattacharyya R, Chakrabarti P. Stereospecific interactions of proline residues in protein structures and complexes. *J Mol Biol* 2003;331:925–940.
33. Zimm BH, Bragg JK. Theory of phase transition between helix and random coil in polypeptide chains. *J Chem Phys* 1959;31:526–535.
34. Lifson S, Roig A. On the theory of helix-coil transition in polypeptides. *J Chem Phys* 1961;34:1963–1974.
35. Huo S, Straub, JE. Direct computation of long time processes in peptides and proteins: reaction path study of coil-to-helix transition in polyalanine. *Proteins* 1999;36:249–261.
36. Bertsch RA, Vaidehi N, Chan SI, Goddard WA. Kinetic steps for  $\alpha$ -helix formation. *Proteins* 1998;33:343–357.
37. Sung SS. Folding simulations of alanine-based peptides with lysine residues. *Biophys J* 1995;68:826–834.
38. Perczel A, Foxman BM, Fasman GD. How reverse turns may mediate the formation of helical segments in proteins: An x-ray model. *Proc Natl Acad Sci USA*. 1992;89:8210–8214.
39. Rooman MJ, Rodriguez J, Wodak SJ. Relation between protein sequence and structure and their significance. *J Mol Biol* 1990;213:337–350.
40. Chou K-C. Prediction of tight turns and their types in proteins. *Analytical Biochem* 2000;286:1–16.
41. Kraulis PJ. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 1991;24:946–950.