PROTEINS: Structure, Function, and Genetics 5:139–148 (1989)

Structural Principles of α/β Barrel Proteins: The Packing of the Interior of the Sheet

Arthur M. Lesk,[1,2] Carl-Ivar Brändén,[2,3] and Cyrus Chothia[1,4]

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England; [2]European Molecular Biology Laboratory, Postfach 1022.09, D-6900 Heidelberg, Federal Republic of Germany; [3]Department of Molecular Biology, Swedish University of Agricultural Sciences, Uppsala Biomedical Center, Box 590, S-75724 Uppsala, Sweden; [4]Christopher Ingold Laboratory, University College London, London WC1H 0AJ, England

ABSTRACT **α/β barrel structures very similar to that first observed in triose phosphate isomerase are now known to occur in 14 enzymes. To understand the origin of this fold, we analyzed in three of these proteins the geometry of the eight-stranded β-sheets and the packing of the residues at the center of the barrel. The packing in this region is seen in its simplest form in glycolate oxidase. It consists of 12 residues arranged in three layers. Each layer contains four side chains. The packing of RubisCO and TIM can be understood in terms of distortions of this simple pattern, caused by residues with small side chains at some of the positions inside the barrel. Two classes of packing are found. In one class, to which RubisCO and TIM belong, the central layer is formed by a residue from the first, third, fifth, and seventh strands; the upper and lower layers are formed by residues from the second, fourth, sixth, and eighth strands. In the second class, to which GAO belongs, this is reversed: it is side chains from the even-numbered strands that form the central layer, and side chains from the odd-numbered strands that form the outer layers. Our results suggest that not all proteins with this fold are related by evolution, but that they represent a common favorable solution to the structural problems involved in the creation of a closed β barrel.**

Key words: protein architecture, packing, evolutionary relationships

INTRODUCTION

Triose phosphate isomerase (TIM) was the first example of a protein constructed from eight β-sheet strand–α-helix units closed into a cylindrical topology suggesting the general shape of a barrel.[1] Similar structures are now known to occur in many other enzymes.[2–13] Here we describe the general principles of this fold, derived from an analysis of the sheet geometry and of the residue packing inside the β-sheets of chicken TIM,[1] spinach glycolate oxidase (GAO),[2] and a domain of *Rhodospirillum rubrum* ribulose-1,5-bisphosphate carboxylase-oxygenase (RubisCO).[3]

The general structure is illustrated in Figure 1, a schematic drawing of TIM. Its secondary structure consists of alternating strands of β-sheet and α-helices. The strands are assembled into a central parallel β-sheet, shaped roughly like a barrel, with helices packed around it. The strands are tilted at angles of approximately −36° to the axis of the structure. The axes of the helices are approximately parallel to the strands, as is usual in α/β structures.[14] Consistent with the observed preference of β–α–β units for a right-handed connection, the chain proceeds around the barrel in a unique direction.[15–17] GAO and RubisCO share all these features of TIM. The cross section of the barrel can vary in eccentricity:[18] of the structures considered here GAO has the most nearly circular sheet; TIM the most elongated. In some cases, a region between successive strands contains an extra helix, in addition to the one packed against the periphery of the sheet. In muconate lactonizing enzyme one of the helices is absent.[4]

The existence of such a set of similar structures, arising from molecules with very dissimilar sequences, poses a number of questions.[19] Why are the folds so similar: Do the proteins have similar folds because they are related by evolution? Or are they similar because there is only one way to satisfy a set of structural requirements? Underlying these problems is the more specific question: What is the nature of the interactions that stabilize this conformation, and how are they consistent with such a wide variety of sequences?

The principles that govern the chain topology,[15–17] the helix–sheet packing,[14] and the geometry of closed β-sheets[18,20] have already been described. What remains to be understood is the nature of the packing within the barrel and this is the main subject of this paper.

Received September 12, 1988, accepted December 20, 1988.

Address reprint requests to Arthur M. Lesk, European Molecular Biology Laboratory, Meyerhofstr. 1, Postfach 1022.09, D-6900 Heidelberg, Federal Republic of Germany.

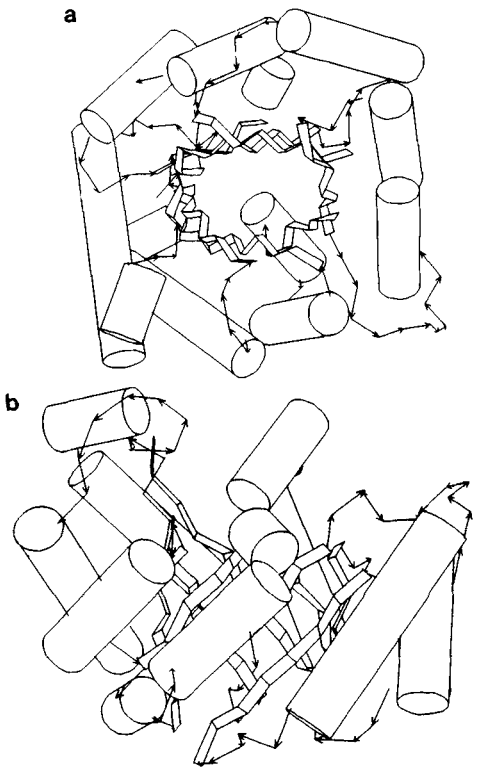

Fig. 1. Triose phosphate isomerase: **a:** Looking into the sheet; **b:** perpendicular to the sheet. The helices are shown as cylinders and the strands of sheet as ribbons.

THE GEOMETRY OF THE β-SHEETS IN α/β BARRELS

From considerations of the geometry of β-sheets a number of distinct folds might exist for α/β barrels.

