Proc. Natl. Acad. Sci. USA
Vol. 74, No. 10, pp. 4130–4134, October 1977
Chemistry

Structure of proteins: Packing of α-helices and pleated sheets

(close packing/preferred conformations/protein secondary and tertiary structure)

CYRUS CHOTHIA*, MICHAEL LEVITT†, AND DOUGLAS RICHARDSON*

* William Ramsay, Ralph Forster, and Christopher Ingold Laboratories, University College London, Gower Street, London, W.C.1., England; and † Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, England

Communicated by Linus Pauling, July 22, 1977

ABSTRACT **Simple models are presented that describe the rules for almost all the packing that occurs between and among α-helices and pleated sheets. These packing rules, together with the primary and secondary structures, are the major determinants of the three-dimensional structure of proteins.**

Twenty-six years ago, Pauling and his colleagues (1, 2) presented the α-helix and the parallel and antiparallel pleated sheets as models for the local folding (secondary structure) of the polypeptide chain in proteins. Since then, various physical techniques (principally x-ray crystal structure analysis) have clearly shown that these secondary structures are almost universally present in protein molecules. We present here models that describe the rules that govern how α-helices and pleated sheets pack together to form the three-dimensional (tertiary) structure of proteins. The rules were developed empirically: *a priori* models were checked and refined by a detailed analysis of the residue-to-residue contacts that occur between and among the α-helices and pleated sheets in 17 proteins. To do this analysis we made extensive use of a computer graphics system for proteins developed by P. J. Pauling and his colleagues (to be published) and of numerical calculations. The models describe almost all of the secondary structure packings we have so far observed. In a later publication they will be used to present a detailed analysis of known structures.

The determining principles

Two principles have a dominating influence on the way in which secondary structures associate.

1. Residues that become buried in the *interior* of a protein close-pack: they occupy a volume similar to that which they occupy in crystals of their amino acid (3, 4).

2. Associated secondary structures retain a conformation close to the minimum free energy conformation of the isolated secondary structures.

This second principle is illustrated by the observation that almost all the protein main-chain torsion angles, φ and ψ, lie in regions of torsion angle space that are free from steric strain—that is, in the normally allowed regions of the Ramachandran map (5). Le Master and Richards, as reported by Richards (6), found that this is also true for the side-chain angle χ_1. In a detailed analysis of the conformational potentials of the residue side chains in bovine pancreatic trypsin inhibitor, Gelin and Karplus (7) showed that their conformations were close to those of the free amino acid.

The static close-packed image of proteins that is implicit in these two principles is, of course, only true of the time-averaged structure. Individual molecules are subject to large transient thermodynamic (and therefore conformational) fluctuations (8).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

The two principles imply that the secondary structures found in protein molecules interact in a manner that gives the maximum van der Waals energy and induces no appreciable steric strain. The rules described below for secondary structure associations arise from these two principles and from the intrinsic geometrical properties of α-helices and pleated sheets.

Helix–helix packing

To illustrate our model for helix–helix packing, we use a graphical construction invented by Crick (9). The residue side chains in a helix with 3.6 residues per turn and radius 5 Å are assumed to have the same size and shape. Two helices are slit down one side in a direction parallel to their axis. Each is opened up and laid flat, one helix face down and one face up. Placing the first lattice over the second is equivalent to bringing the outside of the two helices into contact (Fig. 1).

The model for helix–helix interactions is shown in Fig. 1. The surface of a helix can be described in terms of rows of adjacent side chains. A residue in a helix, i, has two neighbors above it, $i + 3$ and $i + 4$, and two neighbors below it, $i - 3$ and $i - 4$. One row is formed by the residues i, $i \pm 3$, $i \pm 6$, . . . $i \pm 3n$ and another by i, $i \pm 4$, $i \pm 8$, . . . $i \pm 4n$. The alignment of the side chains so they point to their ± 3 *or* ± 4 neighbors means that the $i \pm 3n$ *or* $i \pm 4n$ residue rows form a ridge. The similar alignment of adjacent residues means in turn that the helix surface consists of one of two series of parallel ridges separated by shallow grooves: one series being formed by the rows $i \pm 3n$, $(i \pm 1) \pm 3n$, . . . and the other by the rows $i \pm 4n$, $(i \pm 1) \pm 4n$, ... Our model for helix–helix packing requires that the surface ridges in the first helix pack into the grooves between the ridges in the second helix and vice versa (Fig. 1).

