Protein Engineering vol.12 no.6 pp.447–455, 1999

Secondary structures without backbone: an analysis of backbone mimicry by polar side chains in protein structures

Narayanan Eswar and C.Ramakrishnan[1]

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. E-mail: ramki@crmbu2.mbu.iisc.ernet.in

[1]To whom correspondence should be addressed

Backbone mimicry by the formation of closed-loop C_7, C_{10} and C_{13} (mimics of γ-, β- and α-turns) conformations through side chain–main chain hydrogen bonds by polar groups is a frequent observation in protein structures. A data set of 250 non-homologous and high-resolution protein crystal structures was used to analyze these conformations for their characteristic features. Seven out of the nine polar residues (Ser, Thr, Asn, Asp, Gln, Glu and His) have hydrogen bonding groups in their side chains which can participate in such mimicry and as many as 15% of all these polar residues engage in such conformations. The distributions of dihedral angles of these mimics indicate that only certain combinations of the dihedral angles involved aid the formation of these mimics. The observed examples were categorized into various classes based on these combinations, resulting in well defined motifs. Asn and Asp residues show a very high capability to perform such backbone secondary structural mimicry. The most highly mimicked backbone structure is of the C_{10} conformation by the Asx residues. The mimics formed by His, Ser, Thr and Glx residues are also discussed. The role of such conformations in initiating the formation of regular secondary structures during the course of protein folding seems significant.

Keywords: backbone mimicry/C_7, C_{10}, C_{13} conformations/hydrogen bonds/polar residues

Introduction

Secondary structures in proteins are characterized by repeating patterns of hydrogen bonds between the main chain atoms. Such regular hydrogen-bonded motifs have been the motivation behind several studies, including the pioneering postulation of the existence of the α-helical structure by Pauling *et al.* (1951). However, these hydrogen bonds between the atoms of the polypeptide backbone do not discriminate between the various residues involved since the backbone is identical for all the naturally coded amino acids (except proline). Discrimination between residues comes into play only when interactions (hydrogen bonded or otherwise) between the side chain and main chain or side chain and side chain atoms are considered (Rose and Wolfenden, 1993). It has been established that hydrogen bonds between the atoms of side chains are usually local in nature, leading to the tightly packed structures of proteins (Baker and Hubbard, 1984; Ippolito *et al.*, 1990; Stickle *et al.*, 1992). In this paper we show that a study of the geometry of side chain to backbone interactions through hydrogen bonds can reveal interesting and crucial features.

Earlier observations have indicated that Asn or Asp (Asx) residues can mimic backbone conformations and thoughtful hypotheses have been made based on this feature (Richardson, 1981; Richardson and Richardson, 1989). One is the formation of the Asx turn which involves a hydrogen bond between the side chain 'peptide-like' group of Asx to one of the neighboring backbone peptide groups (Rees *et al.*, 1983). The first occurrence of such a conformation was reported as the 'pseudo turn' owing to its close similarity to the β-turn (Tainer *et al.*, 1982). Another example of such a mimic occurs at the start sites of α-helices. One such feature has been described at helix capping sites where the hydrogen bonds from flanking polar residues are shown to satisfy the otherwise unsatisfied hydrogen bonding potential of backbone amides (Presta and Rose, 1988; Richardson and Richardson, 1988).

The hydrogen bonding interactions of side chains with backbone amide groups have been the focus of many early investigations. Ptitsyn and co-workers (Finkelstein and Ptitsyn, 1977; Finkelstein *et al.*, 1977) had analyzed the 'class of local contacts', namely those conformations of the polypeptide chain which have side chain to backbone hydrogen bonds but lack hydrogen bonds between the backbone atoms. The results were then used to evaluate the thermodynamic parmeters governing the formation of such substructures. A more recent analysis was one of Asx–Pro turns where there is a pseudo-turn followed by a regular β-turn (Wilson and Finlay, 1997). Questel *et al.* (1993) in an analysis described closed ring structures in which the side chain peptide groups hydrogen bond to backbone groups in a way similar to the hydrogen bonding pattern of β-hairpins.

This analysis focuses on backbone mimics which closely resemble the standard secondary structures observed in proteins. Here we show that side chain–main chain hydrogen bonds involving 7, 10 or 13 atoms or, in other words, mimics of γ-, β- or α-turns is common in protein structures. Such formations may be crucial to understanding the mechanism of formation of secondary structures during the course of protein folding.

Materials and methods

A set of 250 non-homologous protein crystal structures refined to high resolution (≤2 Å) and derived from the Brookhaven Protein Data Bank (PDB) (Bernstein *et al.*, 1977) was used for the analysis. [see Gunasekaran *et al.* (1997) for list of proteins]. In cases where the PDB entry had more than one chain which was homologous, only one of the them was used. Only those residues which had temperature factors of 30 $Å^2$ or less was used for the analysis to identify the hydrogen bonds without ambiguity.

Hydrogen bonds were identified based on the well known length and angle criteria (Ramakrishnan and Prasad, 1971). The donor–acceptor distance, l (N···O), was required to lie in the interval defined by 2.4 Å ≤ l ≤ 3.6 Å. In the case of a hydrogen bond between the side chain carbonyl and the main chain amide, the hydrogen atom was geometrically fixed on

Table I. Number of examples of different types of side chain–main chain hydrogen bonds in protein structures leading to the conformations which mimic the backbone

Residue	No. of residues in data set	No. of residues with low *B*-factors[a]	Side chain donor/acceptor	Main chain acceptor/donor	No. of atoms in hydrogen-bonded ring	No. of examples of hydrogen bond
Ser	3463	2413	$O^{\gamma 1}{}_i$	O_{i-1}	7	51
				O_{i-2}	10	13
				O_{i-3}	13	217
Thr	3226	2560	$O^{\gamma 1}{}_i$	O_{i-1}	7	157
				O_{i-2}	10	32
				O_{i-3}	13	204
Asn	2469	1566	$O^{\delta 1}{}_i$	N_{i+1}	7	32
				N_{i+2}	10	191
				N_{i+3}	13	96
			$N^{\delta 2}{}_i$	O_i	7	72
				O_{i+1}	10	13
				O_{i+2}	13	5
Asp	3172	1913	$O^{\delta 1}{}_i$	N_{i+1}	7	69
				N_{i+2}	10	277
				N_{i+3}	13	118
Gln	1862	962	$O^{\varepsilon 1}{}_i$	N_i	7	30
				N_{i-1}	10	0
				N_{i-2}	13	2
Glu	2920	1180	$O^{\varepsilon 1}{}_i$	N_{i-1}	7	88
				N_{i-2}	10	5
				N_{i-3}	13	6
His[b]	1089	817	$N^{\delta 1}{}_i$	O_i	7	10
				O_{i+1}	10	21
				O_{i+2}	13	2

[a]Examples with temperature factors of each atom in the residue being smaller than 30 Å^2.
[b]Only the protonated state of $N^{\delta 1}$ is considered for the analysis.

the nitrogen atom of the backbone using planar geometry and the angle H–N–O was computed. The limits used for this angle were $0° \leqslant \theta \leqslant 45°$ (Mitra and Ramakrishnan, 1981). In the case of a hydrogen bond between the side chain amide and main chain carbonyl, as in the case of Asn and His, the ambiguity in the positions of the amide hydrogens was circumvented by calculating the angle C^{γ}–N^{δ}–O. The limits used for identification in these cases were $75° \leqslant \theta \leqslant 165°$.

