

E. Small Disulfide-Rich or Metal-Rich Domains

The last major category (shown in Figs. 85 and 86) consists of small (usually less than 100 residues) domains whose structures seem to be strongly influenced by their high content either of disulfide bonds (S) or of metal ligands (M). These S–M proteins often look like distorted versions of other, more regular, proteins. The disulfide-rich ones include many toxin and enzyme-inhibitor structures. For most of the disulfide-rich proteins it is known experimentally that they are completely unstable if the disulfides are broken (in contrast to larger disulfide-containing proteins, for which disulfides merely provide additional stabilization for an already-determined structure). Figure 103 shows pancreatic trypsin inhibitor as an example of a disulfide rich protein, and Fig. 104 shows cytochrome c as an example of a metal-rich protein. Most of the S–M proteins are single-domain and monomeric: only wheat germ agglutinin has multiple domains, and only insulin has multiple subunits in the molecule.

The only subgroup of similar structures within the S–M proteins is the toxin-agglutinin folds of the snake neurotoxins and the domains of wheat germ agglutinin (see Fig. 85). They are made up of extended-chain loops with an almost identical topology of -1, +3, -1, +2x rather like a series of half-hitch knots (the β structure is extremely minimal in wheat germ agglutinin) strongly linked by a core of four disulfides, three of which are equivalent (see Drenth et al., 1980). High-potential iron protein and ferredoxin share a local loop structure that binds the iron-sulfur cluster, but otherwise are different.

Most of the metal-rich proteins form approximately cylindrical two-layer structures with either an up and down (rubredoxin, cytochrome c) or a Greek key (ferredoxin) topology, but in which the elements forming the cylinder are a mixture of helices, β strands, and more or less extended portions of the backbone. Cytochrome c_3 is perhaps the ultimate example of an S–M protein, with four hemes in just over a hundred residues, and essentially no secondary structure at all except for one helix.

One way of considering these proteins is as distorted versions of the other structural types. Most S–M proteins can fairly clearly be grouped as either distorted helix clusters (phospholipase, cytochrome c, cytochrome b_5), distorted β barrels (rubredoxin, high-potential iron protein), or distorted open-face sandwiches (erabutoxin, wheat germ agglutinin, pancreatic trypsin inhibitor, or ferredoxin). Figure 105 shows an example of each of these relationships. In fact, one reasonable taxonomy would do away with this fourth major category altogether and place all the S–M proteins as irregular examples of either an α or a β category. We have not chosen that approach, however, because several of the structures (crambin, insulin, and cytochrome c_3) are rather difficult to place in one of the other categories, and also because these small proteins influenced by nonpolypeptide interactions appear to share important features, especially in terms of the probable complexity of their folding process (see Section IV,C).

Another suggestive fact is that there are no small, irregular structures related to the parallel α/β category. Per-

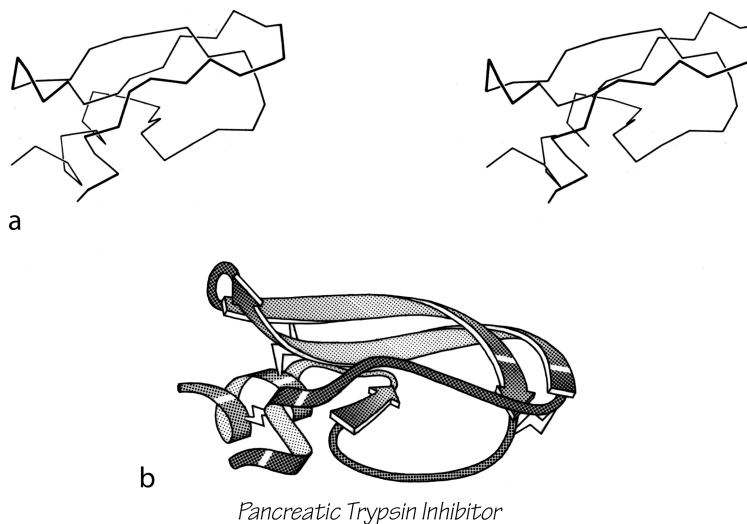


Fig. 103. Basic pancreatic trypsin inhibitor as an example of a small disulfide-rich structure. (a) α -Carbon stereo; (b) backbone schematic, viewed as in a. with disulfides shown as zig-zags. Figure 2 shows an all-atom stereo of this protein with side chains.

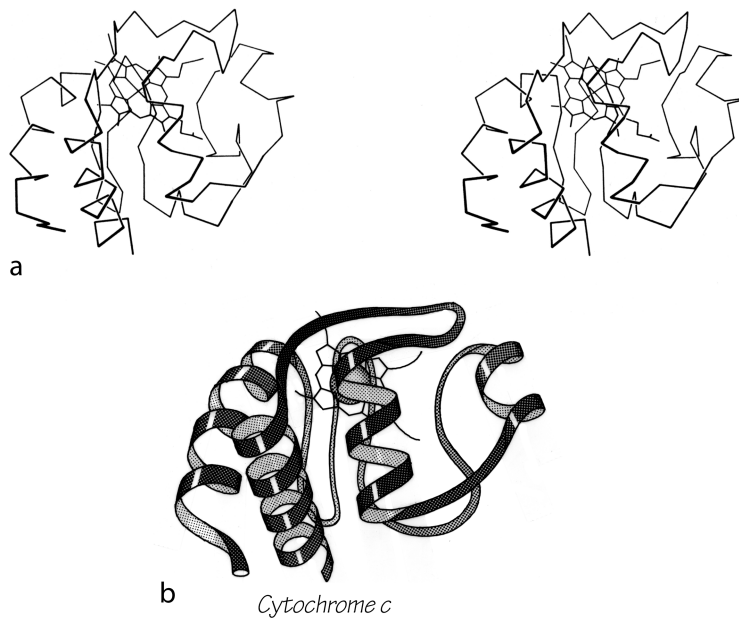


FIG. 104. Cytochrome *c* as an example of a small metal-rich protein. (a) α -Carbon stereo, with heme; (b) backbone schematic, viewed as in a. The backbone forms an approximate up and down cylinder with the heme tucked into the center, but the elements forming the cylinder are a mixture of helices and extended strands.