As is apparent from Fig. 1, this model gives three classes of interaction that differ in the residues they bring into contact and in the angle between the helix axes (Ω)‡ (Table 1). Real helix–helix packings will deviate from this simple model for two reasons. First, the angle between the helix axis and the rows of residues will depend upon the exact twist and radius of the helix. Thus, if the two ideal helices of radius 5 Å each have 3.4 residues per turn, the Ω values will be $-105°$, $-81°$, and $-3°$, and if they each have 3.8 residues per turn the Ω values will be $-66°$, $-32°$, and $+40°$. Second, the side chains do not all have the same size and shape but vary between 0 and 10 atoms.

From the atomic coordinates we have calculated the contacts that occur between the helices and in the α subunit of hemo-

‡ We use Ω to describe the relative orientation of two pieces of secondary structure in contact. Ω is defined as the angle between the strands of the pleated sheet and/or helix axes when projected onto their plane of contact. We ignore the direction of individual α-helices and strands so Ω is defined between $-90°$ and $+90°$ rather than $-180°$ and $+180°$. The angle is negative ($0 \rightarrow -90°$) if the near helix or strand is rotated in a clockwise direction relative to the far helix or strand. If this rotation is anticlockwise, the angle is positive ($0° \rightarrow +90°$).

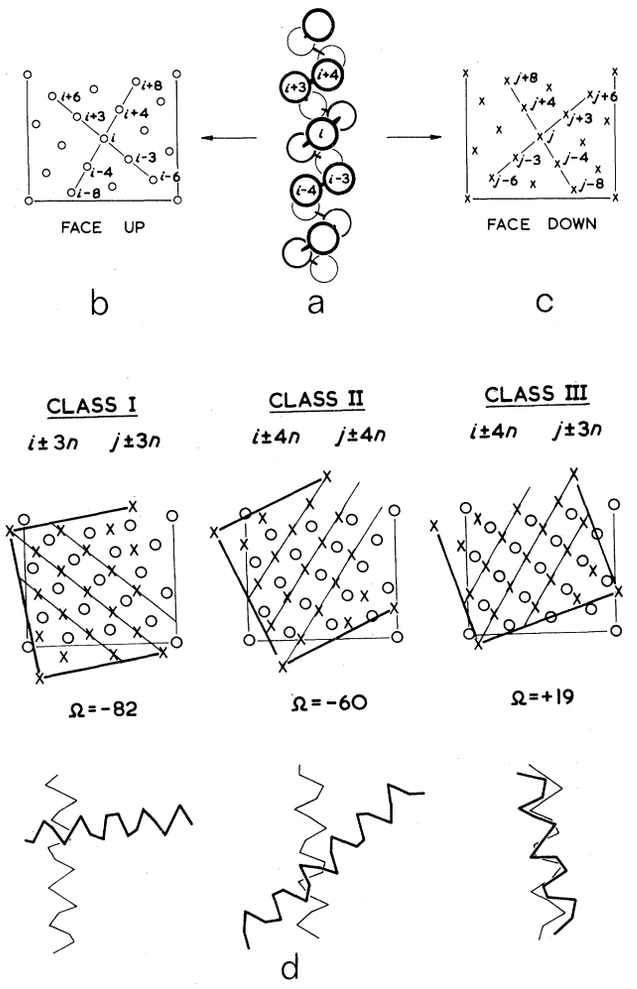


FIG. 1. The model for helix–helix packing: The residues in a helix (*a*) are represented by large circles. This helix is opened up to give two flattened projections (see *text*) shown face-up in *b* and face-down in *c*. The residue rows $i \pm 3n$, $i \pm 4n$, $j \pm 3n$ and $j \pm 4n$ are labeled. (*d*) The three classes of "ridges into grooves" packing. The Ω values are for two ideal helices with 3.6 residues per turn and radius 5 Å. Lighter lines indicate the $j \pm 3n$ and $j \pm 4n$ rows of residues.

