W-AM-Sym I-1

STRUCTURAL DATABASES IN PROTEIN MODELLING AND DESIGN

T. Alwyn Jones, Department of Molecular Biology, BMC, Box 590, S-751 24 Uppsala, SWEDEN

Protein molecules adopt limited energetically preferred conformations. One approach to modelling is therefore to restrict, force, or persuade users to work with these preferred conformations. This can be implemented with databases for side chains and main chains and can help identify sequence dependent conformations. Cluster analysis of protein structures allow the use of a limited subset of fragments for most modelling purposes. These fragments contain little single residue sequence dependence on structure.

W-AM-Sym I-2

ZINC FINGERS

Jeremy M. Berg, Department of Chemistry, The Johns Hopkins University, Baltimore, MD 21218

In recent years a large family of gene regulatory proteins has been discovered that is characterized by the presence of one or more sequences of the form (Tyr,Phe)-X-Cys-$X_{2,4}$-Cys-X_3-Phe-X_5-Leu-X_2-His-$X_{3,4}$-His. Each of these sequences appears to bind a zinc ion via the invariant cysteine and histidine residues to form a small structural domain that has been termed a "zinc finger". A prediction for the three-dimensional structure of these domains was developed based on the discovery of recurring substructures in crystallographically characterized metalloproteins. The structure consists of a two stranded antiparallel beta sheet followed by an alpha helix. The three highly conserved hydrophobic residue pack together to form a hydrophobic core adjacent to the metal coordination unit. Recent experimental results have revealed that this prediction is essentially correct. The knowledge of the structure of single domains has allowed development of models for the structures for arrays of tandemly repeated domains which is the form active in site-specific DNA binding. These models make strong predictions about the mode of interaction between the zinc finger proteins and nucleic acids.

W-AM-Sym I-3

STRUCTURE OF ALPHA1-12, A DESIGNED SYNTHETIC PROTEIN MODEL.

David Eisenberg, Christopher P. Hill, Daniel H. Anderson & Morgan Wesson, Molecular Biology Institute, UCLA, Los Angeles, CA 90024 & William F. De Grado, Central Research & Development, E.I. du Pont de Nemours & Co., Wilmington, DE 19898. **X-ray diffraction reveals that the structure of a designed protein model is more complex than the design. The structure is formed from non-covalent self-association of a 12 residue fragment of a longer peptide designed to form an amphiphilic α-helix with a ridge of Leu residues along one helical face; by interdigitation of the leucines of four such helices, the design called for self-association of four helices into a structure of the four α-helical bundle class. In the actual structure, α-helical tetramers are present, but there are also hexamers with a hydrophobic core of 12 leucine residues. These results indicate that it may be easier to design an amino acid sequence that folds into a compact protein-like packing than to design a specific folding pattern.**

W-AM-Sym I-4

Abstract : Antibody combining sites : prediction and design

Dr. Anthony R. Rees*, Andrew C.R. Martin, David Webster, Janet C. Cheetham and Sally Roberts

* on leave at IGEN, Inc. Rockville, MD USA from Laboratory of Molecular Biophysics, University of Oxford, England)

Algorithms for predicting the three-dimensional structure of CDRs from sequence data alone are being developed by several laboratories. Some procedures are based on the so-called maximum overlap method (MOP) and have been used independently by a number of groups whose approaches differ in the way in which CDR loops are selected as starting points. In general, modeling by homology alone is not reliable enough to generate the correct conformation of all CDRs even when the procedures incorporate the "key residue" modifications of Chothia and Lesk. We have developed a general solution to the prediction problem which combines the resource of the complete protein structure database (the knowledge based component) with conformational search algorithms (the ab initio component) and which requires no arbitrary decisions to be made by the operator. Details of the methodology and of recent developments will be presented.

In addition, progress in the development of algorithms that enable the user to dock the epitope on the surface of an antigen to its antibody combining site by a computational procedure will be discusssed. This requires some knowledge of those antibody residues involved in antigen contact. The manner in which protein engineering can be used to supply this information will be described.