MACROMOLECULAR STRUCTURE
AND SPECIFICITY:
COMPUTER-ASSISTED
MODELING AND APPLICATIONS Vol. 439
Reprinted from
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Engineering Aspects of Protein Structure[a]

F. R. SALEMME

Genex Corporation
Gaithersburg, Maryland 20877

X-ray crystallographic studies have revealed proteins to be highly organized structures of marvelous complexity. A protein's three-dimensional organization is required for biological activity because it provides the spatial framework necessary for accurate positioning of amino acid residues involved in ligand binding or substrate catalysis. Nevertheless, many studies indicate that the functional arrangements manifest in crystal structures have only marginal stability in solution.[1] As a result, proteins exhibit a wide range of dynamical behavior at ambient temperatures.[2-5] Although dynamical excitation of a protein may play a role in catalysis, the continual statistical interchange of kinetic energy between a protein and its solvent environment can lead to transient or irreversible disruptions of structure and function. The basic question addressed here concerns how the long-range structural integrity of a protein is preserved in a dynamically active solution environment.

Information relevant to this question comes from comparative studies of the large number of protein structures presently known from X-ray crystallography.[6] These studies have shown that proteins can be structurally classified and that features of extended structural organization recur among proteins having no apparent sequence or evolutionary relationship. The recurrence of these structural patterns must then reflect an accomodation to some basic physical requirements of facile folding or stability, for which there apparently exist a finite number of structural solutions.[7-8] What we wish to illustrate here is how some recurrent long-range architectural features of proteins achieve and maintain their structural integrity in a kinetically active environment.

It has long been recognized that extended features of protein tertiary organization are generally associated with the formation of hydrogen-bonded secondary structures. Owing to the periodic and structurally invariant nature of the polypeptide backbone, long regions of chain can form regular lattices stabilized by hydrogen-bonded interactions. These extended lattice

[a] This work was supported by research grant nos. GM 30393 and GM 33325 from the National Institutes of Health.

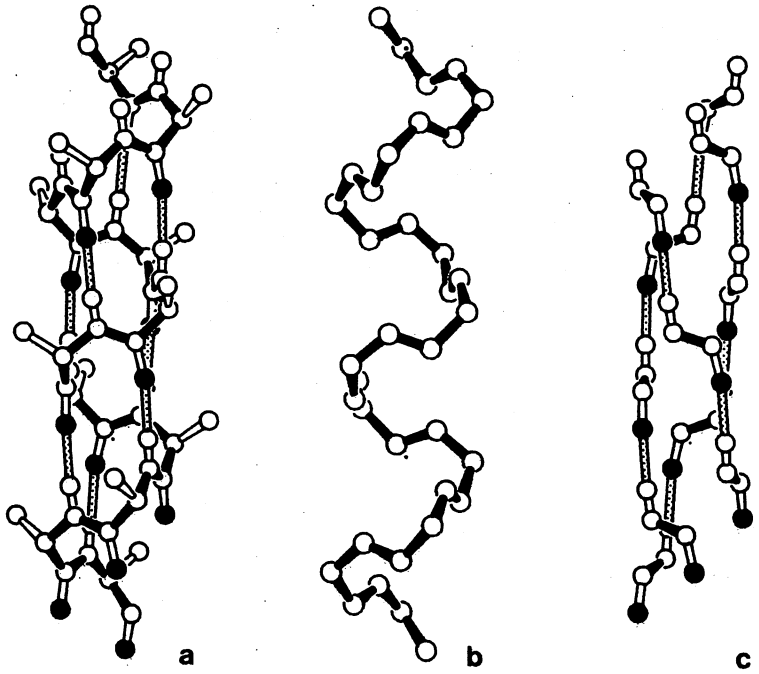

FIGURE 1. A molecular model of an α-helix (**a**). As illustrated in (**b**) and (**c**), the structure is organized as a cylindrical lattice formed of a right-handed backbone helix integrally braced by three left-handed hydrogen-bonded helices.

structures provide the basis for long-range structural organization in the protein.