globin, thermolysin, lysozyme, calcium-binding protein, subtilisin, and staphylococcal nuclease. We found 26 cases of helix–helix packing in which three or more residues from one helix are in contact with three or more residues from the other. We have also calculated, for each of the 26 pairs, the angle between the helix axes. Twenty-five of these can be related to our model for helix–helix packing (in the 26th case, the helices are joined covalently by a disulfide bridge). Inspection of the pattern of residue contacts shows that 3 of these interactions are of a type we have defined as class I, 16 are class II, and 6 are class III. For class I, the observed values of Ω are −80°, −85°, and −95°. In class II the angles vary between −20° and −70°, although 11 are in the range −50° ± 10° (Fig. 2).

Table 1. Classes of interaction

Class	Residues forming rows: First helix	Residues forming rows: Second helix	Ω*
I	$i \pm 3n$	$j \pm 3n$	−82°
II	$i \pm 4n$	$j \pm 4n$	−60°
III	$i \pm 4n$	$j \pm 3n$	+19°

* Angle between helix axes, for two helices of 3.6 residues per turn and radius 5 Å.

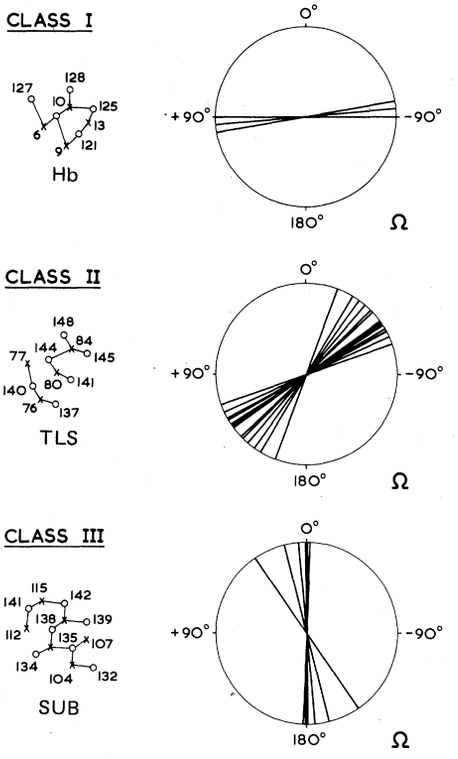


FIG. 2. The observed helix–helix packing in the α-subunit of hemoglobin (Hb), thermolysin (TLS), lysozyme, staphylococcal nuclease, calcium-binding protein, and subtilisin (SUB) (see *text*). We show Ω for the 25 observed helix–helix packings and one example, for each class, of the residue–residue contacts: O, side chains on the far helix; ×, side chains of the near helix. Lines joining the symbols indicate the residue–residue contacts (see also Fig. 1*d*).

Inspection by computer graphics of the six class III interactions shows that three of them have a systematic deviation from the simple model with Ω close to 0° and each helix tilted about 10° away from the plane of contact. Details of this will be published later. The other three members of this class have Ω of +5°, +15°, and +35° (Fig. 2).

The 25 helix–helix packings involve 50 different helix surfaces. Thirty-eight of these have the ridges formed by the $i \pm 4n$ residue rows. This is due to the side chain, i, preferring to point toward the $i - 4$ neighbor.

