Protein Science (1999), *8*:573–586. Cambridge University Press. Printed in the USA.

Analysis of interactive packing of secondary structural elements in α/β units in proteins

BOOJALA V.B. REDDY,[2] HAMPAPATHALU A. NAGARAJARAM,[1] AND TOM L. BLUNDELL[1]

[1]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Old Addenbrooks Site, Cambridge CB2 1GA, United Kingdom
[2]Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad-500 007, India

(RECEIVED July 9, 1998; ACCEPTED November 9, 1998)

Abstract

An α-helix and a β-strand are said to be interactively packed if at least one residue in each of the secondary structural elements loses 10% of its solvent accessible contact area on association with the other secondary structural element. An analysis of all such 5,975 nonidentical α/β units in protein structures, defined at $\leq$2.5 Å resolution, shows that the interaxial distance between the α-helix and the β-strand is linearly correlated with the residue-dependent function, $\log[(V/nda)/n\text{-}int]$, where V is the volume of amino acid residues in the packing interface, *nda* is the normalized difference in solvent accessible contact area of the residues in packed and unpacked secondary structural elements, and *n-int* is the number of residues in the packing interface. The β-sheet unit (βu), defined as a pair of adjacent parallel or antiparallel hydrogen-bonded β-strands, packing with an α-helix shows a better correlation between the interaxial distance and $\log(V/nda)$ for the residues in the packing interface. This packing relationship is shown to be useful in the prediction of interaxial distances in α/β units using the interacting residue information of equivalent α/β units of homologous proteins. It is, therefore, of value in comparative modeling of protein structures.

Keywords: α-helix; α/β packing; β-strand; comparative modeling; packing geometry; protein structure prediction

Alpha-helices and β-sheets are the secondary structural elements that form the cores of most protein structures (Levitt & Chothia, 1976; Richardson, 1981). Similar three-dimensional structures are observed in the families of homologous proteins. Sequence variations in each family, resulting from insertions, deletions, and substitutions, are mostly found at the surface regions of the structures, but some mutations are also accommodated within the core (Lesk & Chothia, 1980, 1986; Bajaj & Blundell, 1984; Chothia & Lesk, 1987; Hilbert et al., 1993). These observations are central to comparative protein modeling of three-dimensional structures of proteins, a technique in which the structure of one or more proteins, defined experimentally, is used to model a homologue (Browne et al., 1969; Greer, 1981; Chothia et al., 1986; Blundell et al., 1987, 1988; Sutcliffe et al., 1987a, 1987b; Havel & Snow, 1991; Sali & Blundell, 1993; Srinivasan et al., 1993; Johnson et al., 1994; Bajorath et al., 1993; Sali, 1995; Rost & Sander, 1996; Sanchez & Sali, 1997). There are numerous predicted models of homologous proteins in the literature using these techniques (see Bajorath et al., 1993; Mosimann et al., 1995; Martin et al., 1997). The models are observed to be comparable to medium resolution X-ray structures where sequence identity between the homologue and target protein is greater than 40% (Srinivasan & Blundell, 1993; Sali et al., 1995). The predictions become more promising and reliable as sequence identity increases.

However, modeling homologous proteins with a significant number of substitutions in the core region leads to considerable changes in residue volumes and other residue-dependent properties of the densely packed region. These residue variations are accommodated by relative shifts and rotations of the secondary structural elements (Lesk & Chothia, 1980; 1986; Chothia & Lesk, 1982). As a consequence the root-mean-square (RMS) differences increase, and the numbers of topologically equivalent residues of the common core decrease, as the sequence differences become larger for pairs of homologous structures (Chothia & Lesk, 1986). Thus, if the protein to be modeled is distantly related to the homologues of known structure, the framework is inevitably biased toward the structure(s) used and the resulting model may be considerably in error. To take care of such distortions in comparative modeling procedures, a method is required to predict relative shifts and rotations in the secondary structural elements (SSEs) as a function of changes in amino acids among the homologous structures. We have, therefore, undertaken a systematic analysis of the quantitative relationships that may exist between residues involved in interactive packing and the geometry of SSEs in protein structures.

Analysis of the packing between β-sheets in the immunoglobulin and the plastocyanin–azurin families (Chothia & Lesk, 1982; Lesk & Chothia, 1982) showed that mutations (insertions, deletions, and substitutions) are accommodated by displacements and rotations of the sheets relative to each other and also through the

Reprint requests to: Tom L. Blundell, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Old Addenbrooks Site, Cambridge CB2 1GA, United Kingdom; e-mail: tom@cryst.bioc.cam.ac.uk.

formation of β-bulges. Similar principles for the accommodation of the mutations are reported for the packing geometry of helices in globins (Lesk & Chothia, 1980). Until our recent studies (Reddy & Blundell, 1993), there were no reports in the literature pertaining to general quantitative relationships between the geometry of packing of SSEs and the residue-dependent parameters in their packing interface.

Studies of various aspects of packing of α-helices in proteins (see Reddy & Blundell, 1993; Mumenthaler & Braun, 1995; Walther et al., 1996) have been reported. From our helix–helix packing analysis, we have been successful in identifying a quantitative relationship that exists between interhelix distance and the volume-dependent function of the residues in the packing interface (Reddy & Blundell, 1993). The analysis shows that such a relationship is useful for predicting interhelix distances when amino acids in the packing interface between homologous helix pairs are substituted. We have also observed a similar packing relationship in the pairs of β-strands and β-sheet units (H.A. Nagarajaram, B.V.B. Reddy, & T.L. Blundell, unpubl. obs.).

The classical work of Levitt and Chothia (1976) and Chothia et al. (1977) on α/β packing describes rules and models for the general arrangements of α-helices and β-sheets in proteins. The structural principles, geometry, and constraints of their packing were also studied extensively (Janin & Chothia, 1980; Cohen et al., 1982; Chothia, 1984). The anatomy and side-chain packing in these units are more ordered in α/β barrel proteins (Lesk et al., 1989; Farber & Petsko, 1990; Raine et al., 1994; Vtyurin & Panov, 1995). The packing energetics between α-helix and β-sheet have been studied by Chou et al. (1985). The structural similarity between α/β proteins and all-β-proteins has been discussed by Efimov (1995). The principles of design of α/β barrels have been explored by Handel (1990) and Lasters et al. (1990). More recently a structural classification of $\alpha\beta\beta$ and $\beta\beta\alpha$ supersecondary structure units in proteins, based on their geometry and connectivity, has been described (Boutonnet et al., 1998).

The complementary twist model (first approximation model) for the geometry of α-helix packing onto parallel or antiparallel pleated sheets prefers a near-parallel orientation (Chothia et al., 1977; Janin & Chothia, 1980; Chothia, 1984). This was confirmed by Cohen et al. (1982), who presented a quantitative study of contact area and shape of surface of α-helix/β-sheet units as a function of their interaxial angles. However, there is no report that quantifies the distance between the α-helix and the β-sheet unit as a function of the residues in the packing interface.

In this paper we report our studies of the quantitative relationship between the α/β interactive-packing distance and the residues in the packing interface. We have investigated interactive packing between the α-helix and the β-strand (α/β) and between α-helix and the β-sheet unit (α/βu) in terms of the distances between the elements, the dihedral angles between the axes, the number and nature of residues involved in packing and the observed relationship between the distance and volume-dependent function of the residues in the packing interface. The analysis also provides a useful basis for predicting the distances between these units as a function of residues involved in the packing interface.

Results and discussion

Lesk and Chothia (1980) define SSEs as close packed when atoms of the two SSEs lie within the distance of the sum of their van der Waals radii plus 0.6 Å. On the other hand, Reddy and Blundell (1993) have defined interactive packing between the secondary structural elements in proteins with respect to the loss of solvent accessible contact area (SACA) (Richmond & Richards, 1978) in the presence of another SSE. We have followed this approach as it accounts for many aspects of interactions (hydrophilic and hydrophobic) involved in cores of proteins. This definition of interactive packing also involves a surface region of SSEs, which is probably more appropriate in a packing analysis. It is supported by free energy considerations, which show that the loss of every 1 Å^2 of SACA contributes about 80 cal to the free energy of hydrophobic association of SSEs (Chothia, 1974; Richmond & Richards, 1978).

The interactive packing in α/β units is analyzed both in terms of interactions between an α-helix and a single β-strand (α/β) (Fig. 1), and interactions of an α-helix and a β-sheet unit comprising of two adjacent, hydrogen-bonded strands (α/βu) (Fig. 2). We, therefore, discuss our observations for these separately.

Alpha-helix/β-strand (α/β) packing

α/β units, which have an α-helix length of seven residues or more and a β-strand length of five residues or more, are considered in

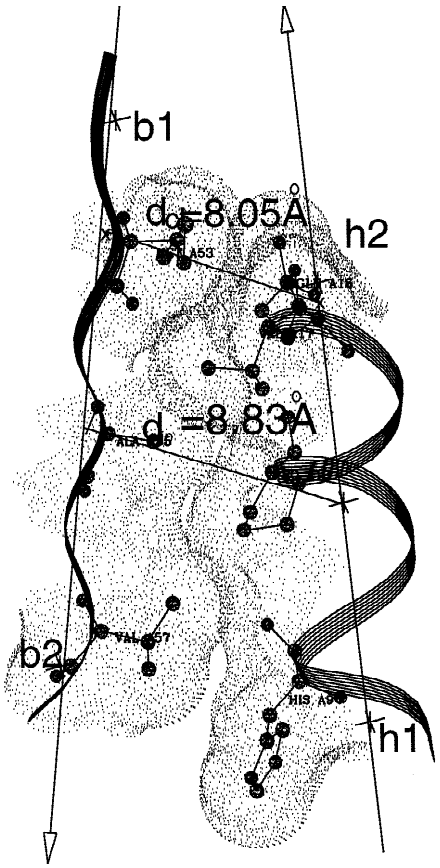

Fig. 1. Interactive packing of an α/β from 2ploA. Axes point in the N- to C-terminal direction. The solvent accessible contact area of interactively packed residues is shown in dots with the corresponding residue atoms in spheres. The extended interacting region on the axis of a β-strand is between *b1* and *b2*, and for a helix is between *h1* and *h2*. The inter-SSE distances, d_{ip} and d_{cl}, are also illustrated.

