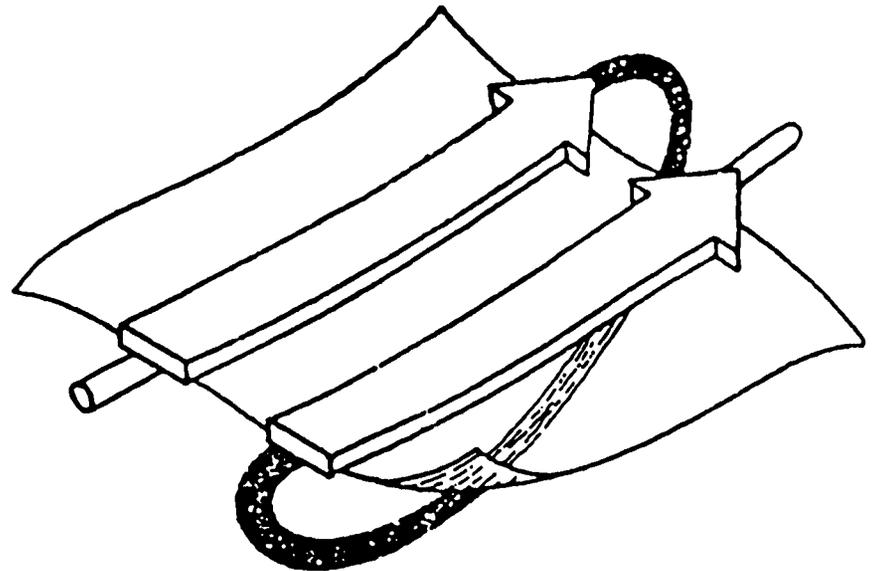
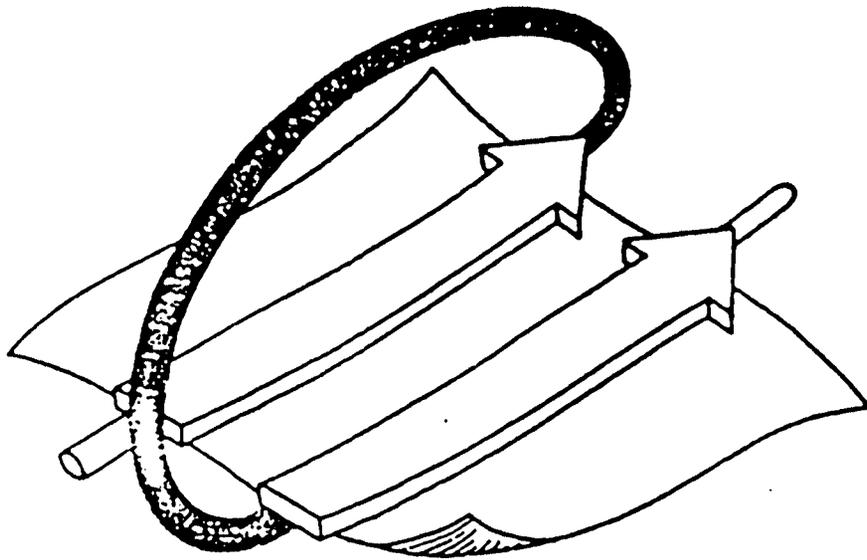
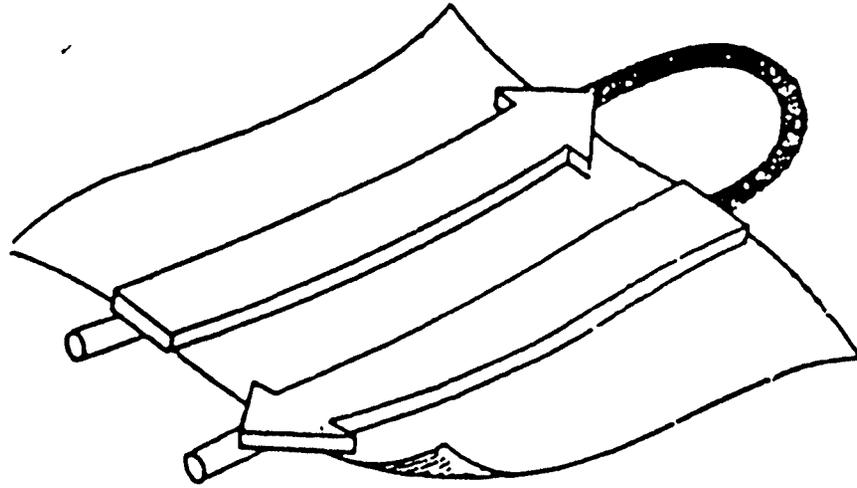
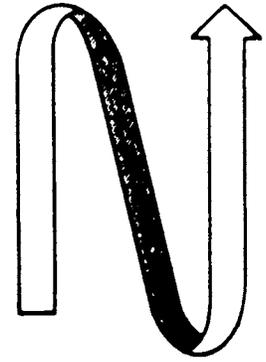
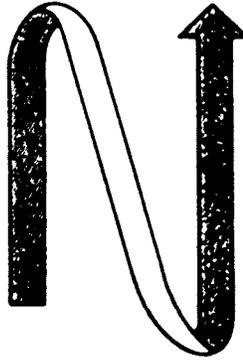
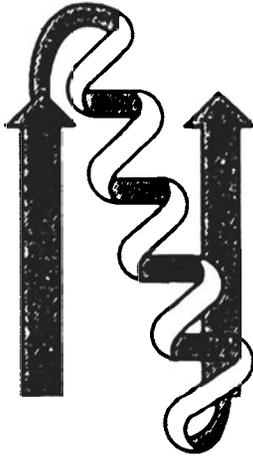


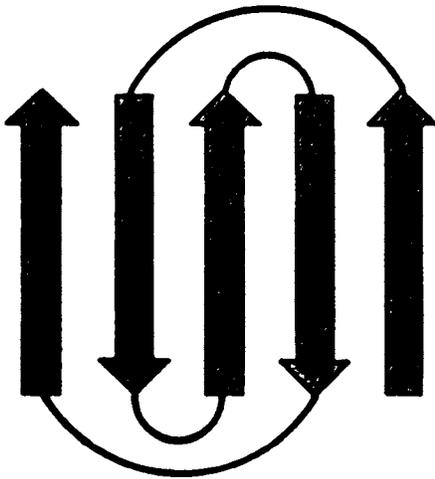
# Proteins: Tertiary Structure, Motifs and Fold Classes

CHEM 430

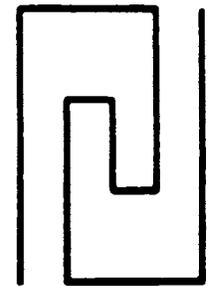
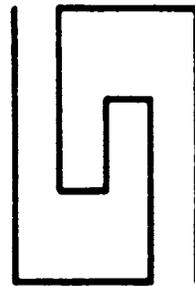




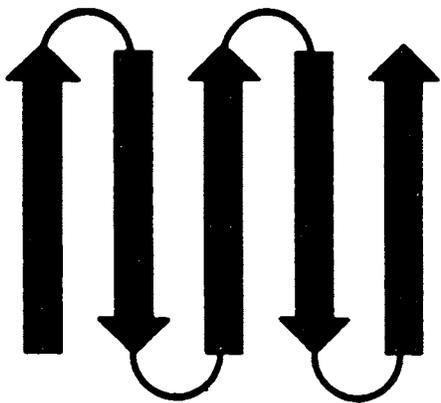
Very rare



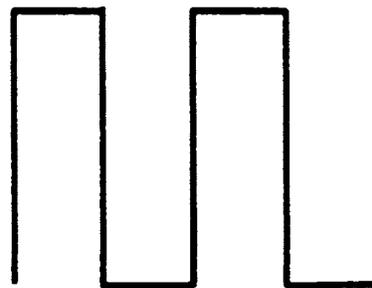
5 2 3 4 1

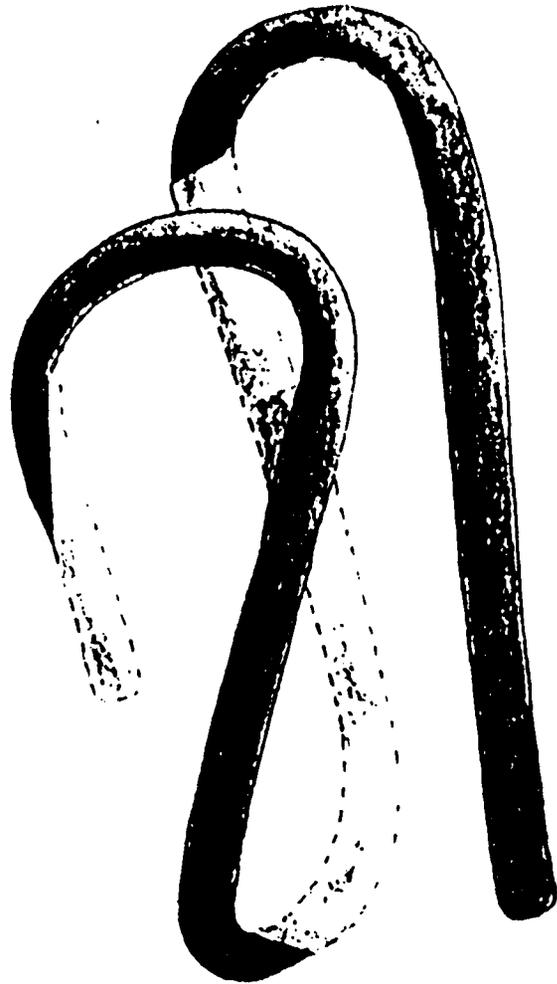


Not observed

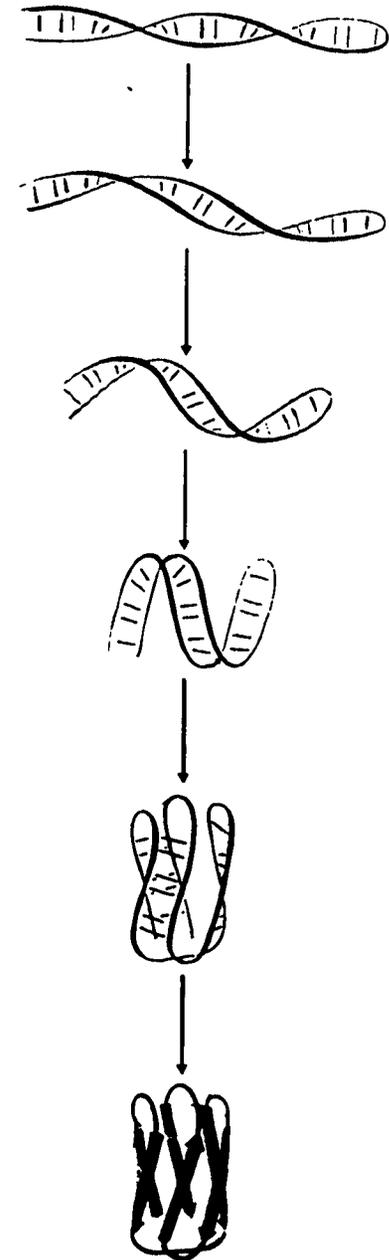


1 2 3 4 5



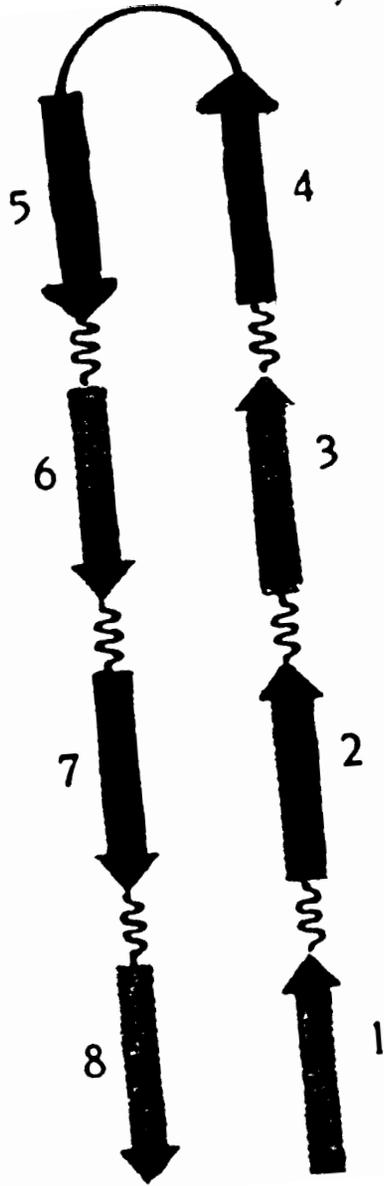


**Figure 5.16** A schematic diagram of a piece of string wrapped around a barrel to illustrate the basic pattern of a jelly roll motif.