FIGURE 1a illustrates an α-helix, one of the most common extended secondary structural arrangements seen in proteins. Theoretical studies of α-helix dynamics are consistent with structural observations indicating that this is a particularly rigid arrangement that tends to resist deformation.[9] The mechanical origins of this ridgidity are readily understood if the backbone covalent and hydrogen-bonded interactions are separated as illustrated in FIGURES 1b and 1c. These make evident that atoms of the sequentially connected backbone form a right-handed helix with 3.6 residues per turn, whereas the contiguous hydrogen-bonded interactions form a triple-strand left-handed helical arrangement. Deformation of any helix by stretching or bending will result in the helix unwinding in a direction opposite to the original helix twist. In the case of the α-helix, the minimum energy structure is one that simultaneously optimizes nonbonded interactions in the right-hand helix, and hydrogen-bonded interactions in the left-hand helices. Deformations of the extended α-helix are strongly resisted because these must involve the simultaneous unwinding of helices of opposite hand. This is physically impossible since the helical interactions in the α-helix are intergrally connected to form a continuous cylindrical lattice. The objective of treating the α-helix from this per-

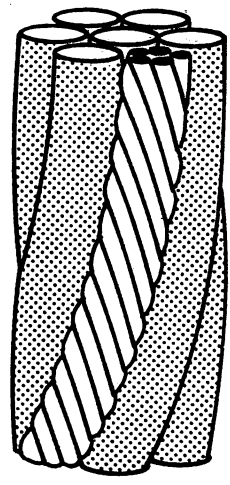

FIGURE 2. A schematic view of a wire cable. Resistance to stretching and twisting is achieved by laying the strands in each wire with opposite twist sense to the lay of the wires in the cable.

spective is to emphasize the distributed nature of the interaction forces that stabilize the extended structure. The basic principle of utilizing the opposition in twist sense to resist extension or disassembly in helical structures has, in fact, been exploited for hundreds of years in the manufacture of rope and cables (FIG. 2).

A second common element of protein secondary structure is the parallel β-sheet. As first described by Pauling and Corey,[10] the classical structure is composed of extended two-fold helical chains arranged with the same N to C polypeptide chain sense. This allows the formation of a regular hydrogen-bonding pattern between chains so that the structure as a whole forms a planar lattice connected by covalent bonds along the chains' direction and hydrogen bonds across the strands of the sheet (FIG. 3). The structure determination of numerous proteins showed that in contrast to the ideal flat model, observed sheets in proteins virtually always formed complex twisted surfaces (FIGS. 4–6), and that the sense of the twist in the sheet was always the same.[11] Subsequent analysis[8] of these structures showed that the overall sheet twist resulted from twisting of its constituent chains from the two-fold helical conformation of the flat sheet, into left-hand helices with 2.1 to 2.5 residues per turn. The polypeptide chain twisting is a consequence of the structures' incorporation of L-amino acids, which energetically prefer left-twisted conformations in extended chains. As the flat structure is one that optimizes interchain hydrogen bonds, it is evident that twisting the sheet to optimize the local chain conformational energy can only occur at the expense of the interchain hydrogen-bond energies. The lattice formed by a twisted sheet is consequently stressed, with the components of conformational energy that drive chain twist acting in opposition to the optimized arrangement of the interchain hydrogen bonds.[8] It is the distributed character of the periodic lattice interactions that give rise to the remarkably uniform curvatures in the observed structures, whereas it is the opposition of helices twisting with op-

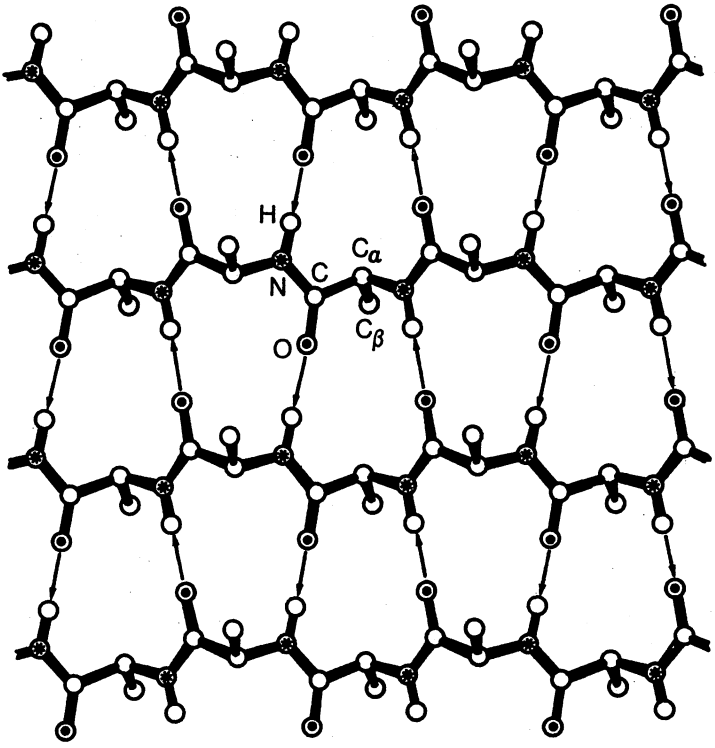

FIGURE 3. A schematic view of a flat, multiple-strand parallel β-sheet. The structure forms a lattice stabilized by covalent interactions along the chain directions and hydrogen-bonded interactions across the chains.

