PROTEINS: Structure, Function, and Genetics 14:191–201 (1992)

Structural Principles for the Propeller Assembly of β-Sheets: The Preference for Seven-Fold Symmetry

Alexey G. Murzin
MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, England

ABSTRACT **Twisted β-sheets, packed face to face, may be arranged in circular formation like blades of a propeller or turbine. This β-propeller fold has been found in three proteins: that in neuraminidase consists of six β-sheets while those in methylamine dehydrogenase and galactose oxidase are composed of seven β-sheets. A model for multisheet packing in the β-propeller fold is proposed. This model gives both geometrical parameters of the β-propellers composed of different numbers of sheets and patterns of residue packing at their sheet-to-sheet interfaces. All the known β-propeller structures have been analyzed, and the observed geometries and residue packing are found to be in good agreement with those predicted by models. It is shown that unusual seven-fold symmetry is preferable to six- or eight-fold symmetry for propeller-like multisheet assembly. According to the model, a six-β-sheet propeller has to have predominantly small residues in the β-strands closed to its six-fold axis, but no strong sequence constraints are necessary for a seven-fold β-propeller.**

Key words: common protein fold, protein architecture, close packing, β-sheet twist, galactose oxidase, influenza virus neuraminidase, methylamine dehydrogenase

INTRODUCTION

A new protein folding motif with seven-fold pseudosymmetry has been discovered recently in methylamine dehydrogenase[1,2] (MADH) and galactose oxidase[3] (GAO). It presents an assembly of seven four-stranded antiparallel β-sheets with the same up-and-down topology. With the first strands running parallel to the seven-fold axis and radially directed hydrogen bonds, twisted β-sheets look like blades of a propeller or turbine (Fig. 1). Previously a similar motif with six-fold pseudosymmetry had been reported for influenza virus neuraminidase[4] (NA). These structures raise two important questions:

1. Is the propeller-like fold (β-propeller) the result of special sequence constraints or is it the consequence of general packing rules?

2. Is the unusual seven-fold symmetry contingent or is it the most expected symmetry for the β-propeller fold?

A simple superposition of 3-D structures (Fig. 2) shows that the relative orientation of two adjacent β-sheets in the β-propeller is the same as in an ordinary β-sandwich protein and can be explained by the model[5,6] for the packing of aligned β-sheets. That general model, however, cannot explain the circular multisheet architecture, but it may easily be extended to this particular case as will be shown here.

Additions to the general model are proposed to explain how β-sheets assemble in a propeller-like fashion. Then geometrical parameters (interstrand angles and distances, radii of inner holes, etc.) for ideal β-propellers with different numbers of β-sheets are calculated. The extended model predicts that the seven-fold β-propeller satisfies average β-sheet geometry best, but β-propellers with six- or eight-fold symmetry are also allowed.

To test the proposed model an analysis of all three known β-propeller structures has been carried out and the observed geometries have been compared with those calculated. The good agreement between the observed and model geometries supports the idea that the (seven-fold) β-propeller presents the new example of a common protein fold like the eight-fold α/β-barrel.

THE MODEL FOR PROPELLER ASSEMBLY OF β-SHEETS

The model proposed here describes both the geometrical parameters of the propeller assembly and the residue packing in its sheet-to-sheet interfaces.

The geometries of β-sheet structure are treated as constants in the present analysis—that is, the half-period, or length of residue in the strand direction, and the distance between two hydrogen-bonded

Received July 19, 1991; accepted December 4, 1991.
Address reprint requests to Dr. Alexey Murzin, Medical Research Council, MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, England.
Dr. Alexey G. Murzin's permanent address is the Computer Research Centre, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia.

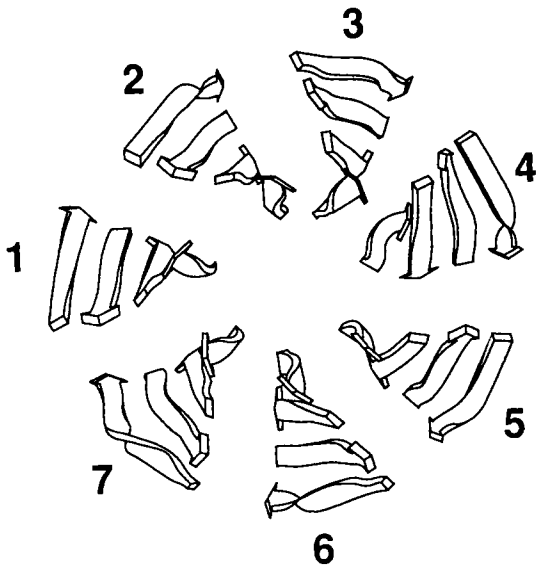

Fig. 1. Computer-generated diagram of β-propeller of galactose oxidase (GAO), viewed along pseudo-7-fold axis. Ribbon arrows represents β-strands that are truncated to the same length (five residues) to give a rectangular shape to all β-sheets. Sharp twist in some arrows is due to β-bulges. β-Sheets are sequentially numbered, beginning from N-terminal one (sheet 7 is composed of N- and C-termini of β-propeller domain).

strands. The distances between opposite strands of two close packed β-sheets and the twist of β-sheets are *not* independent in the propeller assembly, but are related to each other and to the number of β-sheets. There is some freedom to choose the distance between β-sheets in the middle of their interface as a constant and to consider the other interstrand distances and the β-sheet twist as variables.

Residue packing restricts variations of the interstrand distances and other geometrical features. It helps to determine the optimal number of β-sheets for the propeller assembly.

For this reason a qualitative description of two modes of residue packing (Fig. 3), determined by periodicity of β-structure, precedes the quantitative description of the geometries of β-propeller folds.

The Periodicity of β-Structure and Two Modes of Residue Packing in an Aligned Sheet-to-Sheet Interface

The main features of aligned β-sheet packing will be outlined. The strands of the top and the bottom sheets are oriented so that side chains from two opposite β-strands are aligned in the sheet-to-sheet interface. Due to the right-handed twist of β-sheets the main chains of these strand are oriented by some negative angle.[5,7] This angle Ω observed between the middle strands of β-sandwich proteins is about −30°.