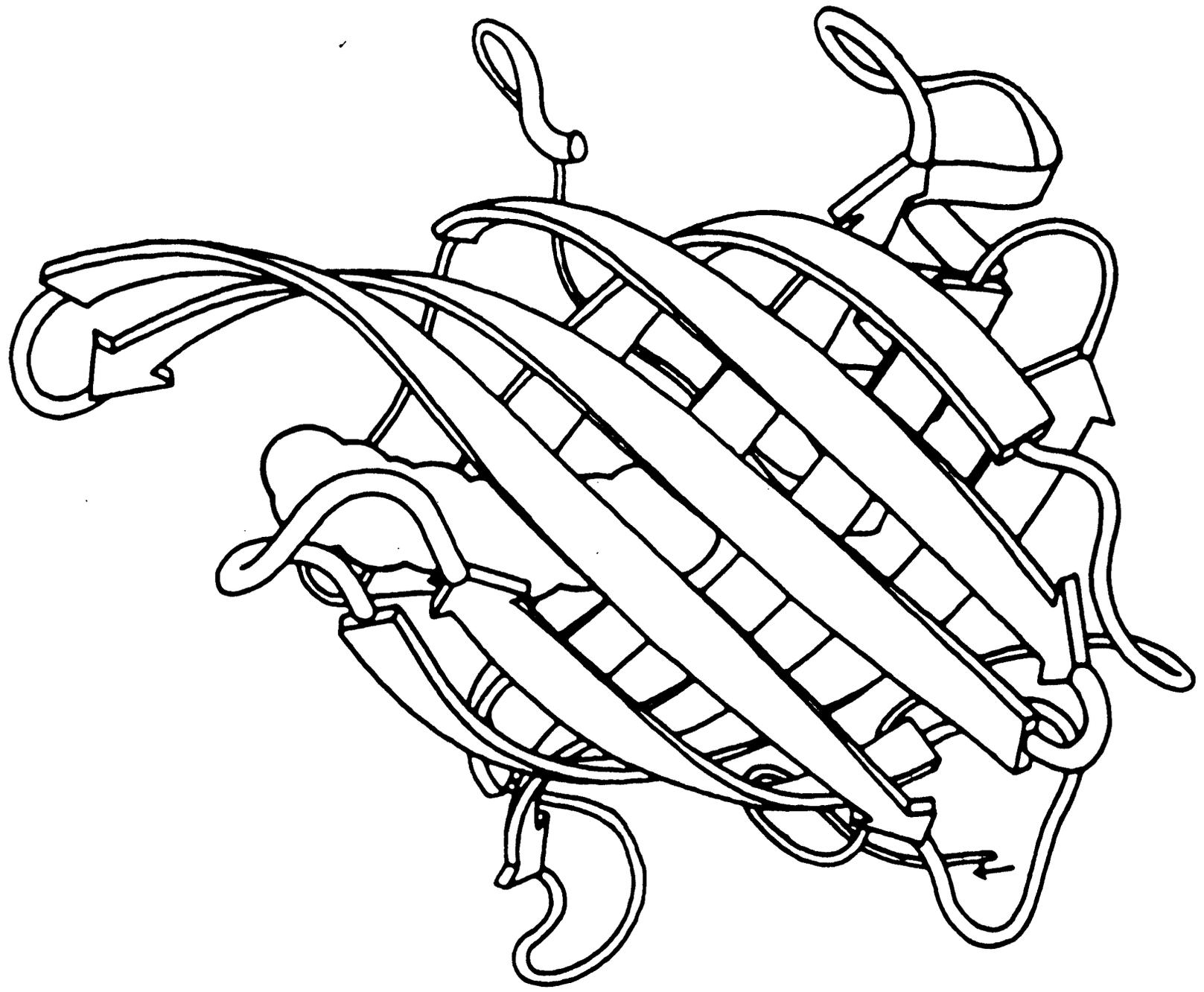
**Figure 5.19** The jelly roll motif may fold from a long antiparallel hairpin structure that serves as a relatively stable folding intermediate. This hairpin structure can then twist and curl as shown in the diagram to form the jelly roll structure of six strands in a way similar to the illustration in Figure 5.17. (Adapted from J. Richardson.)

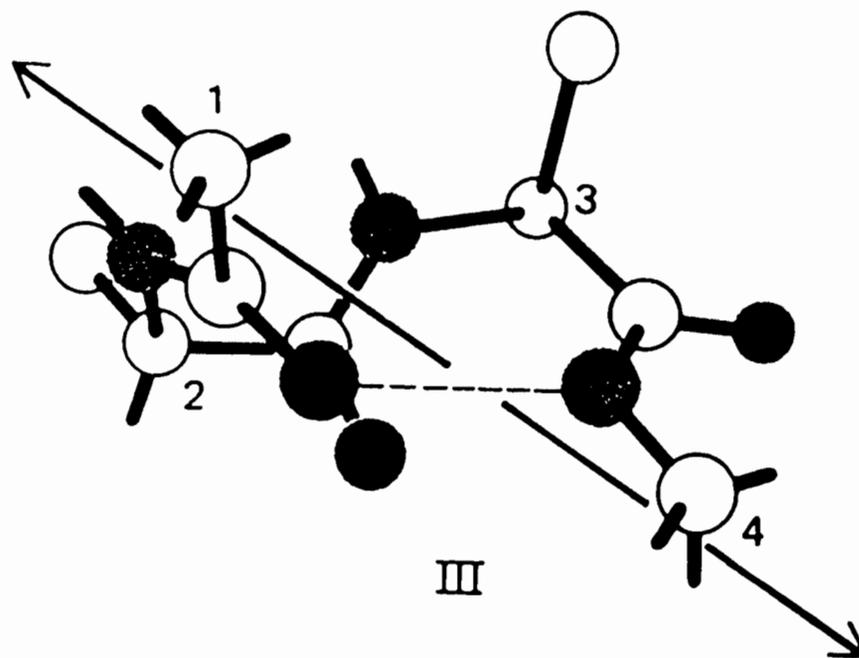
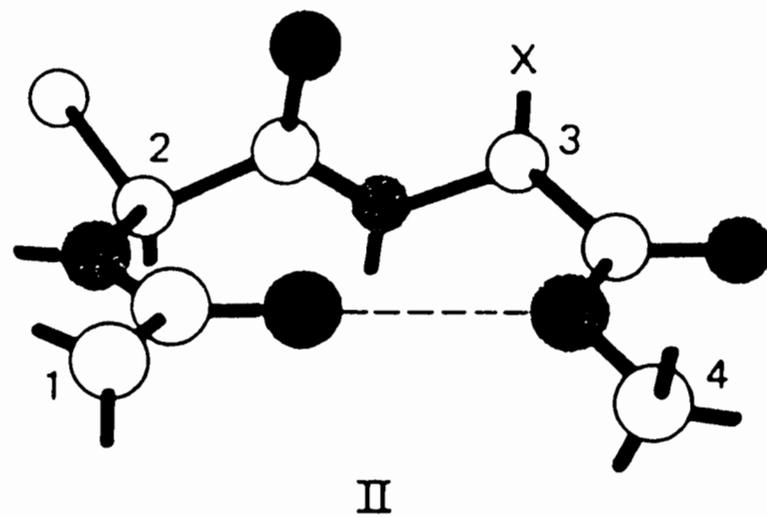
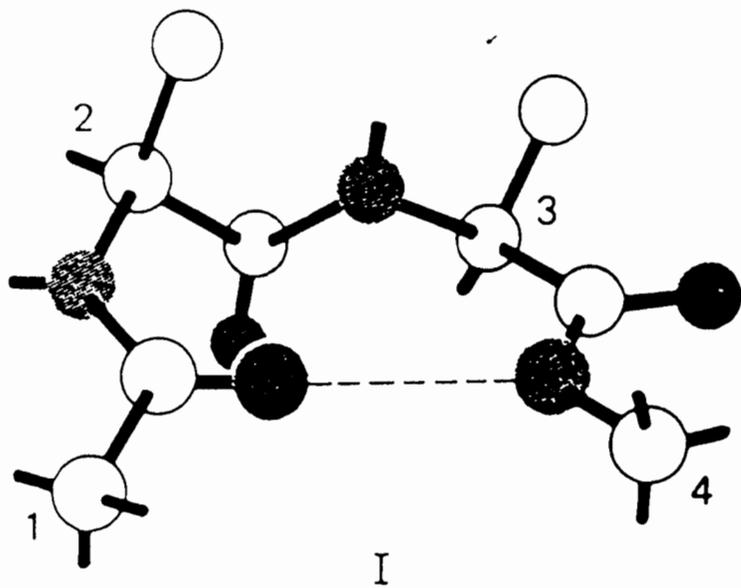
(a)



(b)







Turn

<----i+1---->

<----i+2---->

$\phi$

$\psi$

$\phi$

$\psi$

$\beta$ -turns

Type I	-60	-30	-90	0
Type I'	60	30	90	0
Type II	-60	120	80	0
Type II'	60	-120	-80	0
Type III	-60	-30	-60	-30
Type III'	60	30	60	30
Type VIa ( <i>cis</i> )	-60	120	-90	0
Type VIb ( <i>cis</i> )	-120	120	-60	0