The model for helix–helix packing presented here differs from that previously described by Crick (9). In his model, residue i from one helix fitted in between residues j, $j + 3$, $j + 4$, and $j + 7$ in the second helix and vice versa. This gives two packing patterns in which the helix axes are inclined at +20° or −70°. Although a few of the helix interactions observed here do occur at angles close to these values, the pattern of residue–residue contacts is not that predicted by Crick's model. Colloquially, his model was described as "knobs into holes"; ours can be described as "ridges into grooves."

Helix–sheet packing

The model for packing an α-helix onto a parallel or antiparallel pleated sheet is illustrated in Fig. 3.

There are two general features of pleated sheets that are important for this model: the first is the packing between neighboring residues within a sheet, and the second is its tendency to have a right-hand twist. If we consider the Cα atoms on the same side of a sheet, the distance between neighbors along the strands is 7 Å and that between those in adjacent strands is 5 Å. Side chain volumes vary between 25 Å^3 (alanine)

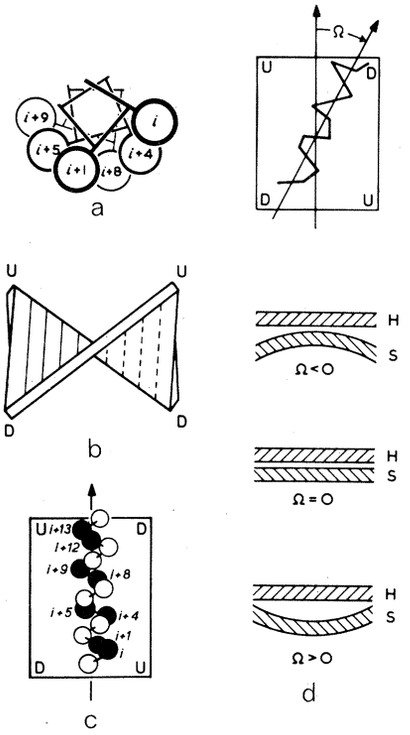


FIG. 3. The model for helix–sheet packing. (*a*) The helical residues $i, i+1, i+4, i+5, i+8, \ldots$ form a surface with a right-hand twist which is complementary to the right-hand twist of a pleated sheet shown in (*b*). (*c*) The helix is shown on top of the twisted sheet. The corners marked U are above the plane of the page and those marked D are below it. (*d*) Sections showing idealized helix–sheet interfaces for different values of Ω (see *text*).

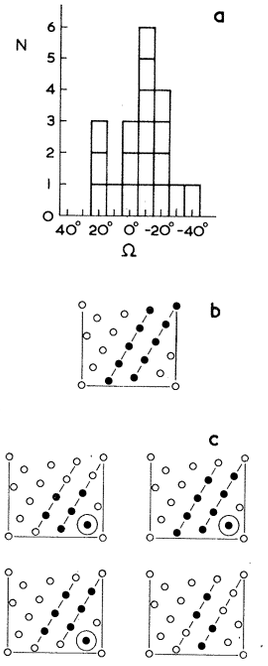


FIG. 4. Observed α-helix–pleated sheet packing. (*a*) Histogram of the observed Ω. The values are for the 19 helices in carboxypeptidase, flavodoxin, triose phosphate isomerase, and subtilisin that have four or more residues in contact with a pleated sheet. The helix is above the sheet. (*b*) The normal contact residues (see *text*) are shown as filled circles in a flattened projection of a helix (see Fig. 1 *a* and *b*). (*c*) The helical residues in contact with the pleated sheet (filled circles) in flavodoxin. Twenty-three of 26 are normal contact residues. The abnormal contacts are ringed for identification.

and 170 Å^3 (tryptophan) (3). This means that most side chains in a pleated sheet will be in contact with their neighbors so the surface will not have ridges and grooves but can be considered as flat with, usually, only small irregular holes and protuberances.