Most side chains in the β-sheet are in contact with their neighbors. A side chain in a twisted β-sheet may contact neighbors in its own β-strand as well as those in adjacent strands. Large side chains can make contact over intermediate strands. Since the surface of β-sheets is formed mainly by side chain atoms, it is essentially smooth with only small irregular protuberances and holes.[8] As a consequence, in sheet-to-sheet interfaces intercalation of side chains from top and bottom sheets is the exception rather than the rule.[5]

For the structures discussed here such exceptions are important since intercalated packing of two aligned β-strands allows their close approach (Fig. 3a). In this type of packing small side chains produce the gaps into which small side chains from the other strand fit. Large residues cannot engage in intercalated packing at the middle of a sheet-to-sheet interfaces but can at its edges. There the strands on different β-sheets can approach closer (about 7 Å) to each other than they can at the middle of interface (about 10 Å[5,7]).

Note, if two aligned β-strands are packed with intercalating residues, they are shifted in the strand direction by an *integer number of the periods of β-structure* (or an even number of residues, Fig. 3a).

Relative arrangement of two aligned nonintercalated strands may seem to have no dependence on the periodicity of β-structure. However, this dependence exists. In the middle of sheet-to-sheet interfaces the residues from the top β-sheet point directly at the residues from the bottom β-sheet. This observation,[5] confirmed by visual inspection of a variety of known β-sandwich protein structures, is accepted here as *the rule*. That is, two aligned, nonintercalated strands are shifted in the strand direction by a *half-integer number of periods of β-structure* (or an *odd* number of residues, Fig. 3b). This rule can be explained, if we consider that aligned, nonintercalated β-strands are 10 Å apart. So on average only C_δ and more remote side chain atoms can intersect the middle surface, which is 5 Å from backbone of each strand (C_γ atoms touch the middle surface and C_β atoms are 0.5 Å below it). Thus irregular, small knobs created by these atoms on β-sheet surfaces are shifted from fixed C_β atom positions, as well as shallow holes on the same surfaces. When these knobs pack into the holes, C_β and C_γ atoms of top and bottom residues stand opposite each other (Fig. 3b).

Geometrical Parameters of β-Propeller Assembles With Ideal *N*-Fold Symmetry

Three parameters are considered here as constants: the length of residue in the strand direction, or half-period of β-structure, a = 3.5 Å; the distance between two hydrogen-bonded strands, b = 4.5 Å; and the distance between two close packed β-sheets in the middle of their interface, d = 10 Å. [Generally, a, b, and d characterize β-structure in all the three dimensions—that is, along the main chain direction, along the hydrogen-bonding direction, and

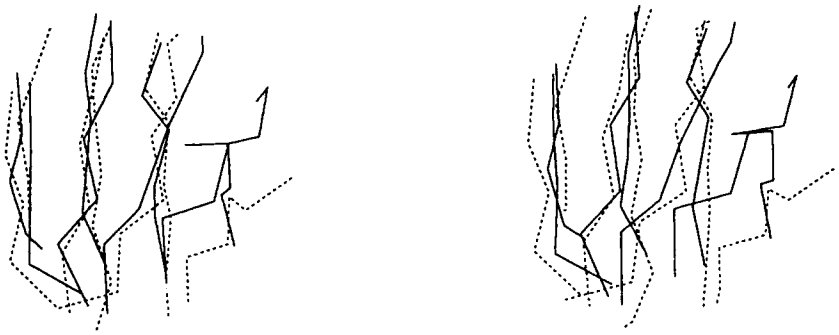

Fig. 2. Stereopicture, illustrating superposition of the first and the second β-sheets of GAO on the plastocyanin two-sheet β-sandwich. β-Strands are represented by their C_α atoms, connected by solid and dotted lines, respectively. This superposition gives root-mean square deviation of 1.5 Å for 38 equivalent C_α atoms, located in seven β-strands of each protein (some strands are superimposed in opposite directions).

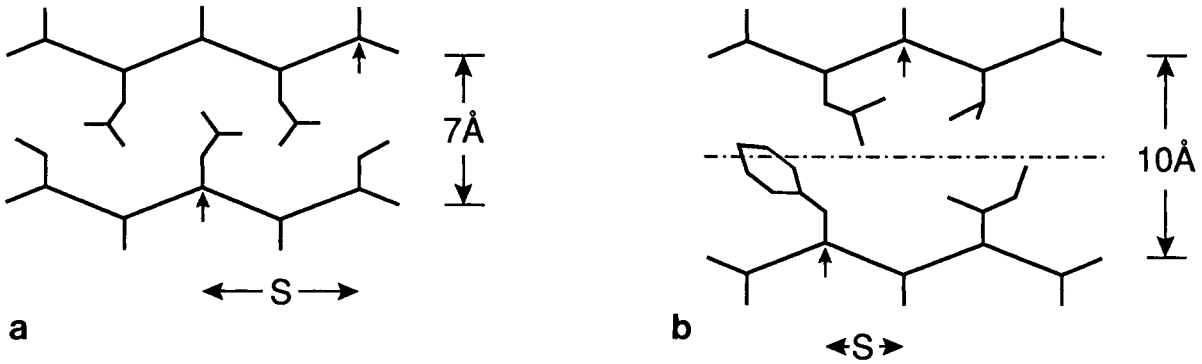

Fig. 3. Scheme of the intercalating **(a)** and the nonintercalating **(b)** packing of residues in two aligned β-strands. **(a)** The 7 Å distance between two intercalating strands results from the contact of the backbone of one strand with C_δ atoms of the opposite strand. This distance would be less than 6.5 Å, if intercalated side chains have no more than γ-atoms. The shift *S* between any two residues, that are in different strands and point in the same direction (marked by short arrows), is equal to integer number of β-structure periods (2 residues). **(b)** For the nonintercalating strands the shift *S* is equal to half-integer number of the periods of β-structure. The 10 Å distance between these strands corresponds to the contacts of C_γ atoms that lie on both sides of the middle surface (dashed line). C_δ (and more remote) side-chain atoms protrude through this surface and fill holes on the opposite β-sheet.

along the side chain (C_α–C_β) direction.] These constants are average values observed in β-sheet proteins and can slightly differ in individual structures.