Results and discussion

The chemical structures of each of the nine polar residues (Arg, Asn, Asp, Gln, Glu, His, Lys, Ser, Thr) were considered as part of a system of six-linked peptide units, with the polar residue at the central C^{α} atom ($C^{\alpha}{}_4$) and checked for all possible side chain–main chain hydrogen bonds leading to the formation of C_7, C_{10} and C_{13} conformations corresponding to the mimics of the γ-turn, β-turn and α-turn, respectively (Toniolo, 1980). Table I gives a number of examples of the various mimics. It can be seen that seven out of the nine polar residues have at least one atom in the side chain which can donate/accept a hydrogen bond leading to the formation of these conformations. The polar atoms of the Arg and Lys side chains cannot form such hydrogen-bonded conformations owing to their greater separation from the main chain. While most of the hydrogen bonds are of the N–H···O type, those of Ser and Thr, which have an oxygen at the γ-position, can form hydrogen-bonded rings of the types stated earlier, only through O–H···O hydrogen bonds. Also, the fact that the number of atoms in the hydrogen-bonded loops are specific makes these bonds highly directional. For instance, the Ser and Thr side chains can form hydrogen bonds of the C_7, C_{10} or C_{13} types only with backbone carbonyl oxygens from residues earlier in the chain. (The possibilities of hydrogen bonds to carbonyl oxygens in the succeeding portion of the chain will result in C_6, C_9 or C_{12} conformations.) Alternatively, the hydroxyl oxygen of Ser/Thr can act as an acceptor for hydrogen bonds from the amide hydrogens of the backbone. This would lead, however, to the formation of hydrogen bonds of the types C_5, C_8 or C_9 with the preceding amides in the chain (N_i, N_{i-1}, N_{i-2}) or of types C_6, C_9 or C_{12} with the succeeding amides (N_{i+1}, N_{i+2}, N_{i+3}). Table I also gives the number of occurrences of the various polar residues in the data set and also those with low *B*-factors. Since this analysis contains results of side chain–main chain hydrogen bonds, it was thought appropriate to use only those residues with low temperature factors for the analysis. Column 3 of Table I gives the number of such residues considered in the analysis.

Out of 4973 Ser and Thr residues considered for the analysis, 674 (13.6%) form hydrogen bonds which mimic the main chain in the one of the three closed-loop conformations. There are a total of 208 examples of the C_7 conformation (51 of Ser and 157 of Thr). The C_{10} conformation does not seem to be very favorable, which is indicated by the relatively fewer examples (a total of 45 with 13 Ser and 32 Thr examples) while the C_{13} conformation appears to be the most favored with 421 examples (217 examples of Ser and 204 examples of Thr).

The δ-atoms of Asn ($O^{\delta 1}$, $N^{\delta 2}$), Asp ($O^{\delta 1}$) and His ($N^{\delta 1}$) are further removed from the main chain by another atom. An increase in the length of the side chain implies an extra freedom about another dihedral angle which in turn implies a greater number of combinations which produce the required conformations. The $O^{\delta 1}$ atoms of the Asn and Asp side chains seem to easily participate in all the three conformations. Almost 22.5% of the 3479 Asn and Asp residues considered form such conformations. There are 101 examples (32 of Asn

and 69 of Asp) of the C_7 conformation, 468 examples (191 of Asn and 277 of Asp) of the C_{10} conformation and 214 (96 of Asn and 118 of Asp) of the C_{13} conformation. Among all the polar residues, these residues show the greatest ability to form the C_{10} conformations. The N^{δ} atoms of Asn and His do not seem to favor the formation of such hydrogen-bonded rings. However, there seems to be a marginal preference for the formation of a C_7 conformation with 72 examples from Asn.

The side chains of Gln and Glu which have polar atoms at the ε-position do not show a significant preference to form such closed-ring conformations. It could be that an extra degree of freedom renders too much flexibility on the side chain which could impede the formation of backbone mimics. However, the O^{ε} atom of Glu and Gln side chain shows only a very marginal preference for the formation of the C_7 conformation while the others are almost absent. It is to be noted that the N^{ε} atom of Gln (or Arg or His) side chain cannot form closed-ringed hydrogen-bonded structures of the types of interest here. Any hydrogen bond to the preceding part of the chain would lead to rings of 9, 12 or 15 atoms whereas those to the succeeding part would lead to 8, 11 or 14 atoms. It is interesting that only the side chains of Asn and Asp are endowed with polar groups at an optimal distance to mimic the backbone. The side chains of other residues are either too constrained for certain conformations or too flexible for others. For the sake of convenience, residues with polar atoms of similar spatial disposition will be grouped together. Ser and Thr will be considered together for interactions through their O^{γ} atom. Asn and Asp will be considered together for hydrogen bonds through the O^{δ} atom, whereas Asn and His for hydrogen bonds through the N^{δ} atom and Glu and Gln for bonds through their O^{ε} atoms.

β-Turn mimic

Classically, a β-turn is characterized by a hydrogen bond between the carbonyl oxygen of residue *i* and the amide nitrogen of residue $i + 3$ forming a closed-ring structure of 10 atoms (Venkatachalam, 1968). Similar C_{10} conformations, forming a ring structure closed by a hydrogen bond, result through a hydrogen bond between the side chain carbonyl group of an Asn or Asp at position i and the backbone amide nitrogen of the residue at position $i + 2$. This conformation has been referred to in the literature as 'pseudo turn' or Asx turn (Tainer *et al.*, 1982; Rees *et al.*, 1983). Figure 1 depicts the remarkable similarity between the standard β-turn and the Asx turn. It can be seen that the side chain of Asx substitutes the backbone with its C^{β} atom in the place of the backbone nitrogen and the peptide plane replaced by the side chain amide group in the case of Asn or the carboxyl group in the case of Asp.