posite senses that results in a rigidly defined extended geometry for the structure as a whole (FIG. 7).

Stressed structures incorporating both the anticlastic curvature and geometrical properties inherent in protein parallel sheets have found wide architectural application in the construction of light and rigid structures such as roofs and power plant cooling towers.[12] Consequently, it is perhaps not surprising to find that nature has used the same engineering principles in the construction of extended and rigid backbone structures in proteins.

The preceding has emphasized how chiral forces ultimately arising from the polypeptide chains' composition of L-amino acids manifest themselves as operative factors stabilizing extended structures in proteins. In the representative examples described above, the opposition of twist components was involved in the maintenance of the extended structural integrity required for the preservation of protein function in a dynamically active environment.

Many additional invariant chiral effects are observed in known proteins

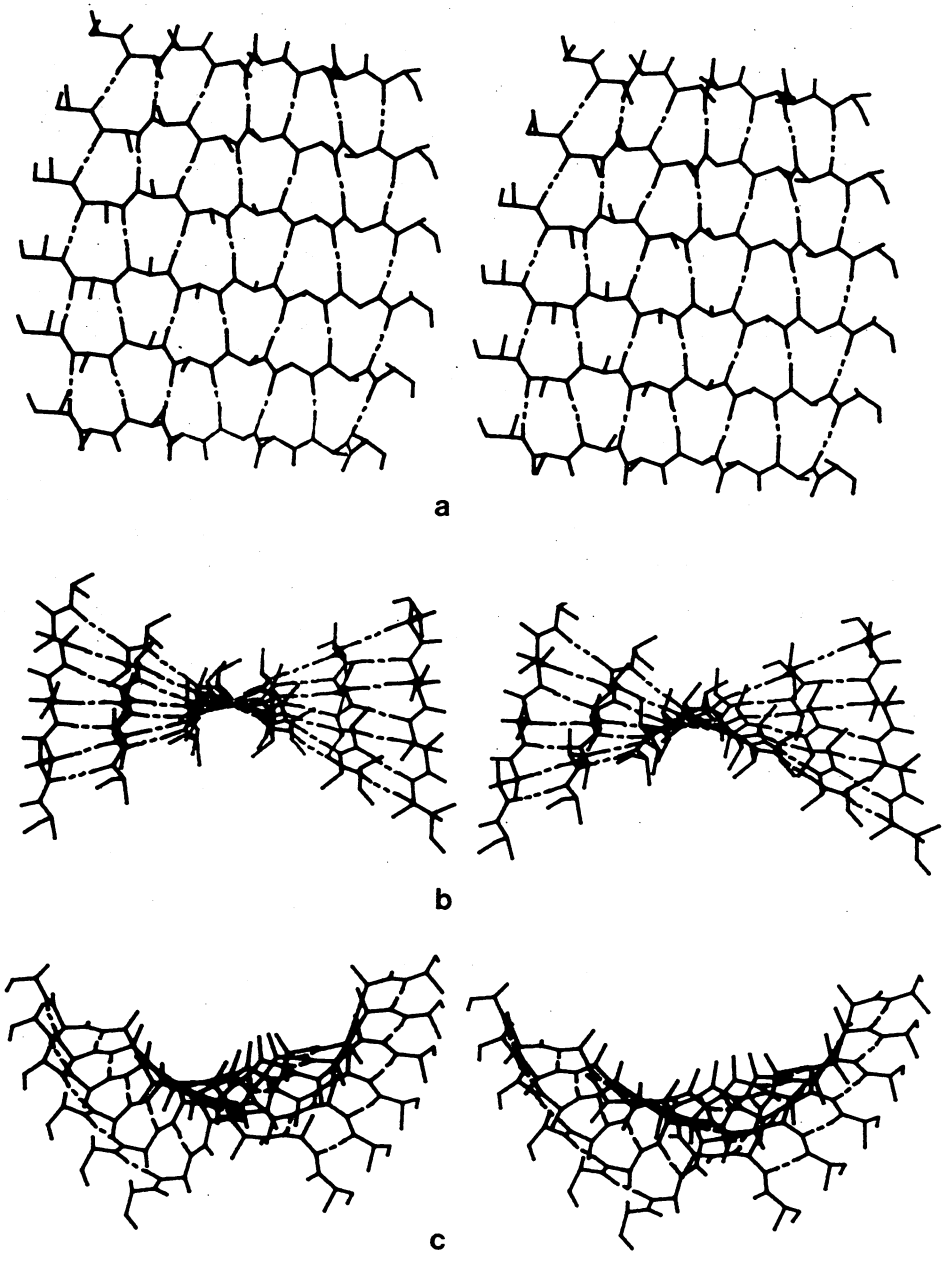

FIGURE 4. Stereo views of a rectangular plan, twisted parallel β-sheet.

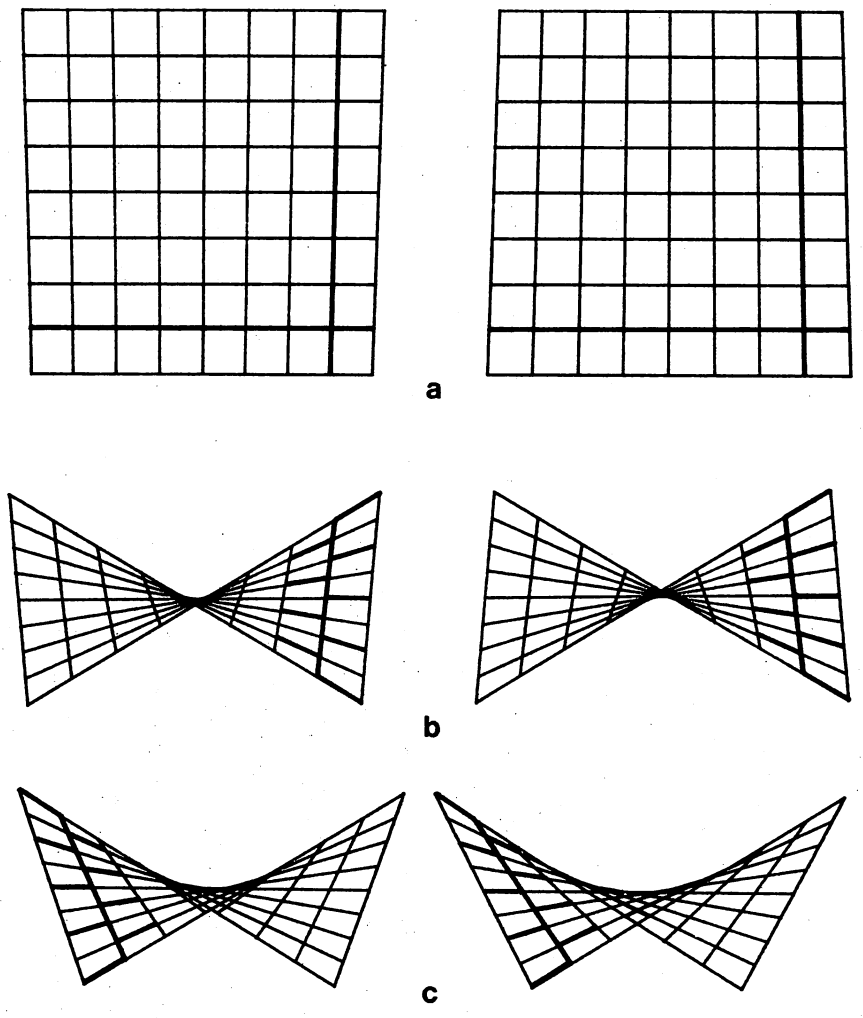

FIGURE 5. Schematic views of the hyperbolic parabolid surface formed by parallel sheets as shown in FIGURE 4. Note that this uniformly curved surface results from stress equilibration among two sets of helical interactions lying respectively along the polypeptide chains and along the contiguous hydrogen bond networks between strands.

involving, for example, the handedness of looping interconnections between stretches of polypeptide chain forming extended secondary structures.[6] Generally these loops form right-handed supercoils, which again may reflect a statistical preference for forming locally left-handed conformations in extended polypeptide chains. Additional manifestations of chiral invariance are present in extended antiparallel β-sheets. However, in this case, the prop-