The β-sheet twist is described here by the angle θ between two hydrogen-bonded strands. The average value of θ in known β-sheets is about −20° with standard deviation of 10°. In a model β-propeller with given number of β-sheets, *N*, it is assumed that the angle θ is the same for the all strand pairs.

Let each β-sheet of a β-propeller have four strands (larger sheets are discussed below). The strands of each sheet are numbered from inner (close to the propeller axis) to outer strand. (The outer strand is the fourth strand in the all known β-propellers.) The second and third strands are referred to as the middle strands.

In an ideal β-propeller all strands with the same number *n* lie on the side surface of cylinder with radius r_n (Fig. 4a).

Let α_n be the dihedral angle between the direction of any strand, number *n*, and the propeller axis. Let Ω_n be the dihedral angle and d_n be the distance between the *n*th strands in two adjacent sheets (Fig. 4a,b). Thus we have the set of coaxial cylinders with obvious relationships between both the radii r_n and the angles α_n:

$$r_{n+1} = r_n + b \tag{1}$$

$$\alpha_{n+1} = \alpha_n + \theta. \tag{2}$$

The interstrand distances d_n and angles Ω_n are given by the following equations:

$$d_n = 2\, r_n \sin(\pi/N)\cos\alpha_n \tag{3}$$

$$\cos\Omega_n = \cos(2\pi/N)\sin^2\alpha_n + \cos^2\alpha_n. \tag{4}$$

It should be noted that Eq. (3) describes distances between *curved* β-strands, lying on the side surface of the corresponding cylinder. It is related to Eq. (5) in ref. 9 that describes the geometries of the β-bar-

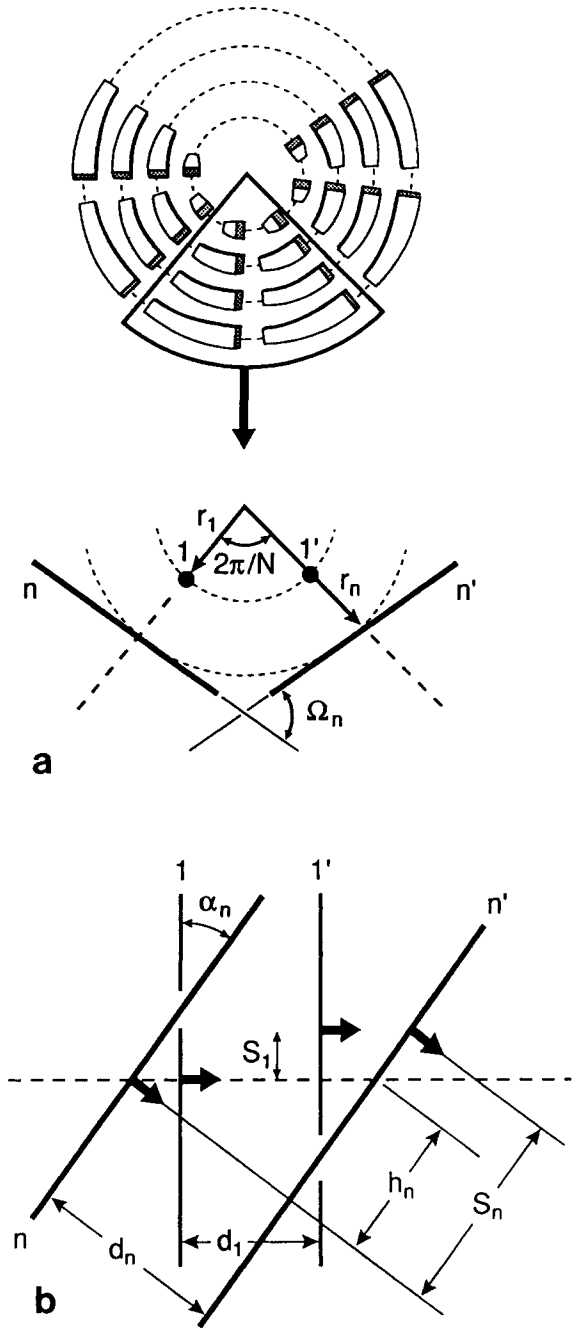

Fig. 4. Scheme describing geometrical features of ideal β-propellers. Two adjacent β-sheets, related by rotation around the *N*-fold axis, are presented with the first and the *n*th strands each on the top view **(a)** and the front view **(b)**. These strands, although represented by straight bold lines (marked 1 and 1', n and n'), are curved and lie on the side surfaces of coaxial cylinders (dashed arches) with the radii r_1 and r_n, respectively. Ω_n is the dihedral angle between the directions of strands n and n'. d_1 is the distance between strands 1 and 1', and d_n is the distance between n and n'. Bold arrows on front view denote residues that are in register in each β-sheet and point in the same direction in each pair of strands. Dashed line corresponds to the plane perpendicular to the β-propeller axis. The strands 1 and 1' are parallel to the propeller axis ($\alpha_1 = 0$) and have no relative displacement ($h_1 = 0$); S_1 is the shift of these strands. α_n is the dihedral angle between *N*-fold axis and the directions of strands, number n. The strands n and n' have the displacement h_n and the shift $S_n = h_n + S_1$.

rels. On the other hand, Eq. (4) describes dihedral angles between *straight* approximations of these strands.