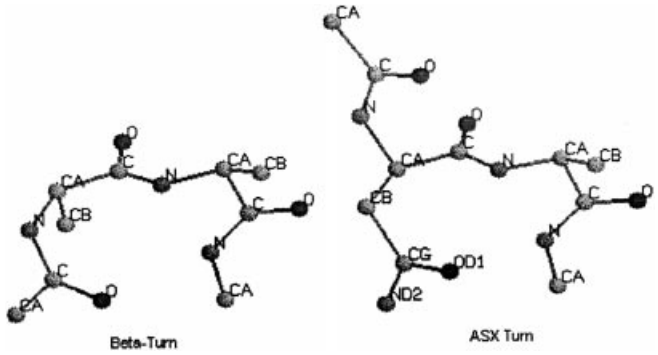

Fig. 1. Similarity of conformation between the β-turn and the Asx-turn. The C^{β} of Asx replaces the N of the backbone amide while the $C^{\gamma}{=}O^{\delta}$ of Asx replaces the backbone carbonyl of the β-turn.

The 468 examples of C_{10} conformations occurring in the data set were analyzed for the distribution of the dihedral angles which characterize the hydrogen bond. Table II shows a 'one to one' comparison of the relevant dihedral angles of the Asx turn *vis-à-vis* the β-turn. In contrast to the β-turn, which is governed by the (ϕ, ψ)s at the central two residues (assuming the planar *trans* conformation for all the intervening peptide units) (Venkatachalam, 1968), the β-turn mimic has freedom about an extra dihedral angle, namely the N^{δ}–C^{γ}–C^{β}–C^{α}, which need not necessarily be planar. The dihedral angles which differentiate the Asx turn from the β-turn are shown in bold in Table II. The angles of the Asx side chain have been defined in a slightly different order from usual in order to permit easy comparison with the β-turn. The distribution of these dihedral angles is shown in Figure 2. The (θ_2, θ_3) plot shows striking similarity to that of β-turns (data not shown) (Venkatachalam, 1968; Wilmot and Thornton, 1990). The points cluster around four regions which are approximately defined by Cluster I → (–60, 120), Cluster II → (–60, –30), Cluster III → (60, –120) and Cluster IV → (60, 30). Maximum population occurs at cluster III. It should be recalled that a cluster in this region would have corresponded to a type II′ β-turn (Wilmot and Thornton, 1990). The clustering of (θ_5, θ_6) points, which corresponds to the (ϕ, ψ) of the residue following Asx, is constrained to the bridge region of the Ramachandran map (Ramakrishnan and Ramachandran, 1965; Ramachandran and Sasisekaran, 1968) between $\psi = \pm 90°$, again very similar to those of β-turns. One can also see from the distribution of θ_1, which corresponds to the planar peptide group in the case of the β-turn, that there are almost no examples in the region $-90° \leqslant \theta_1 \leqslant +90°$, indicating another facet of near similarity to the β-turn.

Figure 3 shows the distribution of the (ϕ, ψ)s of Asx and its following residue and the side chain torsion angles of Asx for the examples which form the C_{10} conformation. It can be seen that the (ϕ, ψ)s of Asx cluster strongly at the extended region of the Ramachandran map with very few examples

Table II. Equivalence of dihedral angles defining the β-turn *vis-à-vis* the Asx-turn

β-Turns	Asx turns	Angle	Remarks
$C^{\alpha}{}_1$–C_1–N_2–$C^{\alpha}{}_2$	$N^{\delta}{}_2$–$C^{\gamma}{}_2$–$C^{\beta}{}_2$–$C^{\alpha}{}_2$	θ_1	Taken as *trans* (~180°) in β-turns
C_1–N_2–$C^{\alpha}{}_2$–C_2	$C^{\gamma}{}_2$–$C^{\beta}{}_2$–$C^{\alpha}{}_2$–C_2	θ_2	ϕ_2 in β-turns
N_2–$C^{\alpha}{}_2$–C_2–N_3	$C^{\beta}{}_2$–$C^{\alpha}{}_2$–C_2–N_3	θ_3	ψ_2 in β-turns
$C^{\alpha}{}_2$–C_2–N_3–$C^{\alpha}{}_3$	$C^{\alpha}{}_2$–C_2–N_3–$C^{\alpha}{}_3$	θ_4	Taken as *trans* in both β-turns/Asx turns
C_2–N_3–$C^{\alpha}{}_3$–C_4	C_2–N_3–$C^{\alpha}{}_3$–C_4	θ_5	ϕ_3 in both β-turns/Asx turns
N_3–$C^{\alpha}{}_3$–C_4–N_4	N_3–$C^{\alpha}{}_3$–C_4–N_4	θ_6	ψ_3 in both β-turns/ Asx turns
$C^{\alpha}{}_3$–C_4–N_4–$C^{\alpha}{}_4$	$C^{\alpha}{}_3$–C_4–N_4–$C^{\alpha}{}_4$	θ_7	Taken as *trans* in both β-turns/Asx turns