$\gamma$  turns

turn	70 to 85	-60 to -70
inverse turn	-70 to -85	60 to 70

I common;  $C_{\alpha}$  are non planar

II Gly in residue 3 (~60% have Gly);  $C_{\alpha}$  are planar

III compare to  $3_{10}$  helix

VI *cis* Pro turn

•buried turns

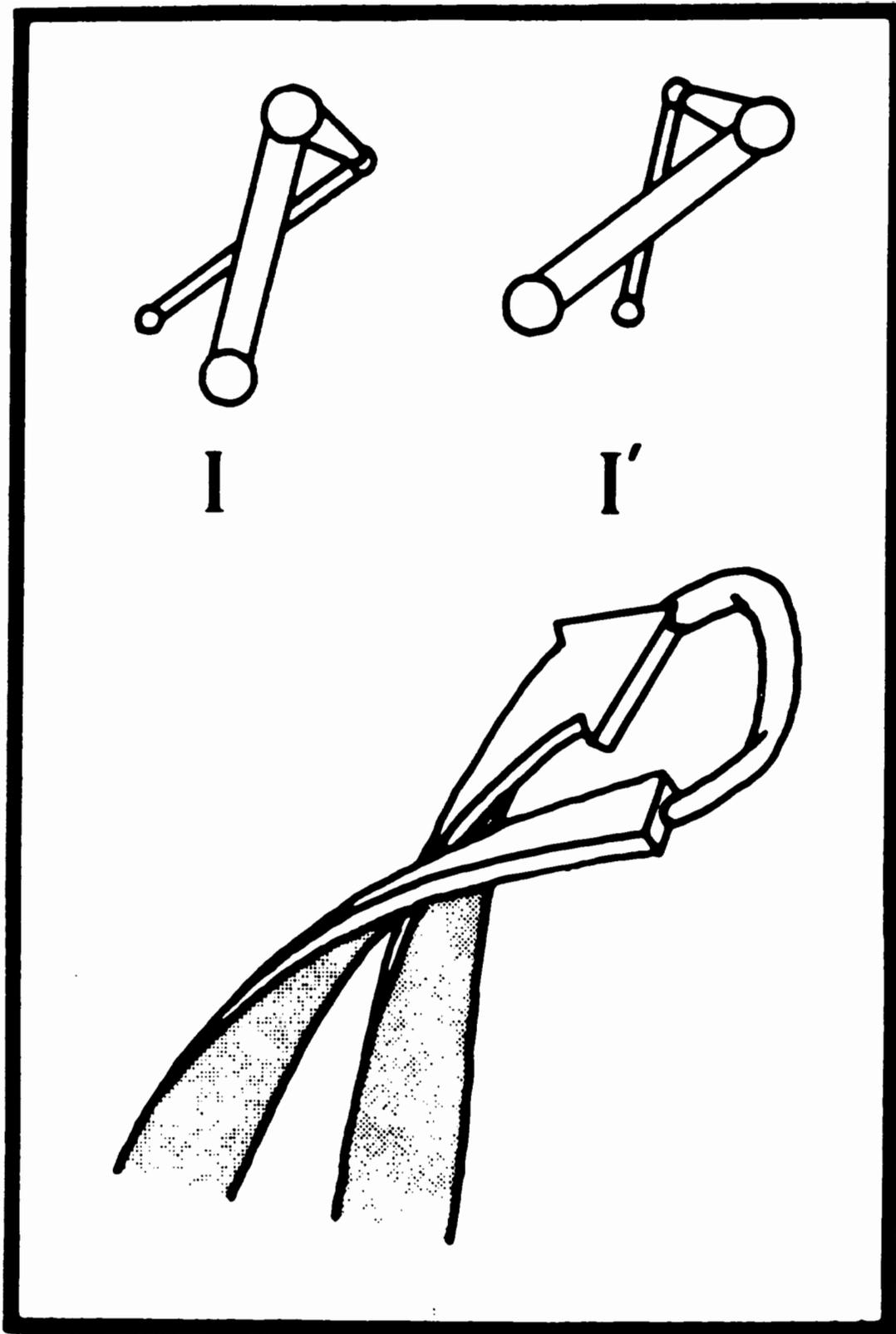
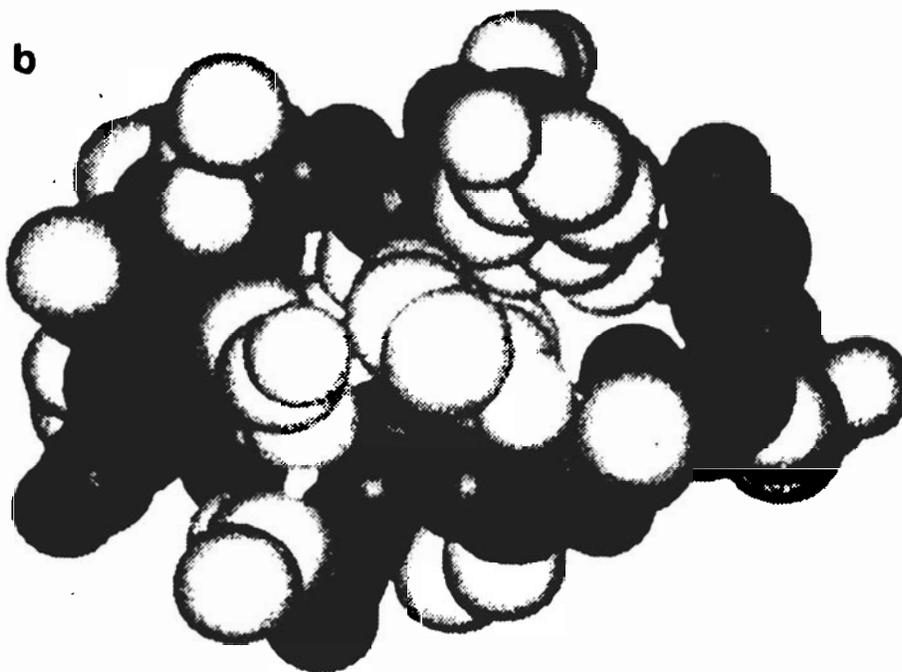
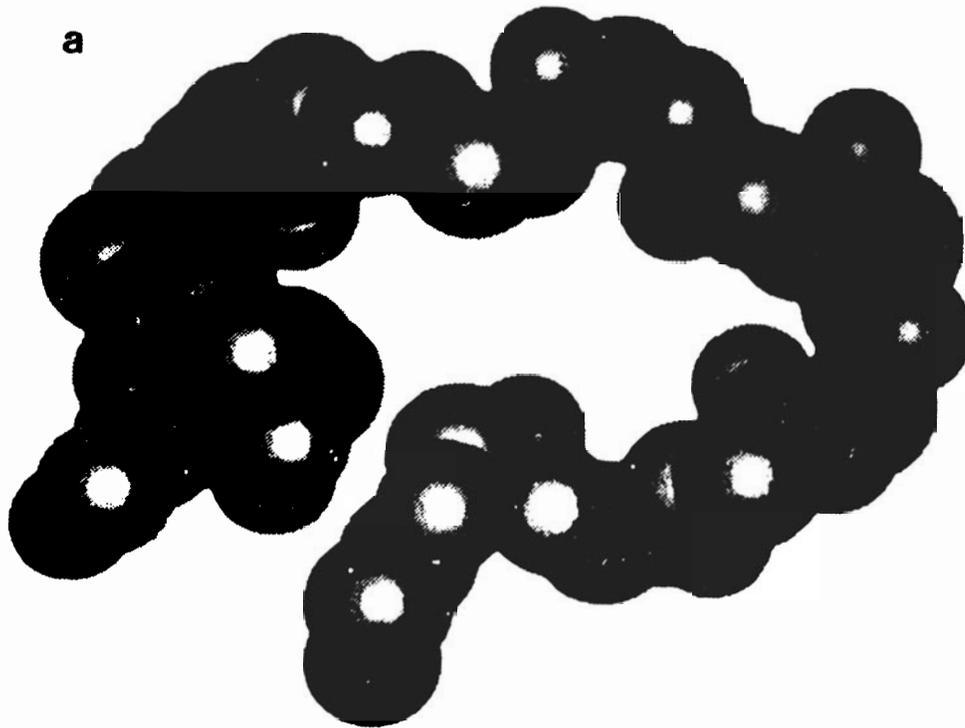


FIGURE 16 Simplified drawings of Type I vs. Type I' tight turns, showing that Type I has a twist incompatible with the twist direction of  $\beta$  hairpins.

# Compact $\Omega$ -Loops in Proteins



**Cytochrome C, Residues 40-54**

Frequency of helical,  $\beta$ , and coil residues in 15 proteins with their conformational parameters  $P_\alpha$ ,  $P_\beta$ , and  $P_c$

Amino acid	$f_\alpha$	$P_\alpha$	$f_\beta$	$P_\beta$	$f_c$	$P_c$
Ala	0.522	1.45	0.167	0.97	0.311	0.66
Arg	0.282	0.79	0.154	0.90	0.564	1.20
Asn	0.263	0.73	0.113	0.65	0.624	1.33
Asp	0.351	0.98	0.137	0.80	0.514	1.09
Cys	0.278	0.77	0.222	1.30	0.500	1.07
Gln	0.421	1.17	0.211	1.23	0.368	0.79
Glu	0.549	1.53	0.044	0.26	0.407	0.87
Gly	0.190	0.53	0.138	0.81	0.668	1.42
His	0.446	1.24	0.122	0.71	0.432	0.92
Ile	0.358	1.00	0.274	1.60	0.368	0.78
Leu	0.480	1.34	0.209	1.22	0.311	0.66
Lys	0.383	1.07	0.126	0.74	0.491	1.05
Met	0.429	1.20	0.286	1.67	0.286	0.61
Phe	0.402	1.12	0.219	1.28	0.378	0.81
Pro	0.212	0.59	0.106	0.62	0.682	1.45
Ser	0.282	0.79	0.124	0.72	0.594	1.27
Thr	0.295	0.82	0.205	1.20	0.494	1.05
Trp	0.409	1.14	0.203	1.19	0.386	0.82
Tyr	0.220	0.61	0.220	1.29	0.560	1.19
Val	0.409	1.14	0.282	1.65	0.309	0.66

Classification of amino acid residues as formers, breakers, or indifferent for  $\alpha$ -helical and  $\beta$ -sheet regions

Helical residues	$P_\alpha$	Classification	$\beta$ -Sheet residues	$P_\beta$	Classification
Glu <sup>(-)</sup>	1.53	H $_\alpha$	Met	1.67	H $_\beta$
Ala	1.45	H $_\alpha$	Val	1.65	H $_\beta$
Leu	1.34	H $_\alpha$	Ile	1.60	H $_\beta$
His <sup>(+)</sup>	1.24	h $_\alpha$	Cys	1.30	h $_\beta$
Met	1.20	h $_\alpha$	Tyr	1.29	h $_\beta$
Gln	1.17	h $_\alpha$	Phe	1.28	h $_\beta$
Trp	1.14	h $_\alpha$	Gln	1.23	h $_\beta$
Val	1.14	h $_\alpha$	Leu	1.22	h $_\beta$
Phe	1.12	h $_\alpha$	Thr	1.20	h $_\beta$
Lys <sup>(+)</sup>	1.07	I $_\alpha$	Trp	1.19	h $_\beta$
Ile	1.00	I $_\alpha$	Ala	0.97	I $_\beta$
Asp <sup>(-)</sup>	0.98	i $_\alpha$	Arg <sup>(+)</sup>	0.50	i $_\beta$
Thr	0.82	i $_\alpha$	Gly	0.81	i $_\beta$
Ser	0.79	i $_\alpha$	Asp <sup>(-)</sup>	0.80	i $_\beta$
Arg <sup>(+)</sup>	0.79	i $_\alpha$	Lys <sup>(+)</sup>	0.74	b $_\beta$
Cys	0.77	i $_\alpha$	Ser	0.72	b $_\beta$
Asn	0.73	b $_\alpha$	His <sup>(+)</sup>	0.71	b $_\beta$
Tyr	0.61	b $_\alpha$	Asn	0.65	b $_\beta$
Pro	0.59	B $_\alpha$	Pro	0.62	b $_\beta$
Gly	0.53	B $_\alpha$	Glu <sup>(-)</sup>	0.26	B $_\beta$

NOTE: H = strong former; h = former; I = weak former; i = indifferent; b = breaker; B = strong breaker.

SOURCE: P. Y. Chou and G. D. Fasman, *Biochemistry* 13:222 (1974).