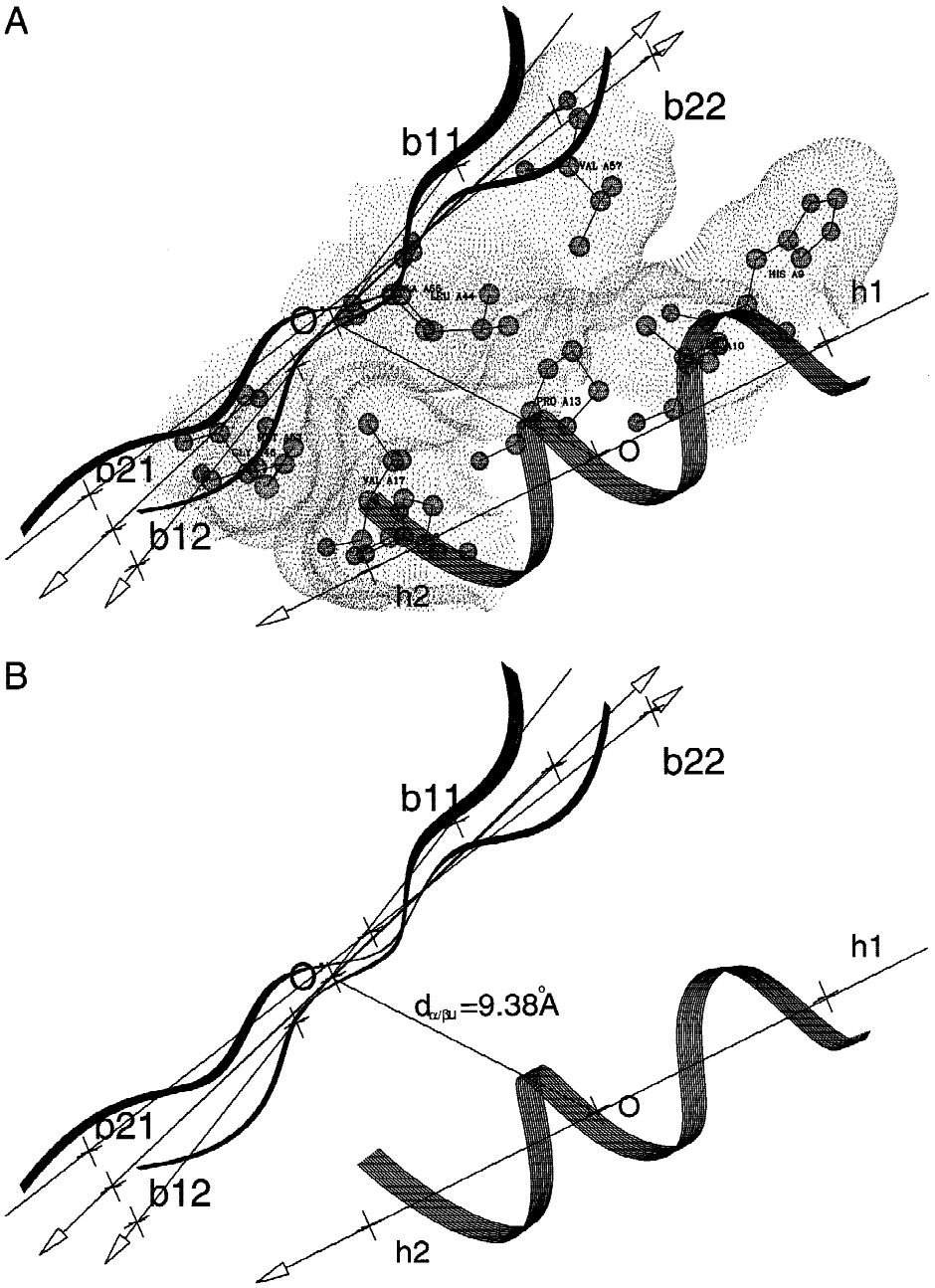

Fig. 2. Interactive packing of an α/βu from 2ploA: (**A**) showing the solvent accessible contact area of interacting residues in dots with the corresponding atoms in spheres and (**B**) showing the main-chain trace, secondary structure axes, ortho-center of βu, and the inter-α/βu distance $d_{\alpha/\beta u}$.

this analysis. For α/β units of identical length and sequence, only that from the structure defined at the best resolution was kept. The data set 11 (*D11*) with 10,362 such α/β units was considered for analysis. The SSEs interacting only through the residues at one of the termini (Fig. 3A-C) were excluded, so reducing the total to 6,163 (*D12*). We also removed the α/βs with at least one distance less than 3.5 Å between a H-bond acceptor on one SSE main chain and a donor on the other SSE main chain (Baker & Hubbard, 1984) (Fig. 3D). Side-chain packing of such α/βs is unlikely to determine the inter-α/β distance and geometry. The remaining 5,975 (*D13*) α/βs were used for further analysis.

Amino acid residue-dependent parameters and inter-α/β distance

The regression coefficient values (r), for data sets *D11*, *D12*, and *D13*, are given in Table 1 for various residue-dependent functions (RDFs) of the amino acids in the packing interface. Only *F1* to *F5*, of the functions that were discussed in Reddy & Blundell (1993) and that gave a regression coefficient value $\geq$0.65, are presented in Table 1. The *nda* and *n-int* of the residues in the interacting region are inversely proportional to the inter-α/β distance(s). There is also a rough inverse correlation between inter-α/β dis-

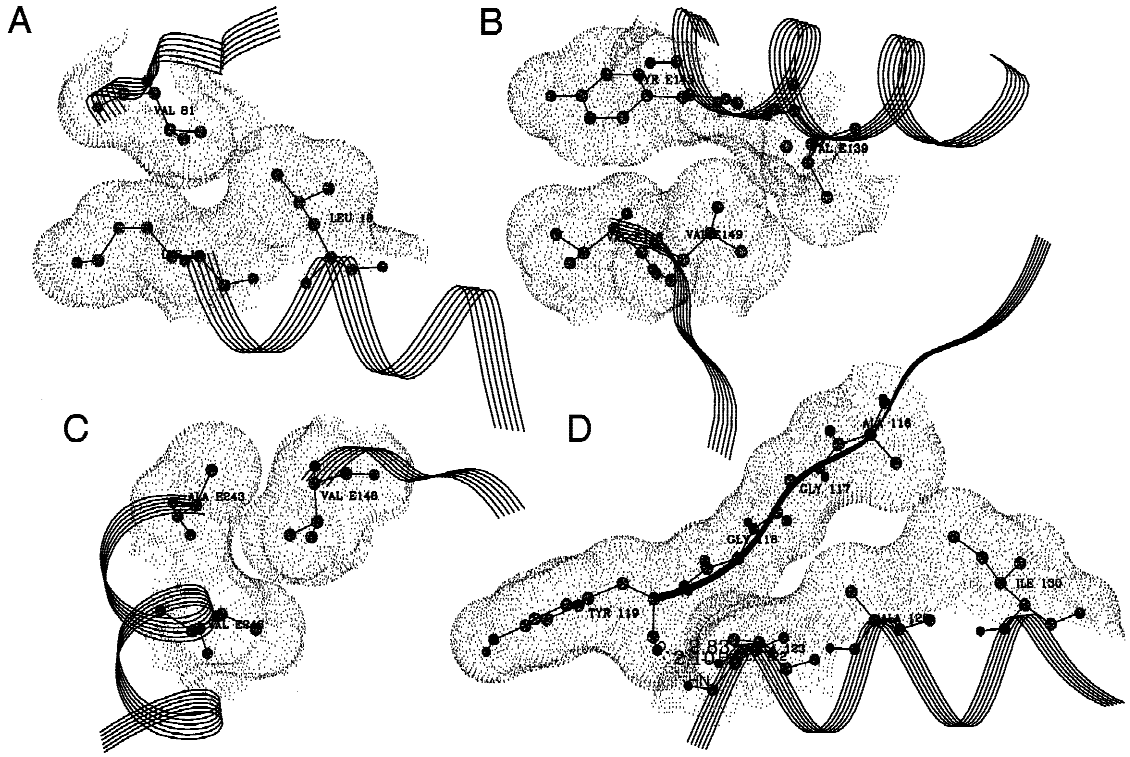

Fig. 3. Representative examples of α/β units: α/βs with interacting residues from the termini of (**A**) α-helix (from 5p21), (**B**) β-strand (from 1cseE), (**C**) both α-helix and β-strand (from 1cseE), and (**D**) interactively packed α/β with a hydrogen bond between the main-chain atoms of two SSEs (from 1cus). Note that all such pairs have been removed from the data set of α/βs (*D11*) and further analysis is carried out on the remaining α/βs (*D13*).

tances and interacting residue volumes (V). These relationships arise because at the smaller separation distances a greater number of residues are likely to interdigitate than at the larger distances. Also, at a given distance the number of interacting residues varies as the angle of packing varies (maximum for near parallel/antiparallel orientations and minimum for crossed orientations). In general it is both the distance and the angle that affect the residue-burial in packing. The functions *F1* to *F5* (derived from *V, nda*, and *n-int*) have a significant correlation with the inter-α/β distance(s), d_{cl} and d_{ip} (Table 1; Figs. 4, 5A). The best correlation(s) is observed for *F4* and the α/β distances d_{cl} and d_{ip}. In the case of *D11, F4* vs. d_{cl} has a higher regression coefficient value than that of *F4* vs. d_{ip}. The regression coefficient values progressively increase as the sample is refined to exclude contacts between termini and those with main-chain hydrogen bonding α/βs. For further analysis, we have considered only the *D13* data set and the α/β distance d_{ip}, as they give best correlation(s) with all the RDFs. The best regression coefficient value is for *F4* vs. d_{ip} (for *D13* r = 0.77), which we denote as the "packing relationship" of α/βs in protein structures (Fig. 5).

Table 1. *Regression coefficient (r) values for various amino acid residue-dependent functions (RDFs) in the interacting region with corresponding α/β distance(s)*[a]

No. α/βs	*D11* (10,362)		*D12* (6,163)		*D13* (5,975)	
Distance in Å	d_{cl}	d_{ip}	d_{cl}	d_{ip}	d_{cl}	d_{ip}
nda	−0.64	−0.58	−0.66	−0.68	−0.67	−0.69
V	−0.43	−0.35	−0.36	−0.36	−0.38	−0.37
n-int	−0.54	−0.47	−0.54	−0.54	−0.55	−0.55
F1 = *V/nda*	0.70	0.71	0.70	0.74	0.71	0.74
F2 = *F1/n-int*	0.69	0.65	0.67	0.67	0.68	0.67
F3 = log(*F1*)	0.71	0.73	0.71	0.76	0.72	0.76
F4 = log(*F2*)	0.73	0.68	0.75	0.77	0.76	0.77
F5 = $(F1)^{1/3}$	0.71	0.73	0.71	0.75	0.72	0.76

[a] *D11* = unfiltered α/βs; *D12* = α/βs after removing interacting residues at the SSE-termini; *D13* = after excluding possible main-chain–main-chain hydrogen bonded α/βs.