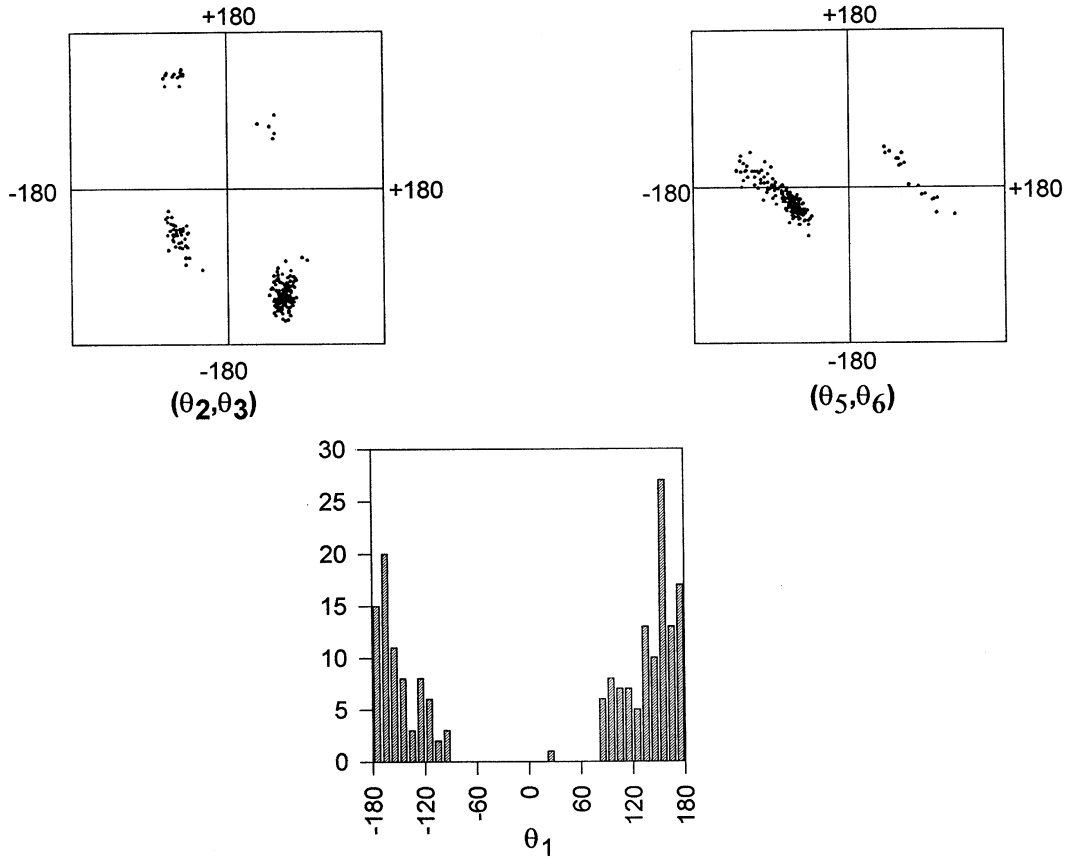

Fig. 2. Distribution of dihedral angles defining the C_{10} conformation (see Table II for definitions). Similarities to the corresponding plots for β-turns can be seen. The largest cluster of (θ_2, θ_3) corresponds to the mimic of the type II′ β-turn.

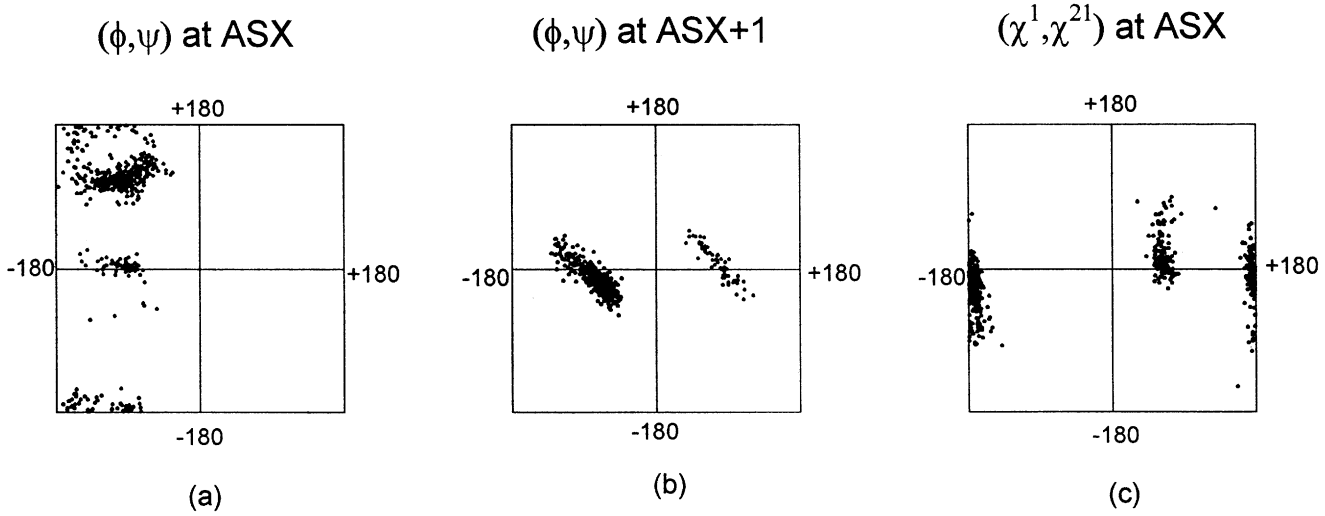

Fig. 3. Distribution of the backbone dihedrals (ϕ, ψ) at (a) Asx and (b) Asx + 1 and (c) the side chain torsion angles χ^1 vs χ^{21} at Asx.

elsewhere. The (χ^1, χ^{21}) points for Asx indicate a preference for a bimodal distribution with the g^- conformer ($\chi^1 \approx -60°$) completely absent. Based on the combinations of these three pairs of dihedral angles, the observed examples can be grouped into four classes. The characteristic range of the angles are shown in Table III. It is interesting to see from that there is a strong distinction in the χ^1 values between classes I and II and also classes III and IV. Class I is characterized by a fully extended state for the Asx. There are 82 examples of this class which fall under the motif described by $\beta'(g^+)\alpha_R$. The β′ here is used to describe the completely extended conformations occurring on the top and bottom edges of the Ramachandran map. The maximum number of examples occur in class II which is described by the motif $\beta(t)\alpha_R$. There are 326 examples of this type which differ from the examples of class I by a small difference in the ψ value but which requires a completely different orientation of the side chain to effect ring closure through the hydrogen bond. While the side chain orientation of the Asx residues of class I are in the g^+ conformation, those in this class are in the *t* conformation. Also, the χ^1

Table III. Parameters differentiating the various classes of C_{10} conformations formed by Asx residues

Class	(ϕ, ψ) at Asx	χ^1 at Asx	χ^{21} at Asx	(ϕ, ψ) at Asx + 1	No. of examples
I	Completely extended	g^+	~0 to +90°	α_R	82
II	Extended	t	~–90 to +90°	α_R	326
III	Bridge R	g^+	~0°	α_L	55
IV	Right-handed α-helical	t	~–90°	α_L	5

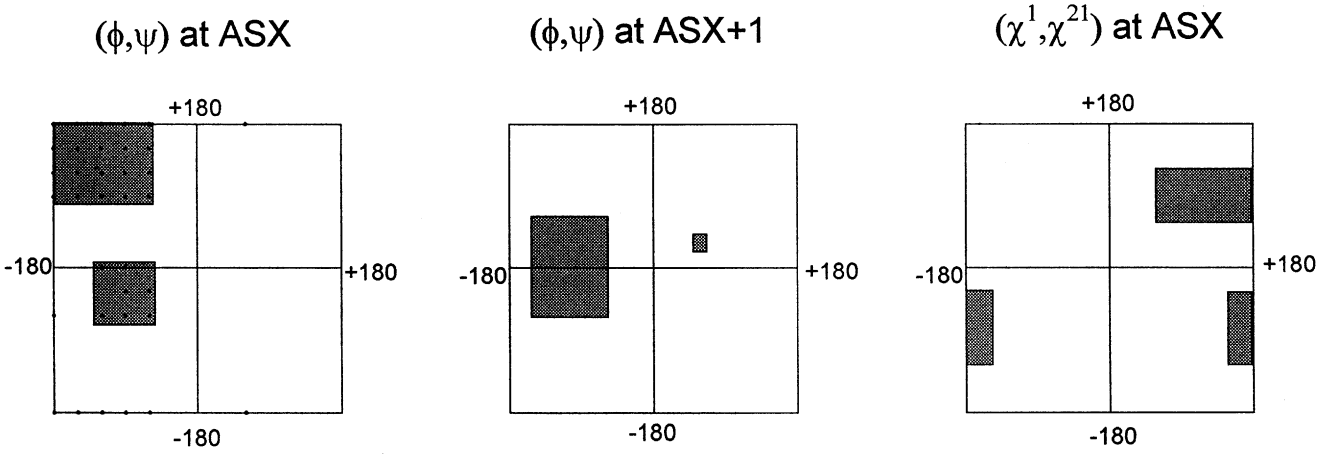

Fig. 4. The stereochemically permitted ranges of the (ϕ, ψ) angles at Asx, Asx + 1 and the side chain rotamers of Asx, calculated using a systematic conformational search technique.