Table II. Directional Information Values:  $I(S_j = X:\bar{X};R_{j+m})$   
for Helical Conformation  $X = H$

Amino acid	Residue position																
	$j-8$	$j-6$	$j-4$	$j-2$	$j$	$j+2$	$j+4$	$j+6$	$j+8$	$j+10$	$j+12$	$j+14$	$j+16$	$j+18$	$j+20$	$j+22$	$j+24$
Gly	0	0	-5	-10	-20	-30	-50	-70	-100	-70	-50	-30	-20	-10	-5	0	0
Ala	0	5	15	20	30	40	45	55	60	55	45	40	30	20	15	5	0
Val	0	0	0	0	-5	-15	-10	-5	0	-5	-10	-10	-5	0	0	0	0
Leu	0	0	5	5	5	10	15	20	25	20	15	10	5	5	5	0	0
Ile	0	0	0	0	0	0	5	10	15	10	5	0	0	0	0	0	0
Ser	0	0	-5	-10	-15	-20	-25	-30	-35	-33	-30	-25	-20	-15	-10	-5	-5
Thr	0	-5	-10	-15	-20	-25	-35	-40	-45	-40	-35	-25	-20	-15	-10	-5	0
Asp	0	0	0	0	10	15	10	5	0	-10	-25	-35	-25	-10	-5	0	0
Glu	25	28	32	35	44	50	60	65	70	55	40	20	10	5	5	0	0
Asn	0	0	0	0	-5	-10	-20	-30	-42	-30	-20	-10	-5	0	0	0	0
Gln	0	0	0	0	0	0	0	0	15	20	25	25	20	10	0	0	0
Lys	5	5	5	5	5	5	10	15	28	40	55	58	50	45	42	40	40
His	0	0	0	0	0	0	0	0	5	15	25	32	36	40	42	43	44
Arg	0	0	-10	-18	-20	-7	10	20	30	36	36	30	20	10	0	0	0
Phe	0	0	0	0	5	15	22	28	30	28	22	15	5	0	0	0	0
Tyr	0	-3	-10	-15	-22	-30	-35	-47	-35	-20	-10	-10	-10	-12	-17	-25	-45
Trp	25	40	30	20	10	42	42	33	20	12	5	3	0	0	0	0	0
Cys	-15	-15	-15	-15	-15	-12	-10	-5	0	0	-10	-25	-35	-43	-46	-47	-37
Met	5	5	8	10	20	30	38	45	50	48	45	42	35	30	20	15	10
Pro	0	-5	-12	-20	-25	-35	-50	-66	-96	-177	-118	-90	-70	-52	-40	-30	-20

position	1	2	3	4	5	6	7	8	
	G	L	S	G	S	G	K	S	<i>E.coli</i> UvrA
	A	G	A	G	S	G	K	T	<i>E.coli</i> UvrD
	G	P	H	G	M	G	K	T	HSV Thymidine Kinase
	G	P	E	S	S	G	K	T	<i>E.coli</i> RecA
	P	S	P	G	T	G	K	T	Phage T4 Gene 44

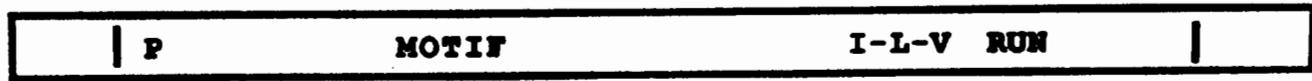
**Fig. 1.** Octapeptides from five out of about 120 known examples of this motif, known as the Walker A-consensus, have been aligned. Many of these proteins are known to bind nucleoside triphosphates (Walker *et al.*, 1982; Gorbalenya *et al.*, 1985; Higgins *et al.*, 1988). The sequence positions have been marked above the alignment, and the names of the parent proteins on the right.

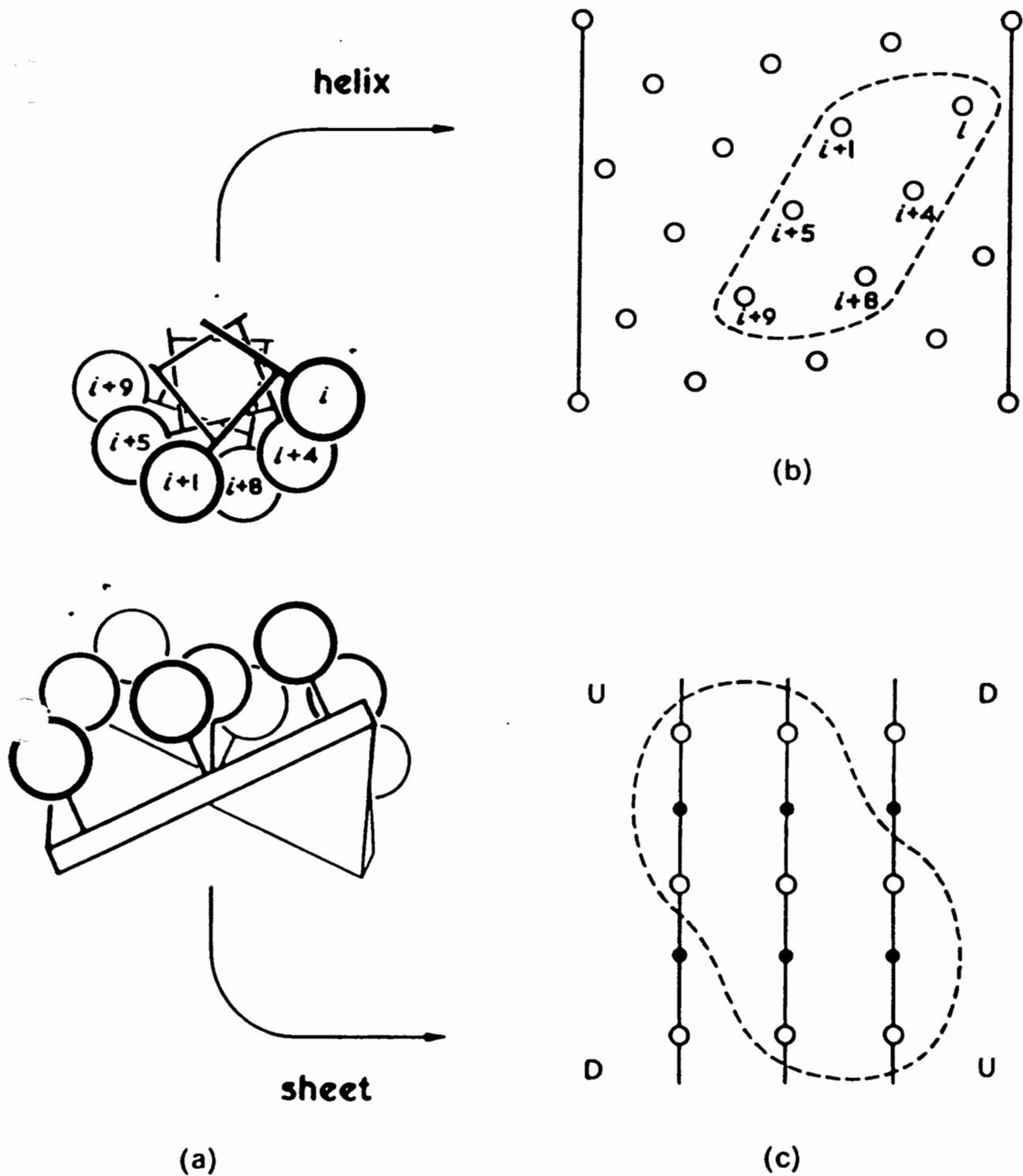
EBV	YV <u>GV</u> LT <u>DGKTL</u> -MKGV
HSV	YIGVIYGGKML-IKGV
VZV	YIGVIYGGKVL-MKGV
CMV	YIGKVEGASGLSMKGV
Phage Phi29	YLRQKTYIQDIYMKEV

**Fig. 3.** The above alignment is from a family of herpes virus DNA polymerases (Davison and Scott, 1986; Kouzarides *et al.*, 1987), and falls between conserved regions which are also found in other virus and cellular DNA polymerases (Wong *et al.*, 1988) – a bacteriophage segment is shown for comparison. The underlined part of the Epstein-Barr Virus sequence matches the Walker A-consensus. However, it is not conserved. The equivalent regions suggest this is merely a loop between two functionally significant parts of the molecule, and is even seven residues longer in the polymerases not shown.

Table I. The sequence pattern in the transmembrane domain of tyrosine kinase growth factor receptors

Position	8	7	6	5	4	3	2	1	0	1	2	3	4																											
Neu	E	Q	R	A	S	P	V	T	F	I	I	A	T	V	V	L	L	F	L	I	L	V	V	V	V	G	I	L	I	K	R	R								
c-erbB-2	E	Q	R	A	S	P	L	T	S	I	V	S	A	V	V	G	I	L	L	V	V	V	L	G	V	V	F	G	I	L	I	K	R	R						
EGFR-H				G	P	K	I	P	S	I	A	T	G	M	V	G	A	L	L	L	L	V	V	A	L	G	I	G	L	F	M	R	R	R						
DER				D	V	N	M	I	F	I	T	G	A	V	L	V	P	T	I	C	I	L	C	V	V	T	Y	I	C	R	Q	K								
v-erb-B				G	S	K	T	P	S	I	A	A	G	V	V	G	G	L	L	C	L	V	V	V	G	L	G	I	G	L	Y	L	R							
PDGFR-A				A	P	T	L	R	S	E	L	T	V	A	A	A	V	L	V	L	L	V	I	V	I	I	S	L	I	V	L	V	I	W	K	Q	K			
PDGFR-B				L	P	F	K	V	V	V	I	S	A	I	L	A	L	V	V	L	T	V	I	S	L	I	I	L	I	M	L	W	Q	K	K					
CSF1R				P	P	D	E	F	L	F	T	P	V	V	V	A	C	M	S	I	M	A	L	L	L	L	L	L	L	L	L	L	L	L	L	Y	K	Y		
v-fms				L	P	D	E	L	L	F	T	P	V	L	L	T	C	M	S	I	M	A	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	Y	K	Y
c-kit				I	H	P	H	T	L	F	T	P	L	L	I	G	F	V	I	V	A	G	M	M	C	I	I	V	M	I	L	T	Y	K	Y					
Ret				P	L	C	D	E	L	C	R	T	V	I	A	A	A	V	L	F	S	F	V	V	S	V	L	L	S	A	F	C	I	H	C	Y				
INS.R	I	A	K	I	I	I	G	P	L	I	F	V	F	L	F	S	V	V	I	G	S	I	Y	L	F	L	R	K	R											
IGF1R				E	N	F	I	H	L	I	I	A	L	P	V	A	V	L	L	I	V	G	G	L	V	I	M	L	Y	V	F	H	R	K						
DILR				Y	A	K	V	F	F	W	L	L	G	I	G	L	A	F	L	I	V	S	L	F	G	Y	V	C	Y	L	H	K	R							
Ros-H				I	P	E	T	S	F	I	L	T	I	I	V	G	I	F	L	V	V	T	I	P	L	T	F	V	W	H	R	R								
Ros-C	S	P	D	I	T	A	I	V	A	V	I	G	A	V	V	L	G	L	T	I	I	L	F	G	F	V	W	H	Q	R										
7LESS				F	V	S	P	E	K	R	G	S	L	V	L	A	I	I	A	P	A	A	I	V	S	S	C	V	L	A	L	V	L	V	R	K	V			
Met				V	Q	P	D	Q	N	F	T	G	L	I	A	G	V	V	S	I	S	T	A	L	L	L	L	G	F	F	L	W	L	K	K	R				
Trk				K	D	E	T	P	F	G	V	S	V	A	V	G	L	A	V	F	A	C	L	F	L	S	T	L	L	L	V	L	N	K	C					
Ltk				P	T	T	A	S	P	L	I	L	M	G	A	V	V	A	A	L	A	L	S	L	L	M	M	C	A	V	L	I	L	V	N	Q	K			