The parallel and antiparallel pleated sheets found in globular proteins have a twist that is right-handed if the sheet is viewed in a direction parallel to the polypeptide chain (10). The effect of this twist is that neighboring chains wind around each other while remaining a constant distance apart. The same thing happens to the ropes in a rope ladder if a far rung is given a right-hand twist relative to a near rung. For a typical pleated sheet, the right-handed twist about an axis parallel to the chain direction is observed to be about 5° per Å.

Now let us consider an α-helix. The residue pairs $(i, i+1)$, $(i+4, i+5)$, $(i+8, i+9)\ldots$ wind around the helix with a right-hand twist (Fig. 3). For a regular α-helix this twist is about 6° per Å. Also, given suitable side chain size and conformation, these residues can form a flattened, although irregular, surface. For reasons that will become apparent below, we shall call these residues the "normal contact residues."

In its simplest form, our model for α-helix–pleated sheet packing can be stated as follows: an α-helix will pack onto a pleated sheet with its axis parallel to the strands of the sheet because, in this orientation, the normal contact residues form a surface complementary to that of the sheet. Such a model would predict that the helix residues in contact with the sheet will be $i, i+1, i+4, i+5, i+8, i+9, \ldots$, and that the angle between the helix axis and the strands in the sheet (Ω) will be 0°.

On a more detailed level we can see that the twist of the pleated sheet means that orientations of the helix away from the parallel position (Ω = 0°) are more likely to occur in the negative (clockwise) direction (the helix is assumed to be above the sheet) (Fig. 3*d*). In this orientation (Ω < 0°), although the exposed ends of the helix move away from the sheet, its center is still able to close-pack. In the opposite orientation (Ω > 0°), the two ends of the helix will pack onto the sheet but they lift its center off the sheet and so create an internal cavity in the protein. Using the same argument, we can show that helices packed on sheets with a large twist will have negative Ω.

A dominant structural feature of the proteins flavodoxin, carboxypeptidase, subtilisin, and triose phosphate isomerase is a large central pleated sheet flanked by α-helices. We have examined the contacts that occur between the residues in these proteins and have found that they contain 19 helices that have four or more residues in contact with the face of a pleated sheet. In total, the 19 helices have 129 residues in contact with the pleated sheets, of which 112 (87%) are what we defined above as "normal contact residues" for sheet–helix packing. In Fig. 4 we show, for flavodoxin, the helical residues that are in contact with the central pleated sheet and the distribution of the values of Ω. All these angles are in reasonable agreement with our model. Thirteen have values in the range $-10° \pm 10°$ and the distribution is skewed toward negative values.

Sheet–sheet packing

We consider in this section the face-to-face packing of two pleated sheets. Inspection of protein structures shows that we can distinguish two classes of double-sheet structures: in the first class they are formed by the packing of two large and essentially independent pleated sheets; in the second class they are formed by the folding over of single sheets.