haps this reflects the fact that domains organized around parallel β sheet are necessarily fairly large and seem to be dependent on large, buried, and quite regular β structure for their stability. There are in fact no hemes (in spite of all the helices) or iron-sulfur clusters in parallel α/β proteins, and no disulfides except for the single active-site disulfides of thioredoxin, glutathione peroxidase, and glutathione reductase.

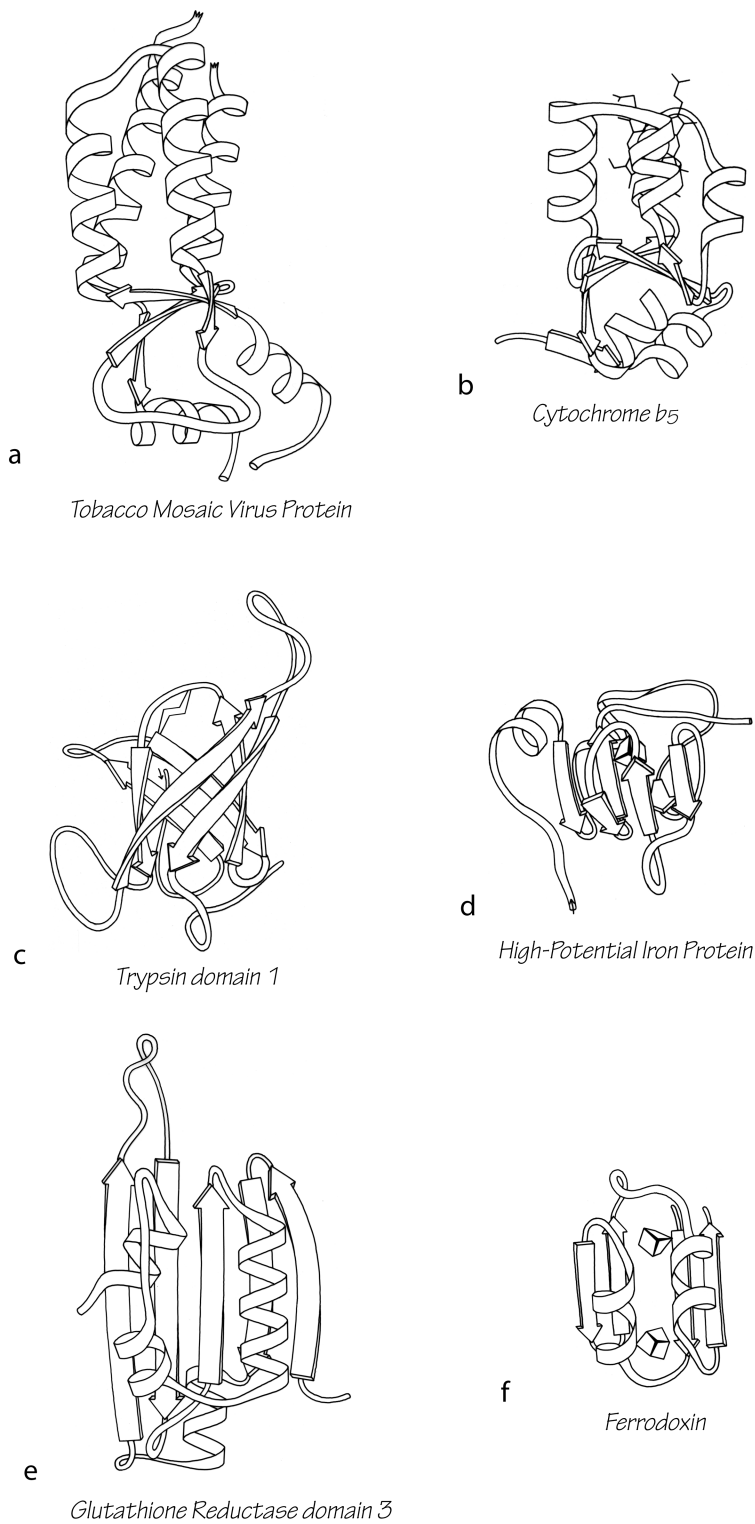


FIG. 105. Examples of small disulfide-rich or metal-rich proteins (shown on the right side) compared with their more regular counterparts in other structural categories (shown at the left). (a) Tobacco mosaic virus protein, an up-and-down helix bundle; (b) cytochrome b₅, a distorted up-and-down helix bundle; (c) trypsin domain 1, a Greek key antiparallel β barrel; (d) high-potential iron protein, a distorted Greek key β barrel; (e) glutathione reductase domain 3, an open-face sandwich β sheet; (f) ferredoxin, a distorted open-face sandwich β sheet.