The most prominent feature of the β-propeller is the hole at its axis. The radius of this hole differs from r_1 only by the van der Waals envelopes of the inner strands and it is to be as small as possible. It follows from Eq. (3) that the internal hole radius r_1 would be minimal when $\alpha_1 = 0$. (In this case the radius r_1 depends only on the distance d_1, which cannot be less than the lower limit of the interstrand distances.) Thus, the β-sheets in the β-propeller are packed so that their inner strands go parallel to the propeller axis, and, from Eq. (2):

$$\alpha_n = (n-1)\theta. \tag{5}$$

The β-sheets in the propeller assembly are shifted relative each other. As shown in Figure 4b, consider two points in strands n and n' that lie in a plane perpendicular to the N-fold axis. Define the strand displacement h_n as the projection of the distance between these points on the strand direction (the side surface of cylinder is thought to be unrolled), divided by the residue length a, and so measured in half-periods like the strand shift S (Fig. 3).

$$h_n = d_n \tan \alpha_n / a. \tag{6}$$

Because of absence of a shift between strands in their own β-sheet, the strand displacements h_n and the strand shifts S_n are related by a simple equation (Fig. 4b):

$$S_n - S_1 = h_n - h_1 \tag{7}$$

where S_1 and h_1 are the shift and the displacement of the inner strands in two adjacent β-sheets. $h_1 = 0$, as follows from Eqs. (5) and (6). Here the term *shift* is reserved to describe relationships between the residues of two aligned β-sheets only; in all other cases the term *displacement* is used. One of these cases is the displacement h_m of two adjacent β-sheets, that is calculated not for any pair of strands but for the middle line between the second and third strand:

$$h_m = d \tan 1.5\theta / a. \tag{8}$$

The shift of two β-strands, described by Eqs. (6) and (7), determines their mode of residue packing. It must be consistent with the interstrand distances described by Eq. (3).

Calculation of β-Propeller Geometries

Expecting the twist angle θ be close to the average value −20°, we are able to estimate the geometries of ideal β-propellers with a given number N of β-sheets (particularly for $N = 6$ and 7). It follows that in the absence of the twist ($\theta = 0$) the intersheet distance, d_n increases from inner strands to outer ones, but for the twisted sheets ($\theta > 0$) the intersheet dis-

tance d_n increases to a maximum for some $n=n^*$ and then decreases. For typical values of twist the maximal distance is expected to be in the middle of the β-sheet-to-β-sheet interfaces. Thus the twist of β-sheets is the major factor governing their assembly into the β-propeller.

The twist of β-sheets determines their relative orientation. The relative orientation of two adjacent β-sheets changes from the inner strands to the outer ones. For the inner strands $\Omega_1=0°$; for the outer ones Ω_4 is less than $-60°$; and the expected Ω for the middle strands of β-propeller sheets is about $-30°$, as it is for β-sandwich proteins.[5,7]

Continuing the comparison between the sheets packing in β-propellers and β-sandwiches, it is possible to calculate the exact value of the model twist angle θ by requiring that the separation of two β-sheets in the middle of their interface be equal to $d=10$ Å, i.e., for the middle strands: $d_2=d_3=d$. Then from Eqs. (1), (3), and (5) it follows:

$$1/\cos 2\theta - 1/\cos\theta = 2\sin(\pi/N)b/d. \qquad (9)$$

Solving Eq. (9), we obtain the twist angle θ for given N and are able to calculate all the geometrical parameters of ideal β-propellers with different numbers of β-sheets. These parameters are presented in Table I.

Four-stranded β-sheets are used to derive the model. It may be thought that different the β-propeller geometries might result from β-sheets with different numbers of strands. This is not the case. β-Sheets with more than four strands cannot form propeller assemblies. Calculations show that additional strands, placed either before the inner strands or after the outer ones, would have unacceptably short distances between them.

β-PROPELLERS WITH DIFFERENT *N*-FOLD SYMMETRY: COMPARISON OF GEOMETRIES AND PATTERNS OF RESIDUE PACKING

The calculated twist θ of β-sheets in the β-propellers is slightly greater than that observed in β-sandwich proteins[5,7] but it decreases with an increase in the number of sheets N. On the other hand the radius of central hole increases with an increase of N. These two effects are counterbalanced at some optimal number of sheets. Two reasons that this number is *seven* are presented below.

Residue Packing in β-Sheet-to-β-Sheet Interfaces: Model

Three model β-propellers listed in Table I can have the residue packing and geometries compatible with those allowed by normal residue composition and standard β-sheet geometries. In all the models the inner strands of two adjacent β-sheets are very close to each other and must have the intercalating mode of residue packing. Thus, the shift S_1 between these strands is equal zero and, from Eq. (7), the shift S_n between other strands are to be equal to their calculated displacements h_n (see Table I). For the second strand the shift is about one residue. For the third strand it is about three residues. The shift values for these internal strands satisfy the condition for nonintercalated residue packing (strands are shifted by half-integer number of the periods of β-structure or odd number of residues). The shift of the outer strands is about five residues, but their closeness requires the intercalating mode of packing. However, this discrepancy is not significant if the outer strands have no more than five residues (typical length of β-strand). In this case they will pack end-to-end with a small interface.

Allowing small deviations of the inner strand shift S_1 we are able to optimize a residue packing in the middle of sheet-to-sheet interfaces. [For example, if in the model propeller with $N=7$ have the shift $S_1=-0.20$ (0.1 period), then, from Eq. (7), the shifts of the middle strands $S_2=1.06$ and $S_3=2.93$ are very close to ideal values 1 and 3.] However, optimizing sequentially (e.g., clockwise) sheet-to-sheet interfaces, we accumulate these small shifts in the last interface and distort the ideal N-fold symmetry.

An optimization of the sheet-to-sheet packing is effective, if propeller assembly satisfy the two conditions:

1. Providing the optimal packing in any particular interface requires the one-period difference between the displacements of the middle strands h_2 and h_3 ($h_3-h_2=2$).
2. Extending the optimal residue packing to all the interfaces requires the sheet displacement h_m to be equal to two ($h_m=2$).

The both values h_3-h_2 and h_m in the *seven-fold* β-propeller are closer to 2 than those in six- or eight-fold ones. Thus, this symmetry satisfies best to inherent periodicity of β-structure.

