B. Antiparallel α Domains

The first major grouping of structures contains domains that are essentially all α-helical. Since there is relatively little other structure besides the helices, the simplest ways of connecting them involve predominantly antiparallel helix interactions, and that is in fact what is observed for these proteins. This category corresponds to Levitt and Chothia’s all-α category, but it has more members both because of a number of new structures and because of helical domains in proteins they classified as α + β (such as thermolysin). [This category is again now called “All-α”, both to match the “All-β” which now has parallel β folds (see below), and also because of the increased role of perpendicular helices (here seen only for the E-F hands of carp Ca-binding protein in Fig. 74) and of multi-chain parallel coils (here seen only for the flu haemagglutinin in Fig. 72).]

Figures 72 through 74 show schematic diagrams of the antiparallel α domains, grouped into subcategories. Almost all of them are two layer structures. The simplest and commonest subgroup looks like a bundle of sticks: usually four helices bundled in a cylinder with simple +1 connections. Most of the helices are quite close to exactly antiparallel, with typically a left-handed superhelical twist of less than 15° relative to the common axis of the bundle. These structures were first described as a group in Argos et al. (1977). Figure 87 illustrates myohemerythrin as an example of this structure type, showing an α-carbon stereo, a schematic drawing, and a topology diagram.

The simple up-and-down helix bundle structures include the hemerythrins (myohemerythrin and the hemerythrin subunits), cytochrome b562, cytochrome c’ uterusoglobin, tobacco mosaic virus protein, staphylococcal protein A fragment, and probably the ferritin subunits [yes]. Tyrosine-tRNA synthetase domain 2 has quite a similar organization, but the last helix tilts away from the bundle (Blow et al., 1977). The uterusoglobin subunit also has its fourth helix out to one side, but in the dimer molecule (Fig. 88) those final helices each complete a compact four-helix bundle with the rest of the opposite subunit. In cytochrome c’ there is a similar but less extreme arrangement in which the first helix lies at a greater angle to the bundle axis and forms the tightest part of the dimer contact. [The uterusoglobin arrangement

![Fig. 87. Myohemerythrin as an example of an up-and-down helix bundle. (a) α-Carbon stereo; (b) schematic drawing of the backbone structure, from the same viewpoint as in a.](image)
is an early example of what is now called a domain-swapped dimer (Eisenberg)

Tobacco mosaic virus protein has a small, highly twisted antiparallel β sheet at the base of the helix bundle, with two more helices underneath the sheet (see Fig. 72). Cytochrome $b_6$ looks remarkably similar (see Fig. 105), but the helices are much shorter. That structure could have been classified as an up-and-down helix bundle, but we have placed it in the small metal-rich proteins because its helix bundle is very small and distorted and the heme interactions appear more important than the direct helix contacts.

All but one of the above structures have four helices in the bundle, with $+1,+1,+1$ connections. For the up and down topology on a cylinder, handedness can be defined by whether the chain turns to the right or to the left at the end of the first structure element (whether it is a helix or a β strand). With an even number of helices, reversing N to C direction of the chain also reverses handedness of the topology; for an odd number of helices or strands handedness is invariant to chain reversal. For $+1,+1,+1$ topologies in general, handedness is not a very robust criterion of similarity, since it reverses on addition or deletion of one of the structure elements at the N-terminus but not at the C-terminus, so that a given five-helix structure could have evolved from either handedness of four-helix structure. Hemerythrin, cytochrome $b_5$, cytochrome c', uteroglobin, and tobacco mosaic virus protein are all right-handed, while cytochrome $b_5$, tyrosine-tRNA synthetase, and staphylococcal protein A fragment are left-handed. [The up & down 4-helix bundle is a very common superfold. But more complex topologies with long connections occur.]

The connectivity is not known for the seven-helix bundle of purple membrane protein (Henderson and Unwin, 1975), but on the basis of its resemblance to other antiparallel α proteins the most likely topologies would be either up-and-down or Greek key (see below). An analysis based on the sequence and the relative electron-densities of the helices (Engelman et al., 1980) considers a left-handed up-and-down topology as the most probable model. [The up & down 7 helix bundle structure was confirmed by higher resolution electron diffraction and x-ray structures, both for purple membrane protein (bacteriorhodopsin) and for rhodopsin.]

Many of the up-and-down helix bundle proteins form large multisubunit arrays. Hemerythrin is an octomer, with the end of one helix bundle butting against the side of the next one around the 4-fold axis (Ward et al., 1975). The 24 ferritin subunits form a hollow spherical shell with the helix bundles approximately tangential to the shell and the subunit interactions around the 3-fold and 4-fold axes rather like the interactions in hemerythrin. Tobacco mosaic virus protein, on the other hand, forms a tightly packed long helix of subunits; the α-helical bundles are aligned radially, with RNA bound at their inner ends. Purple membrane protein spans the membrane, forming a two dimensional crystal-line array with the helix bundles perpendicular to the membrane and parallel to each other around the 3-fold axis.