Schematic plans of the β-sheets in TIM and GAO are shown in Figure 2. For each protein the first strand of the β-sheet is drawn twice: at both sides of the diagram. To make a model of the eight-stranded barrel structure corresponding to that observed in proteins, the sheet must be folded back and the two images of first strand glued over each other: superposing A onto A′ and B onto B′.

The longer the strands, the greater would be the "depth" of the barrel. In the observed structures the strands vary in length. However, all strands, in the three structures considered here, contain three residues at the same height relative to the barrel axis and these form a continuous hydrogen bonded net around the barrel (as indicated in Fig. 2). We will show in the next part of the paper that residues from this section of the sheets pack to form the interior of the barrel. Here we discuss the geometry of these common regions.

A simple theoretical model for the geometry of barrels formed from β-sheets was presented some time ago by McLachlan.[20] Recently Lasters et al.[18] presented a more detailed model. McLachlan defined barrel geometries in terms of the following:

1. The number of strands.
2. The shear number, S, a measure of the stagger of the strands. Starting from any residue in the sheet, a pathway traced on the sheet perpendicular to the strand direction will reintersect the starting strand at a residue displaced by S residues from the starting one (Figure 2).
3. The tilt of the strands with respect to the barrel axis.
4. The twist of the strands, or the average angle between adjacent strands.

McLachlan also described the mathematical relationships among these quantities (see Table I).

Because the geometry of the sheet can vary only within fairly narrow limits (the distance between successive Cα's, projected onto the cylinder surface, is 3.4 Å, and the perpendicular interstrand distance is 4.5 Å), the choice of the parameters of strand number and shear determines the angle between the strands and the cylinder axis. McLachlan's description accounts quantitatively to high accuracy for the features of TIM, GAO, and RubisCO (Table I). For these and all known protein structures showing the closed α/β barrel fold there are 8 strands, the shear (S) is 8, and the tilt of the strands to the barrel axis is approximately −36°. The negative sign of the tilt arises from the right-handed twist of the β-sheets. The average twist angle between successive strands is 26°, a value similar to that observed in open β-sheets.[14]

Lasters et al.,[18] following the work of Novotny et al.,[21] developed a hyperboloid model for the geometry of β-barrel structures. Hyperboloids are fitted to the coordinates of the sheet Cα atoms. Parameters derived from this the model are then used to determine the relationships among the geometrical features. Lasters et al.[18] analyzed the β barrels in nine different protein structures. Their results showed that although the mean radius of the barrels is fairly constant, the eccentricity of the central cross section varies, with axial ratios between 1.0 and 1.48.

THE PACKING OF RESIDUES INSIDE β-BARRELS

The structure of an assembly of α-helices or β-sheets in a protein is determined by the available low-energy conformations and the way in which they can achieve a close packing. In this section we describe how the close packing of residues in the interior of the barrel is achieved.

Glycolate Oxidase

Glycolate oxidase exhibits most clearly the pattern of packing. It is the most nearly circular in cross section, and shows the most symmetrical ar-