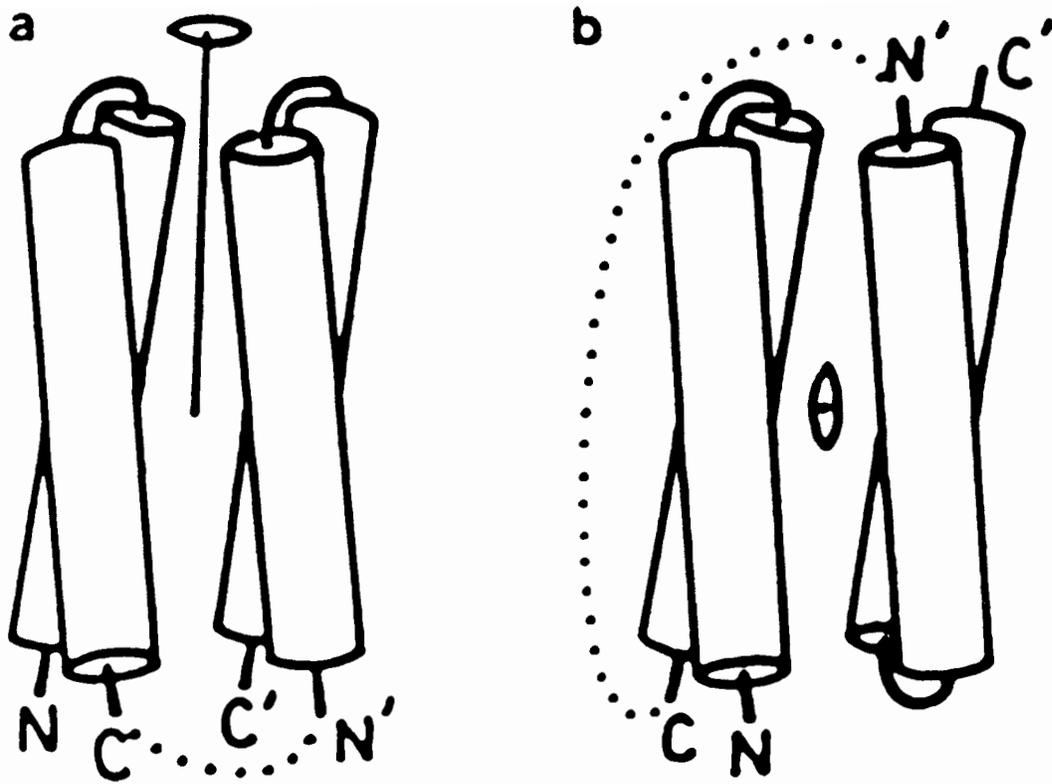




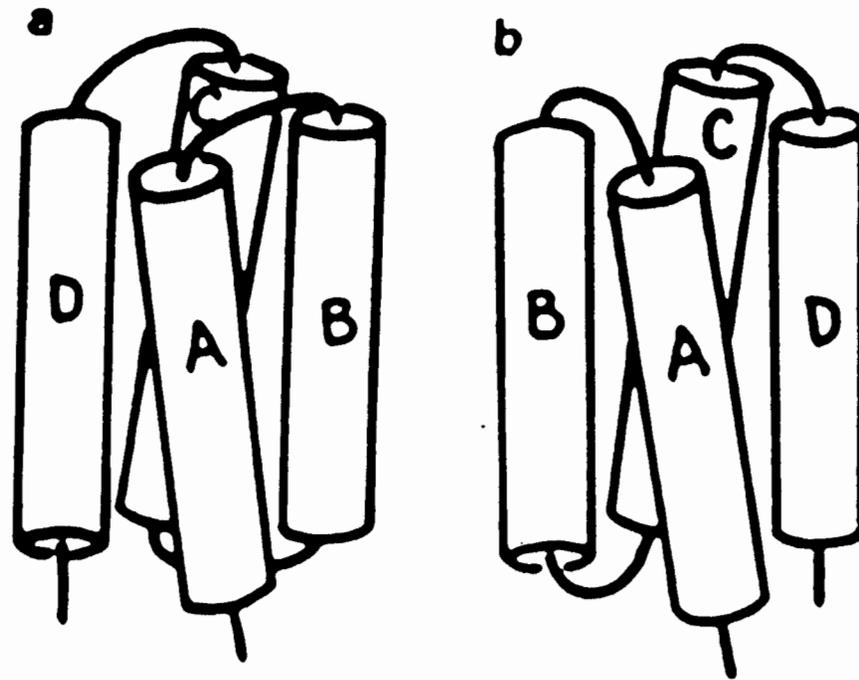
**Figure 9** The complementary twist model for  $\alpha$ -helix to  $\beta$ -sheet packing. Side chains on one side of an  $\alpha$  helix and a twisted  $\beta$  sheet are shown as open circles (a). When the helix axis is parallel to the strand direction, the helix residues  $i, i+1, i+4, i+5, i+8, i+9, \dots$  form a twisted surface that is complementary to that of the twisted  $\beta$  sheet. The patterns formed by the residues in contact at the interface are shown on flattened projections of the  $\alpha$  helix (b) and  $\beta$  sheet (c). U and D mark the corners of the  $\beta$  sheet that move up and down from the plane when the  $\beta$  sheet is twisted. Note that the  $\beta$ -sheet contact residues cluster about a line joining U to U, the concave diagonal of the  $\beta$  sheet. [Reprinted with permission from (44).]

**Table 4.1. Classification of protein topologies**

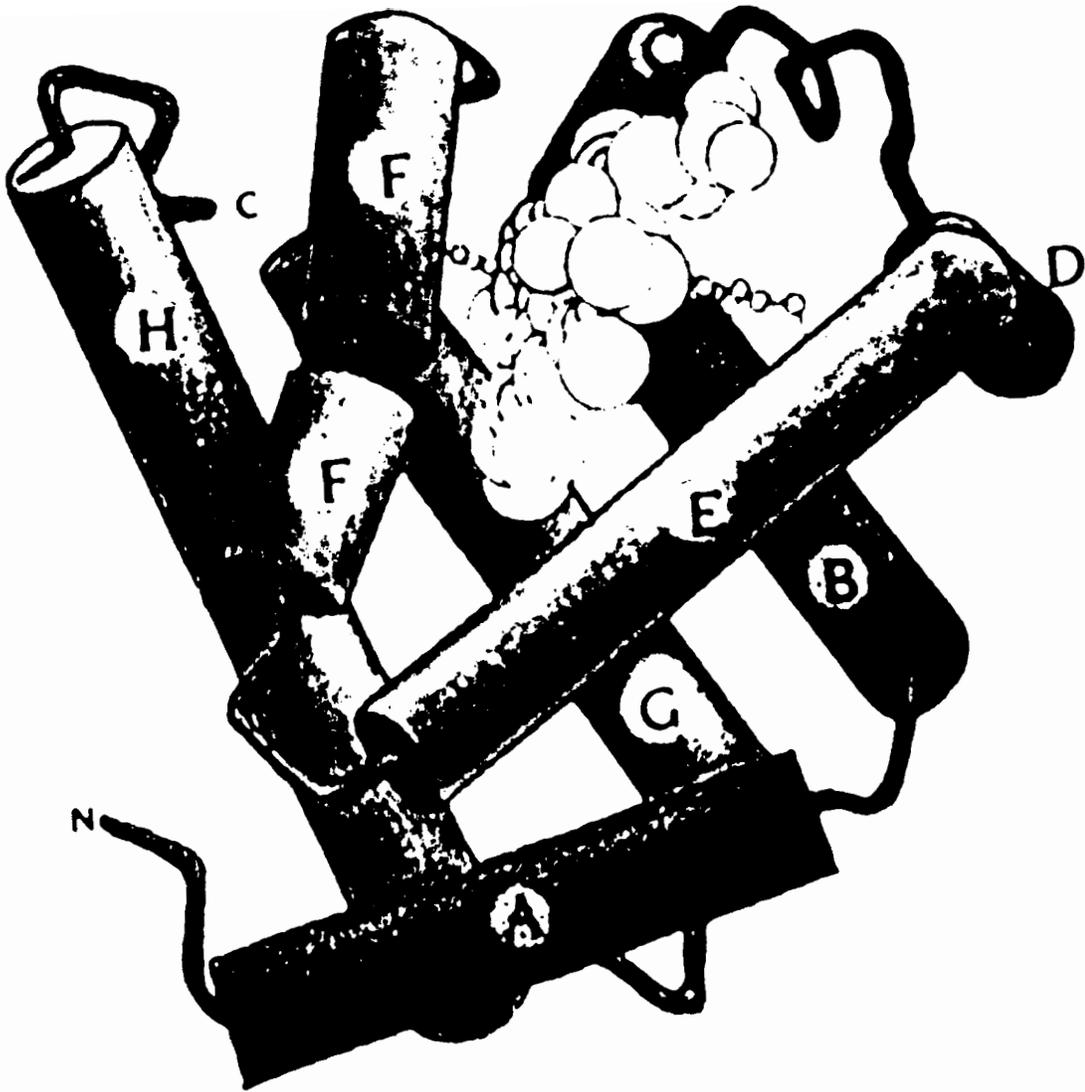
<b>Property</b>	<b>Class</b>	<b>Characteristic</b>	<b>Examples</b>
<i>Secondary structure content</i>	1. $\alpha$ -helical	secondary structure almost exclusively $\alpha$ -helical	myoglobin, cytochrome c, citrate synthase
	2. $\beta$ -sheet	secondary structure almost exclusively $\beta$ -sheet	chymotrypsin, immunoglobulin domain
	3. $\alpha + \beta$ and $\alpha/\beta$	secondary structure contains both $\alpha$ -helix and $\beta$ -sheet	papain, alcohol dehydrogenase, triose phosphate isomerase
<i>Tertiary structure</i>	2.1 parallel $\beta$ -sheet	double $\beta$ -sheet sandwich, strands roughly parallel	immunoglobulin domain
	2.2 orthogonal $\beta$ -sheet	double $\beta$ -sheet sandwich, strands of different sheets roughly perpendicular	chymotrypsin domains
	2.3 $\beta$ -sheet (other)	other types of spatial relationships between sheets	neuraminidase, interleukin-1 $\beta$
	3.1 $\alpha + \beta$	$\alpha$ -helices and strands of $\beta$ -sheet separated in different parts of molecule. Absence of $\beta$ - $\alpha$ - $\beta$ supersecondary structure	papain, staphylococcal nuclease
	3.2 $\alpha/\beta$	helices and sheet assembled from $\beta$ - $\alpha$ - $\beta$ units—strands of sheet parallel	alcohol dehydrogenase, triose phosphate isomerase
	3.2.1 $\alpha/\beta$ -linear	line through centres of strands of sheet roughly linear	alcohol dehydrogenase
	3.2.2 $\alpha/\beta$ -closed	line through centres of strands of sheet roughly circular	triose phosphate isomerase