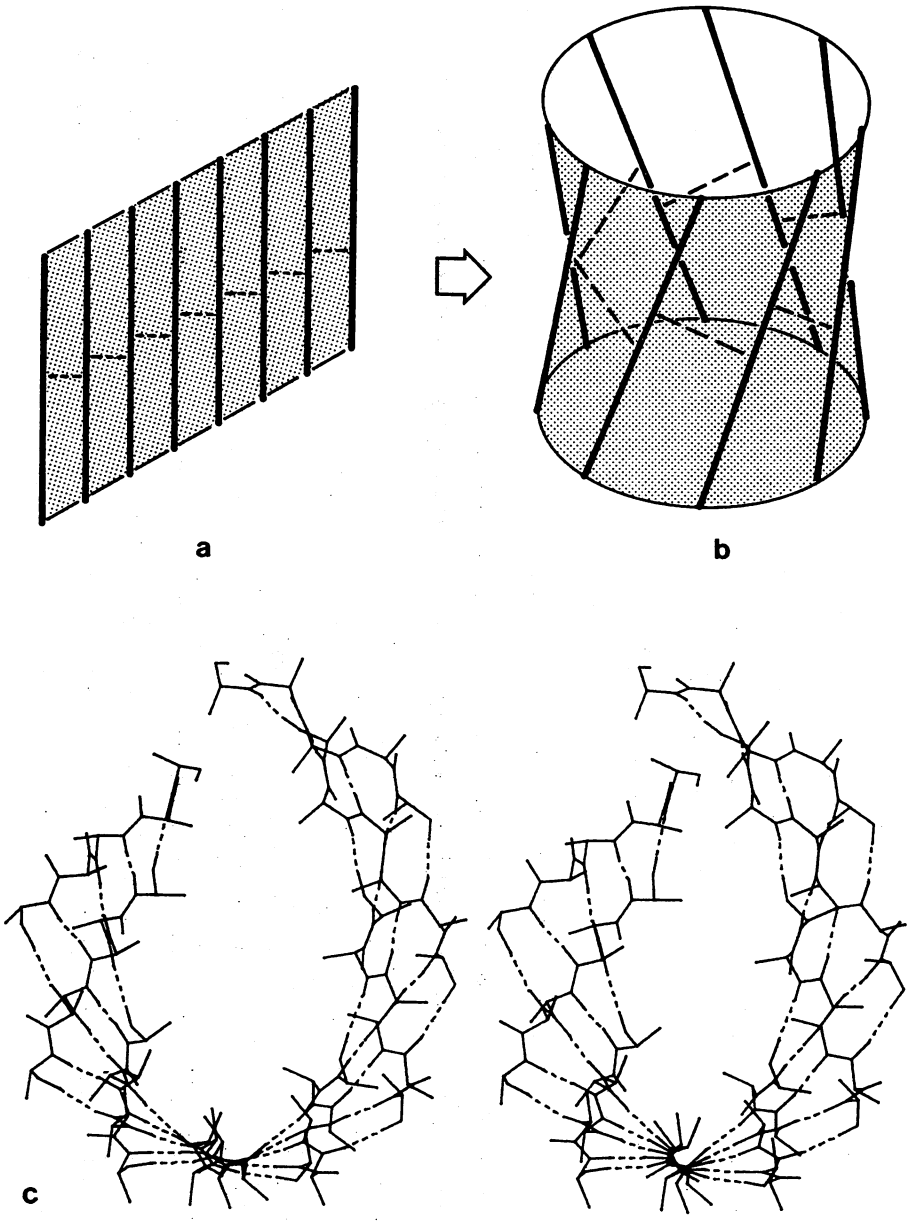

FIGURE 6. (a, b) The formation of parallel sheet structures with cylindrical curvature results from the pattern of hydrogen-bonded interactions between adjacent chains. (c) A stereo view down a helical chain axis of a cylindrically curved sheet with a staggered-chain hydrogen bonding pattern. Both the structures shown in FIGURES 5 and 6b are locally curved in opposite directions (anticlastically curved), an arrangement that makes such stressed lattice structures particularly rigid.