values of class II seem to have a greater freedom indicated by the larger spread of the distribution shown in Figure 3(c). The class III which has 55 examples is characterized by the (ϕ, ψ) values of the Asx lying in the bridge region of the Ramachandran map. The side chain conformer observed for these examples is the g^+ conformer. This along with the examples of class IV require the residue following Asx to be in the α_L conformation. There are just five examples in class IV. Interestingly, as in the case of the first two classes, a small difference in the ψ-values at Asx leads to entirely different orientations of the side chain between the examples of class III and IV.

In order to explore other stereochemically allowed conformations with no observed examples in the data set, a theoretical conformational search was performed by systematically varying each of the six dihedral angles in intervals of 30° and checking for the required hydrogen bond. The output of the search algorithm was screened by eliminating conformations with positive non-bonded energies or with (ϕ, ψ)s lying outside the allowed regions of the Ramachandran map. Figure 4 shows the regions of the Ramachandran map which allow the formation of the C_{10} conformation, for Asx and its following residue. The comparison of Figures 3 and 4 shows that theoretical results agree with the observed data. Three points of importance arising out of the results are that (1) at the position of the Asx while both the α_R and β regions of the Ramachandran map favor formation of such hydrogen bonds, the α_L region is completely disfavored; (2) the conformation of the residue following Asx is strongly restricted to the region of the map enclosed between $\psi = \pm 90°$; and (3) the possibility of a g^- conformer is indeed completely ruled out.

Recently, Wilson and Finlay (1997) reported an analysis on such C_{10} conformations followed by a 4→1 backbone–backbone hydrogen bond and a proline following Asx, which they termed the Asx–Pro turn. From the present analysis, it is very evident that since most of the examples have an α_R following Asx, the constrained conformation of proline makes it an obvious choice to occur following the Asx. The propensity of amino acids to occur at the position following Asx, shown in Figure 5, indicates that though Pro shows an overwhelming preference to be at this position, Gly too shows a moderate preference. Asn, Lys, Met, Ser, Thr and Trp with propensity values close to 1 indicate neither favor nor disfavor.

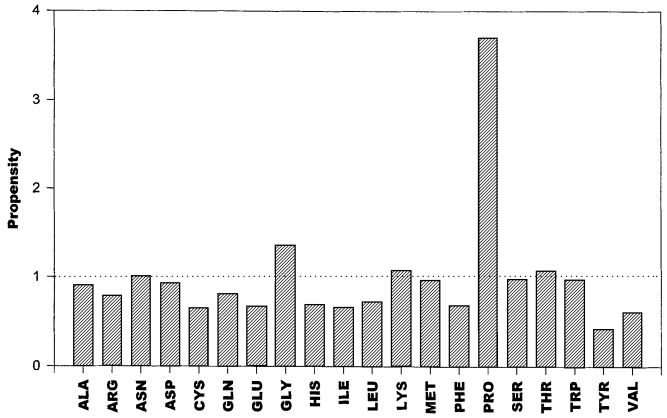

Fig. 5. Propensity of amino acid residues that follow Asx residue when the side chain mimics a C_{10} conformation.

Richardson and Richardson (1989) had indicated that a possibility to form such hydrogen-bonded structures would make Asx a less likely residue at position 2 of β-turns, since it would then compete with the backbone carbonyl group to form a hydrogen bond with the amide group of residue 4 (see Figure 1). In complete agreement with this observation Asx has been shown to have extremely low propensities for position

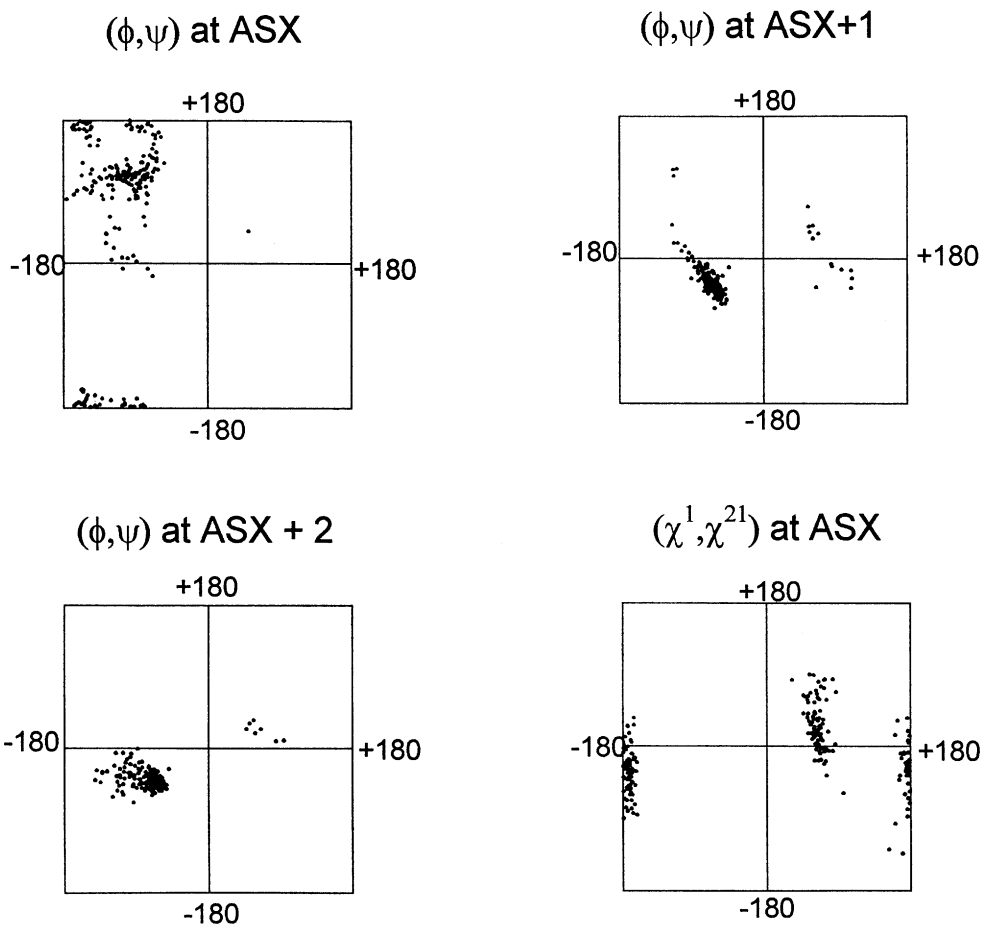

Fig. 6. Distribution of (ϕ, ψ) angles at Asx, Asx + 1, Asx + 2 and the side chain rotamers of Asx residues adopting a C_{13} conformation.

