Structure of proteins: Packing of \( \alpha \)-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 \( \alpha \)-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 \( \alpha \)-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 \( \alpha \)-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 \( \alpha \)-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, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 the second principle is illustrated by the observation that the protein main-chain torsion angles, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 the second principle is illustrated by the observation that the protein main-chain torsion angles, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 the second principle is illustrated by the observation that the protein main-chain torsion angles, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 the second principle is illustrated by the observation that the protein main-chain torsion angles, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 the second principle is illustrated by the observation that the protein main-chain torsion angles, \( \phi \) and \( \psi \), 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). LeMaster and Richards, as reported by Richards (6), found that this is also true for the side-chain angle \( \chi_1 \). In a detailed analysis of the conformational potentials of the residual 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 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; X, side chains of the near helix. Lines joining the symbols indicate the residue-residue contacts (see also Fig. 1d).

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 ±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.”

**Table 1. Classes of interaction**

<table>
<thead>
<tr>
<th>Class</th>
<th>Residues forming rows</th>
<th>First helix</th>
<th>Second helix</th>
<th>Ω*</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>i ± 3n, j ± 3n</td>
<td>i ± 3n</td>
<td>j ± 3n</td>
<td>−82°</td>
</tr>
<tr>
<td>II</td>
<td>i ± 4n, j ± 4n</td>
<td>i ± 4n</td>
<td>j ± 4n</td>
<td>−60°</td>
</tr>
<tr>
<td>III</td>
<td>i ± 4n, j ± 4n</td>
<td>i ± 3n</td>
<td>j ± 3n</td>
<td>+19°</td>
</tr>
</tbody>
</table>

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

**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 Ca 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 Å³ (alanine)
and 170 Å² (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 (Fig. 3). For a regular a-helix this twist is about 5° 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.”

Now let us consider an a-helix. The residue pairs (i, i+1), (i+4, i+5), (i+8, i+9) ... wind around the helix with a right-hand twist (Fig. 3). For a regular a-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 a-helix-pleated sheet packing can be stated as follows: an a-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, ... 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. 3d). 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 a-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° ± 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
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, \( \Omega_c \), is given by the expression:

\[
\Omega_c = -2(|T_1 - T_2|) \tag{1}
\]

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, \( \Omega_c \). Using the atomic coordinates, we have also measured the actual rotation, \( \Omega \) for the immunoglobulin fragment \( V_{REI} \), \( \Omega_c = -32^\circ \) and \( \Omega = -35^\circ \); for prealbumin, \( \Omega_c = -36^\circ \) and \( \Omega = -35^\circ \); for concanavalin A, \( \Omega_c = -82^\circ \) and \( \Omega = -30^\circ \); and for superoxide dismutase, \( \Omega_c = -52^\circ \) and \( \Omega = -50^\circ \) (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

\begin{table}
<table>
<thead>
<tr>
<th>Protein</th>
<th>Twist of pleated sheets</th>
<th>Relative orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bence-Jones protein ( V_{REI} )</td>
<td>58° 42° -32° -35°</td>
<td></td>
</tr>
<tr>
<td>F'realbumin</td>
<td>64° 46° -36° -35°</td>
<td></td>
</tr>
<tr>
<td>Superoxide dismutase</td>
<td>37° 63° -52° -50°</td>
<td></td>
</tr>
<tr>
<td>Cloncanavalin A</td>
<td>29° 70° -82° -30°</td>
<td></td>
</tr>
</tbody>
</table>

* 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 \( \Omega \).
of this in the immunoglobulin fragment V_{HRI} 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. 6a. This local super-twist 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. 6c), 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. 6b and the particular example of the first elastase domain (Fig. 6c).

If the two sheets in Fig. 6b were independent of each other, they would only have two corners in van der Waals contact, the upper left and lower right corners 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. 6b 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. 6c).

Double-sheet structures joined at two corners, as in elastase (Fig. 6c), 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 a-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 a-helix or is part of a pleated sheet (12). Thus, the important intramolecular contacts that occur in a protein are between its constituent a-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 Dr. Peter Pauling and Joel 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 CECAI workshop on Protein Dynamics. We thank 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 RO1 NS12021-02 BBCA) and the Medical Research Council.