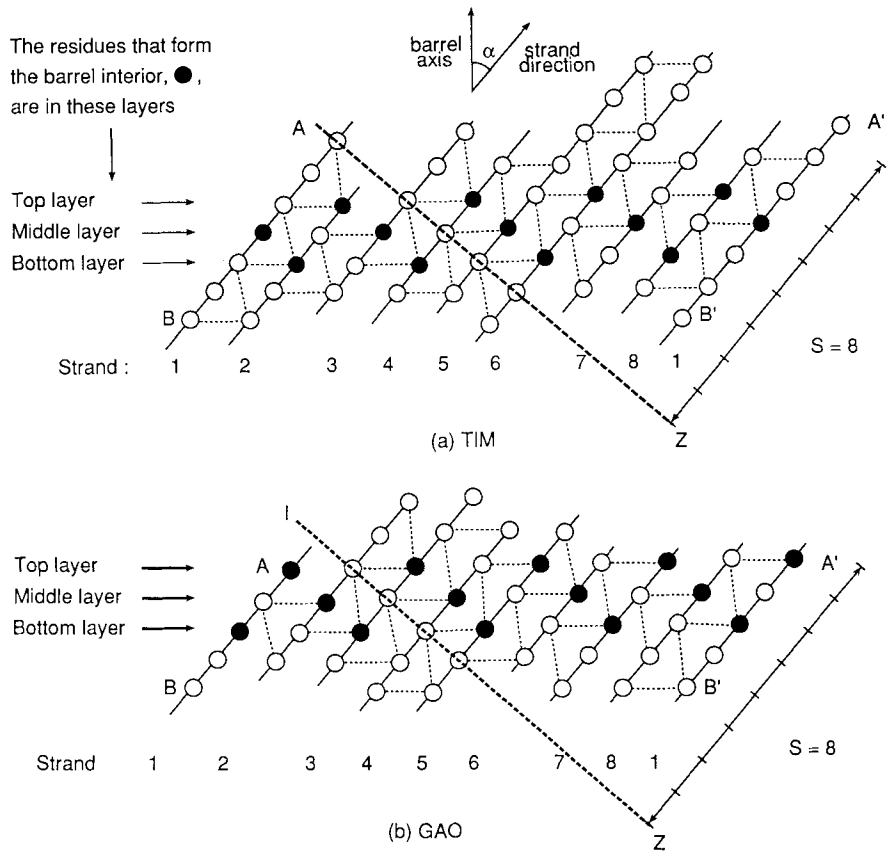

Fig. 2. Schematic diagrams of the parallel β-sheets in **a:** triose phosphate isomerase and **b:** glycolate oxidase. Each circle represents a residue. Broken lines represent hydrogen bonds. Observed structures contain eight strands. Here nine strands are shown, as the first stand is drawn on both the left and right edges of the sheets. To make the eight-stranded barrel structure, the leftmost strand must be superimposed onto the rightmost by curling the paper back and the glueing of residues A and B to residues A′ and B′. (If the paper is curled forward we get a structure in which the sheet twist has the wrong hand, the inside residues are outside and the strand tilt relative to the barrel axis is in the wrong direction.) Superimposing residues A and B onto A′ and B′ creates a barrel with strands tilted with respect to its axis, by an angle α. To form an eight-stranded barrel with strands *parallel* to its axis, A in **(a)** would have to be glued onto point Z. The shear number *S* is a measure of this stagger of the β-sheet strands.[20] For the structures illustrated, it is given by the number of residues that displace Z from A′. Here *S* is 8. Other closed structures might be constructed by glueing A and B to other pairs of residues. Obvious restrictions are that one may not glue a residue with side chains pointing into the page onto a residue with side chains pointing out of the page. This means that for all barrels of this type, whether they contain an odd or even number of strands, the shear must be even. At the center of the barrel the side chains of 12 residues pack together. These are indicated on the β-sheet plans by shaded circles. The 12 residues are in arranged in three sets of four, with the residues in the same set being at the same height relative to the barrel axis. If one or more of the strands were made antiparallel to the others the pattern formed by the internal residues would be the same.

rangement of side chains in the region inside the sheet.

Figure 3a shows the hydrogen-bonding net that arises by "rolling out the barrel." Nine strands are shown; the one at the edge is duplicated. On each strand of sheet, alternate side chains point toward the region inside the sheet and out toward the helices. The packing inside the barrel is formed by the interactions of the 12 residues that have their side chains pointing inward. Two residues are contributed by each of the first, third, fifth, and seventh strands and one residue by each of the second, fourth, sixth, and eighth strands. Residues with letters identifying the side chains indicate the positions of these inward-pointing residues. Note that residues at the same height along the axis of the sheet (vertical in Figs. 2 and 3a) are not nearest neighbors on adjacent strands; this is because of the tilt of the strands with respect to the barrel axis.

The packing of these side chains in the barrel interior is shown in Figure 3b, a side view of the sheet of GAO, pruned to three residues per strand. The side chains occupy three tiers of layers, with almost perfect segregation. Each layer contains side chains from four alternate strands. The four strands that contribute residues pointing "in" on one level have their side chains pointing "out" on the next. The layering is a consequence of the tilting of the strands with respect to the axis. The tilt angle of $-36°$ that produces the layering is a consequence of the sheet geometry; in particular, of the number of strands and the shear.

TABLE I. The Fit of the Observed β Barrel Geometries to McLachlan's Model[20,*]

McLachlan's formulas for the predicted values of
α, the slope of the strands to the barrel axis,
R, the mean radius of the barrel, and
τ, the twist of the β-sheet = the number of turns by which the plane of the β-sheet twists on moving from one residue to the next along a strand, are

$$S = nb \tan \alpha/a, \qquad R = b/[2 \sin(\pi/n)\cos \alpha], \qquad \text{and } \tau = a \sin \alpha \cos \alpha/2\pi R$$

where
a = Cα–Cα distance along the strands,
b = distance between neighboring strands,
n = number of strands, and
S = shear of the sheet (Fig. 2)
For all values listed below that are designated predicted, the values of a and b were those determined for the particular structure, and $S = n = 8$. For each structure the sheet was pruned to three residues per strand.

(a) Glycolate oxidase		
Residues: 74–76 103–105 125–127 152–154 228–230 249–251 283–285 305–307		
a = Average Cα-Cα distance along strands:		3.2±0.2 Å
b = Average interstrand distance:		4.5±0.4 Å
α:	Predicted value = arctan (3.24/4.46) =	36.0°
	Observed value	35.9±8.0
Radius:	Predicted value = 4.46/[2 sin(π/8) cos(35.96)] =	7.2 Å
	Observed value =	7.3 Å
Twist:	Predicted value = 3.24 sin α cos α/2πR =	0.034
	Observed value =	0.034
(b) *Rhodospirillum rubrum* RubisCO		
Residues: 161–163 189–191 228–230 259–261 284–286 319–321 365–367 389–391		
a = Average Cα–Cα distance along strands:		3.3±0.2 Å
b = Average interstrand distance:		4.4±0.3 Å
α:	Predicted value = arctan (3.27/4.41) =	36.5°
	Observed value	35.0±11.4°
Radius:	Predicted value = 4.41/[2 sin(π/8) cos(34.98)] =	7.0 Å
	Observed value =	7.3 Å
Twist:	Predicted value = 3.24 sin α cos α/2πR =	0.035
	Observed value =	0.035
(c) Chicken triose phosphate isomerase		
Residues: 8–10 40–42 61–63 90–92 123–125 161–163 206–208 228–230		
a = Average Cα–Cα distance along strands:		3.4±0.2 Å
b = Average interstrand distance:		4.2±0.6 Å
α:	Predicted value = arctan (3.38/4.20) =	38.8°
	Observed value	36.5±5.5°
Radius:	Predicted value = 4.20/[2 sin(π/8) cos(38.8)] =	7.0 Å
	Observed value =	6.5 Å
Twist:	Predicted value = 3.38 sin α cos α/2πR =	0.037
	Observed value =	0.038

*The procedure used to derive the predicted values was as follows: the values of a and b were determined from the chosen residues and averaged. To determine b, the distance between neighboring strands was computed by finding the least-square line through the N, Cα, and C atoms of each strand, and calculating the perpendicular distance between the lines fitting neighboring strands. These values of **a** and **b** were substituted in McLauchlan's formulas. The observed values of α, R, and the twist were derived as follows. The axis of each structure was determined by fitting the chosen Cα atoms to a right circular cylinder. The radius of the sheet was taken to be the radius of this cylinder, that is, the average radius of the Cα atoms in the structure. The angle of inclination α of each strand was taken to be the angle between the axis of the structure and the least-squares line fitted to the N, Cα, and C atoms of the strand. The value of the twist was computed from McLachlan's formula using observed values of the angle α and the radius.
Because McLachlan's formulas do not distinguish enantiomorphs, the values of the twist angles appearing in this table are given as unsigned rather than negative numbers.