2 of any of the β-turns except the type I′ which requires an α_L at position 2 of the turn (Wilmot and Thornton, 1988; Hutchinson and Thornton, 1994; Ramakrishnan *et al.*, 1996). Interestingly, it can be seen from Figure 4 that an Asx in an α_L conformation cannot form such a hydrogen bond. Also, it has been shown that Asx residues have a great tendency to occur at position 1 of type I β-turns (Hutchinson and Thornton, 1994; Srinivasan *et al.*, 1994). These turns require an α_R conformation at position 2 and from the data presented above one can then guess that it presents a favorable situation for the formation of the C_{10} conformation. This hydrogen bond would help stabilize the turn by satisfying the hydrogen bond potential of the central peptide unit. Indeed, 294 out of the 468 examples analyzed are at position 1 of type I turns (or the very much similar type III) and 68 examples out of the 294 fall in the category of Asx–Pro turns as described by Wilson and Finlay (1997).

α-Turn mimic

The α-turns are characteristic of the helical segments of proteins. It is a three-residue turn requiring the description of the middle three (ϕ, ψ)s for its definition and is also characterized by a hydrogen bond between the backbone carbonyl and amide groups of residue *i* and *i* + 4, respectively (C_{13} conformation) (Nataraj *et al.*, 1995). The mimicry of such turns is frequently formed by Asx residues involving a hydrogen bond between the side chain carbonyl oxygen of Asx and the main chain amide of residue Asx + 3, as indicated in Table I. The defining parameters of such mimics would then be the (ϕ, ψ)s at Asx and its two following residues and also the (χ^1, χ^{21}) at Asx. The distributions of these parameters are shown in Figure 6. It can be seen again that the Asx residues cluster at the extended region of the (ϕ, ψ) space while the (ϕ, ψ)s of the following two residues predominantly cluster at α_R. The side chain torsion angles show a bimodal distribution, similar to that of the β-turn mimics, with the g^- conformer completely absent. A description of the motifs formed and the number of examples of each kind is presented in Table IV. Out of a total of 210 examples, almost as many as 119 examples fall under the motif described by $\beta(t)\alpha_R\alpha_R$. The $\beta'(g^+)\alpha_R\alpha_R$ motif containing 75 examples differs from the previous motif at the conformation of Asx by a change in its ψ value and also its side chain conformer.

Table IV. Motifs of various C_{13} conformations formed by Asx and the number of examples of each motif

Motif	No. of examples
$\beta'(g^+)\alpha_R\alpha_R$	75
$\beta'(g^+)\beta\alpha_L$	1
$\beta(t)\alpha_R\alpha_R$	119
$\beta(t)\beta\alpha_L$	2
$\beta_R(g^+)\alpha_L\alpha_R$	9
$\beta_R(g^+)\alpha_L\alpha_L$	4

Several analyses have shown that such a hydrogen-bonded motif is characteristic of helix start sites (Presta and Rose, 1988; Richardson and Richardson, 1988; Doig *et al.*, 1997;

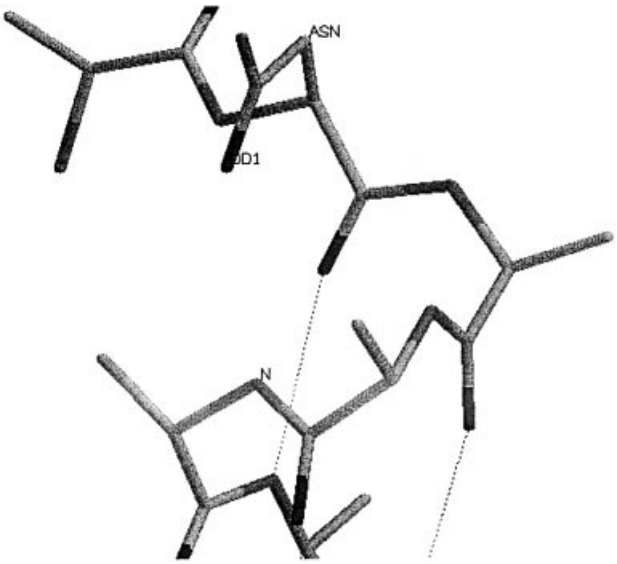

Fig. 7. An example of a helix-initiating site by an Asx residue mimicking an α-turn (1ABE, Asn205).

Aurora and Rose, 1998). Out of the 210 examples of C_{13} conformations observed, 129 examples initiate a helix at position Asx + 1. At helix start sites, the C_{13} conformation mimics the backbone to form a turn of the helix, but in effect leads the backbone away from helical conformation as shown in Figure 7. Among the 81 examples which do not initiate an α-helix, it was found that in 26 examples the Asx residue is at the starting position of an α-turn. This is also in agreement with calculations on the preference of various amino acids to occur at each of the five positions of α-turns, which indicates that Asn is a preferred residue at position 1 of α-turns (Nataraj *et al.*, 1995).

Alternatively, 13-atom mimics can also be formed by an O–H···O hydrogen bond between the hydroxyl oxygen of a Ser or a Thr with the backbone carboxyl oxygen of residue $i - 2$. There are 421 examples of this kind as shown in Table I. The number of dihedral angles defining such conformations is less than that in the Asx case. The distributions of the (φ, ψ)s of the residues in the C_{13} ring and the χ^1 of Ser/Thr are shown in Figure 8. The χ^1 angles are clustered around the g^+ conformation, with the *t* and g^- conformers almost completely absent. It can be seen from the plots that most of the (φ, ψ) points cluster around the region of the right-handed α-helix.