The packing relationship in α/β units

Figure 5 shows that the interaxial distance varies from 5.4 to 17 Å, with a mean of 10.5 Å. More than 80% α/βs lie between 8.5–12 Å . The corresponding correlated *F4* {= [(*V*/*nda*)/*n-int*]} shows a variation from 1.4 to 3.0 with a mean *F4* of 2.0, with greater than 80% of α/βs in the range of 1.7 to 2.3. The total number of interacting residues can be as many as 16, with an average of 5.7. Removal of the α/βs with only two interacting residues or with more than 10 residues in the interacting interface has no effect on the regression coefficient (in fact it decreases in both cases). The average volume of residues in the interacting interface is about 858 Å^3, with a range of 247 to 2,575 Å^3, and the average volume per pair of interactively packed residues is about 303 Å^3. As there may be other α/β geometry and residue-dependent

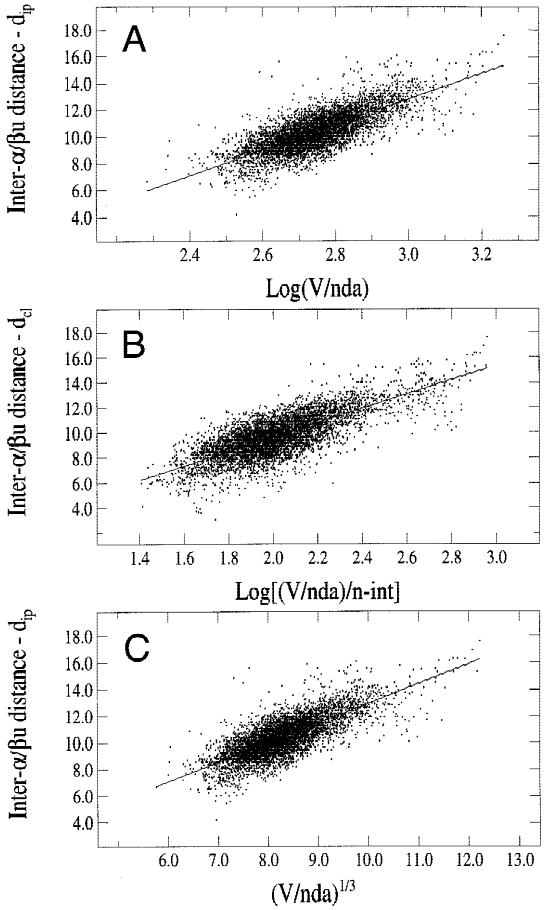

Fig. 4. Scatter plots showing correlation for the *D13* data set (5,975 α/βs) for some of the residue-dependent functions (RDFs) vs. inter-α/β distance(s). (**A**) Regression coefficient (r) = 0.76, standard deviation (σ) = 0.94; equation of regression line: $y = (9.56)x - 15.85$; (**B**) $r = 0.76$; $\sigma = 1.09$; $y = (5.71)x - 1.79$ and (**C**) $r = 0.76$; $\sigma = 0.94$; $y = 1.47 - (1.77)x$.

Table 2. *Regression coefficient (r), other regression line parameters, and average number of residues (n-int) and volumes (V) in the interacting interface in different intervals of the interaxial angle (Ω)*

Interaxial angle (Ω)	No. α/βs	Ave. *n-int*	Ave. *V* (Å^3)	*r* (*F4* vs. d_{ip})	Slope	Intercept
D12	5,977	5.7	858	0.77	4.99	0.38
Antiparallel	4,392	5.7	855	0.79	5.14	0.10
Parallel	1,585	5.7	866	0.72	4.60	1.08
−180 to −150	1,032	5.6	836	0.79	4.86	0.58
−180 to −170	553	5.7	840	0.81	5.01	0.33
−170 to −150	479	5.5	479	0.75	4.69	0.88
−150 to −120	185	5.1	185	0.77	5.84	−1.59
−120 to −90	153	5.3	803	0.76	5.99	−2.09
−90 to −60	212	5.4	820	0.80	5.55	−1.03
−60 to −30	466	5.9	882	0.79	5.20	−0.13
−30 to 0	426	6.1	913	0.69	3.67	3.06
0 to 30	232	5.4	819	0.60	4.42	1.46
30 to 60	142	5.4	858	0.62	4.62	0.87
60 to 90	106	5.1	106	0.81	5.99	−2.11
90 to 120	262	5.2	262	0.87	6.48	−3.02
120 to 150	972	5.6	847	0.84	5.64	−0.91
120 to 135	329	5.4	820	0.84	5.73	−1.24
135 to 150	643	5.7	861	0.84	5.64	−0.84
150 to 180	1,789	5.9	888	0.77	4.74	1.06
150 to 160	558	6.0	893	0.82	5.04	0.50
160 to 170	605	5.9	882	0.74	4.42	1.73
170 to 180	626	5.9	888	0.74	4.61	1.27

parameters that may increase the correlation with the inter-α/β distance when incorporated into RDFs, we studied the dependence of the packing relationship on the interaxial angular orientations (Ω) of α/βs.

Packing relationship and α/β interaxial angular orientations

The interaxial angle (ignoring the vector sense of the axes) between the SSEs of α/βs shows angular preference from −30° to +10° with a range of −60° to +20° (Fig. 5E). This angular distribution agrees well with the earlier observations (Janin & Chothia, 1980; Cohen et al., 1982) that α-helices prefer to pack onto β-strands in a near parallel way. If the vector sense is taken into consideration, there are about 73% (4,392) of α/βs with antiparallel (−90° to −180° and +90° to +180°) packing (Fig. 5B). These pairs show a higher regression coefficient value than the remaining 27% (1,585) with parallel (−90° to +90°) packing (Table 2). The regression coefficient value of α/βs in the angular orientations (Ω) in the region 60° to 160° is significantly higher than that of the total *D13* data set. The α/βs in Ω intervals −180° to −170° and −90° to −30° also have a higher regression coefficient than the total *D13*. These angular regions are also more populated regions of α/β packing. Thus, the packing relation is better for α/βs in more preferred Ω intervals. The differences in the slope and the *Y*-intercept of regression lines at different interaxial angles can be seen in Table 2.

Packing relationship and distortions in SSEs geometry

Ideally, the geometry of an α-helix is treated as a regular cylindrical rod with a groove on the surface running along the right-hand twist of the helix. In a similar way, a β-strand is treated as an extended structure with a small right-handed twist of its flat surface along the axis. However, in real protein structures α-helices can have curves, bends, and kinks (Blundell et al., 1983; Barlow & Thornton, 1988) depending on the nature of the side chains. The geometry also depends on the packing relations with other elements, such as metal ions, enzyme prosthetic groups, solvent, and the other SSEs in the structure. Similarly, the individual β-strands are super-twisted, bent, coiled, and often have bulges (Chothia, 1984). The side chains of amino acid residues of α/βs could pack interactively through any part of such a distorted surface region of SSEs (Fig. 6). We have, therefore, computed a parameter Δax (see Materials and methods) that quantifies such distortions on the axes of SSEs from the ideal geometry. We have then examined the locations of α/βs that have a high value of Δax on the regression plot (Fig. 5A). As expected, a majority of these SSEs lie well above or below the regression line. We have therefore studied the dependence of the regression coefficient on the distortion of SSE geometry, as measured by Δax, for some of the best-correlated RDFs (Table 3).

The data set *D13* is classified into different groups depending on the Δax value for the α-helix, β-strand, and Δax of both the SSEs

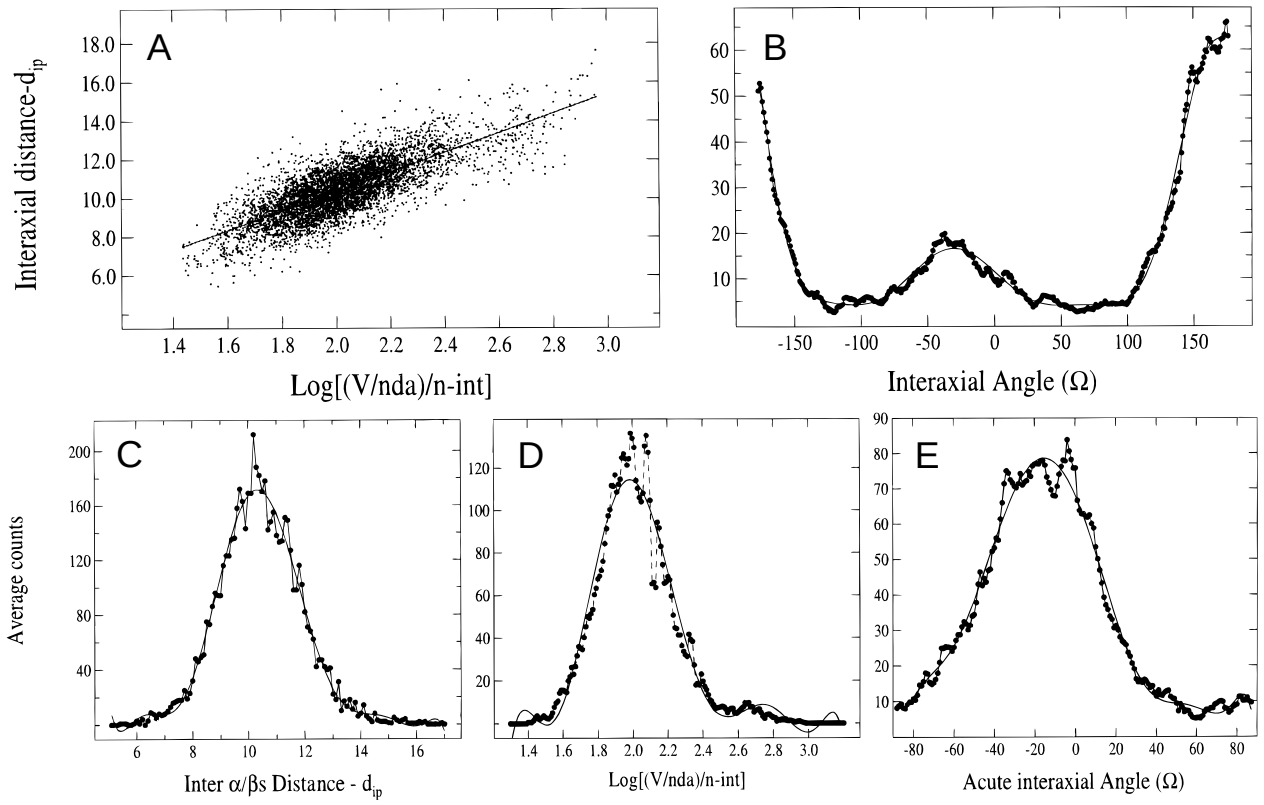

Fig. 5. Packing relationship in α/βs: (**A**) Scatter plot showing correlation between $F4 = \log[(V/nda)/n\text{-}int]$ vs. inter-α/β distance (d_{ip}), $r = 0.77$, $\sigma = 0.91$, $y = (5.06)x + 0.25$. Average number of α/βs: (**B**) in 9° window size of interaxial angles and in small intervals of (**C**) inter-α/β distance $- d_{ip}$, (**E**) *F4* values and in (**E**) acute interaxial angle (Ω).