In this packing there is a double alternation similar to that of a chessboard: the side chains that point "in" are on odd-numbered levels on odd-numbered strands, and on even-numbered levels on even-numbered strands.

It is possible to read directly from the unrolled sheet in Figure 3a the residues that form each layer in the interior of the sheet. Thus, the topmost (C-terminal) layer should contain side chains from residues Ala-76, Gln-127, Lys-230, and Asp-285. The middle layer should contain side chains from residues Thr-104, Ala-153, Ile-250, and Phe-306. The bottom layer should contain side chains from residues Met-74, Phe-125, Leu-228, and Phe-283. The actual packing of these residues is seen in Figure 4, which shows serial sections cut through a space-filling model of the GAO sheet. The sections are at levels corresponding to the three layers. In each case, three slices separated by 1 Å are shown. Atoms from odd-numbered strands are outlined by broken

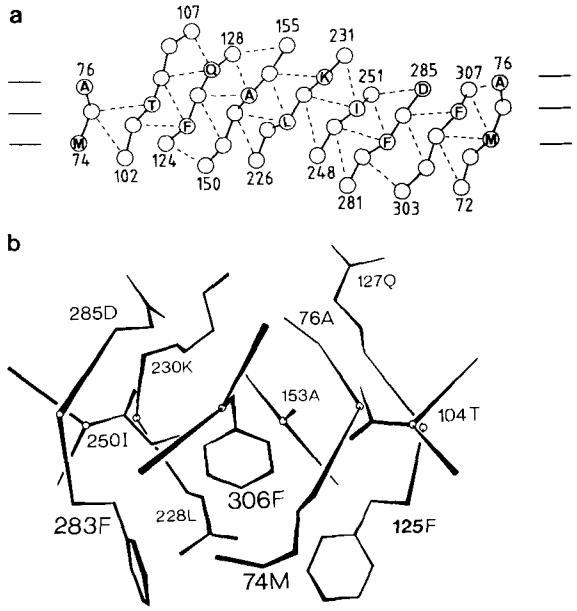

Fig. 3. Glycolate oxidase. **a:** The β-sheet of GAO, drawn by unrolling the barrel. This figure and Figures 5 and 7 are cylindrical projections drawn from the atomic coordinates and therefore give an accurate picture of the residue positions. Each circle represents a residue; those with letters identify the residues whose side chains pack in the central region of the barrel. Broken lines represent hydrogen bonds. Note that nine strands are shown; the one at the edge is duplicated. The axis of the sheet is vertical in this figure. **b:** Residue packing at the center of the barrel. Strands of the GAO β-sheet are pruned to three Cα per strand and the side chains that fill the center of the barrel.

lines, and atoms from even-numbered strands are outlined by solid lines. Each slice shows the packing of four residues on each tier. In the outer layers the four packed side chains are outlined by broken lines, indicating that they arise from alternate strands. In the central layer, the four packed side chains are outlined by solid lines, indicating that they come from the other four alternate strands.

Inspection of the space-filling drawings suggests that there is a cavity in the center of the barrel (Fig. 4b). Calculations of atomic volumes using the Voronoi polyhedron program of Richards[22] show that the total volume of the residues in the interior of the barrel is within 3% of that expected from normal packing densities of protein interiors. The overall packing density is not significantly less than normal.

The packing inside the sheet is limited to three layers. The four residues that might form another layer at the N-termini of the strands are Ile-102, Ala-151, Gly-248, and Gly-304. These residues, which include two glycines and an alanine, are too small to form a proper layer. But even if this were not the case their close approach would be prevented by the protrusion of the side chains that form the bottom layer within the barrel. The protrusion of all

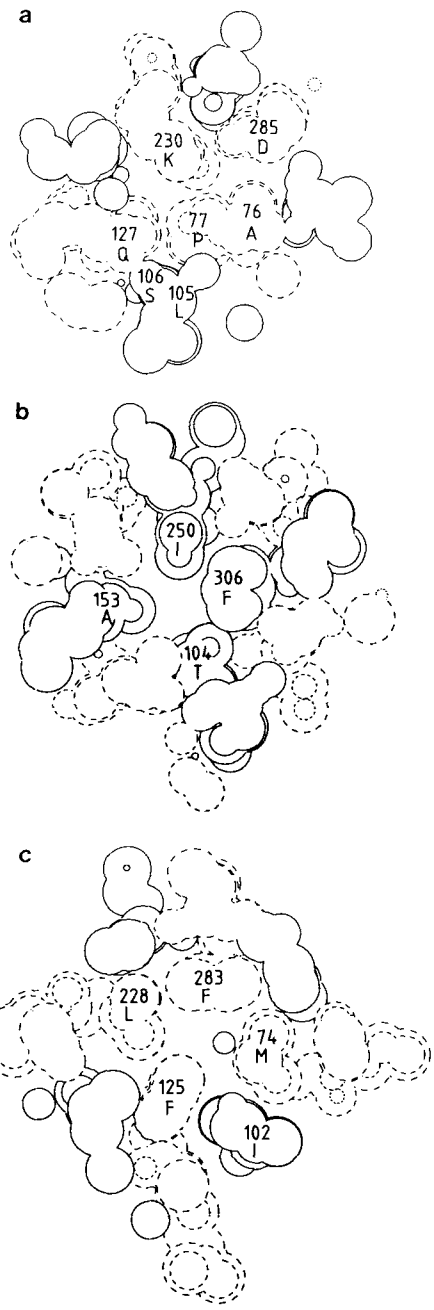

Fig. 4. **a,b,c:** Serial sections cut through a space filling model (van der Waals slices) of the three layers of residues packing inside the barrel of spinach glycolate oxidase. In each drawing three slices separated by 1 Å are shown. Atoms from alternate strands are drawn by solid and broken lines.