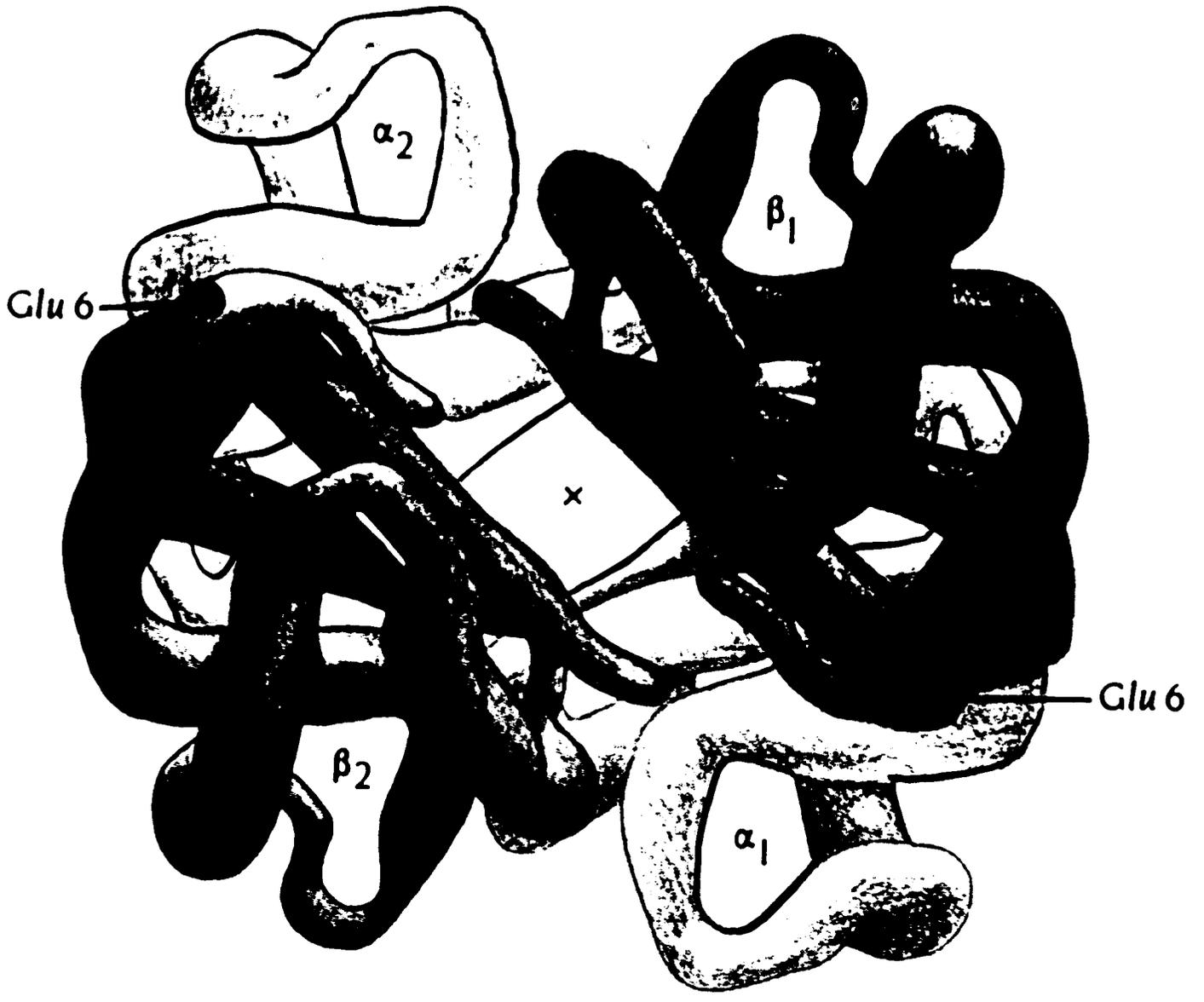
**Figure 37.** Diagram of two alternative arrangements for helix-pair dimers, with dotted lines to show the type of connection needed to join them into a single covalent chain. (a) The dimers are related by a lengthwise twofold axis, and they can be joined by a short connection into an up-and-down helix bundle (like those in Fig. 36). (b) The dimers are related by a crosswise twofold axis (as is the case for the ROP dimer), and they can only be joined by a long crossover connection of the type seen in ferretin and growth hormone.



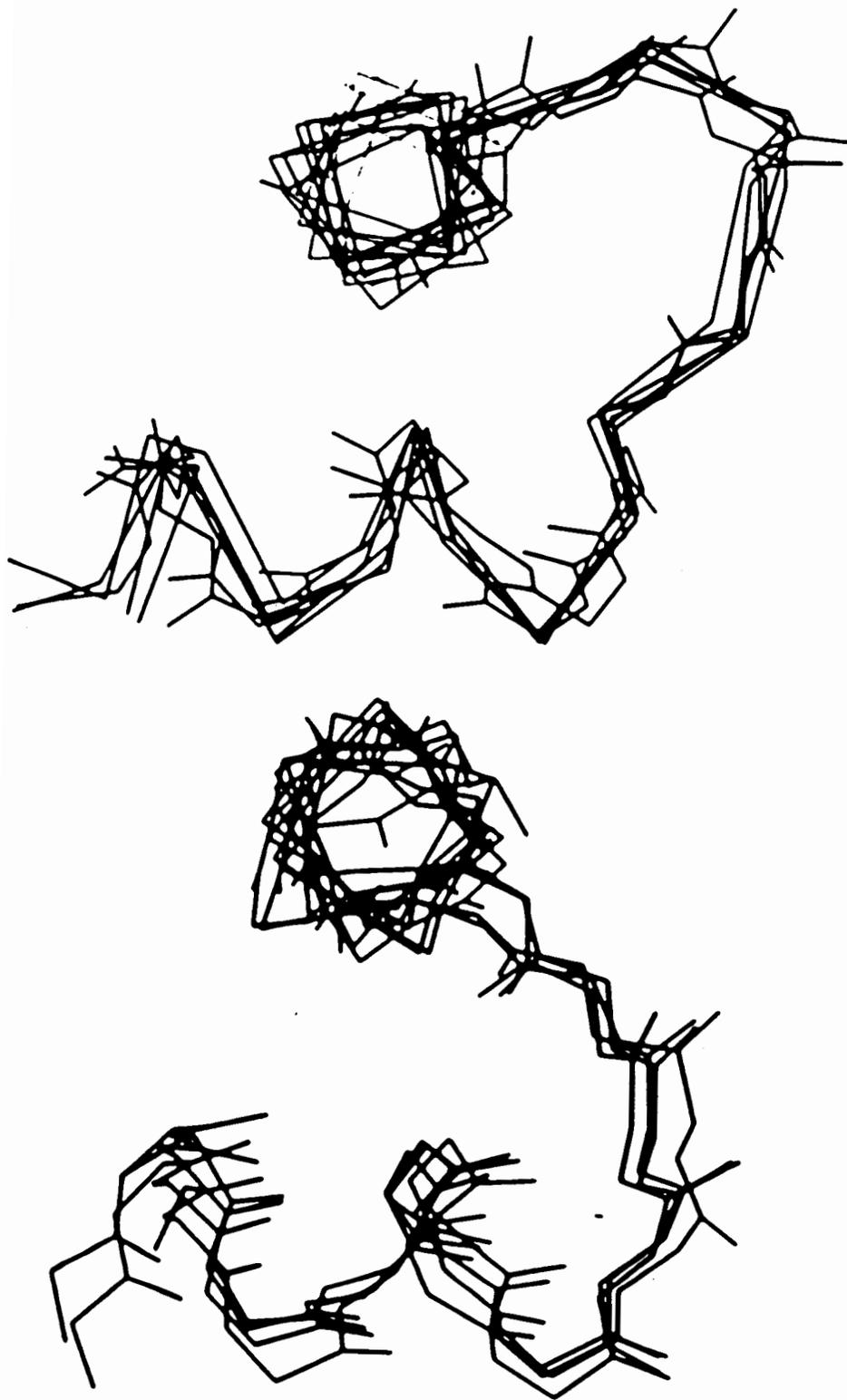
**Figure 39.** Diagrams of the two directions in which an up-and-down helix bundle can pack: (a) the more common version turns right at the top of the first helix, as in the proteins of Fig. 36; (b) the alternative version turns left at the top of the first helix, as seen in cytochrome  $b_5$  and ROP protein.



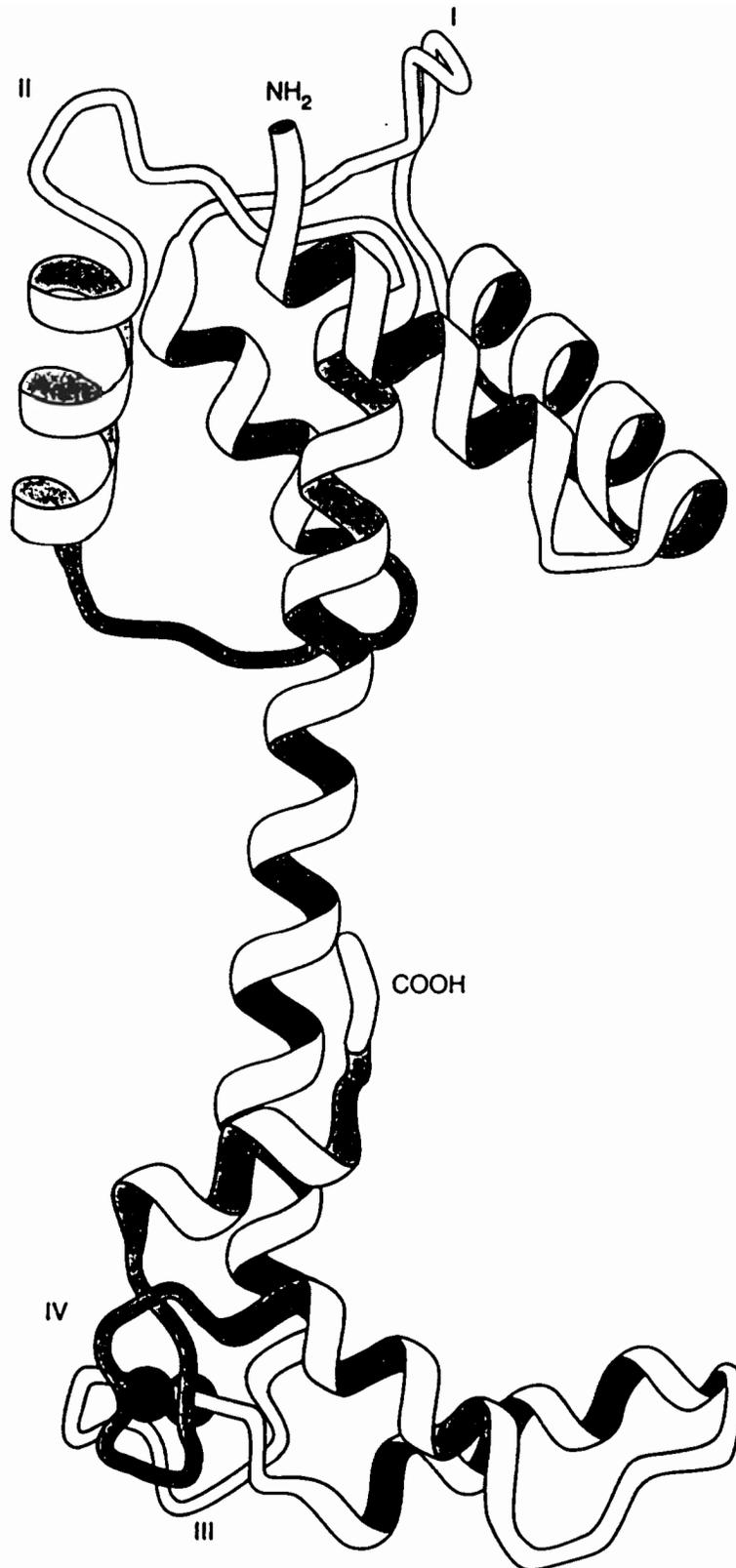
**Figure 3.4** Schematic drawing of the globin domain. The eight  $\alpha$  helices are labeled A to H. A–D are red, E and F green, and G and H blue. (Adapted from originals provided by A. Lesk.)