together (see Table 3). The group of α/βs with lower Δax values have better regression coefficients. Among the classified groups of *D13* in Table 3 the best correlation (0.82) is observed for *F4* vs. d_{ip} having the Δax of the β-strand ≤ 0.7 and the Δax of α-helix ≤ 0.4.

Packing relationship: Geometry distortions and angular preferences

We have classified *D13* into three further groups with α/βs that lie (1) above the regression line (ARL), (2) closer to the regression line (CRL), and (3) below the regression line (BRL). They are grouped in the ratio of one, two, and one with total numbers for ARL (1,494), CRL (2,987), and BRL (1,494), respectively. We have then calculated average values of *V, n-int, nda*, and Δax values for these groups of α/βs (see Table 4). The average volume of residues in the interacting region for the ARL group is higher than that of the CRL or BRL groups, indicating that ARL deviations are predominantly due to involvement of larger volumes in the packing interface. Similarly, the BRL deviations could be due to smaller volumes in the packing interface. The volumes are dependent on *n-int* and the composition of larger and smaller size residues in the packing interface. The average values, *V* and *n-int*, given in columns 3 and 4 of Table 4, indicate that deviations arise for both reasons. The average *nda* value of ARL is marginally higher than that of the CRL or the BRL values, indicating that ARL α/βs have more of closely packed pairs (Fig. 7). The average Δax for CRL is marginally lower than that of ARL or BRL. This indicates that a majority of α/βs with larger Δax values is in the ABL (ARL + BRL). However, CRL do have a significant number of α/βs with high Δax value. This indicates that the higher Δax values of SSEs are one of the reasons for the deviations from packing relationship.

We have examined the normalized occurrence of α/βs as a function of Ω in ARL, BRL, and CRL groups of α/βs (see Fig. 8). These values show that highly populated regions, having near parallel or antiparallel orientations, allow the maximum number of residues to interact with the maximum volume of residues in the packing interface. In many α/βs such a close packing results in

Table 3. *The regression coefficients for different α/βs; groups on the basis of the distortions of the SSEs as quantified by Δax* [a]

	Δax of SSEs							
	Δax-α-helix			Δax-β-strand		α and β axes distortions		
RDFs	*D13*	≤0.4	>0.4	≤0.7	>0.7	*D14*	*D15*	*D16*
F1	0.74	0.78	0.69	0.79	0.71	0.82	0.63	0.75
F3	0.76	0.79	0.71	0.81	0.72	0.82	0.66	0.76
F4	0.77	0.80	0.73	0.81	0.74	0.82	0.69	0.77
No. α/βs	5,975	3,719	2,256	2,910	3,065	2,161	1,013	2,801

[a] *D14* = Δax of α-helix ≤ 0.7 and of βs ≤ 0.4; *D15* = Δax of α-helix > 0.7 and of βs > 0.4; *D16* = remaining α/βs.

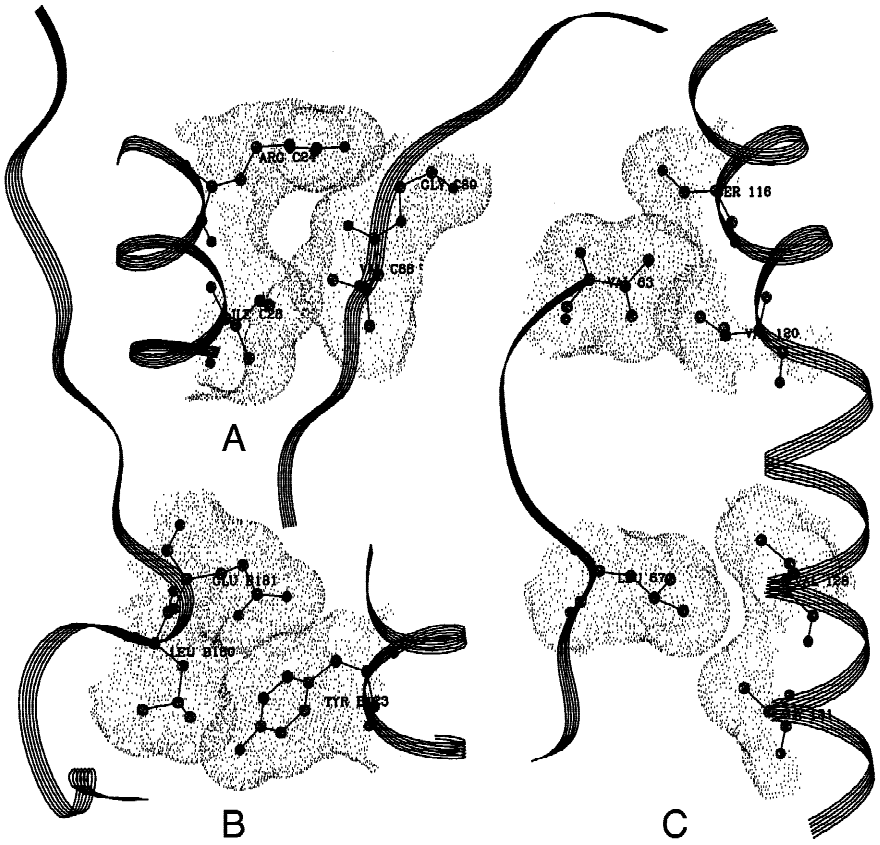

Fig. 6. Representative examples of α/βs with significant distortions. Interaction with residues from (**A**) curved β-strand (1ecpC), (**B**) β-bulge (1gggB), and (**C**) curved α-helix with a distorted β-strand (1tib).

inclusion of some of the residues facing away from the interacting interface (Fig. 7B), thereby giving rise to a major deviation from the packing relationship. This comparative study further indicates that packing relationships are better followed by the α/βs with orientations in near parallel geometry, as these have less distortions. Closely packed α/βs with small residues and twisted α/βs with large residues, packing at the optimum angles, also violate the packing relationship.

Table 4. *Average values of residue-dependent parameters for different classified groups of α/βs*[a]

	Average values							
α/βs data sets	No. α/βs	*V* (Å^3)	*n-int*	*a/b* (Å^3)	*nda*	Δax-α	Δax-β	*c*[b]
D13	5,975	858	5.69	150.8	1.58	0.89	0.97	1.86
ARL	1,494	953	6.10	156.2	1.61	0.89	1.08	1.97
CRL	2,987	849	5.66	151.6	1.57	0.88	0.90	1.78
BRL	1,494	780	5.34	147.2	1.57	0.93	0.98	1.91
ABL[c]	2,988	867	5.72	151.6	1.59	0.91	1.03	1.94

[a] Classified by whether they are above the regression line (ARL), below the line (BRL), and close to it (CRL).
[b] c = mean average of Δax-α + Δax-β.
[c] ABL = ARL + BRL.

Alpha helix/β-sheet-unit (α/βu) packing

A detailed analysis of packing geometry and residual surface suggests that α-helix/β-sheet interactions are often confined to the two adjacent β-strands of a β-sheet as a topological unit (Cohen et al., 1982). Therefore, we have also analyzed the α-helix/β-sheet unit (α/βu) packing by redefining the RDFs and geometry parameters for βu as a packing SSE (see Materials and methods). To calculate the axis of βu, we have used only the backbone atoms of the extended interacting region of the β-strands (Fig. 2). Where an extended region is not available in a particular α/βu, we have confined the analysis to the interacting region.

In this analysis we have considered α-helices and β-strands with five or more residues. Only the regions of βu that have hydrogen-bonding interactions are considered. Since we define the ortho-center of βu and mid-point of the helix axis to be within the interacting interface, we have not removed the terminally interacting α/βu. There are 1,749 nonidentical α/βu units in the data set considered for analysis.