Composition of the Inner Strands of the β-Propellers

The 7.0 Å distance d_1 between the adjacent inner strands of seven-fold β-propellers corresponds to contact of a C_δ atom of a top-strand residue with main chain atoms of the bottom strand (Fig. 3). This distance will be provided by intercalated packing of average-size residues. However, to provide the shorter 6.5 Å distance in the six-fold β-propeller, the inner strands must be composed of smaller residues.

ANALYSIS OF KNOWN β-PROPELLER FOLDS

To test the accuracy of the model, its predictions for the geometrical features of β-propeller folds and for residue packing in sheet-to-sheet interfaces were

TABLE I. The Geometrical Parameters of Model and Observed β-Propeller Structures

Number of sheets N	Structure	Twist angle, θ (deg)	Radius of hole, r_1 (Å)	Separation of inner strands, d_1 (Å)	Separation of outer strands, d_4 (Å)	Constants of model*			Strand and sheet displacements†			
						a(Å)	b(Å)	d(Å)	h_2	h_3	h_4	h_m
6	Model	25.0	6.5	6.5	5.2	3.50	4.50	10.0	1.33	3.40	5.52	2.19
	NA	22.5 ± 7.4	6.6 ± 1.8	6.5 ± 0.3	6.8 ± 2.8	3.44 ± 0.35	4.46 ± 0.16	10.2 ± 0.9				
7	Model	23.8	8.1	7.0	6.0				1.26	3.13	5.09	2.05
	MADH	22.6 ± 4.2	8.2 ± 0.6	7.1 ± 0.6	7.4 ± 1.4	3.35 ± 0.15	4.43 ± 0.17	10.2 ± 0.7				
	GAO	23.6 ± 3.9	8.2 ± 0.2	7.1 ± 1.3	6.3 ± 1.8	3.46 ± 0.35	4.41 ± 0.10	10.1 ± 0.8				
8	Model	22.8	9.7	7.4	6.5				1.20	2.92	4.71	1.94

*a is the average length of residue in the strand direction, b is the average distance between two hydrogen-bonded strands, and d is the average distance between two close packed β-sheets in the middle of their interface; they are the constants of the model and are compared here with those in the observed structures.

†h_2, h_3, and h_4 are the displacements of the second, third, and fourth strands, respectively; h_m is the displacement of two β-sheets in the middle of their interface. The displacements are used to describe the residue packing in β-sheet-to-β-sheet interfaces of the model β-propellers, that is compared with residue packing in the observed structures (see text).

compared with geometries and packing observed in protein structures.

Observed Geometries: Sources and Methods

Data of protein structures were kindly made available by authors: coordinates of galactose oxidase[3] by N. Ito and S.E.V. Phillips, coordinates of neuraminidase[4] by P.M. Colman, and C_α-atom coordinates of methylamine dehydrogenase[2] by W.G.J. Hol. The geometrical parameters of these β-propeller structures were then calculated with a computer program[10] written by A.M. Lesk.

The long β-strands in all three structure have been truncated to five residues to place all the β-sheets in the same rectangular framework (bulged strands were truncated to six residues). In NA, some β-sheets do not cover the full five-residue rectangular frame, because several strands are shorter than five residues.

A line was fit through backbone atoms (or only the C_αs in MADH) of each strand, and all interstrand angles and distances have been calculated as angles and distances between these lines.

The results are summarized in Table I. As evident, the observed geometries are very close to those calculated from the model presented above.

The distance a is half the average distance between two adjacent C_β atoms (or C_α in MADH) situated on the same side of the β-strand. The distance b and the twist angle θ are average values from all pairs of hydrogen-bonded strands. The distances d_1 and d_4 are average values from all pairs of adjacent inner and outer strands.

The intersheet distance d is the average value for all β-sheet-to-β-sheet interfaces. To eliminate the effect of the radial displacement of β-sheets (see below), the individual d values for each interface are taken as the average of the two shortest distances in all four pairs of the middle strands (2–2′; 2–3′; 3–2′; 3–3′).

The internal radius r_1, has been calculated from the distance between C_α atoms in the inner strands. These C_α atoms, which are in the middle of each inner strand and lie approximately in the same plane passing perpendicular to the β-propeller axis, have been selected using FRODO[11] on Evans & Sutherland ESV Workstation. For the six-folded β-propeller of NA, the radius r_1 is half the average distance between C_α atoms of opposite sheets (1,4; 2,5; 3,6). For the seven-folded β-propellers of GAO and MADH the radius r_1 is the average radius of the circles that outline triangles made from three C_α atoms, two of them from adjacent sheets and third from opposite one (i.e., 1,2,5; 2,3,6; 3,4,7; 4,5,1; 5,6,2; 6,7,3; 7,1,4). (See the numbering of the propeller β-sheets in ref. 1 and 4 and in legends to Figures 1 and 5.)

Despite the agreement observed geometries with model ones, the overall architecture of real β-propeller folds is not quite perfect, especially in the six-fold β-propeller of NA. The essential deviations of real β-propeller structures from the ideal model are described below.

β-Sheet-to-β-Sheet Interfaces: Radial and Axial Displacements and Residue Packing Patterns

According to the proposed model, there is a relative displacement h_m between two β-sheets at their interface. Although it occurs along the β-strands, this displacement is roughly in the direction of the propeller axis and so will be denoted as an axial displacement. However, there may be another type of relative sheet displacement, that is not described in the model of the ideal β-propeller. This is a displacement of two β-sheets in their hydrogen-bonding direction. With respect to the circular symmetry of the sheet assembly, this is roughly in the radial direction and so will be denoted as a radial displacement.

The radial displacement in any sheet-to-sheet interface moves one β-sheet toward the N-fold axis and fills partially the central hole of the β-propeller. However, this displacement reduces this particular interface. In the observed β-propeller folds the displacements are small for all but one interface, where the β-sheets are shifted by one strand. The small displacements are not random, but systematic, so that axial projections of real β-propellers look like spirals rather than sets of concentric circles (Fig. 5).

The observed distortion of the ideal N-fold symmetry may result from the packing conditions in sheet-to-sheet interfaces. As mentioned above, the residue packing cannot be optimized with conserving a circular symmetry and is expected to have a different pattern at least at one interface of N.