In all these proteins the sheet has a right-hand twist when

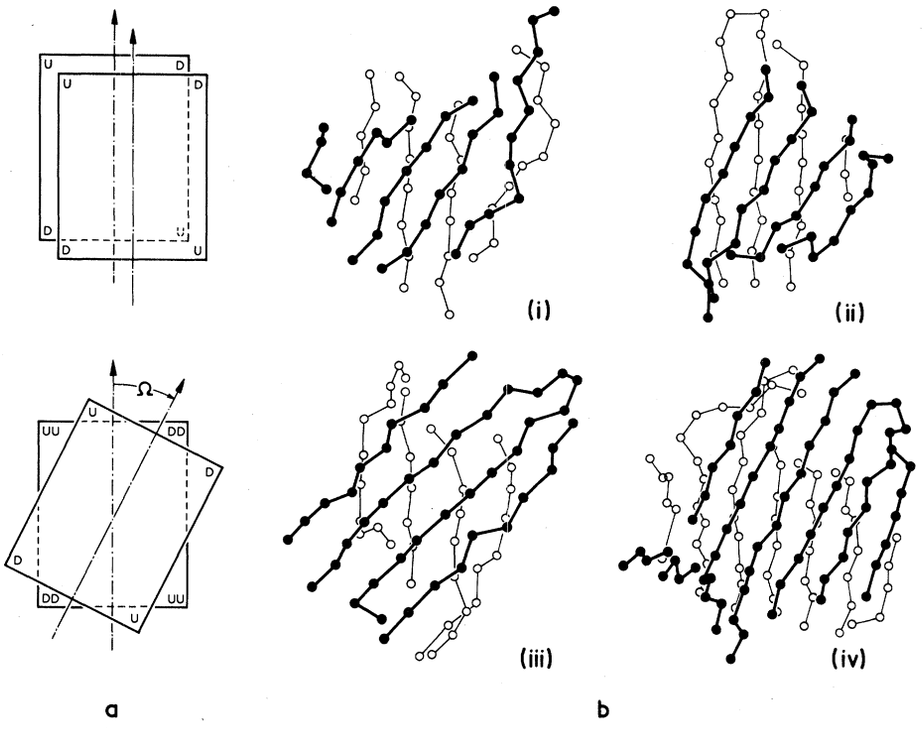


FIG. 5. The packing of two independent pleated sheets. (*a*) Two pleated sheets are represented by two thin smooth sheets. Arrows indicate the chain direction. Each sheet has a right-handed twist, so the corners are either above the plane of the paper (U or UU), or below the plane (D or DD). The sheets are placed face-to-face. Two sheets with the same twist close-pack when their strands are parallel. To close-pack two sheets with different twists, one sheet is rotated in a negative direction relative to the other. (*b*) The Cα atoms of the double-pleated sheet structures in the immunoglobulin fragment V_{REI} (i), prealbumin (ii), superoxide dismutase (iii), and concanavalin A (iv). Filled circles denote atoms in the upper sheet and open circles those in the lower sheet. Note that the top sheet of concanavalin A is bent so that the bottom left-hand corner points down perpendicular to the plane of the page.

viewed along the strands. The extent of this twist varies from sheet to sheet. Thus, the general problem here is how sheets of different twist and residue composition associate so that the central part of the contact surface is close-packed. We shall first discuss the packing of two independent pleated sheets (Fig. 5).

Let two sheets have the same degree of twist and the same residue composition. It is obvious that these close-pack when face to face with the strands in one sheet parallel to those in the other sheet (Fig. 5). Now, let one sheet have a greater twist than the other. This difference in twist can be due to a real difference in the main chain torsion angles of the two sheets or just be the surface effect of a difference in their residue compositions, or both. If these two sheets are placed face to face with their strands parallel, only two opposite corners will form contacts. Rotation of the top sheet in a clockwise (negative) direction will allow the centers of the sheets to pack more closely (Fig. 5). A rotation in the opposite (anticlockwise) direction will force the centers of the sheets to be even further apart.

Computer model building shows that if T_1 and T_2 are the twists of the two smooth sheets, the rotation angle that gives the best packing, Ω_c, is given by the expression:

$$\Omega_c = -2(|T_1 - T_2|) \qquad [1]$$

The immunoglobulin domains, concanavalin A, prealbumin, and superoxide dismutase contain two large antiparallel pleated sheets of different twist packed face to face. A projection of these double-sheet structures is given in Fig. 5. We have measured the overall twist of each of the sheets and, using Eq. **1**, have calculated the expected relative rotation, Ω_c. Using the atomic coordinates, we have also measured the actual rotation, Ω. For the immunoglobulin fragment V_{REI}, $\Omega_c = -32°$ and $\Omega = -35°$; for prealbumin, $\Omega_c = -36°$ and $\Omega = -35°$; for concanavalin A, $\Omega_c = -82°$ and $\Omega = -30°$; and for superoxide dismutase, $\Omega_c = -52°$ and $\Omega = -50°$ (Table 2). The failure of the relationship for concanavalin A is due to the pleated sheets of this protein being bent as well as twisted (Fig. 5).

