AN ILLUSTRATED MUSEUM OF PROTEIN STRUCTURES

Jane S. Richardson, John A. Tainer, and David C. Richardson, *208 Medical Sciences IA, Duke University, Durham, North Carolina 27710 U.S.A.*

This poster presents schematic drawings of most of the distinctly different domains that occur in the known protein structures, all using the same scale, the same conventions for representing secondary structure, and the same viewing direction within each group of similar structures (see Fig. 1). The drawings are organized according to the classification scheme presented in reference 1 and briefly outlined below.

I. ANTIPARALLEL α
- 1. 2-layer helix bundles
  - A. Up and down helix bundles (e.g., hemerythrin)
  - B. Greek key helix bundles (e.g., hemoglobin)
- 2. 3-layer helix bundles (e.g., citrate synthetase)
- 3. Miscellaneous antiparallel α (e.g., carp muscle Ca-binding protein)

II. PARALLEL α/β
- 1. 4-layer, singly-wound parallel β barrels (e.g., triose phosphate isomerase)
- 2. 3-layer parallel α/β
  - A. Doubly-wound parallel β sheets (e.g., lactate dehydrogenase domain 1)
  - B. Other 3-layer α/β (e.g., carboxypeptidase)
- 3. 5-layer doubly-wound α/β (e.g., phosphorylase domains)

III. ANTIPARALLEL β
- 1. 2-layer β barrels
  - A. Up and down β barrels (e.g., soybean trypsin inhibitor)
  - B. Greek key β barrels (e.g., immunoglobulin domains)
    - a. Jellyroll Greek key β barrels (e.g., tomato bushy stunt virus domains)
  - C. Other and partial antiparallel β barrels (e.g., pancreatic ribonuclease)
- 2. Miscellaneous antiparallel β (e.g., bacteriochlorophyll protein)

IV. SMALL IRREGULAR PROTEINS
- 1. Disulfide-rich small proteins
  - A. Toxin-agglutinin fold (e.g., erabutoxin)
  - B. Other disulfide-rich (e.g., pancreatic trypsin inhibitor)
- 2. Metal-rich small proteins
  - A. Up and down ligand cages (e.g., cytochrome *c*)
  - B. Greek key ligand cages (e.g., ferredoxin)
  - C. Other metal-rich (e.g., cytochrome b_5)