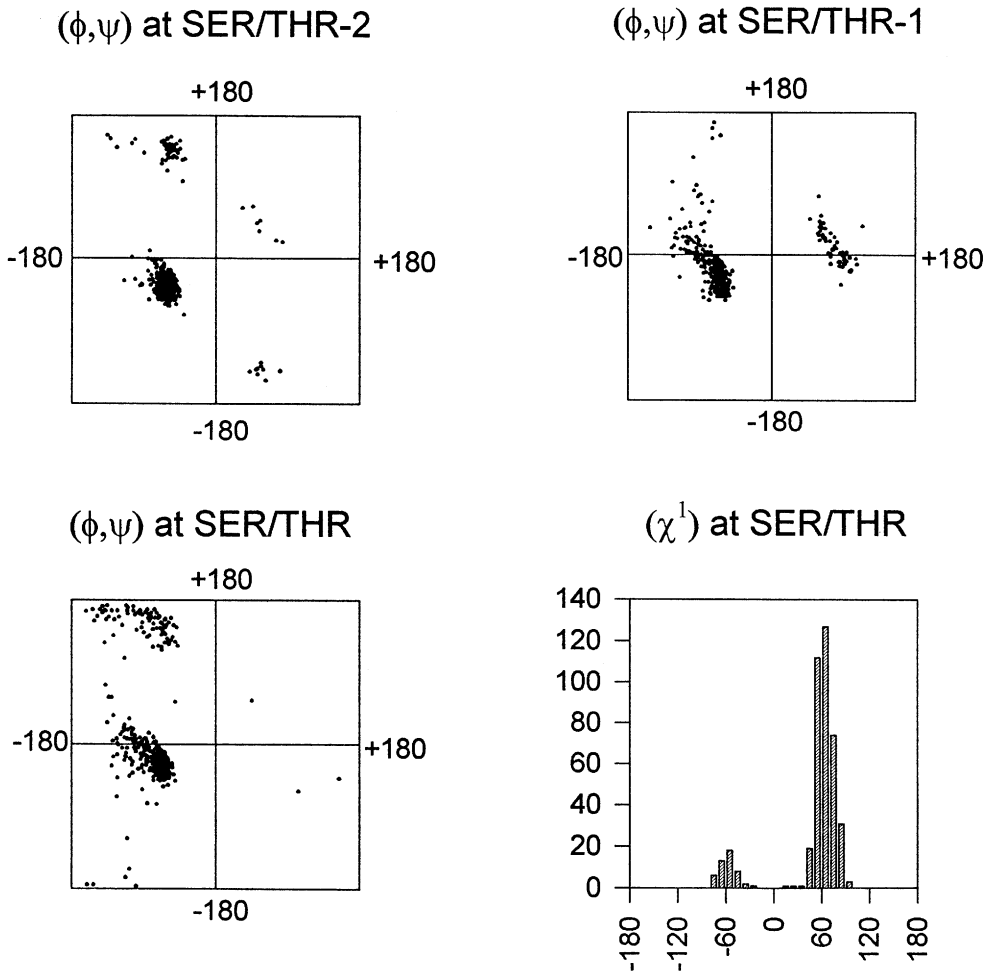

Fig. 8. Dihedral angle distributions of the Ser/Thr examples which form C_{13} conformations. These examples are characterized by an O–H···O hydrogen bond to the preceding part of the chain.

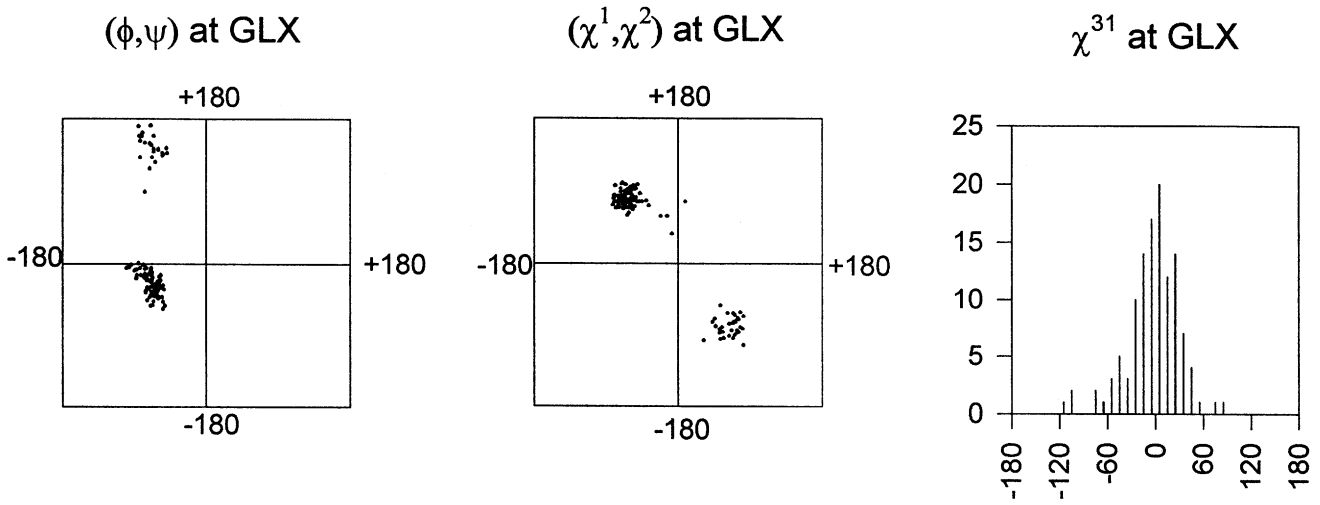

Fig. 9. Distributions of the (ϕ, ψ) angles of Glx residues forming C_7 conformations along with the distributions of the side chain rotamers.

Indeed, 223 out of the 421 examples are part of regular α-helices. This indicates that these hydrogen bonds are mostly the effect of the helical nature of the surrounding residues, offering stability to the helical conformation of the polypeptide chain (Stickle *et al.*, 1992).

γ-Turn mimics

The formation of the C_7 conformation [a mimic of the γ-turn (Milner-White, 1990; Milner-White *et al.*, 1988)] also seems to be a fairly common occurrence in protein structures. From Table I it can be seen that there are 101 such conformations formed by the O^{δ} atom of Asx residues, 82 examples formed by the N^{δ} atom of Asn or His and 118 examples formed by O^{ε} atom of Glx residues. Apart from these N–H···O hydrogen bonds there are a further 208 examples from Ser/Thr residues in which the O^{γ} atom is hydrogen bonded to the main chain carboxyl oxygen of the previous residue. The plots of the relevant dihedral angles for the 118 examples of the Glx residues are shown in Figure 9. Here the O^{ε} atom of the Glx side chain is hydrogen bonded to the main chain amide nitrogen of the same residue. Therefore, the dihedral angles which determine conformation are only the (ϕ, ψ) at Glx and the three side chain torsion angles. It can be seen from Figure 9 that the (ϕ, ψ)s are strongly clustered at the region of the α-helix with only a very sparse population at the β-region. The (χ^1, χ^2) plot reveals that clustering is very strong around the regions (±60°, ±60°). The χ^{31} distribution shows almost a normal distribution centered around the 0° value. The other types of C_7 conformation do not show any strong features. The dihedral angle points, in these cases, are spread out in the (ϕ, ψ) space, indicating that they might be artifacts of the polypeptide chain geometry.