Residue-dependent parameters and inter-α/βu distance

The regression coefficient values for the inter-α/βu distances ($d_{\alpha/\beta u}$) and various RDFs in Table 5 indicate that there is an equally interesting packing relationship in these units. The residue-dependent functions, *F1, F3*, and *F5*, have regression coefficient values $\geq$0.8 (see Fig. 9 for *F1* and *F5*; Fig. 10A for *F3*). In the case of α/βs, *F4* shows the highest regression coefficient value. How-

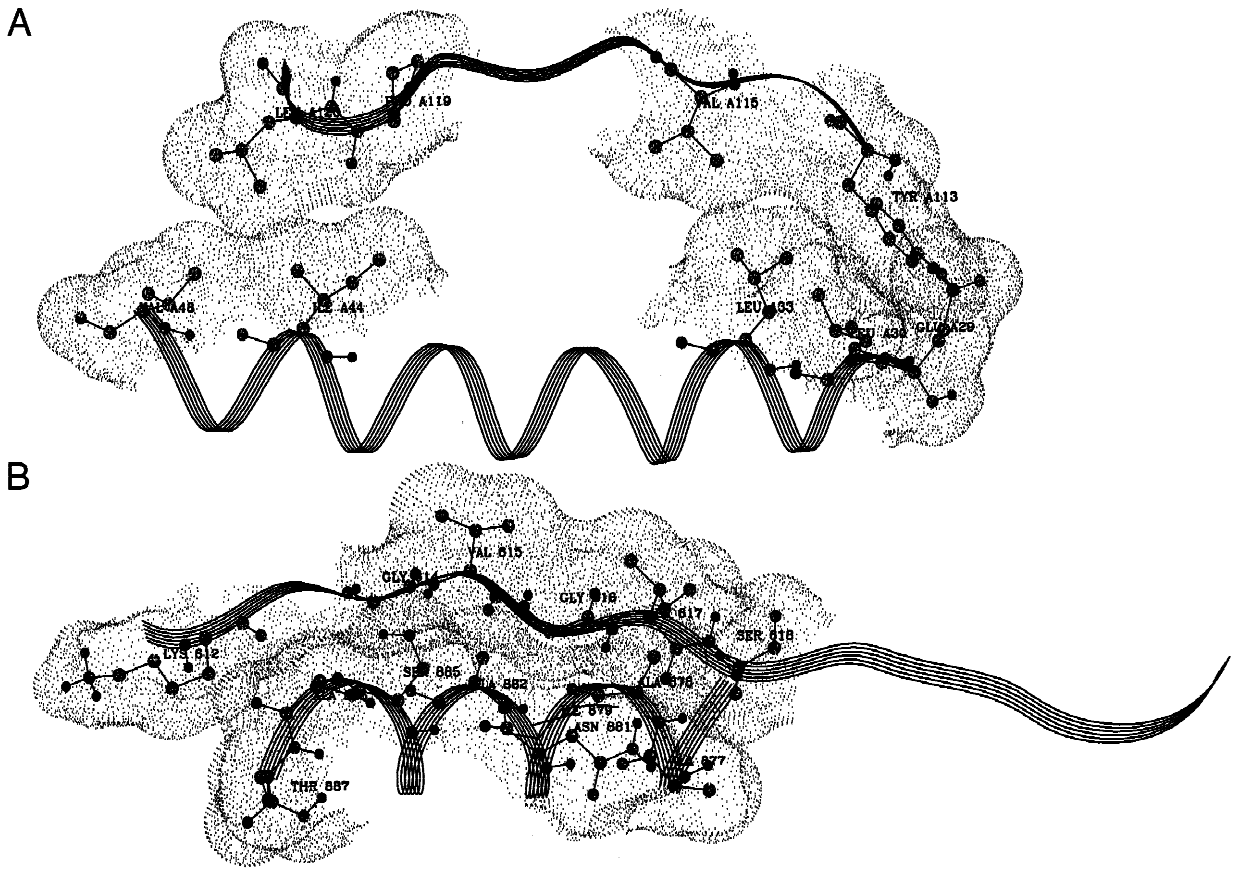

Fig. 7. Representative example of (**A**) α/β from 1ftaA, significantly above the regression line, showing the interactive packing through the side-chain groups of larger residues with curvature away from the interacting interface, and (**B**) α/β from 1alo, significantly below the regression line, interacting through small residues with optimal complementary twist between α-helix and β-strand.

ever, this function has no good correlation for α/βus. In the data set there are approximately equal numbers of parallel (863) and antiparallel (886) β-sheet units. The α/βu units with antiparallel β-sheet units are better correlated than the ones with parallel β-sheet units. The function *F3* and *F5* shows the best correlation of 0.82 with $d_{\alpha/\beta u}$. We, therefore, denote the correlation between *F3* and $d_{\alpha/\beta u}$ (and also *F5* vs. $d_{\alpha/\beta u}$) as the packing relationship in the α/βu units.

Table 5. *Amino acid residue-dependent functions (RDFs) of the interacting region and corresponding correlation coefficients for interactive packing of α/βu*

RDF	α/βu	pll-βu[a]	apl-βu[b]
nda	−0.66	−0.60	−0.69
n-int	−0.50	−0.41	−0.55
V	−0.32	−0.23	−0.38
F1 = V/nda	0.80	0.79	0.81
F2 = F1/n-int	0.65	0.64	0.67
F3 = log(*F1*)	0.82	0.80	0.82
F4 = log(*F2*)	0.73	0.69	0.75
F5 = $(F1)^{1/3}$	0.81	0.80	0.82
Total α/βu	1,749	863	886

[a]pll-βu = parallel hydrogen-bonded β-strands in βu.
[b]apl-βu = antiparallel hydrogen-bonded β-strands in βu.

The packing relationship in α/βu units

Figure 10A shows the distribution of α/βu interaxial distances $d_{\alpha/\beta u}$ with the best-correlated function *F3*. The inter-α/βu distance $d_{\alpha/\beta u}$ varies from 12–14 Å, with a mean distance of 10.37 Å, and more than 80% of α/βu in the range 8.0–12.0 Å. The value of *F3* varies from 2.4 to 3.1, with an effective range of 2.6 to 2.9, having the highest number of α/βu at an average value of 2.7. The number of residues in the interacting region varies from 3 to 20 (effective range 3–11) with an average of 8.13. The α/βus packed through parallel and antiparallel β-sheet units show similar distribution patterns. The preferred interaxial angle in α/βu is about −5° (near parallel orientations) with an effective range of −40° to +20° (Fig. 10B), shown in our α/β packing studies, and reported by earlier investigators (Janin & Chothia, 1980; Cohen et al., 1982).

Dependence of packing relationship on α/βu angular orientation

Angles for the α/apl-βu units peak at about −26° (average Ω of −8.5°) with wider angular distribution compared to the α/pll-βu units (Fig. 10B). The α/pll-βu have higher occurrence at −11° and an average value of −1.6°, with an effective range from −40° to 30°.

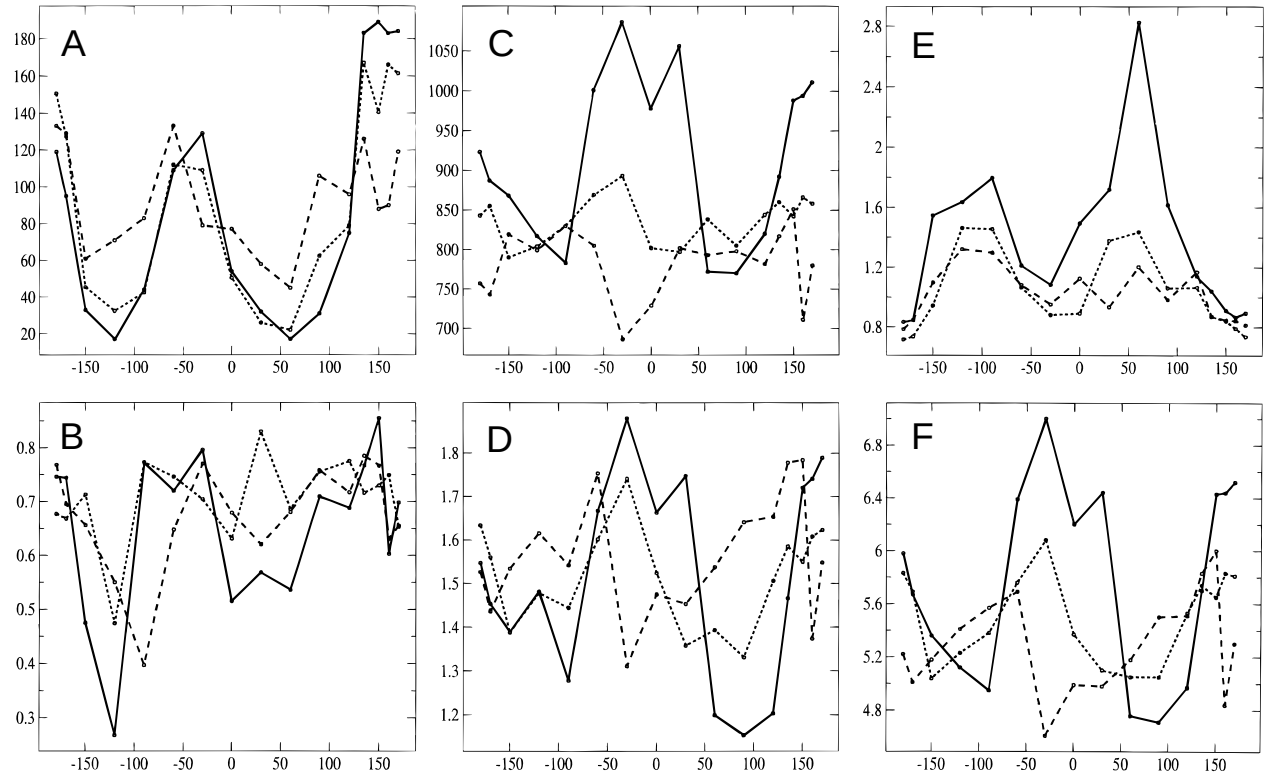

Fig. 8. The variation in average values of residue-dependent parameters for α/β packing units in: ARL (solid line), CRL (small dashed), and BRL (large dashed) in 30° intervals of their interaxial angle. (**A**) Average occurrence; (**B**) regression coefficient (r); (**C**) average volume (V); (**D**) average normalized difference in accessibility (*nda*); (**E**) average distortion in SSE geometry (α-Δax + β-Δax); and (**F**) average number of residues (*n-int*) in the interacting interface.

Figure 11 shows some of the α/βu that lie very much below the regression line. Such deviations mostly occur in (1) closely packed interfaces with smaller residues or (2) terminally interacting units where the ortho-center of βu is much closer to the helix (Fig. 11A,B). Those above the regression line occur where (1) there are more extended side-chain orientations of larger interacting residues, (2) terminally interacting residues of α/βu with larger side-chain groups, or (3) interaction only of the residues from both termini (Fig. 11C,D). It is interesting to note that, for interactions of α/βu, the packing relations are better maintained than for α/βs, and the interaxial distances are within the standard deviation of 0.71 Å from the regression line.