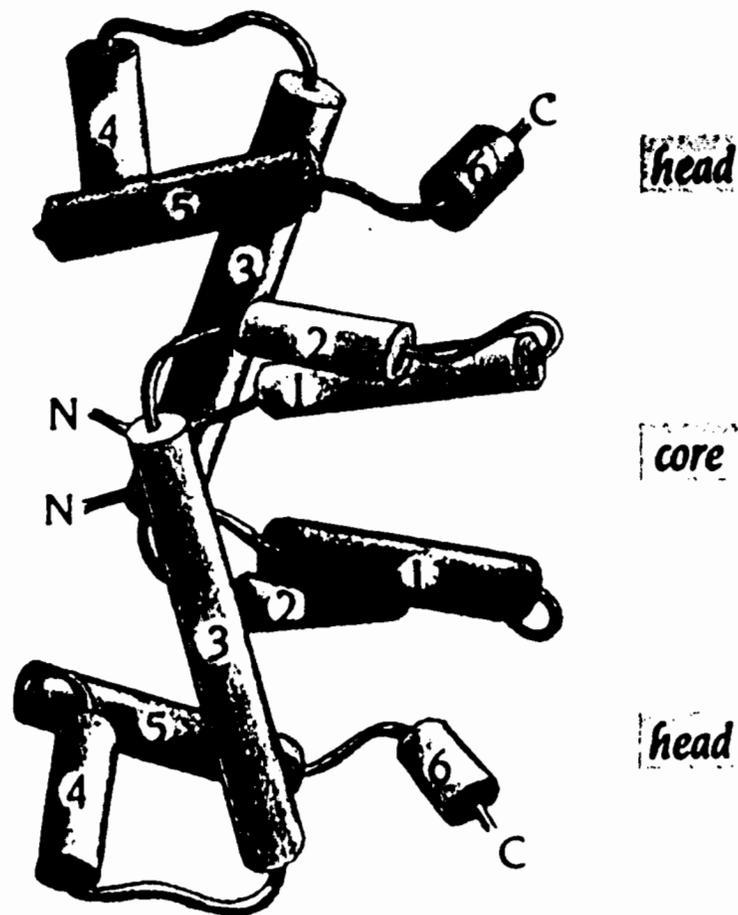


(a) six superimposed examples of Ca-binding

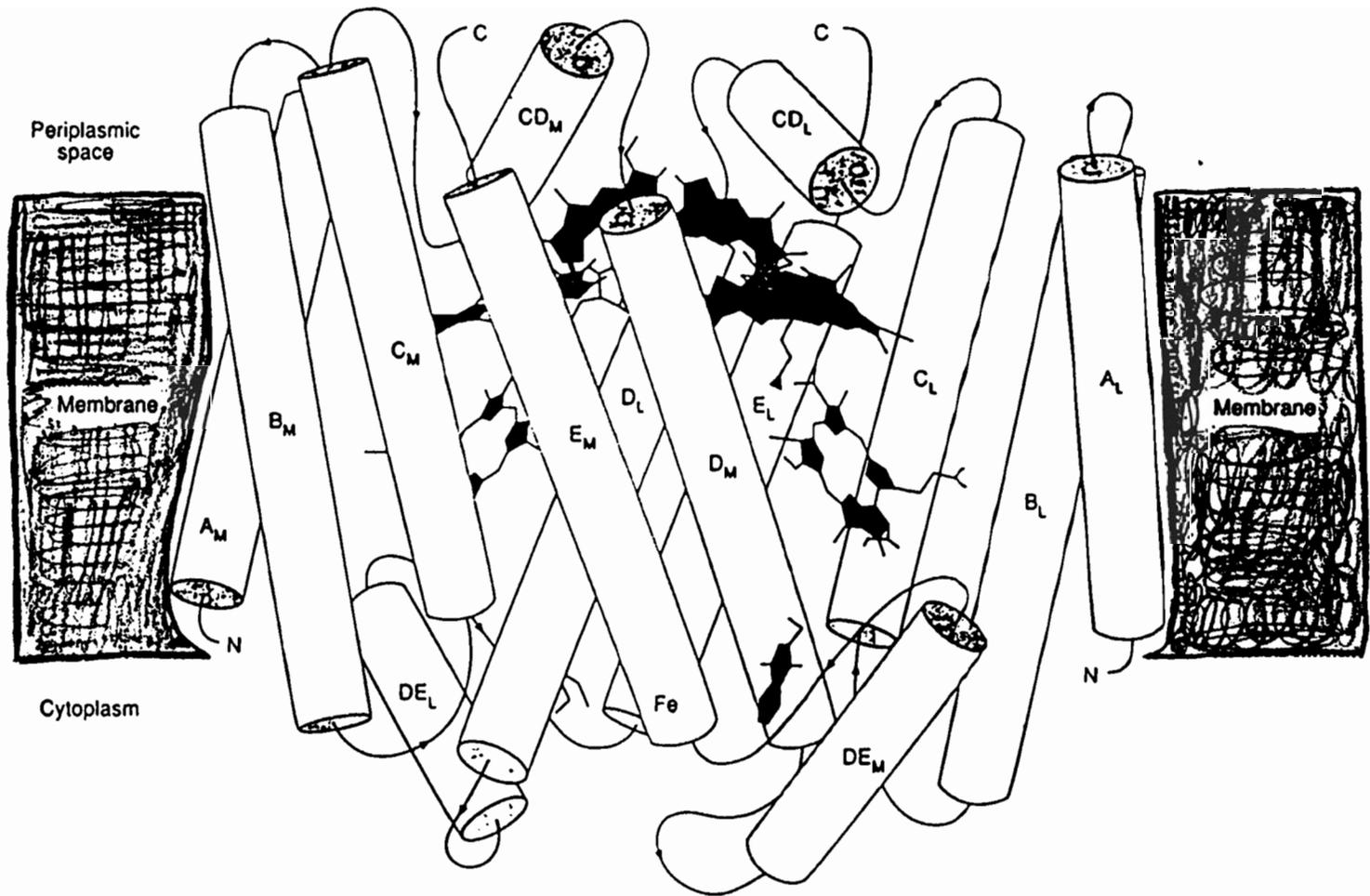


**Figure 4.1.** The structure of troponin C with  $\text{Ca}^{2+}$  bound at two of its four EF hands. Sites II and IV have been highlighted in orange. Note the differences in their conformation due to the presence of  $\text{Ca}^{2+}$  at sites III and IV and its absence at sites I and II. Reproduced with permission from Herzberg, O. and James, M.N.G. (1988) *J. Mol. Biol.*, 203, 761.