One of the most important and interesting antiparallel α structures is the globin fold, which has been found in the three-dimensional structures of a large group of related proteins including myoglobin and the hemoglobins of various mammals, glycera, lamprey, insect and even legume root nodules. The globin fold is a good example of how there may be several alternative useful ways of describing a given structure. To someone studying hemoglobin function the relevant level of description includes all the structural detail that can be made comprehensible, or perhaps generalized to include what is common to all the globin structures. On the other hand, if one is concerned, as we are here, with obtaining a memorably simple description of the whole structure and relating it to other protein structures, then the issue is deciding which features are most important to include in the simplification and with which if any other proteins it can meaningfully be compared. Classifying the globins as all-α proteins is obviously true and useful, but Levitt and Chothia’s (1976) scheme of representing the globin topology does not suggest similarities to any of the other all-α proteins, even when the more recent structures are included. Argos and Rossmann (1979) have suggested an interesting similarity of structure around the heme pocket for the globins, cytochrome $b_5$, and cytochrome $c_551$. Their description is probably the most relevant one for trying to understand how heme-binding pockets are organized, but it does not seem suitable as a general structural description since the omitted halves of the three structures are all extremely.
different and do not form separable domains.

Figure 89 illustrates two different tries at simplified representation of the globin structure. For reference, Fig. 89a shows the hemoglobin β chain in stereo. Figure 89b shows the globin structure schematically as two layers of helices with the elements in one layer approximately perpendicular to those in the other layer; this can be contrasted with a possible description of the up-and-down helix bundles as two layers with their elements approximately parallel to each other. The perpendicular layers provide a rather successful simple schema for the globin structure, but unfortunately there are no other proteins that can be adequately described as two perpendicular layers of helices. Also, specification of the topology in this scheme is cumbersome, since the chain skips back and forth between layers.

Figure 89c schematizes the globin structure as a twisted cylinder of helices, analogous to the antiparallel β barrels to be discussed in Section III,D. The up-and-down helix bundle structures are of course also readily described as cylinders, so that this schema makes the majority of antiparallel α structures directly analogous to the majority of antiparallel β structures. Their topologies can be conveniently specified by the simple nomenclature listing connection types (see Section II,B). The major irregularity of the globin fold when considered as a cylinder is that one element (the A and B helices) bends sharply to close the cylinder; this feature is also seen in five- and six-stranded β barrels such as trypsin. But perhaps the most satisfying
feature of schematizing the globin fold on a cylinder is that it can then be grouped with other structures (thermolysin d2, T4 phage lysozyme d2, papain d1, and cytochrome c peroxidase d2) which also show the “Greek key” (see Richardson, 1977) topology of + 3,- 1,- 1. Papain domain 1 also shows the diagnostic feature of Greek key structures by containing a non-nearest-neighbor connection which skips across the end of the cylinder, however, most of its helices are short and they form a rather irregular bundle. Papain domain 1 contains two disulfides; we will find repeatedly that increasing disulfide content goes along with decreasing regularity of both secondary and tertiary structure.

These four structures then form the second major subgrouping of antiparallel α domains, which we will call Greek key helix bundles (see Fig. 73). The helix elements lie on an approximate cylinder (see Fig. 89d for an end view), with 0 to 45˚ right-handed twist relative to the cylinder axis; they are connected with a Greek key topology which can have either a counterclockwise (globins) or a clockwise (thermolysin d2 and T4 lysozyme d2) swirl when viewed from the outside.

The remaining structures in this category (carp Ca-binding protein, egg lysozyme, citrate synthase, catalase d2, and ρ-hydroxybenzoate hydroxylase d3) are miscellaneous helical domains. However, there is good evidence from sequences and from functional resemblances (Kretsinger, 1976) that carp Ca-binding protein exemplifies a whole group of proteins that are constructed of “E-F hands” (see Section II,F) and that regulate or are regulated by changes in Ca\(^{2+}\) concentration. [This is now a large group typified by calmodulin. The “T-shaped” helix-turn-helix motifs of many transcription factors are superimposable onto E-F hands in the reverse sequence direction (Richardson).] Citrate synthetase may be the first example of a group of larger helical domains with three layers. [The peroxidases (e.g. 2CYP) are examples, but the layers are usually not distinct.] Irregular helical structures with a moderate number of disulfides can be classified either here or as small SS-rich. We have classified egg lysozyme here (with only 4 disulfides in 129 residues), while phospholipase A2 (with 7 disulfides in 123 residues) is classified with the small SS-rich proteins.

[The most distinct new all-α folds are the cylindrical or crescent-shaped repeats of helix hairpins, both of which are exemplified by the two chains of farnesyl transferase.]