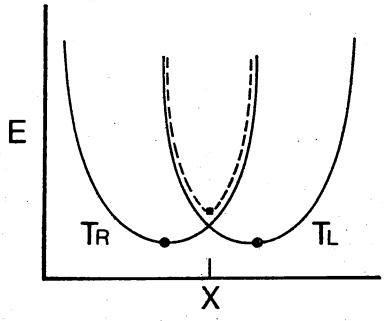

FIGURE 7. A plot of potential energy versus displacement in extended stressed structures. The structural stabilization of extended lattice structures can be usefully viewed as a result of the interaction of potential surfaces that reflect delocalized twist components in the structure as a whole (*solid curves*). The stressed configuration is constrained by the potential (*dashed*) defined by the opposing twist components. Although the individual twist potentials may be relatively soft (*e.g.*, small changes in potential energy occur for relatively large variations in displacement), the dashed potential produced by twist forces acting in opposition results in larger variations in energy with displacement, and consequently a more rigid structural configuration.

erties observed suggest that long-range cooperative interactions associated with sheet hydrogen-bond formation play a role in structure folding.

Multiple-strand antiparallel sheet domains virtually always appear to have been assembled by a folding pathway that first involves the formation of a long double-strand "hairpin" sheet. Subsequent assembly then appears to involve the concerted right-handed supercoiling of the double-strand sheet to produce the final structural domain (FIG. 8). This hypothetical mechanism[6] for folding sheets is of interest in the present context because it infers that the hairpin sheet is a highly flexible structure with a built-in tendency to form right-handed supercoils. Detailed studies of the conformational flexibility of double-strand antiparallel sheets indeed show them to possess an unusual extent of cooperative flexibility.[8] Most interestingly, the structure can be shown to supercoil perferentially with a right-handed sense to produce structures that are topologically equivalent to the observed domains in proteins (FIG. 9). These results indicate that the periodic hydrogen bonding constraints shown above to be important in the protein's final structural organization may, in the case of antiparallel sheet domains, additionally play a determinative role in defining a highly cooperative domain folding pathway. In a mechanical sense, the cooperative supercoiling of the double-strand hairpin sheet arises from its kinematic properties, *i.e.*, how the structure undergoes characteristic cooperative motions as defined by the nature and geometry of its mechanical linkages. Alternatively, this kinematic behavior can be viewed as the result of the superposition of delocalized vibrational modes of the structure as a whole.[5] In this case, initial formation of the hydrogen-bonded interactions in the double-strand sheet serve to couple the structure, thus resulting in the partitioning of the structure's kinetic energy into a new set of vibrational modes characteristic of the structure as a whole.[8] It is the su-

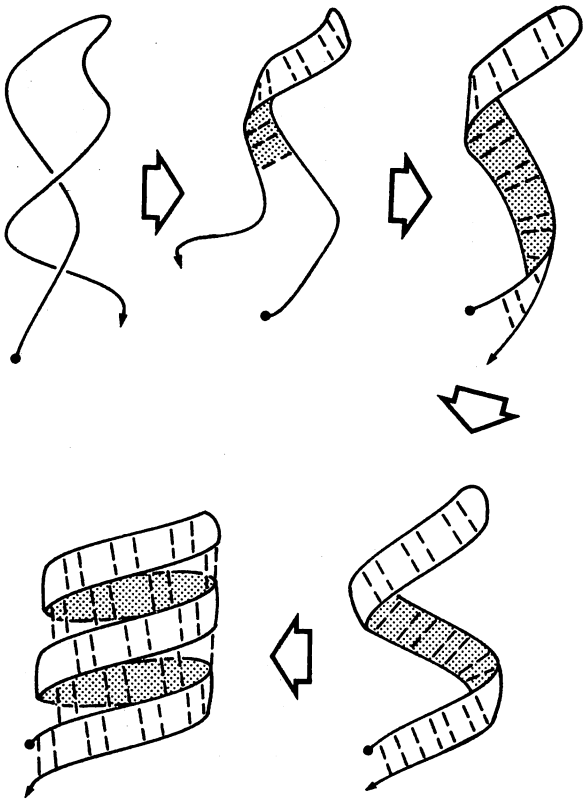

FIGURE 8. A schematic illustration of a hypothetical folding pathway for an antiparallel β-roll domain as observed in some proteins of known structure (modified after an illustration by Jane Richardson). Structure folding proceeds by the right-hand supercoiling of an initially formed double-strand, "hairpin" antiparallel sheet. The chiral folding trajectory appears to be determined by the cooperative folding properties of the double-strand structure (FIGURE 9).

perposition of these low-energy modes that reflect themselves in the structure's kinematic properties and apparently must be involved in the process of folding β-sheet domains.