Prediction of interaxial distance using α/β packing relationship

To predict the interaxial distance between interactively packed β-strands and α-helices using the packing relationship described

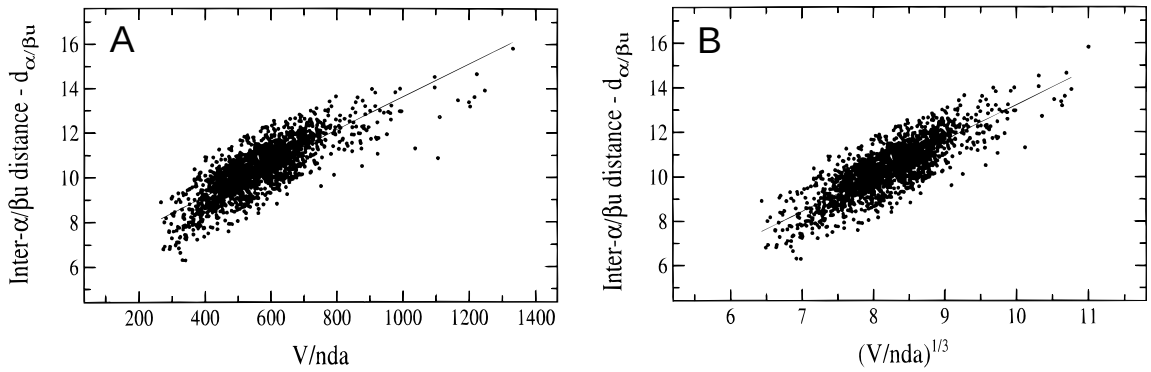

Fig. 9. Scatter plots for 1,749 α/βu units, showing correlation of $d_{\alpha/\beta u}$ with (**A**) $F1 = V/nda$; $r = 0.80$, standard deviation (σ) = 0.72; equation of regression line $y = (0.01)x + (6.18)$; and (**B**) $F5 = (V/nda)^{1/3}$; $r = 0.814$; $\sigma = 0.71$; $y = (1.62)x - 2.74$.

Table 6. *Standard solvent accessible contact areas (a_{std}), volumes (v_{std}), and nda for the amino acids (AA) in α-helix (α-nda$_i$)and in β-strand (β-nda$_i$)*[a]

AA	a_{std} (Å^2)	v_{std} (Å^3)	α-nda$_i$	β-nda$_i$
A	33.27	92	0.48(0.21)	0.35(0.15)
C	41.90	118	0.55(0.17)	0.37(0.13)
D	39.38	125	0.29(0.11)	0.25(0.14)
E	48.53	155	0.31(0.16)	0.25(0.12)
F	60.81	203	0.38(0.18)	0.26(0.12)
G	23.75	66	0.47(0.29)	0.47(0.23)
H	54.98	167	0.31(0.14)	0.27(0.11)
I	55.52	169	0.43(0.18)	0.27(0.11)
K	61.45	171	0.32(0.16)	0.21(0.11)
L	56.53	168	0.40(0.17)	0.26(0.11)
M	61.86	171	0.38(0.18)	0.27(0.11)
N	40.91	135	0.30(0.16)	0.32(0.14)
P	44.58	129	0.39(0.17)	0.29(0.17)
Q	51.50	161	0.34(0.12)	0.24(0.11)
R	72.58	202	0.26(0.15)	0.22(0.08)
S	33.77	99	0.48(0.25)	0.34(0.16)
T	41.85	122	0.40(0.18)	0.29(0.14)
V	47.71	142	0.42(0.19)	0.29(0.13)
W	75.22	238	0.35(0.13)	0.23(0.10)
Y	62.17	204	0.34(0.15)	0.22(0.10)

[a]The associated standard deviations are given in parentheses.

above, we need information about the possible interacting residues in that pair. This can be obtained from the sequence alignment of the target with its homologue of known structure (template). The target residues equivalent to the interacting residues in the template are used to calculate the value of *F4* and the distance between the target SSE pair is calculated using an appropriate packing relationship depending on the angle of packing in the template pair. The total interacting residue volumes are calculated using the values given by Chothia (1975) (Table 6). The total normalized difference in solvent accessible contact areas (*nda*) is either derived from the representative values for *nda*s of the 20 amino acid residues (Table 6) calculated as the average values (d_{pred1}) from the interacting residues in nonidentical α/β units in a set of nonhomologous protein structures or from the *nda*s of equivalent residues in the template (d_{pred2}).

To test the usefulness of the packing relationship to predict interaxial distances between α-helices and the β-strands, we have chosen four families of proteins. They are: (1) the disulfide oxidoreductases (10 structures), (2) the lactamase/malate dehydrogenases (7 structures), (3) the periplasmic binding proteins (3 structures), and (4) the β-lactamases (3 structures) (Table 7). In a family each protein is considered as the target and distance predictions are made using every other protein in that family. To obtain information about possible interacting residues for the targets, we have used structure-based alignments, obtained by

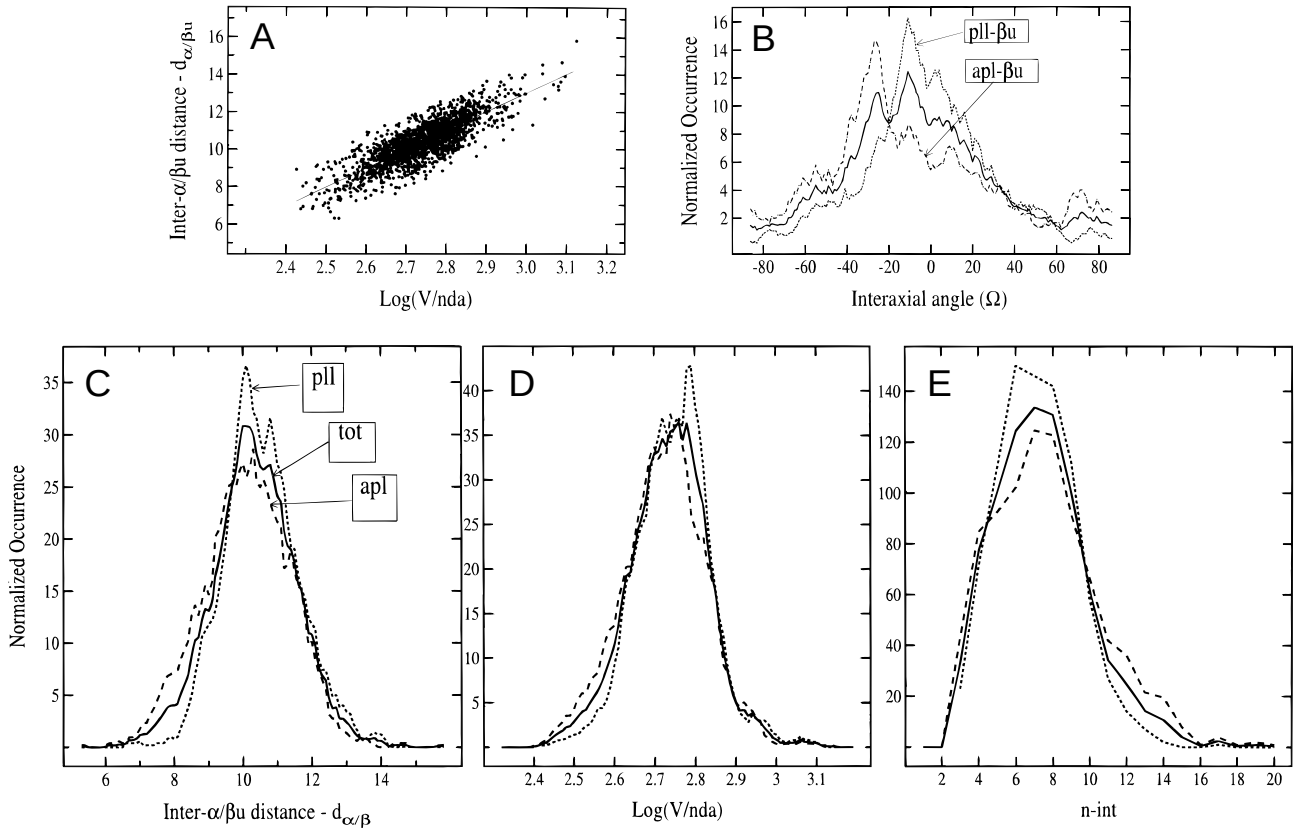

Fig. 10. Packing relation in α/βu units. (**A**) Scatter plot for 1,749 α/βu units, showing correlation of $d_{\alpha/\beta u}$ with $F3 = \log(V/nda)$; $r = 0.82$; $\sigma = 0.71$; $y = (10.09)x + (-17.26)$; (**B**) normalized occurrence of α/βu units versus interaxial angle (Ω). Note the small difference in optimal preferred angles between pll-βu and apl-βu in their interactive packing with α-helix. Normalized occurrence of α/βu vs. (**C**) $d_{\alpha/\beta u}$, (**D**) $F3 = \log(V/nda)$, and (**E**) number of residues in interactive interface (*n-int*).

Table 7. *The four families of proteins used in the prediction of interaxial distances*

Family	PDB structures (chain)	Average sequence length	Average sequence identity
(1) Disulfide oxidoreductases	2ptr(A),1nda(A), 3grs,1ger(A),1ojt, 1ebd(A),3lad(A), 1lpf(A),1lvl,1npx	465	32%
(2) Lactamase/malate dehydrogenases	1mld(A),2cmd, 1ldn(A),1lld(A), 9ldb(A),4mdh(A), 1bdm(A)	320	30%
(3) Periplasmic binding proteins-sugar	2dri, 2gbp,1abe	295	21%
(4) β-lactamases	4blm(A),3blm, 1btl	259	37%

using COMPARER (Sali & Blundell, 1990) and deposited in the in-house data-base HOMSTRAD (*http://cryst-bioc.cam.ac.uk/~homstrad*).

Since we have considered all possible pair-wise alignments in a family, a pair of target SSEs can have as many predicted values as the number of the basis structures in that family. The quality of these predictions is assessed by comparing the observed values with both the predicted distances and the template distances. We define an error in the predicted distance (or template distance) as equal to the difference between the predicted distance (or template distance) and the observed distance.

