of the peptide backbone have been smoothed out to form a continuous space curve. It takes about 2 h of VAX equivalent time to do the smoothing and about 2 min on the S1-100 array processor. When compared to the carbon α-representation which is too simple and harsh or the normal peptide backbone representation which is too complex, the smoothed space curve representation has the airy qualities of the original Richardson diagrams. Using Jane Richardson's monograph as a guide and the Protein Data Bank from Brookhaven National Laboratory, 125 proteins and/or domains have been extracted. Each space curve is coloured to show helices and sheets and is oriented to bring the N-terminus and/or the C-terminus of the protein to the front. Using stereo pairs of images it is possible to begin to think about the late stages of the folding of proteins. In reality what one does is imagine the unfolding of the protein from the crystal structure, hence the name of the kit. There are some real surprises. You can really imagine the sequence of motions by means of which the protein folded. The space curve representation makes it easy to think about 'tugging' loops, peptide strands and domains in 3D. In work to follow, I will systematically pull apart all of these proteins. The Protein Unfolding Kit will be distributed in stereo slides, video tape (3/4 inch industrial, VHS and Beta formats), and as coordinate datasets on magnetic tape.

5

Modelling of related proteins

A M Lesk and C Chothia
MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The general response of protein structures to mutations, insertions and deletions is conformational change. A 'core' of the structure, that retains its basic folding pattern, comprises major elements of secondary structure and residues flanking them, including active-site peptides. The core may be as little as 40% of the structure for distantly related proteins, but is 90% or more for proteins with amino-acid sequence homologies of 50% or more. There is a direct relation between the structural difference of the main-chain atoms of the core residues of a pair of proteins and the overall amino-acid sequence homology. The deviation reflects the shifts and rotations of packed secondary structures with respect to one another. Successful model building of an unknown protein depends on knowing the structure of a reasonably close relative. We describe the prediction and test of a model for the V_L and V_H domains of the antilysozyme antibody D1.3.

6

An algorithm which predicts the conformation of short lengths of chain in proteins

J Moult and M N G James
MRC Group in Protein Structure and Function, Biochemistry Department, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada

Graphics aided model building of protein molecules homologous to those of known structure is often successful for the conserved core of the protein structure, but fails to predict correctly the conformation of regions where there is low homology between the amino-acid sequences. Since such regions are often the principle determinants of functional differences between proteins this is a serious deficiency. We report here an algorithm which can be used to examine a large number of possible conformations of such regions, using a series of filters so that only a few of these need be subjected to full energetic evaluation. The loop sequence to be considered is divided into two halves. Then each half is grown from its fixed end on the rest of the protein independently, using backbone atoms only, and generating a set of internally allowed conformations. All such conformations which penetrate more than a specified distance below the surface of the protein are rejected as the chain grows. Next, pairs are selected from the two sets of independent half chains for which the end-to-end distance is less than a specified amount, and their ends joined. Side chains are then built onto each of the accepted full loops, generating a set of internally allowed conformations, and again rejecting those for which there is excessive protein penetration. Finally, remaining conformations are energetically evaluated, including solvent effects by the use of image charges and the size of the exposed hydrophobic area. This procedure has been applied to the prediction of the conformation of surface loops of *Streptomyces griseus* trypsin, up to 6 residues in length. For these cases, the number of possible conformations (including the side chain combinations) is restricted by the filtering to a few thousand, and the lowest energy conformation amongst each set has nearly the lowest RMS to the X-ray structure. The best RMS values are between 0.5 and 1.1 Å.

7

Development of software tools for protein structure design

H Nakamura*, T Yamazaki*, H Abe, T Noguchi†, Y Seno††, H Wako‡ and N Go‡‡**
*Department of Applied Physics, Faculty of Engineering, University of Tokyo, 7-2-1 Hongo, Bunkyo-ku, Tokyo, 113 Japan
**Department of Physics, Nishinippon Institute of Technology, Japan
†Faculty of Engineering, Kyushu University, Hakozaki, Higashi-ku, Fukuoka-shi, Japan
††Department of Physics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan
‡Juntendo University, Japan
‡‡Department of Physics, Kyushu University, Hakozaki, Higashi-ku, Fukuoka-shi, Japan

Several software tools for designing new proteins by site-directed mutagenesis have been developed by our working group.
Molecular graphics tool: Mild (Molecular ILlustration and Design) is a molecular graphic program on a raster 3D graphic system (Seiko GR3000). Interactive usage with different models including solid models is possible. Real time amino-acid substitution is available on the