four of the side chains that form the bottom layer, Met-74, Phe-126, Leu-228, and Phe-283 (see Fig. 3b), causes the N-termini of the strands to bend away from the barrel axis. Similarly the extension of the side chains on the top layer, Gln-127, Lys-230, and Asp-285, prevents the formation of a fourth layer of residues at the C-termini of the strands of the sheet.

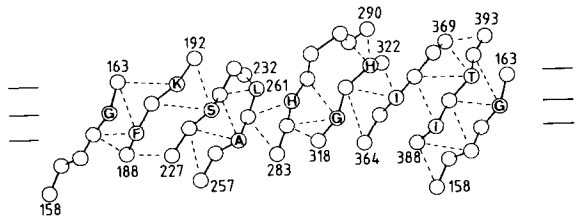

Fig. 5. The β-sheet of RubisCO.

These results suggests a simple model for the packing of residues inside the sheet of glycolate oxidase:

1. On each strand residues alternate in pointing into and out from the interior of the barrel.
2. The tilt of the strands relative to the axis of the sheet and the twist of the sheet places the residues in layers, each containing four side chains from alternate strands. The central region of the barrel is filled by 12 side chains from three of these layers (Fig. 3).
3. Qualitatively, the packing is a layered a–b–a type structure. The geometric operation relating successive layers is a rotation by 45° around the axis of the barrel and a translation along this axis by approximately 3 Å (Fig. 4).
4. The formation of a fourth layer is prevented by the protrusion of the side chains from the top and bottom layers (Fig. 3b).

RubisCO

The packing pattern in RubisCO is similar to that in GAO. From the unrolled β-sheet (Fig. 5) one expects the topmost layer to contain the side chains of residues Lys-191, Leu-261, His-321, and Thr-391, the middle layer to contain side chains of residues Gly-162, Ser-229, His-285, Ile-366, and the bottom layer to contain side chains from residues Phe-189, Ala-259, Gly-319, and Ile-389. The top layer fits the picture given by GAO most closely. On the middle and bottom layers, however, one of the residues expected to contribute an inward-pointing side chain is a glycine. This produces a distortion of the regular packing arrangement, because the hole in the structure is filled by another side chain. Consequently the segregation of side chains into layers is not as simple as that in glycolate oxidase.

The residue packing is shown in Figure 6: a set of serial sections cut through a space filling model of the RubisCO barrel. The top layer (Fig. 6a) is fairly regular, although the packing is less symmetrical than that of GAO because of the side chain of Lys-191 that packs right across the barrel interior. In the central layer (Fig. 6b) the region that would correspond to the side chain of residue 162, a glycine in RubisCO, is occupied by the side chain atoms of residues Phe-189 and Ile-389, both from the next

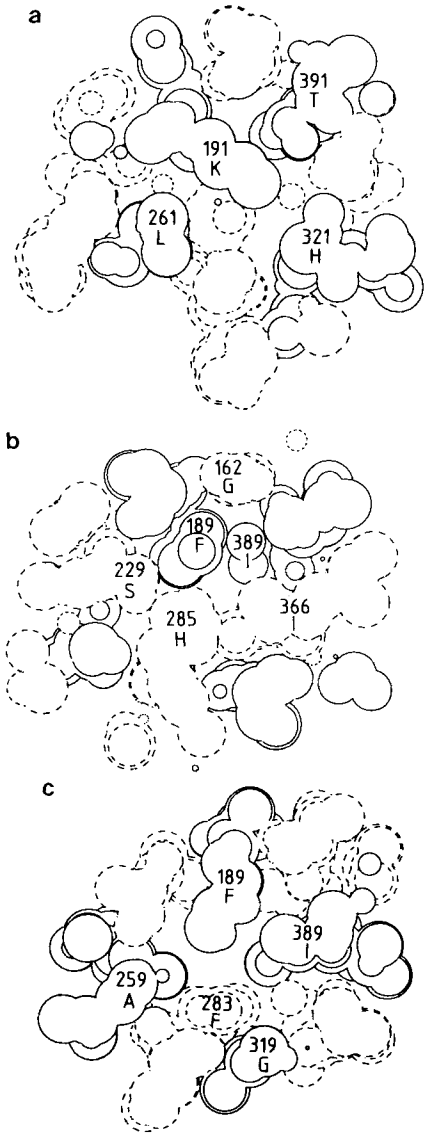

Fig. 6. **a,b,c:** van der Waals slices through the three layers of residues packing inside the sheet of RubisCO.

layer. In the bottom layer (Fig. 6c) the region expected to be occupied by the side chain of residue 319, were it not a glycine, is occupied by residue Phe-283 from the region outside the bottom layer.

The RubisCO barrel only contains three well-packed layers within the sheet. As in GAO, the formation of a fourth layer is prevented by the protrusion of residues from the top and bottom layers: His-321 and Ile-389.