This relative rotation of two rectangular sheets of different twist means that their corners either lift off the packed interface or stick out at the side (Fig. 5). In real structures the corners that lift off the interface are sometimes removed by shortening the sheet at these corners. This gives the sheets a rhomboidal rather than rectangular shape and makes the double-sheet structure appear elliptical when viewed from the side in certain directions (for example, see figure 3 in ref. 11). The strands that stick out at the side sometimes fold under their own sheet and hydrogen bond to become part of the second sheet. There is one example

Table 2. Twist and relative orientation of packed pleated sheets

Protein	Twist of pleated sheets*		Relative orientation†	
	Upper (T_1)	Lower (T_2)	Ω_c‡	Ω§
Bence-Jones protein V_{REI}	58°	42°	−32°	−35°
Prealbumin	64°	46°	−36°	−35°
Superoxide dismutase	37°	63°	−52°	−50°
Concanavalin A	29°	70°	−82°	−30°

* The upper and lower sheets are defined by Fig. 5.
† Defined in the text and in Fig. 5.
‡ $\Omega_c = -2(|T_1 - T_2|)$.
§ The observed value of Ω.

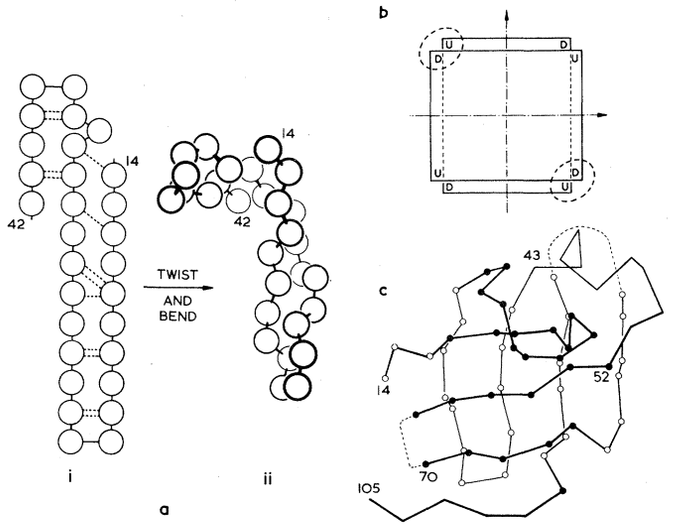


FIG. 6. The packing of pleated sheets with a right-handed supertwist. (*a*) Residues 14–42 of elastase are represented by open circles. These residues form a triple-stranded pleated sheet whose hydrogen bonds are shown as dashed lines (- - - - -) in i. The main chain conformation of this sheet is shown in ii; note its right-handed twist and supertwist. (*b*) Two pleated sheets are schematically treated as two thin smooth sheets. The strand direction in each is indicated by an arrow. They are placed face to face with their strand directions at right angle. Each sheet has a right-handed twist so that corners marked U are above the plane and those marked D are below. Note that the upper left and lower right corners are in contact while the other two splay apart. (*c*) The double-pleated sheet structure of first domain in elastase. The Cα atoms of each sheet residue are indicated by a circle. Filled circles show those in the upper sheet and open circles, those in the lower. Note that the two sheets are covalently joined at the upper left and lower right corners. Where the sheets splay apart, residues 43–51 and 103 are inserted to maintain the packing. Compare with *b*.

of this in the immunoglobulin fragment V_{REI} and two examples in superoxide dismutase.

Double-sheet structures that are formed by folding over a single sheet involve two related problems: how does a pleated sheet fold over, and how do the two halves pack together? Due to the combined effect of the steric limitation on possible main chain conformations (5), the right-handed twist of pleated sheets (9) and the necessity to retain interchain hydrogen bonds, pleated sheets must fold over by means of a local right-handed supertwist as illustrated in Fig. 6*a*. This local supertwist puts the strand directions in the two halves at approximately right angle to each other.