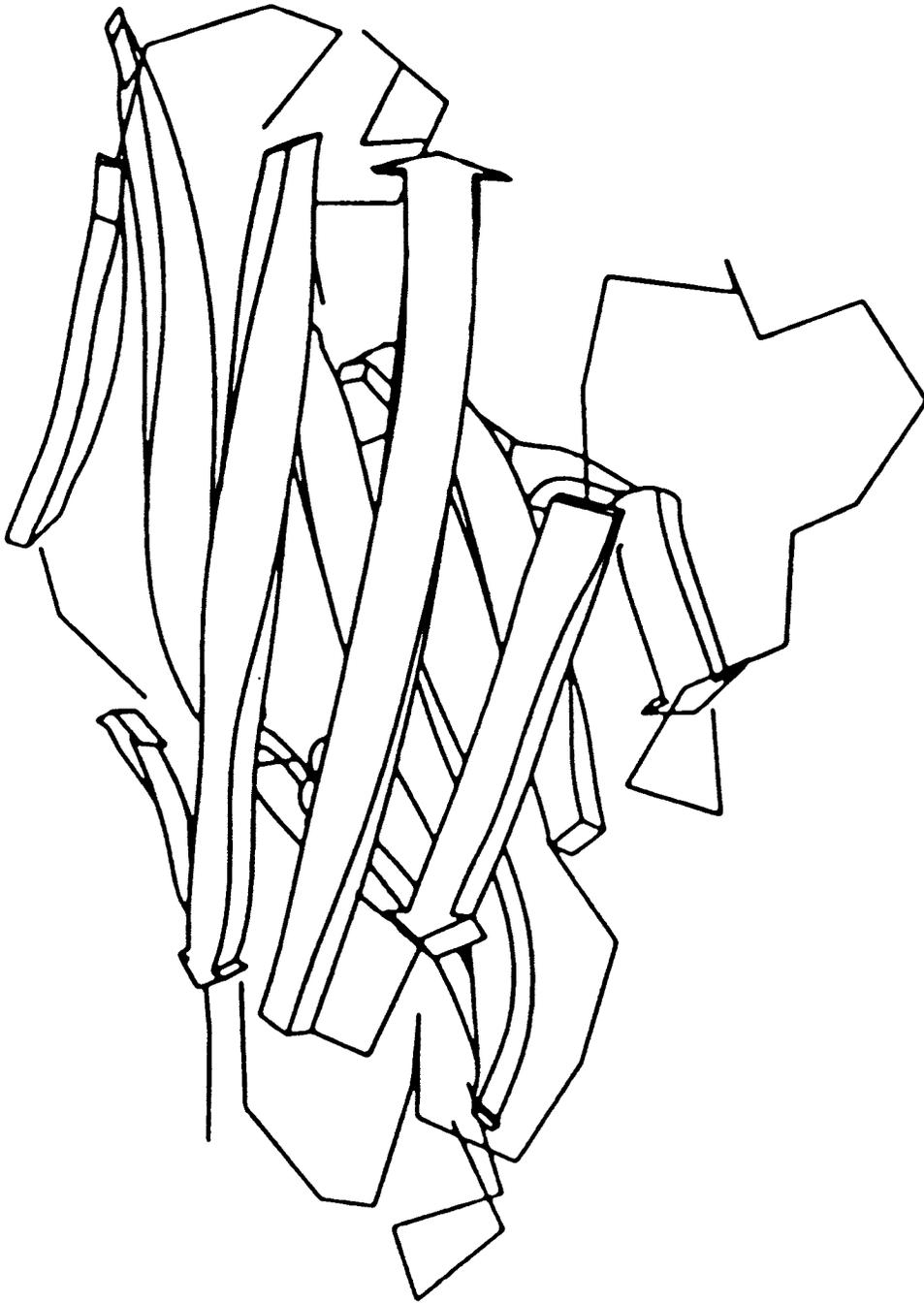




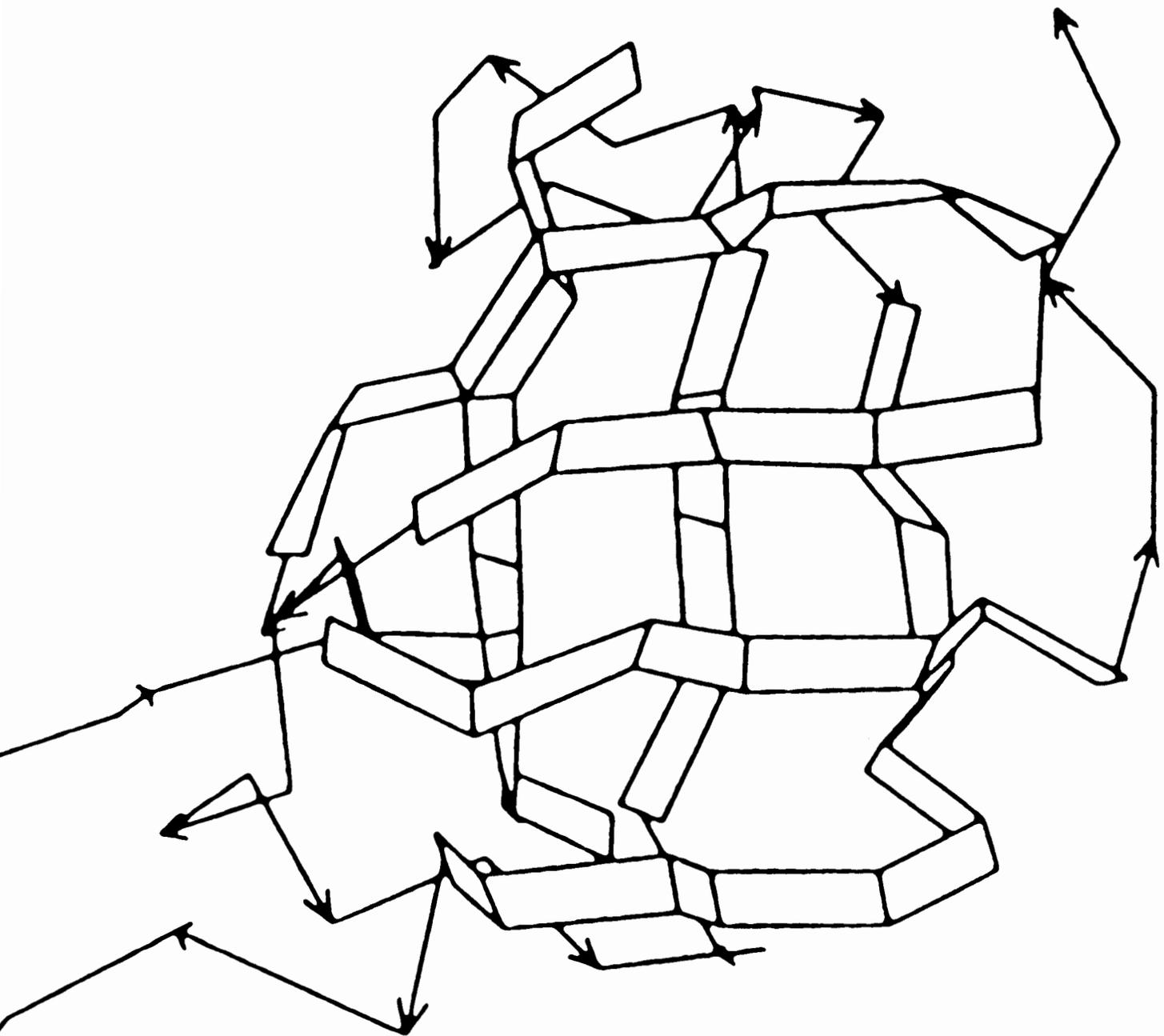
**Figure 7.23** The  $\alpha$  helices of the N-terminal region of the *trp* repressor are involved in subunit interactions and form a stable core in the middle of the dimer.  $\alpha$  helices 4–6, which include the helix-turn-helix motif, form two “head” regions at the two ends of the molecule.  $\alpha$  helix 3 connects the core to the head in both subunits. (Adapted from Schevitz et al., *Nature* 317: 782, 1985.)



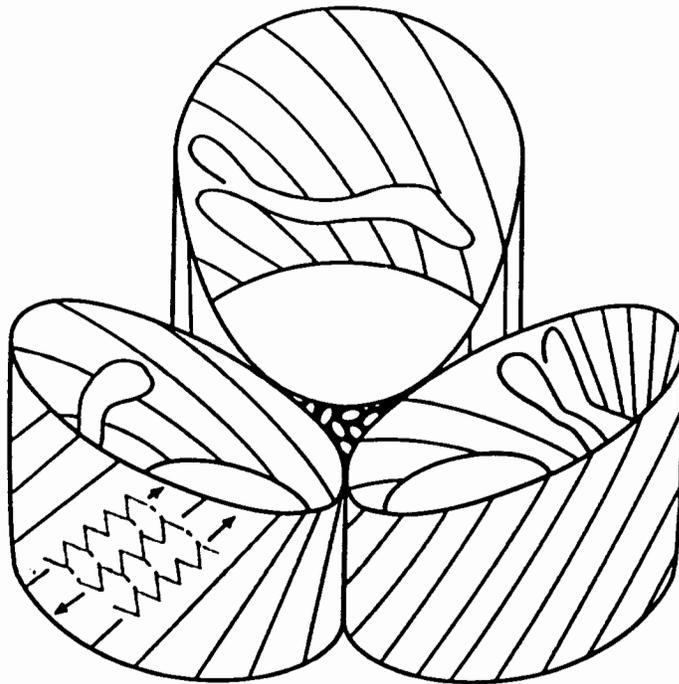
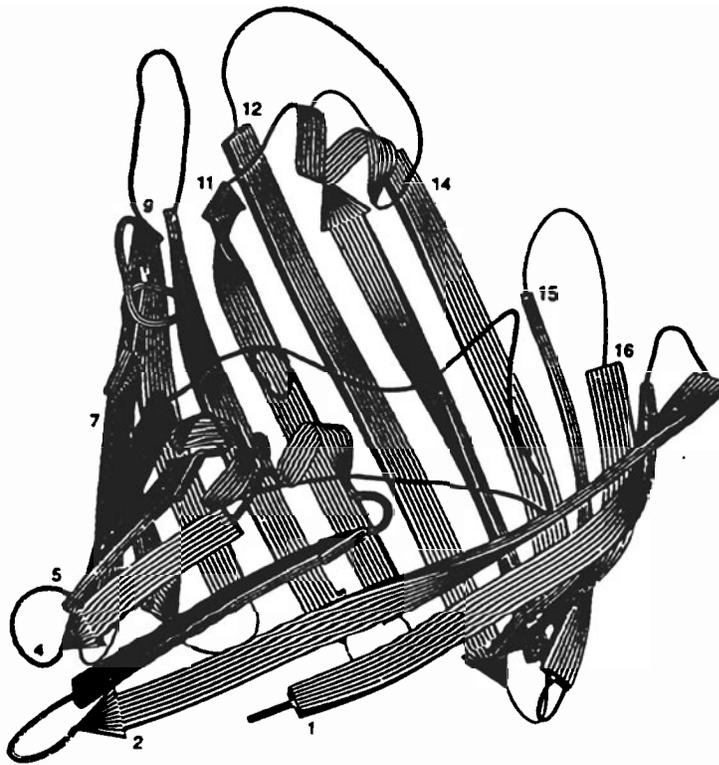
**Figure 3.1** Schematic diagram of the L (grey) and M (pale orange) subunits of the photosynthetic reaction centre; the major  $\alpha$ -helices are labelled A to E for each subunit; minor helices are labelled by the major helices they connect. Note the 2-fold pseudo-symmetry axis; the L and M subunits probably arose from a common ancestral protein that functioned as a dimer. The H subunit of the protein (not shown) is mainly located on the cytoplasmic face of the membrane, but also has one transmembrane helix. A cytochrome subunit binds to the periplasmic side of the membrane by protein-protein interactions. Co-factors and prosthetic groups are shown in orange. Reproduced with permission from Michel, H. and Dessenhofer, J. (1988) *Biochemistry*, 27, 1.



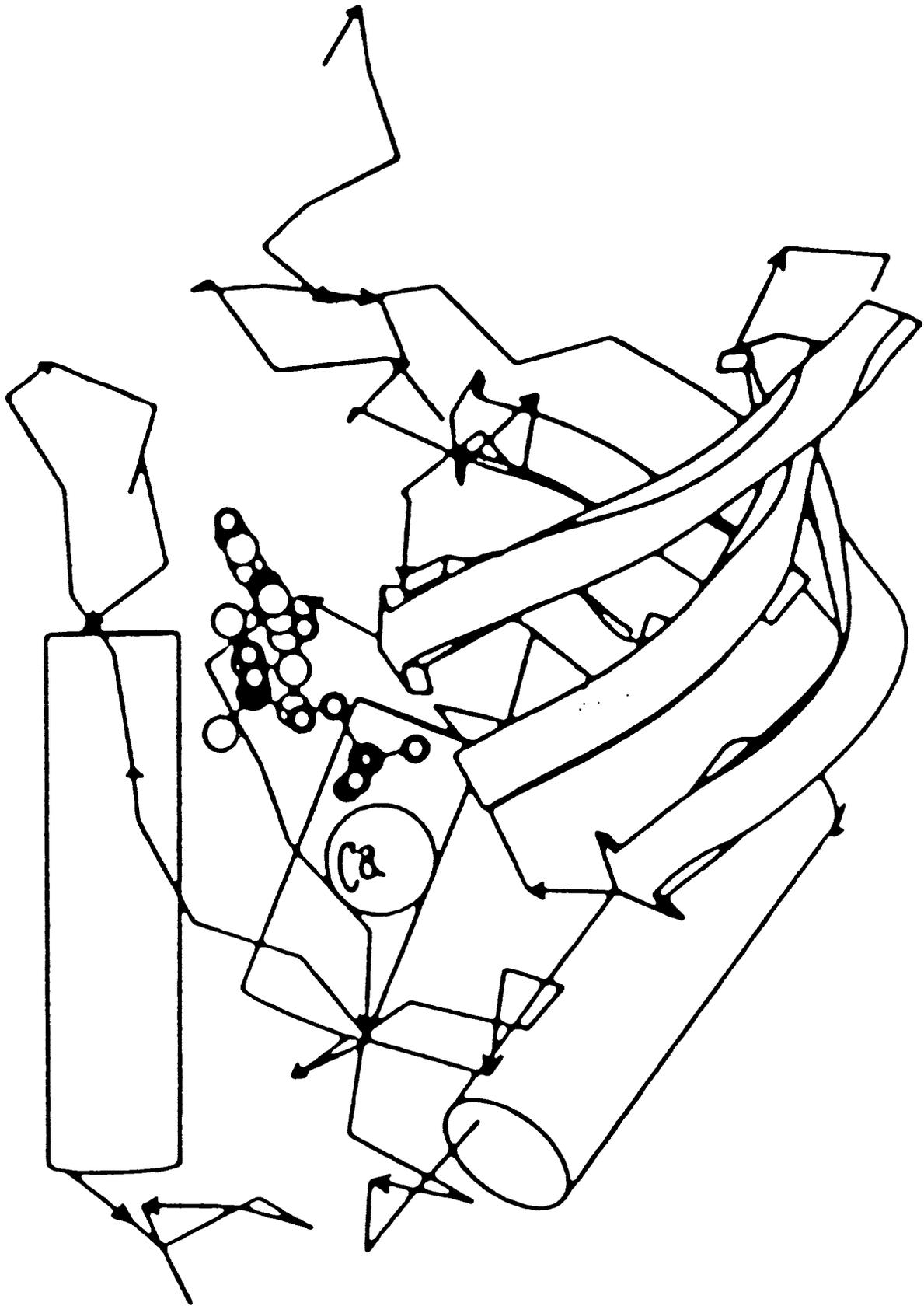
immunoglobulin domain (V<sub>H</sub> KOL)



N-terminal domain of  $\gamma$ -chymotrypsin



**Figure 3.2.** (a) Schematic view of one subunit of porin, which consists of 16 anti-parallel  $\beta$ -strands, each 6–17 residues long, linked together by short lengths of  $\alpha$ -helix or by loops to form a barrel. One loop of 44 residues (shown in orange) linking an  $\alpha$ -helix after strand 5 to strand 6 protrudes into the barrel and restricts the channel. The bottom rim of the barrel faces the periplasm and is relatively flat, while the top rim with its longer and more irregular connections is slanted and also less regular. Reproduced, with permission from ref. 14. (b) Schematic representation of a porin trimer, with the channel-restricting loop shown in orange. Reproduced with permission from Schirmer, T. and Rosenbusch, J.P. (1991) *Curr. Opin Struct. Biol.*, 1, 539.



Staphylococcal nuclease