General nature of the errors in the observed distances d_{pred1} and d_{pred2}

Figure 12 shows the distribution of the errors associated with d_{pred1} (shown as broken line) and d_{pred2} (shown as solid line). The distribution of error with d_{pred1} (E_{dpred1}) is skewed toward the negative side (mean $= -0.7(1.00)$) with 64% of the cases in the error range of -1.0 and $+1.0$ Å, whereas the errors of d_{pred2} (E_{dpred2}) are almost centrally distributed about the origin (mean $= -0.2(0.8)$) with 77% of them within the range of -1.0 and $+1.0$ Å. The skewness in the E_{dpred1} distribution indicates that a large number of distances are underpredicted. This could be due to

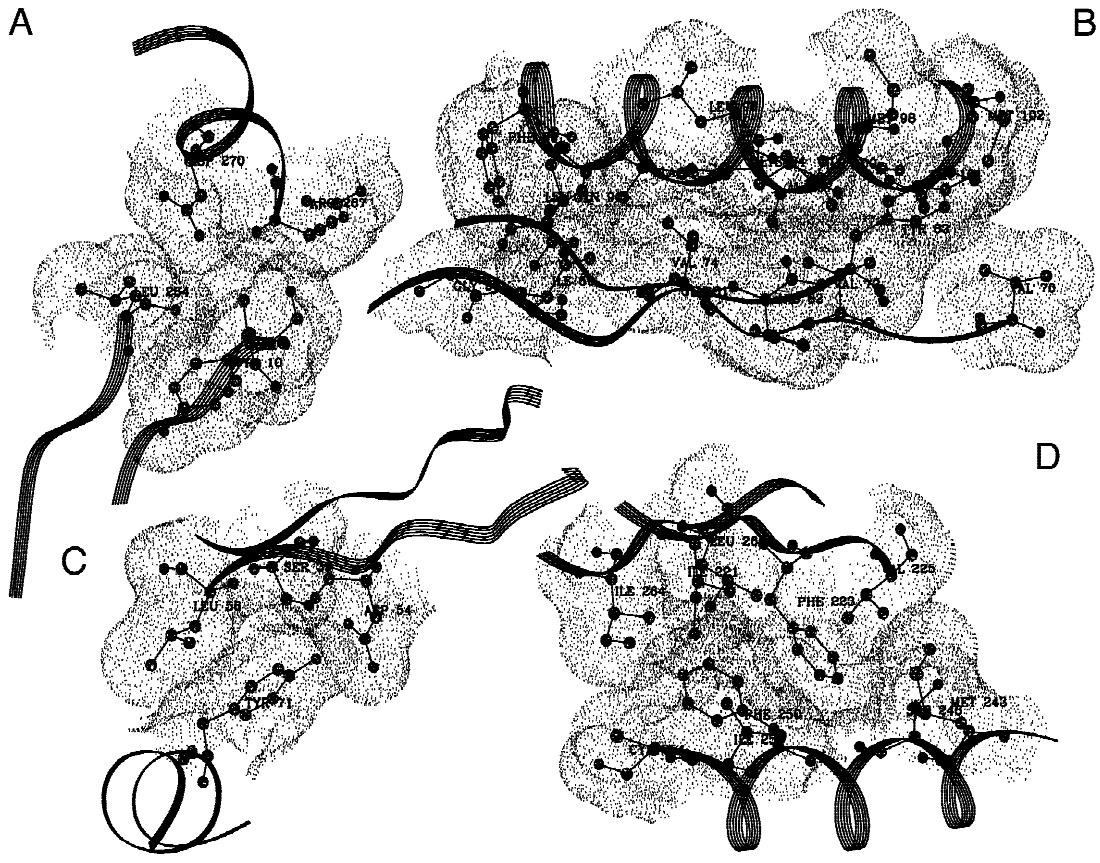

Fig. 11. Representative examples of α/βu units: (**A**) from structure 1cnv, interacting through the residues at the termini of βu where the ortho-center of the α/βu come closer to the center of the interacting region of the α-helix, and (**B**) from structure 1tfe, the α-helix has close packing interactions through small residues. These α/βu have packing relationships that fall significantly below the regression line. Note that the residues facing away from the interacting interface are also defined as interacting as they lose ≥10% of their SACA upon packing; (**C**) α/βu from structure 4q21 and (**D**) from structure 1gia. These α/βu units are significantly above the regression line (ARL) having the interactive packing through side-chain groups of larger residues.

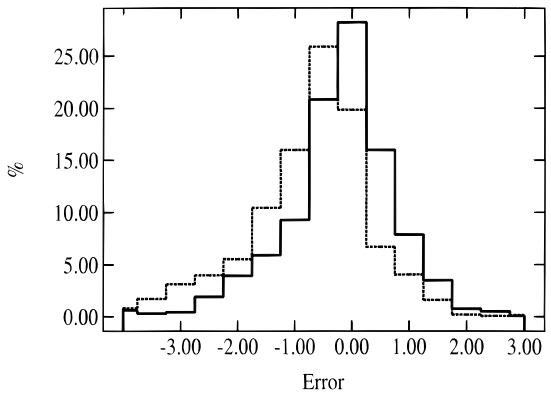

Fig. 12. Step plots showing the distribution of the error (Å) in the two sets of predicted distances: E_{dpred1} (shown as broken line) and E_{dpred2} (shown as solid line).

taking the average *nda* values in the calculation of the values of *F4*. A comparison of the three *nda* values (average, target, and template) revealed that in most cases template values are closer to the target values (data not shown) than the average values and that the average values are higher than the template values. Hence it seems that for the families considered in the present study the *nda* values borrowed from the template give rise to better predictions than the average values. Owing to the fact that d_{pred2}s are closer to the observed distances than d_{pred1}s we will consider only d_{pred2}s for further investigations.

Comparison of the template distances and the predicted distances with the observed distances

One of the aims in the present study is to use the predicted distances to improve the inter-SSE distances for a given target. The predicted distances should be closer to the observed distances than the template distances. We, therefore, compared the errors associated with the predicted distances (E_{dpred2}) to those associated with the template distances (E_{temp}).

For 47% of the 1,629 predictions made, E_{dpred2} was found smaller than E_{temp} (d_{pred2} closer to observed distance than $d_{\text{temp.}}$). In the remaining cases, E_{dpred2} was either greater than (31%) or equal (22%) to E_{temp}. We also compared the errors in the subset of predictions (281) made using the closest homologues of the targets. Of these predictions the percentages showing $E_{\text{dpred2}} < E_{\text{temp}}$, $E_{\text{dpred2}} > E_{\text{temp}}$, and $E_{\text{dpred2}} = E_{\text{temp}}$ were, respectively, 41%, 33%, and 26% showing that d_{pred2} is often closer to the observed distance and indicating an advantage in using predicted distances over the template distances in comparative modeling.

Thus, our investigations show that the predictions of interaxial distances are more useful than those taken directly from the homologues. In fact, the number of predicted distances can be as many as the number of templates, so the weighted average is generally used in modeling of the target. The weights can be made a function of the inverse of square of sequence differences between the template and the target (Srinivasan & Blundell, 1993).

Conclusion

We have presented an analysis of packing in α-helix/β-strand and α-helix/β-sheet units, in terms of the interaxial distance and a residue-dependent function of the amino acids in the packing interface. The interaxial distances in the α/β and α/βu units are best correlated with the residue-dependent functions, $\log[(V/nda)/n\text{-}int]$ and $\log(V/nda)$, respectively. The packing of these units shows a preference for near antiparallel arrangements between the SSEs. The structural distortions in the geometry of α-helices and β-strands have significant effects on their packing relationships. The α-helix shows a better packing relationship with antiparallel-βu than the parallel-βu, with a small difference in their preferred interaxial angles (Ω). The packing relationship in α/β is shown to be useful for prediction of interaxial distances using the interacting residue information from the equivalent α/β units of homologous proteins.

In our analyses the interaxial distances for helix pairs have shown the best correlation with $\log(V/nda)$ values of the residues in the packing interface (Reddy & Blundell, 1993). In the case of pairs of β-strands, the distance between the axes of two SSEs at the projection of closest C^{α}-atoms of the SSEs gives best correlation with the $\log[(V/nda)/n\text{-}int]$ value of the residues in the packing interface (H.A. Nagarajaram, B.V.B. Reddy, & T.L. Blundell, unpubl. obs.). In the case of α-helix/β-strand packing, the interaxial distance is better correlated with $\log[(V/nda)/n\text{-}int]$ value of residues in the packing interface. The predicted inter-SSE distances between these different combinations of SSEs are shown to be significantly correlated with the observed distances. A detailed assessment and analysis of their usefulness to improve models generated by comparative modeling procedures will be discussed elsewhere.

Materials and methods

Three-dimensional co-ordinates of 6,531 protein chains, defined by X-ray analysis at 2.5 Å or better (PDB; Bernstein et al., 1977), are used in the analysis. The method of Kabsch and Sander (1983), as implemented by Smith (1989) in his SSTRUC program, is used to identify secondary structural elements. In most cases we follow our earlier procedures (Reddy & Blundell, 1993) for calculating geometrical parameters of SSEs, for defining packing residues between SSEs and for calculating amino acid residue-dependent parameters in the packing interfaces.

Solvent accessibility contact area and interactive packing

(1) Solvent accessible contact areas (SACA) for individual residues, both in an isolated SSE (a_i) and in the presence of an interacting SSE (a_c), are calculated using the method of Richmond and Richards (1978) as implemented by Sali (1991) in his PSA program. The percentage difference in solvent accessible contact area of each residue is calculated as $pda_1 = (a_i - a_c) \times 100/a_{std}$, where a_{std} is the total SACA of residue in Gly-X-Gly form (Table 6). An interacting residue of an SSE is one that loses 10% or more solvent accessible contact area in the presence of its interactively packed SSE.

(2) An interacting region of a SSE is defined as the continuous region along the axes, covering the C^{α}-projections of the first and the last interacting residue.

(3) An extended-interacting region on the axis is defined as the interacting region plus the projections of two additional C^{α} on either side of the interacting region (see Fig. 1).

(4) An α-helix and a β-strand are said to be interactively packed only if at least one residue from each SSE has a difference in

SACA greater than 10% ($pda_i \geq 10\%$) between isolated and interacting SSEs. An α-helix and βu are said to be interactively packed only if each of the β-strands of the βu are independently packed with α-helix.

Geometrical parameter

(1) α-helix and β-strand axes are calculated using the method of Blundell et al. (1983). For each residue at i, a standard probe α-helix (or β-strand) is superposed to give a least-squares fit for atoms C'_{i-1}, N_i, C_i^{α}, C'_i, N_{i+1}, along the polypeptide chain of the SSE. After each fit the co-ordinates are calculated for the projection of the (α-helix or β-strand) atom C_i^{α} onto the probe axis. The set of points thus generated is used to fit a least-squares line as an approximate linear axis (Fig. 1).

(2) The RMS deviation of real axis points from the corresponding points on the approximated linear axis is computed as the backbone distortion parameter Δax of the SSEs.

(3) The dihedral angle (Ω) between the axes of the α-helix and the β-strand is calculated taking the N- to C-terminus directional vector of axes and the line joining the mid-points of their axes. The interaxial angle (Ω), therefore, can vary from $-180°$ to $+180°$.