Pleated sheets folded over by a right-handed supertwist occur in the trypsin family (Fig. 6*c*), staphylococcal nuclease, triose phosphate isomerase, thermolysin, alcohol dehydrogenase, ribonuclease, and papain. The way in which the two halves of the sheet pack together varies, but it is a variation on one theme. To illustrate this we use the schematic double sheet in Fig. 6*b* and the particular example of the first elastase domain (Fig. 6*c*).

If the two sheets in Fig. 6*b* were independent of each other, they would only have two corners in van der Waals contact, the upper left and lower right in our diagram. If, however, the double sheet is formed by folding over a single sheet, a local right-handed supertwist occurs at one or both of these two corners (compare Fig. 6 *b* and *c*). This covalent link is shorter than a van der Waals contact and it pulls the two halves of the sheet together so that they are actually in contact along the diagonal joining the two close corners. As we move away from this diagonal, the twist of the sheets makes them splay apart and we often find that a short piece of polypeptide is inserted to maintain the packing (Fig. 6*c*).

Double-sheet structures joined at two corners, as in elastase (Fig. 6*c*), are found in the trypsin family, staphylococcal nuclease, triose phosphate isomerase, and thermolysin. Double-sheet structures joined at one corner only occur in alcohol dehydrogenase, ribonuclease, and papain.

Conclusion

The packing models we have presented derive from the general geometrical properties of α-helices and pleated sheets and from the necessity to close-pack without inducing appreciable steric strain. The models implicitly ignore the exact size and shape of individual side chains. Their success shows that this is reasonable. However, our approach can be extended to include the effects of individual side chains: this might allow an exact understanding of particular packings and of those few cases that are not described by our models.

In protein molecules the polypeptide chain tends to run back and forth across the molecule. Each run usually contains an α-helix or is part of a pleated sheet (12). Thus, the important intramolecular contacts that occur in a protein are between its constituent α-helices and/or its pleated sheets. Because the models that we have presented describe almost all the secondary structure packings that we have so far analyzed, the packing rules given by these models are the important determinant of the three-dimensional structure of proteins.

This paper owes much to the critical encouragement of Drs. Peter Pauling and Joël Janin. The work was initiated at the Service de Biochimie Cellulaire of the Institut Pasteur in Paris and part of it was done at the 1976 CECAM workshop on Protein Dynamics. We thank Dr. Richard Feldmann, Prof. D. C. Phillips, and Dr. C. Blake for atomic coordinates. The work was supported by grants to Dr. Peter Pauling from the Institute of Neurological and Communicative Diseases and Stroke (1 R01 NS12021-02 BBCA) and the Medical Research Council.

1. Pauling, L., Corey, R. B. & Branson, H. R. (1951) *Proc. Natl. Acad. Sci. USA* **37**, 205–211.
2. Pauling, L. & Corey, R. B. (1951) *Proc. Natl. Acad. Sci. USA* **37**, 235–285.
3. Richards, F. M. (1974) *J. Mol. Biol.* **82**, 1–14.
4. Chothia, C. H. (1975) *Nature* **254**, 304–308.
5. Ramachandran, G. N. & Sasisekharan, V. (1968) *Adv. Protein Chem.* **23**, 284–437.
6. Richards, F. M. (1977) *Adv. Biophys. Bioeng.* **6**, in press.
7. Gelin, B. & Karplus, M. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2002–2006.
8. Cooper, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2740–2741.
9. Crick, F. H. C. (1953) *Acta Crystallogr.* **6**, 689–697.
10. Chothia, C. H. (1973) *J. Mol. Biol.* **75**, 295–302.
11. Richardson, J. S., Thomas, K. A. & Richardson, D. C. (1975) *Biochem. Biophys. Res. Commun.* **63**, 986–992.
12. Levitt, M. & Chothia, C. H. (1976) *Nature* **261**, 552–558.