Triose Phosphate Isomerase

Of the three proteins considered here, the barrel of TIM has the most asymmetric cross section, that is, the furthest from circular. It also shows the greatest deviation from the paradigm layered structure of

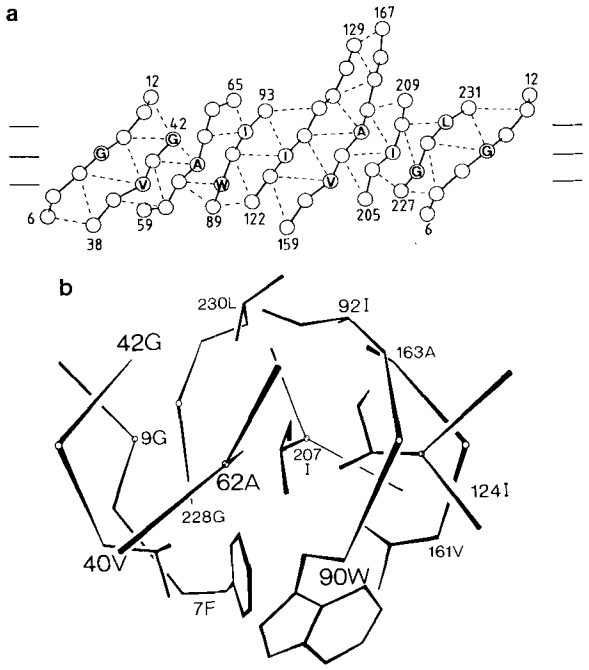

Fig. 7. Chicken triose phosphate isomerase. **a:** The β-sheet of TIM. **b:** Residue packing at the center of the barrel. All but the first strand of the TIM β-sheet are shown pruned to the central three Cα and the side chains that fill the center of the barrel. The first strand also includes Phe-7 whose side chain fills the cavity next to Gly-228.

the residues in its interior. In chicken TIM, of the 12 residues the side chains of which are expected to form the barrel interior, three are glycines and two are alanines (Fig. 7). Although the substantial reduction in the size of the sheet residues that pack the interior creates distortion, the structure can still be understood in terms of the same basic pattern of residue packing.

Figure 7a shows the unrolled β-sheet of TIM, on which we indicate the 12 residues that pack at the center of the barrel. Figure 8 shows slices through the barrel of TIM. In the top layer the interior sheet residues are Gly-42, Ile-92, Ala-163, and Leu-230. The cavity next to Gly-42 is filled by the side chain of residue Asn-11, from above (Fig. 8a). The side chain of this residue is buried and its position is stabilized by hydrogen bonds.

In the central layer, the interior region is expected to be occupied by the side chains of residues Gly-9, Ala-62, Ile-124, and Ile-207. In this structure, the cavity next to Gly-9 is partly filled by the backbone movement that produces the noncircular cross section of the barrel. This is only partly successful and the structure has a cavity adjacent to this residue (Fig. 8b); such large cavities are rare but not unknown in other protein structures.[23]

In the bottom layer, the region inside the sheets is expected to be occupied by the side chains of Val-40,

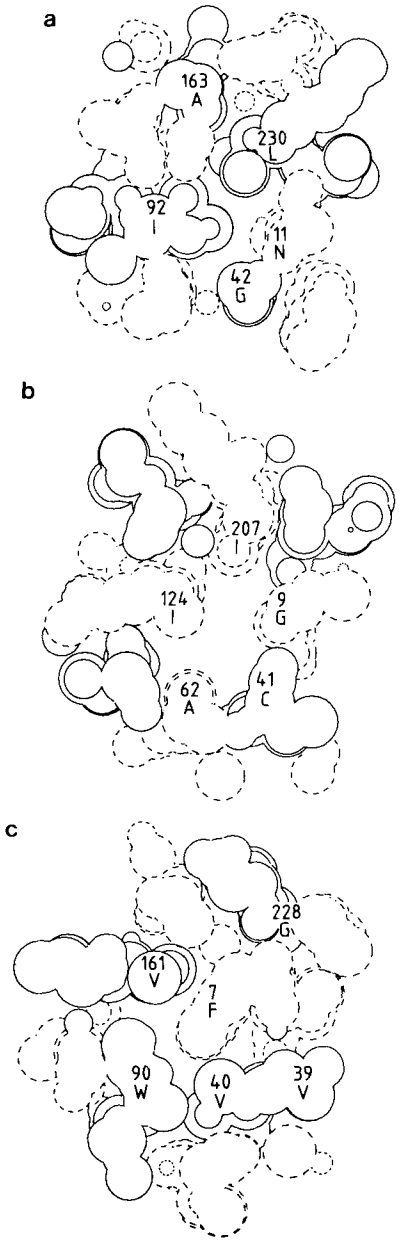

Fig. 8. **a,b,c:** van der Waals slices through the three layers of residues packing inside the sheet of TIM.

Trp-90, Val-161, and Gly-228. The region that Gly-228 would be expected to occupy is filled by the side chain of Phe-7, from below (Fig. 8a).

TIM contains only three well-packed layers within the sheet. The formation of a fourth layer is prevented at the top by the protrusion of Trp-90 and at the bottom by the protrusion of Ile-92 and Leu-230.

In summary, the packing of the region inside the sheet in GAO, RubisCO, and TIM can be understood in terms of a three-tiered arrangement in which alternate strands contribute side chains to alternate layers. When interior residues are small—glycines

or alanines—part of a residue on an adjacent layer or a residue from outside the barrel inserts its side chain into the layer; or the backbone may be deformed. The formation of a fourth layer is prevented by the protrusion of side chains from the top and bottom layers.

DISCUSSION

Two Classes of α/β-Barrels

Underlying the similarity of folding of α/β barrels, there are really *two* distinct classes. In the continuous hydrogen-bonded girdle of three residues per strand common to the three structures, alternate strands contribute one or two side chains, respectively, to the region within the sheet. Thus, if the strands are numbered 1 through 8 from the N-terminus of the sequence, we call class 1 structures those for which the N-terminal strand (and the other odd-numbered strands) contribute one residue; and we call class 2 structures those for which the N-terminal strand (and the other odd-numbered strands) contribute two residues (see Fig. 2). RubisCO and TIM are both class 1 structures; GAO is in class 2.