The one large radial displacement occurs between the third and fourth β-sheets in MADH. The first strand of third sheet appears opposite the second strand of fourth sheet. This displacement produces an impression that the MADH β-propeller is half an ideal six-fold β-propeller and half an ideal eight-fold one (Fig. 5), that is in a good agreement with observed deviations[1] from ideal seven-fold symmetry in this structure.

The β-propeller domain of GAO has a unique interface with the C-terminal domain of this molecule. A β-hairpin of C-terminal domain fills the internal hole, and the top of this hairpin is essential in the active site of GAO, which is located near the opposite end of the β-propeller channel.[4] Because its internal hole is better filled, the radial displacements in the GAO β-propeller are less and its seven-fold symmetry is more perfect than in the MADH propeller. The most significant radial displacement occurs between the first and the second β-sheets of GAO, moving the second sheet toward the seven-fold axis (Fig. 1). The best superposition of the GAO

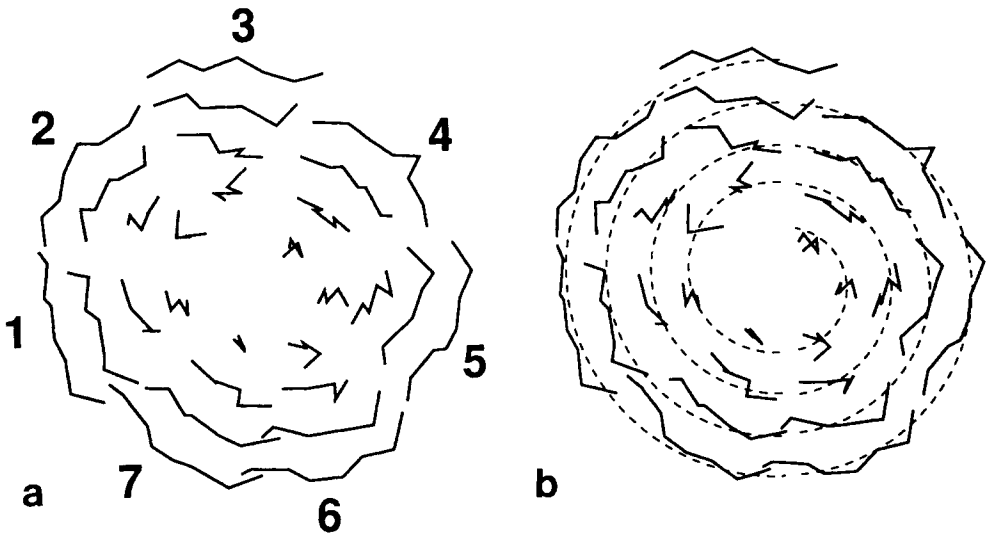

Fig. 5. The β-propeller of MADH viewed along seven-fold axis. **(a)** C_α trace for the β-strands only is presented with solid lines. β-Sheets are sequentially numbered[1] (in the same manner as β-sheets of GAO on Fig. 1). The large radial displacement between β-sheets 3 and 4 distorts ideal seven-fold symmetry. **(b)** This distortion can be approximated by combination of half a six-fold β-propeller (sheets 4, 5, and 6) and half an eight-fold β-propeller (sheets 7, 1, 2, and 3); dashed concentric half-circles outline each half. **a** and **b** together may be treated as two parts of stereopicture.

and MADH β-propellers has the first sheet of GAO correspond to the third sheet of MADH. That is, when their large radial displacements coincide.

In NA the six-residue bulge in the inner strand of the third sheet fills the central channel, however, the largest radial displacement occurs in the interface between the fourth and the fifth sheets, with the fourth sheet moving towards the six-fold axis. This large displacement is unlike those in MADH and GAO, where the β-sheets closer to the propeller axis have the greater sequential numbers.

It has been suggested above that the axial displacement of β-sheets in real β-propellers should correspond to the periodicity of β-structure. The calculated displacement h_m in the seven-fold β-propeller is very close to one β-structure period, and, as result, all seven β-sheet-to-β-sheet interfaces in GAO and MADH are very similar (with some differences in the interfaces with large radial displacements). A visual inspection of the residue packing in the well-determined structure of GAO shows good agreement with the model. The pattern of residue contacts repeats at each interface, and the strand shifts correspond to calculated ones (Fig. 6). Residues in the inner strands are intercalated, while those in the middle strands point directly at each other and so have nonintercalated packing.

On the other hand, the calculated displacement h_n in six-fold β-propeller is 10% greater than a β-structure period (more than one residue per assembly). In NA this excess is not distributed uniformly within its six β-sheet-to-β-sheet interfaces, but accumulates in the interface between the fourth and the fifth β-sheets. Due to the large axial and radial displacements these two sheets have a few contacting residues. The long folded loop connecting the fourth and fifth sheets assists in forming their odd interface. All other five sheet-to-sheet interfaces are similar to those in GAO and have the predicted pattern of residue contacts.

Packing and Composition of Inner Strands

As predicted, all the inner strands in the NA and GAO β-propeller are packed with intercalation of both large and small residues (Fig. 7). Although the full atomic structure of MADH is not available, relative orientations of its inner strand backbones can be seen to correspond to the intercalcalating mode of packing. Only 28% of the residues in the inner strands of NA are large (i.e., residues with C_δ atoms plus Val) in comparison with 52% in GAO. The amino acid sequence of the β-propeller subunit of MADH is not known yet, but a tentative X-ray sequence shows a significant amount (more than 50%) of large residues in its inner strands.