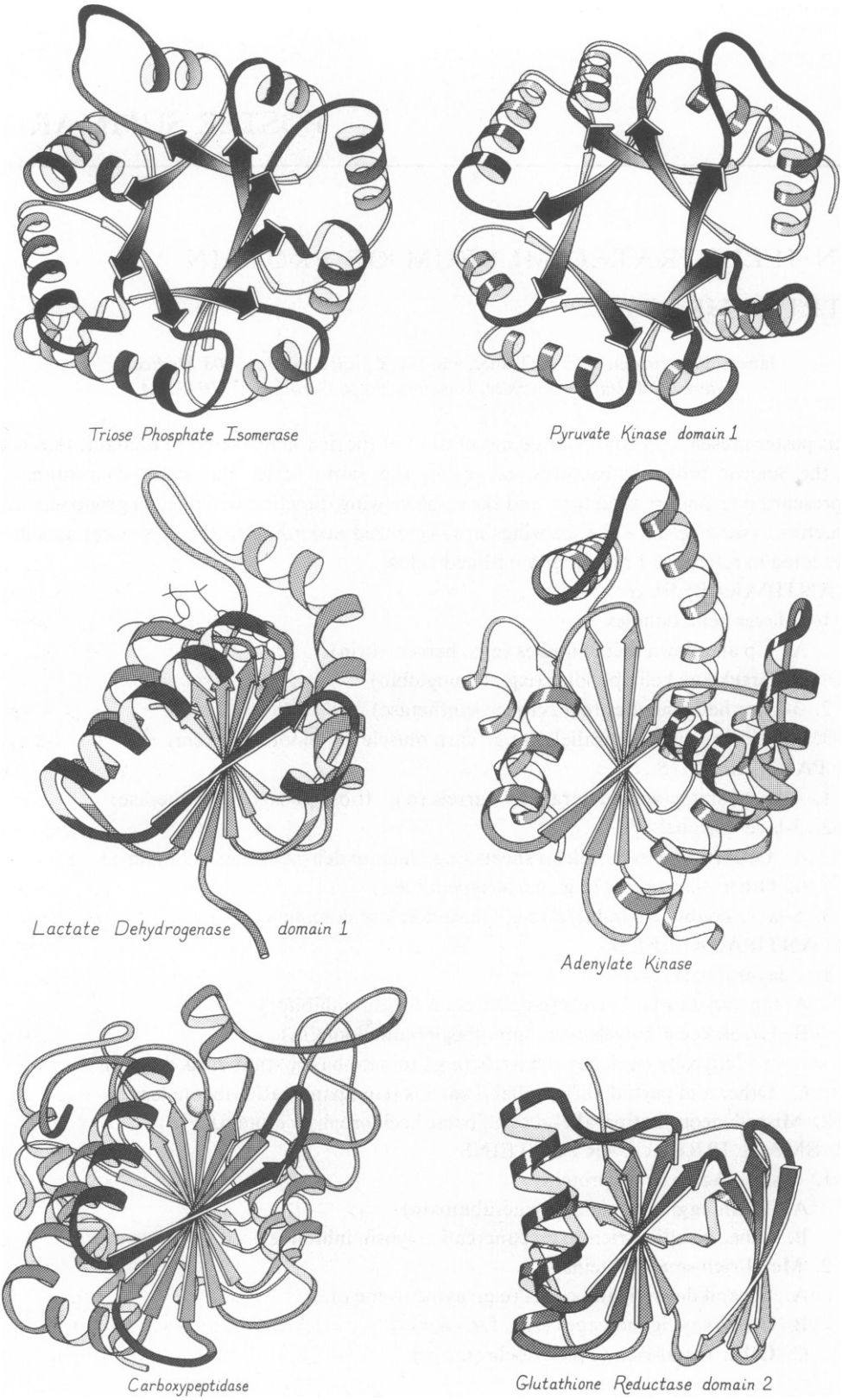

Figure 1 Schematic backbone drawings of a few of the parallel α/β protein domains. Top two structures: singly-wound parallel β barrels. Bottom four structures: 3-layer α/β domains with varying degrees of resemblance to the classic doubly-wound β sheet of lactate dehydrogenase.

Received for publication 17 December 1979.

REFERENCES

1. Richardson, J. S. 1979. The anatomy and taxonomy of protein structures. *Adv. Prot. Chem.* In press.

QUANTITATIVE ANALYSIS OF STRUCTURAL DOMAINS IN PROTEIN

M. N. Liebman, *The Institute for Cancer Research, Philadelphia, Pennsylvania 19111 U.S.A.*

X-ray crystallographic studies have provided the three-dimensional structures of more than 100 proteins, including both structurally and functionally related families of macromolecules. To attempt to understand the relationship between structure, function, evolution, and macromolecular recognition and specificity, it has been of interest to compare the structures of related proteins or polypeptide folding domains. The statistical, root-mean-square deviation has provided a semiquantitative measure of structural agreement after the superposition of the segments to be compared. A new method is reported which permits the quantitative separation and comparison of the contributions of secondary, tertiary, and quaternary structure without the requirement of direct superposition technique.

The observation of several polypeptide folding patterns (~40–100 residues in length) reveals both intrinsic functional properties (e.g., nucleotide binding domain), and apparent structural stability (e.g., β-barrels, immunoglobulin fold, hemerythrin fold). It is of interest to be able to compare these analogous features at all structural levels: primary, secondary (structural elements of the domain), tertiary (intra-domain organization), and quaternary (inter-domain and intermolecular packing). The representation of the protein structures by distance matrix methods (1, 2, 3) has already provided qualitative methods for indicating structural domains (4), intra- and inter-molecular symmetry (5), quantitative assignment of structural insertions and deletions in the polypeptide (5), and protein:protein interactions(5). Quantitative examination of idealized secondary and tertiary structural interactions have also used this method (5).

Distance representation involves the construction of a square matrix of n cells, where n is the number of amino acids in the protein. The elements of this matrix, (i–j), contain the distance between the i-th and j-th alpha carbon along the polypeptide. Selective contouring of this matrix reveals levels of structural organization by pattern recognition (1, 2, 3, 4, 5). Comparison between structures using this representation can be achieved without superpositions of the three-dimensional coordinates because the distance matrices are internally referenced and thus independent of molecular rotation or translation (5).

It has been recently shown (6) that the comparison of two protein structures by use of a root-mean-square statistic is highly dependent on the nature of the secondary structures within the proteins. This reflects the correlated nature of the polypeptide chain caused by the chemical linkage, and also is indicative of the difference between topological and topographical identities. Thus it is inadequate to describe the difference between sperm whale