If structures in these two classes did not arise independently, there must be a pathway between them that evolution could follow. It is difficult to imagine a simple mechanism by which structures in these two classes could be interconverted by a succession of point mutations and the adaptive mechanisms normally observed in protein evolution.[24] The most direct pathway would involve an intermediate structure with two or four layers packing in the barrel interior. As discussed above the formation of a fourth layer is usually prevented by the protrusion of residues from the top and bottom layers. It is not clear that an intermediate with only two layers would be stable.

The active site in all the known the α/β barrel proteins is found at the carbonyl end of the barrel and involves residues that are part of the β-sheet or immediately adjacent to it. This means that the evolutionary interconversion between the two classes by a succession of point mutations would require a complete reconstruction of the active site.

Interconversion between classes by mechanisms involving permuting either the gene or the protein cannot be ruled out by these arguments.[25]

The Quantitative Structural Similarity of β-Barrels Measured by Least-Squares Fitting

The structural features that we have discussed in this paper provide a much more reliable guide to the structural relationships between GAO, TIM, and RubisCO than the simple least-squares fitting of atomic coordinates. Although root mean square (rms) deviations of atomic positions derived from least-squares fitting of a set of core residues is a guide to the closeness of the relationships in homologous structures, their utility depends on being able to establish the correct alignment of the sequences and to determine the regions that have similar folds.[26,27] In the case of the β-barrel structures, the identification of the common packing pattern permits an assignment of equivalent residues in the sheet for superposition calculations.

For TIM and Rubisco whose barrel structures are in the same class, the superposition of the 24 residues that form the common sheet structure, the three in each of the eight strands, gives an rms difference in the position of main chain atoms of 2.0 Å. Superposition of the 38 residues common to the β-sheets of these particular two proteins (Figs. 5 and 7) gave an rms difference in mainchain atomic position of 2.1 Å. (It should be noted that these fits of atomic coordinates involve sets of residues that are smaller than those used in deriving relationships between homologous proteins.[27] This means that the rms differences derived from the two types of calculations are not comparable.)

Because the barrel in GAO is not in the same class as those in TIM or RubisCO, it can be superposed only by allowing the odd-numbered strands of GAO to correspond to the even–numbered or TIM or RubisCO. If strands 1, . . . , 8 of TIM are superposed on strands 2, . . . , 8, 1 of GAO, the rms difference in the position of the main chain atoms of the 24 common sheet residues is 1.7 Å. For RubisCO and GAO, the value is 1.0 Å. This value is half that found from the fitting of TIM and RubisCO, barrels of the same class. The lower value for the GAO–RubisCO comparison arises from the similarity of their overall shapes: they have nearly circular cross sections, whereas the cross-sectional axial ratio in TIM[18] is 1.48.

Alternative Closed Barrel Structures

All the closed barrel structures so far found in α/β proteins are formed by sheets with eight strands with a shear number of eight. This structure involves optimal sheet geometry and normal sheet twist.[18,20] We have shown here that at the center of the barrel this structure places 12 side chains in a regular array and in a volume close to that of the average side chain. The packing arrangement is flexible and robust: very large residues in the barrel are accommodated by the protrusion of side chains at the top or bottom of the barrel. Small side chains are usually accommodated by the insertion of side chains from outside the barrel.

Why do the closed barrels so far observed in proteins all share the same structure? Alternative structures would involve differences in the β-sheet geometry or residue packing or both. The geometrical analyses of McLachlan[20] and Lasters et al.[18] show that closed barrel structures close to, but different from, the observed one can be formed with optimal β-sheet geometries and twists. This sug-

gests that the major reason for the preference for the observed topology must come from the residue packing.

Lasters et al.[18] pointed out that the radius of a 10-stranded barrel, about 8.9 Å, would have an interior too large to be filled by a normal set of residues. Similarly, a barrel of six strands, with a radius of about 5.5 Å, would have an interior that is too small.

Consider barrels with seven or nine strands. In a β-sheet strand alternate residues form hydrogen bonds on the same side of a strand. This means that the shear number of a barrel has to be even (see Fig. 2). Odd shear numbers would require the residues in one strand to form hydrogen bonds in both directions. (Another way to see that the shear must be even is to recognize that one must superpose residues that have side chains pointing in the same direction with respect to the local plane of the sheet, and that on any strand the side chains that point in the same direction are separated by an even number of residues.) This means that the structures closest to that observed are a seven stranded barrel with a shear of six or eight or a nine stranded barrel with a shear of eight or ten.

Seven-stranded barrels with shears of six or eight would have irregular arrangements of the interior residues. This can be seen if in the sheet diagram of TIM shown in Figure 2 one superposes the buried (shaded) residue in the first strand onto the top or bottom buried residues in the eighth strand. Such a barrel, if unrolled as in Figures 3a, 5, and 7a, would not show the regular pattern of alternation of interior residues seen in GAO, RubisCO, and TIM. The same result is obtained if we consider a nine-stranded barrel with shears of eight or ten. In addition, the seven-stranded barrel with a shear of six has a rather small radius, 6.2 Å, and the nine-stranded barrel with a shear of ten a rather large one, 8.6 Å.

Thus, though alternative barrel structures, with strands and shear numbers close to the natural ones cannot be ruled out, the irregular arrangement of side chains that would occur in their interiors means that they would require special sets of residues, and therefore be unstable to evolution.