Packing of External Strands: β-Bulges

As is expected from the model, external strands have displacements of about five residues. If the external β-strands had only five residues each, they would form end-to-end contacts with their neighbors. Longer strands would produce a larger interface. In this case, the short distance between these strands requires an intercalating mode of packing, but, as predicted by the model, the shift of these strands corresponds to a nonintercalating mode. This conflict may be resolved by introducing β-bulges[12] into the long external strands. A β-bulge

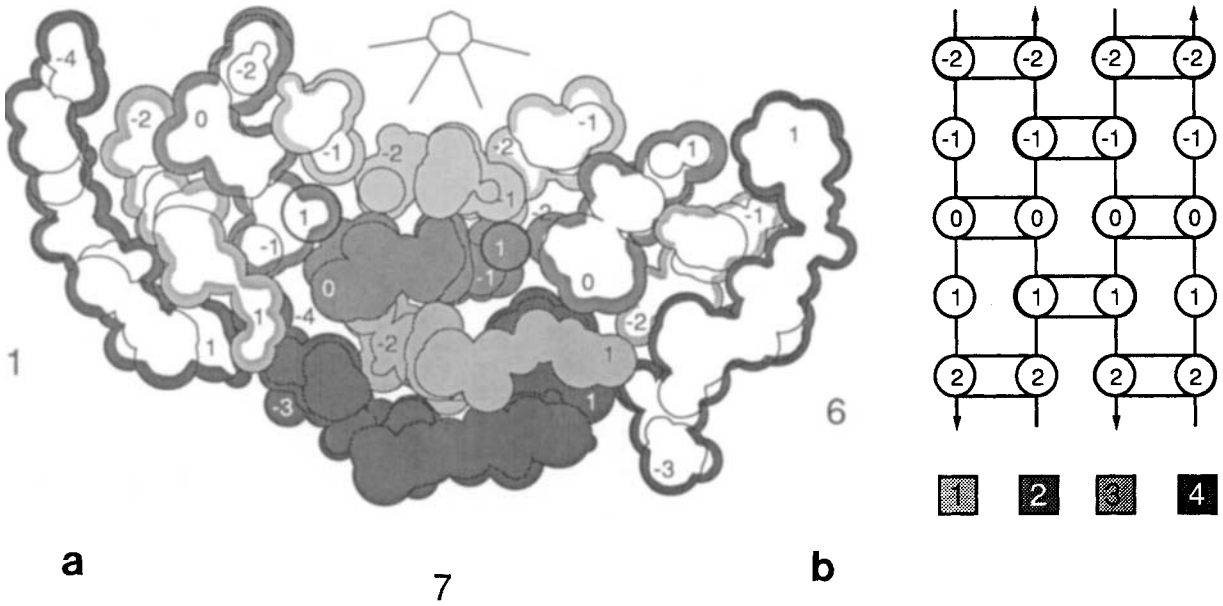

Fig. 6. **(a)** Serial section cut through a space filling model (van der Waals slices) of three consequent β-sheets of GAO (6, 7, and 1, see also Fig. 1). Sections were cut perpendicular to the seven-fold axis (presented by the regular heptagon). Five slices separated by 1 Å are shown. Atoms that are in each β-strand are correspondingly shaded in the middle β-sheet and outlined by wide shaded border in the adjacent β-sheets. **(b)** Scheme of rectangular β-sheet framework, explaining the shading of each strand and the numbering of residues on **(a)**. Only the residues forming sheet-to-sheet interfaces are marked on **(a)** to outline the strand shifts S_n, which are $S_1 \approx 0$ for first strand, $S_2 \approx 1$ for the second, $S_3 \approx 3$ for the third. Strand 4 in sheet 7 has a β-bulge (two residues) in position 0, so the observed shift S_4 of this strand is about 4 residues, but its displacement $h_4 \approx 5$.

(additional residue) inserted in the β-strand does not affect the displacement h_n of this strand relative to its aligned counterparts (see Fig. 6). However, it affects the shift S_n of these strands [one residue should be subtracted from the shift value predicted by Eq. (7)] and changes their packing mode.

In GAO there is a β-bulge in the middle of long outer strand in the seventh β-sheet, and this strand is in contact with the long outer strand in the first β-sheet. Two outer strands in GAO have only five residues and pack end-to-end. Three have similar C-terminal extension after the fifth residue: a β-bulge and one β-structural hydrogen bond; see Fig. 1). In MADH, where practically all external strands have more than five residues, β-bulges occur in six of the seven strands. A similar situation appears in NA, where five of the six external strands have β-bulges.

CONCLUSION

Structural principles for β-propeller assemblies, reported here, do not require special sequence constraints, except for small residues on inner strands in six-fold β-propellers. Thus, we should not expect sequence similarity between individual β-sheet motifs or between different propellers to be greater than that required by the nonpolar nature of sheet-to-sheet interfaces.

On the other hand, proteins with a β-propeller fold might be created by a gene duplication that had repeated the sequence coding an individual β-sheet motif several times and joined these repeats in a single chain.

This work does not reject such explanation but does provide a structural context for such an origin. For example, in the creation of a β-propeller protein, the β-sheet motif could be repeated more than seven times. However, the β-propeller fold accommodates seven motifs and any additional β-structure would not be included in the propeller domain but would be lost in evolution or evolve in other domains.

Without strong sequence constraints on the fold the β-sheet motifs in β-propeller proteins could diverge during evolution so as to lose obvious sequence homology.

Although β-propeller proteins may originate from identical β-sheet motifs and may have a common ancestor, their architecture is provided not by the specific sequence repeats, but by the general features of β-sheet structure such as periodicity, twist, and amino acid composition. This paper shows that there is a preference for a seven-fold β-propeller over six- or eight-fold ones, and it might be expected that this type of fold will be found again in other proteins, more than the two others.