(4) The distance between two SSEs is calculated in two ways: (a) as the distance between the mid-points of the linear axes of the two SSEs (d_{ip}); (b) as the closest distance between C^{α} projections of the interacting residues from their corresponding linear axes (d_{cl}).

(5) The distance between an α-helix and a β-sheet unit ($d_{\alpha/\beta u}$) is calculated as the distance from mid-point of the extended-interacting region of the helix axis to the ortho-center of the β-sheet unit formed by two adjacent parallel or antiparallel β-strands (β-ladder) (see Fig. 2). The axis of a βu is a straight line, equidistant to the axes of both the β-strands through the ortho-center of the βu. The interaxial angle of α/βu (Ω) is the dihedral angle of their axes through the mid-point of the helix axis and the ortho-center of the βu. Since the βu consists of either parallel or antiparallel hydrogen-bonded β-strands, there is no directional vector taken into consideration for the axis of βu, and the angle between the axes of α/βu varies only from $-90°$ to $+90°$.

Amino acid residue-dependent parameters of the packing interface

The total volume of residues, $V = \Sigma v1_i + \Sigma v2_i$, in the interacting region is calculated as the sum of volumes of every residue that shows $pda_i \geq 10\%$ on interaction with its partner SSE. $v1_i$ and $v2_i$ are the standard volumes of the interacting residues in the α-helix and the β-strand (or βu), respectively (see Table 6). The total number of residues involved in interactive packing is computed as *n-int.*

The sum of the fractional loss in SACA of each interactively packed residue, with respect to its total standard SACA, is computed as the normalized difference in SACA, $nda = \Sigma nda_j$, where $nda_j = (a_i - a_c)/a_{std}$ and j is the interactively packed residue. Values of several functions using *V*, *n-int*, and *nda* have been calculated (see Table 1) to test their correlations with inter-SSE distance (d_{ip}, d_{cl}, and $d_{\alpha/\beta u}$).

Acknowledgments

BVBR acknowledges the Department of Science and Technology (India) for a research grant, the Indian National Science Academy and the Royal Society (UK) for the fellowship under international exchange program. H.A.N. is recipient of a fellowship funded by a grant from Oxford Molecular Ltd.

References

Bajaj M, Blundell TL. 1984. Evolution and the tertiary structure of globular proteins. *Ann Rev Biochem 13*:453–492.

Bajorath J, Stenkamp R, Aruffo A. 1993. Knowledge-based model-building of proteins—Concepts and examples. *Protein Sci 2*:1798–1810.

Baker EN, Hubbard RE. 1984. Hydrogen bonding in globular proteins. *Prog Biophy Mol Biol 44*:97–179.

Barlow DJ, Thornton JM. 1988. Helix geometry in proteins. *J Mol Biol 201*:601–619.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Blundell TL, Barlow D, Borkakoti N, Thornton J. 1983. Solvent-induced distortions and the curvature of α-helices. *Nature (London) 306*:281–283.

Blundell TL, Carney D, Gardine S, Hayes F, Howlin B, Hubbard T, Overington J, Singh DA, Sibanda L, Sutcliffe M. 1988. Knowledge-based protein modelling and design. *Eur J Biochem 172*:513–520.

Blundell TL, Sibanda BL, Sternberg MJE, Thornton JM. 1987. Knowledge-based prediction of protein structures and design of novel molecules. *Nature (London) 323*:347–352.

Boutonnet NS, Kajava AV, Rooman MJ. 1998. Structural classification of $\alpha\beta\beta$ and $\beta\beta\alpha$ supersecondary structure units in proteins. *Proteins 30*:193–212.

Browne WJ, North ACT, Philips DC, Brew K, Vanaman TC, Hill RL. 1969. A possible three-dimensional structure of bovine β-lactalbumin based on that of hen's egg white lysozyme. *J Mol Biol 42*:65–86.

Chothia C. 1974. Hydrophobic bonding and accessible surface area in proteins. *Nature 248*:338–339.

Chothia C. 1975. Structural invarients in protein folding. *Nature 254*:705–708.

Chothia C. 1984. Principles that determine the structure of proteins. *Ann Rev Biochem 53*:537–572.

Chothia C, Lesk AM. 1982. Evolution of proteins formed by β-sheets: 1. Plastocyanin and azurin. *J Mol Biol 160*:325–342.

Chothia C, Lesk AM. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J 5*:823–826.

Chothia C, Lesk AM. 1987. The evolution of protein structures. *Cold Spring Harbour Symp Quant Biol 52*:399–405.

Chothia C, Lesk AM, Levitt M, Amitt AG, Mariuzza RA, Phillips SEV, Poljak RJ. 1986. Predicted structure of immunoglobulin D1.3 and its comparison with the crystal structure. *Science 233*:755–758.

Chothia C, Levitt M, Richardson D. 1977. Structure of proteins: Packing of α-helices and pleated sheets. *Proc Natl Acad Sci USA 74*:4130–4134.

Chou KC, Nemethy G, Rumsey S, Tuttle RW, Scheraga HA. 1985. Interactions between an alpha-helix and a beta-sheet. Energetics of alpha/beta packing in proteins. *J Mol Biol 186*:591–609.

Cohen FE, Sternberg MJE, Taylor WR. 1982. Analysis and prediction of the packing of α-helices against a β-sheet in the tertiary structure of globular proteins. *J Mol Biol 156*:821–862.

Efimov AV. 1995. Structural similarity between 2-layer alpha/beta-proteins and beta-proteins. *J Mol Biol 245*:402–415.

Farber GK, Petsko GA. 1990. The evolution of alpha-beta-barrel enzymes. *TIBS 15*:228–228.

Greer J. 1981. Comparative model-building of the mammalian serine proteinases. *J Mol Biol 153*:1027–1042.

Handel T. 1990. De novo design of an alpha-beta-barrel protein. *Protein Eng 3*:233–234.

Havel TF, Snow ME. 1991. A new method for building protein conformations from sequence alignments with homologues of known structure. *J Mol Biol 217*:1–7.

Hilbert M, Bohm G, Jaenicke R. 1993. Structural relationships of homologous proteins as a fundamental principle in homology modeling. *Proteins 17*:138–151.

Janin J, Chothia C. 1980. Packing of α-helices onto β-pleated sheet and the anatomy of $\alpha\beta$-proteins. *J Mol Biol 143*:95–128.

Johnson MS, Srinivasan N, Sowdhamini R, Blundell TL. 1994. Knowledge-based protein modelling. *CRC Crit Rev Biochem Mol Biol 29*:1–68.

Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen bonded and geometrical features. *Biopolymers 22*: 2577–2637.

Lasters I, Wodak SJ, Pio F. 1990. The design of idealized α/β-barrels—Analysis of beta-sheet closure requirements. *Proteins 7*:249–256.

Lesk AM, Branden CI, Chothia C. 1989. Structural principles of alpha-beta-

barrel proteins—The packing of the interior of the sheet. *Proteins 5*:139–148.
Lesk AM, Chothia C. 1980. How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *J Mol Biol 136*:225–270.
Lesk AM, Chothia C. 1982. Evolution of proteins formed by β-sheets: II. The core of immunoglobulin domains. *J Mol Biol 160*:325–342
Lesk AM, Chothia C. 1986. The response of protein structures to amino acid sequence changes. *Phil Trans Roy Soc Lond 317*:345–356.
Levitt M, Chothia C. 1976. Structural patterns in globular proteins. *Nature 261*:552–558.
Martin ACR, MacArthur MW, Thornton JM. 1997. Assessment of comparative modelling in CASP2. *Proteins S1*:14–28.
Mosimann S, Meleshko R, James MNG. 1995. A critical-assessment of comparative molecular modeling of tertiary structures of proteins. *Proteins Struct Funct Genet 23*:301–317.
Mumenthaler C, Braun W. 1995. Predicting the helix packing of globular-proteins by self-correcting distance geometry. *Protein Sci 4*:863–871.
Raine ARC, Scrutton NS, Mathews FS. 1994. The evolution of alternate core packing in eight fold alpha/beta-barrels. *Protein Sci 3*:1889–1892.
Reddy BVB, Blundell TL. 1993. Packing of secondary structural elements in proteins: Analysis and prediction of inter-helix distances. *J Mol Biol 233*:464–479.
Richardson JS. 1981. The anatomy and taxonomy of protein structure. *Adv Prot Chem 34*:167–339.
Richmond TJ, Richards FM. 1978. Packing of α-helices: Geometrical constraints and contact areas. *J Mol Biol 119*:537–555.
Rost B, Sander C. 1996. Bringing the protein sequence-structure gap by structure predictions. *Ann Rev Bioph and Biomole Str 25*:113–136.
Sali A. 1991. Modelling of three-dimensional structure of proteins from their sequence of amino acid residues [Ph.D. thesis]. University of London.
Sali A. 1995. Modelling mutations and homologous proteins. *Curr Opin Biotech 6*:437–451.
Sali A, Blundell TL. 1990. The definition of topological equivalence in homologous and analogous structures: A procedure involving a comparison of local properties and relationships. *J Mol Biol 212*:403–428.
Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol 234*:779–815.
Sali A, Potterton L, Yuan F, Van Vlijmen H, Karplus M. 1995. Evaluation of comparative protein modelling by MODELLER. *Proteins 23*:318–326.
Sanchez R, Sali A. 1997. Advances in comparative protein-structure modelling. *Curr Opin Struct Biol 7*:206–214.
Smith D. 1989. SSTRUC: A program to calculate secondary structural summary. England: Department of Crystallography, Birkbeck College, University of London.
Srinivasan N, Blundell TL. 1993. An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure. *Protein Eng 6*:501–512.
Srinivasan S, March CJ, Sudarshanam S. 1993. An automated-method for modeling proteins on known templates using distance geometry. *Protein Sci 2*:277–289.
Sutcliffe MJ, Haneef I, Carney D, Blundell TL. 1987a. Knowledge-based modelling of homologous proteins. Part I: Three dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng 1*:377–384.
Sutcliffe MJ, Hayes FRF, Blundell TL. 1987b. Knowledge-based modelling of homologous proteins. Part II: Rules for the conformations of substituted side-chains. *Protein Eng 1*:385–392.
Walther D, Eisenhaber F, Argos P. 1996. Principles of helix-helix packing in proteins—The helical lattice superposition model. *J Mol Biol 255*:536–553.
Vtyurin N, Panov V. 1995. Packing constraints of hydrophobic side-chains in alpha/beta (8) barrels. *Proteins 21*:256–260.