CONCLUSIONS

We have described a common structural principle for the packing of residues in the interior of α/β barrels. It provides a means for identifying residues that play common roles in the different structures and, therefore, for the comparison of the relations of their functional residues. These results, together with the previous geometrical analyses,[18,20] show how the unique fold found in α/β barrel proteins provides a particularly favorable solution to the various structural problems involved in the creation of a closed β barrel.

ACKNOWLEDGMENTS

We thank Dr. P. Artymuik for the refined coordinates of chicken triose phosphate isomerase and The Royal Society for support.

REFERENCES

1. Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D., Waley, S.G. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. Nature (London) 255:609–614, 1975.
2. Lindqvist, Y., Brändén, C.-I. Structure of glycolate oxidase from spinach. Proc. Natl. Acad. Sci. U.S.A. 82:6855–6859, 1985.
3. Schneider, G., Lindqvist, Y., Brändén, C.-I., Lorimer, G. 3-dimensional structure of ribulose-1,5-bisphosphate carboxylase-oxygenase from *Rhodospirillum rubrum* at 2.9 Å resolution. EMBO J. 5:3409–3415, 1986.
4. Goldman, A., Ollis, D.L., Steitz, T.A. Crystal-structure of muconate lactonizing enzyme at 3 Å resolution. J. Mol. Biol. 194:143–153, 1987.
5. Stuart, D.I., Levine, M., Muirhead, H., Stammers, D.K. Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. J. Mol. Biol. 134:109–142, 1979.
6. Mavridis, J.M., Hatada, M.H., Tulinsky, A., Lebioda, L. Structure of 2-keto-3-deoxy-6-phosphogluconate aldolase at 2.8 Å resolution. J. Mol. Biol. 162:419–444, 1982.
7. Matsuura, Y., Kusunoki, M., Harada, W., Kakudo, M. Structure and possible catalytic residues of taka-amylase A. J. Biochem. 95:697–702, 1984.
8. Carrell, H.L., Rubin, B.H., Hurley, T.J., Glusker, J.P. X-ray crystal-structure of D-xylose isomerase at 4 Å resolution. J. Biol. Chem. 259:3230–3236, 1984.
9. Lim, L.W., Shamala, N., Mathews, F.S., Steenkamp, D.J., Hamlin, R., Xuong, N.H. 3-dimensional structure of the iron-sulfur flavoprotein trimethylamine dehydrogenase at 2.4 Å resolution. J. Biol. Chem. 261:15140–15146, 1986.
10. Xia, Z.X., Shamala, N., Bethge, P.H., Lim, L.W., Bellamy, H.D., Xuong, N.H., Lederer, F., Mathews, F.S. Three-dimensional structure of flavocytochrome-B2 from bakers-yeast at 3.0–Å resolution. Proc. Natl. Acad. Sci. U.S.A. 84:2629–2633, 1987.
11. Priestle, J.P., Grutter, M.G., White, J.L., Vincent, M.G., Kania, M., Wilson, E., Jardetzky, T.S., Kirshner, K., Jansonius, J.N. 3-dimensional structure of the bifunctional enzyme N-(5′-phosphoribosyl) anthranilate isomerase-indole-3-glycerol-phosphate synthase from *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 84:5690–5694, 1987.
12. Hyde, C.C., Ahmed, S.A., Padlan, E.A., Miles, E.W., Davies, D.R. Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. J. Biol. Chem. 263:17857–17871, 1988.
13. Lebioda, L., Stec, B. Crystal-structure of enolase indicates that enolase and pyruvate kinase evolved from a common ancestor. Nature (London) 333:683–686, 1988.
14. Janin, J., Chothia, C. Packing of α-helices onto β-pleated sheets and the anatomy of α/β proteins. J. Mol. Biol. 143:95–128, 1980.
15. Sternberg, M.J.E., Thornton, J.M. On the conformation of proteins: The handedness of the β-strand-α-helix-β-strand unit. J. Mol. Biol. 105:367–382, 1976.
16. Richardson, J.S. Handedness of crossover connections in β sheets. Proc. Natl. Acad. Sci. U.S.A. 73:2619–2623, 1976.
17. Sternberg, M.J.E., Cohen, F.E., Taylor, W.R., Feldman, R.J. Analysis and prediction of structural motifs in the glycolytic enzymes. Phil. Trans. Roy. Soc. (London) B293: 177–189, 1981.
18. Lasters, I., Wodak, S.J., Alard, P., van Cutsem, E. Structural principles of parallel β-barrels in proteins. Proc. Natl. Acad. Sci. U.S.A. 85:3338–3342, 1988.
19. Chothia, C. Protein structure—the 14th barrel rolls out. Nature (London) 333:598–599, 1988.
20. McLachlan, A.D. Gene duplications in the structural evolution of chymotrypsin. J. Mol. Biol. 128:49–79, 1979.
21. Novotny, J., Bruccoleri, R., Newell, J. Twisted hyperboloid (*strophoid*) as a model of β-barrels in proteins. J. Mol. Biol. 177:567–573, 1984.

22. Richards, F.M. The interpretation of protein structures: Total volume, group volume distributions and packing density. J. Mol. Biol. 82:1–14, 1974.
23. Rashin, A.A., Iofin, M., Honig, B. Internal cavities and buried waters in globular proteins. Biochemistry 25:3619–3625, 1986.
24. Chothia, C., Lesk, A.M. The evolution of protein structures. Cold Spring Harbor Symp. Quant. Biol. LII:399–405, 1987.
25. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., Kirschner, K. Correct folding of circularly permuted variants of a αβ barrel enzyme *in vivo*. Science 243:206–210.
26. Lesk, A.M., Chothia, C. The response of protein structures to amino acid sequence changes. Phil. Trans. Roy. Soc. London 317:345–356, 1986.
27. Chothia, C., Lesk, A.M. Relationship between the divergence of sequence and structure in proteins. EMBO J. 5: 823–826, 1986.