ACKNOWLEDGMENTS

I am very grateful to C. Chothia for help and useful discussions, to A.M. Lesk for computer programs and Figures 1, 2, 5, and 7, and to P.M. Colman,

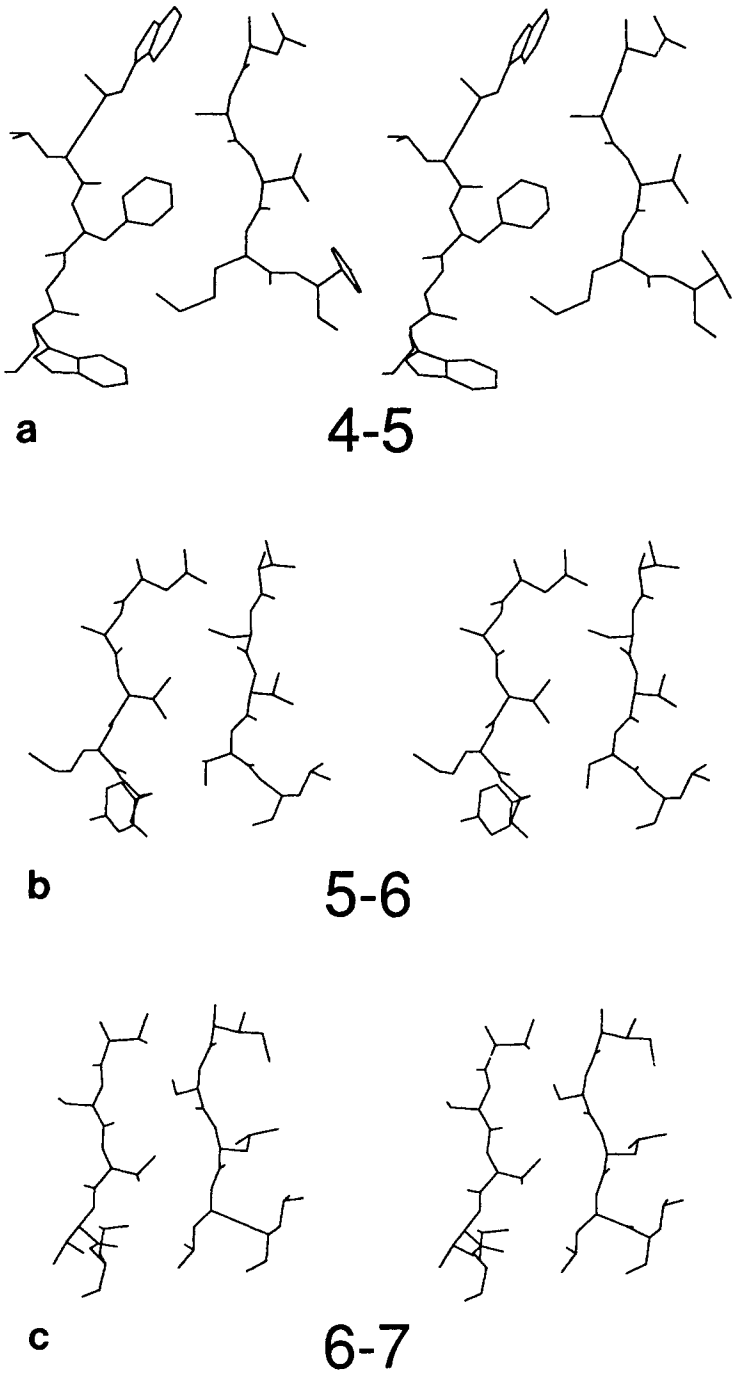

Fig. 7. Stereopictures of the intercalated inner strands in the β-propeller of GAO. Strands from the sheets 4 and 5 **(a)**, 5 and 6 **(b)**, and 6 and 7 **(c)** are presented with all nonhydrogen atoms. The interstrand distances are in pair **a**, 7.7 Å; in **b**, 7.5 Å; in **c**, 6.1 Å.

W.G.J. Hol, N. Ito, and S.E.V. Phillips for coordinates prior to their availability in the Protein Data Bank. A long-term fellowship from the European Molecular Biology Organization (ALTF 287-1990) is gratefully acknowledged. My thanks go to M.B. Gerstein for valuable comments on the manuscript.

REFERENCES

1. Vellieux, F.M.D., Huitema, F., Groendijk, H., Kalk, K.H., Frank, J., Jzn., Jongeigan, J.A., Duine, J.A., Petratos, K., Drenth, J., Hol, W.G.J. Structure of quinoprotein methylamine dehydrogenase at 2.25 Å resolution. EMBO J. 8: 2171–2178, 1989.
2. Vellieux, F.M.D., H., Kalk, K.H., K. Drenth, J., Hol,

W.G.J. Structure determination of quinoprotein methylamine dehydrogenase from *Thiobacillus versutus*. Acta Crystallogr. B46:806–823, 1990.
3. Ito, N., Phillips, S.E.V., Stevens, C., Ogel, Z.B., McPherson, M.J., Keen, J.N., Yadav, K.D.S., Knowels, P.F. Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. Nature (London) 350:87–90, 1991.
4. Varghese, J.N., Laver, W.G., Colman, P.M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature (London) 303:35–40, 1983.
5. Chothia, C., Janin, J. Relative orientation of close-packed β-pleated sheets in proteins. Proc, Natl. Acad. Sci. U.S.A. 78:4146–4150, 1981.
6. Chothia, C., Finkelstein, A.V. The classification and origins of protein folding patterns. Annu. Rev. Biochem. 59: 1007–1039, 1990.
7. Cohen, F.E., Sternberg, M.J.E., Taylor, W.R. Analysis of the tertiary structure of protein β-sheet sandwiches. J. Mol. Biol. 148:253–272, 1981.
8. Chothia, C., Levitt, M., Richardson, D. Structure of proteins: Packing of α-helices and pleated sheets. Proc. Natl. Acad. Sci. U.S.A. 74:4130–4134, 1977.
9. McLachlan, A.D. Gene duplications in the structural evolution of chymotrypsin. J. Mol. Biol. 128:49–79, 1979.
10. Lesk, A.M. Integrated access to sequence and structural data. In: "Biosequences: Perspectives and user services in Europe." Saccone, C., ed. Bruxelles: EEC, 1986:23–28.
11. Jones, T.A. Interactive computer graphics: FRODO. Methods Enzymol. 115:157–171, 1985.
12. Richardson, J.S., Getzoff, E.D., Richardson, D.C. The β-bulge: A common small unit of non-repetitive protein structure. Proc. Natl. Acad. Sci. U.S.A. 75:2574–2578, 1978.