We have shown that many aspects of long-range order in proteins reflect the distribution of forces over extended lattice structures. In the case of α-helices or parallel β-structures, these distributed forces effectively serve to prestress the structure, so conferring rigidity in a dynamically active environment. Although extended lattice interactions play similar roles in the organization of antiparallel sheet domains,[8] here they additionally play a determinative role in the process of protein folding. Although many questions remain concerning the origins of protein structural stability, we have illustrated here that many of their properties plainly reflect adherence to well-known principles of structural engineering.

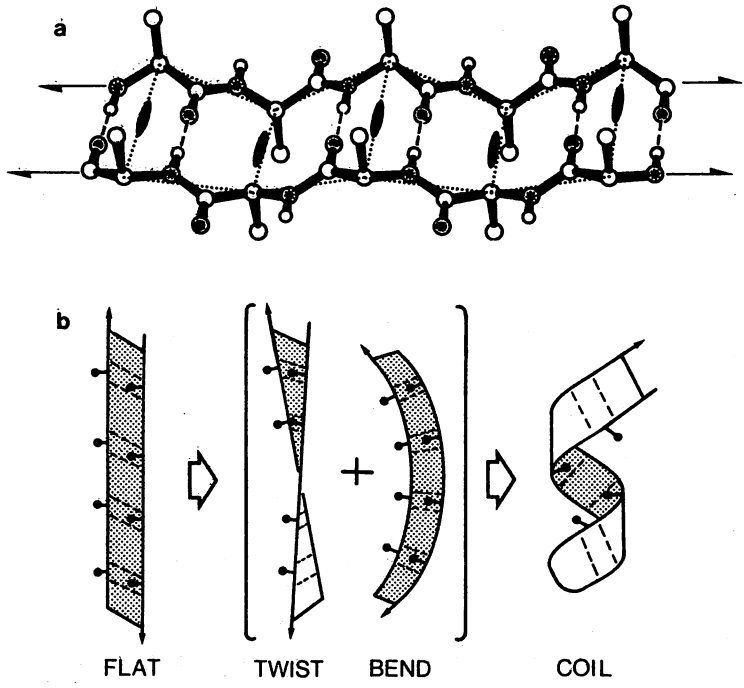

FIGURE 9. A section of double-strand antiparallel sheet illustrating that opposite surfaces of the sheet differ (**a**); *e.g.,* one side is populated by substituents from 10-membered hydrogen-bonded rings, and the other by substituents from 14-membered rings. Analysis of the low energy deformation pathways for the flat sheet show that the observed right-supercoiling results from coupled superposition of twist and bend components (**b**). Both twist and bend have an energetically preferred sense, so that resulting double-strand coils always have small ring substituents located on the coil interior. This feature is also preserved in the observed structures illustrated in FIGURE 8 (*shaded surfaces*), and so provides primary evidence of the role of cooperative double-strand coiling as a pathway for their assembly.

REFERENCES

1. PRIVALOV, P. L. 1979. Adv. Prot. Chem. **33**: 167–241.
2. CARERI, G., P. FASELLA & E. GRATTON. 1979. A. Rev. Biophys. Bioeng. **8**: 6997.
3. MCCAMMON, J. A., B. R. GELIN & M. KARPLUS. 1977. Nature **277**: 585–590.
4. COOPER, A. 1976. Proc. Nat. Acad. Sci. U.S.A. **73**: 2740–2741.
5. PETICOLAS, W. L. 1979. Meth. Enzymol. **61**: 425–456 (part H).
6. RICHARDSON, J. S. 1981. Adv. Prot. Chem. **34**: 167–339.
7. WEBER, P. C. & F. R. SALEMME. 1980. Nature (London) **287**: 82–84.
8. SALEMME, F. R. 1983. Prog. Biophys. Molec. Biol. **42**: 95–133.
9. LEVY, R. M. & M. KARPLUS. 1979. Bipolymers **18**: 2465–2995.
10. PAULING, L. & R. B. COREY. 1951. Proc. Nat. Acad. Sci. U.S.A. **37**: 729–740.
11. CHOTHIA, C. 1973. J. Molec. Biol. **75**: 295–302.
12. OTTO, F. 1969. Tensile Structures. MIT Press. Cambridge, Mass.