Conclusion

Formation of secondary structure-like conformations through hydrogen bonds from the side chain to the main chain plays very important roles in protein structures. As Lewis *et al.* (1973) suggested, the β-turns might be the nucleation center for the reversal of chain direction, in which case the stability of the turn conformation becomes extremely crucial. In such an event, the β-turn mimic or the C_{10} conformation could play an important role providing extra stability to the turn in assisting chain reversal. The importance of the C_{13} conformation at the initiation sites of α-helices is already well known. It serves an important purpose of initiating the helix by keeping the following residues in an α-helical conformation through the hydrogen bond. The participation of Asx residues in such mimicry also displays the elegance with which nature has selected the 20 amino acids. From the data presented earlier, it is very clear that neither the shorter side chains of Ser/Thr nor the longer side chains of Glx are favorable for β-turn and α-turn mimicry. The optimally placed polar groups of Asx seem to be the best candidates for the purpose. The Asx/Glx residues additionally show favorability to form γ-turn mimics. These mimics might be essential in stabilizing the conformation of single residues (Karplus, 1996).

Arguments similar to those presented here have been used to model the Asn-rich regions of the circumsporozite coat protein of *Plasmodium falciparum* in the absence of its experimental structure (Wilson *et al.*, 1997). The results presented here can also be used to identify the side chain carbonyl oxygen and amides which are otherwise sometimes ambiguously assigned (McDonald and Thornton, 1994). Further, owing to the close resemblance of the Asx/Glx side chains to the backbones of the non-standard β/γ-amino acids, the results presented here might augment *de novo* design of peptides containing these (Appella *et al.*, 1997; Hintermann and Seebach, 1997).

Acknowledgments

The authors thank Dr N.Srinivasan for useful discussions and suggestions in the preparation of the manuscript.

References

Appella,D.H., Christianson,L.A., Klein,D.A., Powell,D.R., Huang,X., Barchi, J.J. and Gellman,S.C. (1997), *Nature*, **387**, 381–384.
Aurora,R. and Rose,G.D. (1998) *Protein Sci.*, **7**, 21–38.
Baker,E.N. and Hubbard,R.E. (1984) *Prog. Biophys. Mol. Biol.*, **44**, 97–179.
Bernstein,F.C., Koetzle,T.G., Williams,G., Meyer,E.,Jr, Brice,M.D., Rogers, J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) *J. Mol. Biol.*, **112**, 535–542.
Doig,A.J., MacArthur,M.W., Stapley,B.J. and Thornton,J.M. (1997) *Protein Sci.*, **6**, 147–155.
Finkelstein,A.V. and Ptitsyn,O.B. (1977) *Biopolymers*, **16**, 469–495.
Finkelstein,A.V., Ptitsyn,O.B. and Kozitsyn,S.A. (1977) *Biopolymer.*, **16**, 497–524.
Gunasekaran,K., Ramakrishnan,C. and Balaram,P. (1997) *Protein Engng*, **10**, 101–111.
Hintermann,T. and Seebach, D. (1997) *Chimia*, **50**, 244–247.

Hutchinson,E.G. and Thornton,J.M. (1994) *Protein Sci*., **3**, 2207–2216.
Ippolito,J.A., Alexander,R.S. and Christianson,D.W. (1990) *J. Mol. Biol*., **215**, 457–471.
Karplus,P.A. (1996) *Protein. Sci*., **5**, 1406–1420.
Lewis,P.N., Momany,F.A. and Scheraga,H.A. (1973) *Biochim. Biophys. Acta*, **303**, 211–229.
McDonald,I.K. and Thornton,J.M. (1994) *Protein Engng*, **8**, 217–224.
Milner-White,E.J. (1990) *J. Mol. Biol*., **216**, 385–397.
Milner-White,E.J., Ross,B.M., Ismail,R., Belhadj-Mostefa,K. and Poet,R. (1988) *J. Mol. Biol*., **204**, 777–782.
Mitra,J. and Ramakrishanan,C. (1981) *Int. J. Pept. Protein Res*., **17**, 401–411.
Nataraj,D.V., Srinivasan,N., Sowdhamini,R. and Ramakrishnan,C. (1995) *Curr. Sci*., **69**, 434–447.
Pauling,L., Corey,R.B. and Branson,H.R. (1951) *Proc. Natl Acad. Sci. USA*, **37**, 205–211.
Presta,L.G. and Rose,G.D. (1988) *Science*, **240**, 1632–1641.
Questel,J.L., Morris,D.G., MacCallum,P.H., Poet,R. and Milner-White,E.J. (1993) *J. Mol. Biol*., **231**, 888–896.
Ramachandran,G.N. and Sasisekaran,V. (1968) *Adv. Protein Chem*. **23**, 283–437.
Ramakrishnan,C. and Prasad,N. (1971) *Int. J. Protein Res*., **3**, 209–231.
Ramakrishnan,C. and Ramachandran,G.N. (1965) *Biophys. J*., **5**, 909–933.
Ramakrishnan,C., Srinivasan,N. and Nararaj,D.V. (1996) *Int. J. Pept. Protein Res*., **48**, 420–428.
Rees,D.C., Lewis,M. and Lipscomb,W.N. (1983) *J. Mol. Biol*., **168**, 367–387.
Richardson,J.S. (1981) *Adv. Protein Chem*., **34**, 168–339.
Richardson,J.S. and Richardson,D.C. (1988) *Science*, **240**, 1648–1652.
Richardson,J.S. and Richardson,D.C. (1989) In Fasman,G.D. (ed.), *Prediction of Protein Structure and the Principles of Protein Conformation.* Plenum Press, New York, pp. 1–98.
Rose,G.D. and Wolfenden,R. (1993) *Annu. Rev. Biomol. Struct*., **22**, 381–415.
Srinivasan,N., Anuradha,V.S., Ramakrishnan,C., Sowdhamini,R. and Balaram,P. (1994) *Int. J. Pept. Protein Res*., **44**, 112–122.
Stickle,D.F., Presta,L.G., Dill,K.A. and Rose,G.D. (1992) *J. Mol. Biol*., **226**, 1143–1159.
Tainer,J.A., Getzoff,E.D., Beem,K.M., Richardson,J.S. and Richardson,D.C. (1982) *J. Mol. Biol*., **160**, 181–217.
Toniolo,C. (1980) *CRC. Rev. Biochem*., 1–44.
Venkatachalam,C.M. (1968) *Biopolymers*, **6**, 1425–1436.
Wilmot,C.M. and Thornton,J.M. (1988) *J. Mol. Biol*., **203**, 221–232.
Wilmot,C.M. and Thornton,J.M. (1990) *Protein Engng*, **3**, 479–493.
Wilson,D.R. and Finlay,B.B. (1997) *Protein Engng*, **10**, 519–529.
Wilson,D.R., Wirtz,R.A. and Finlay,B.B. (1997) *Protein Engng*, **10**, 531–540.

Received November 30, 1998; revised February 19, 1999; accepted March 3, 1999