Structure of a lectin

Crystallographic results

from C. C. F. Blake

ALTHOUGH the major effort in protein crystallography is at present centred on enzyme structure, structural studies on certain non-enzyme proteins may provide information of somewhat wider biological importance. One such group of proteins is the lectins. These molecules, found largely in plants, and particularly in seeds, have a number of interesting and unusual biological properties that derive from their ability to interact with cell surfaces. Their characteristic property is the ability to agglutinate erythrocytes, sometimes in a blood-group specific manner. Other cells may also be affected and most interestingly some lectins are preferentially able to agglutinate cells that have been transformed by viruses, carcinogens or radiation. In addition certain lectins are mitogenic; that is, they are able to stimulate the transformation of lymphocytes from small resting cells to large actively growing cells that may ultimately undergo mitotic division. The specific inhibition of the activity of many lectins by simple sugars has led to the conclusion that lectins bind specifically to saccharide chains on the surfaces of cells. The study of the structure and activity of these proteins could lead to an understanding of the molecular basis of their sugar specificity and the mechanism of their action on cells, which could in turn provide an insight into the structure and behaviour of the cells themselves.

Because of its availability and ease of extraction the most extensively studied lectin is concanavalin A from the jack bean. It is also one of the more interesting members of the group, because in addition to being haemagglutinating, it is also mitogenic and capable of specifically agglutinating malignant cells. The concanavalin A molecule is composed of polypeptide chains that have a molecular weight of 25,500; below pH 6 the molecular species is a dimer, and above pH 6.5 it is a tetramer. The protein apparently binds to glucosyl and mannosyl residues at the non-reducing end of oligosaccharides, and it is most potentially inhibited by α -methyl-D-manno- and glucopyranosides. Like other lectins concanavalin A requires the ordered binding of manganese and calcium ions before it can exhibit its sugar binding capacity.

Since concanavalin A was crystallised by J. B. Sumner as long ago as 1919, it is not surprising that it has more

recently been the subject of extensive X-ray diffraction studies. The general features of the molecular structure were reported independently three years ago by Edelman's group at the Rockefeller University (Edelman *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **69**, 2580; 1972) and by Hardman and Ainsworth (*Biochemistry*, **11**, 4910; 1972) at the Argonne National Laboratory. Recently a series of four papers has come from the Rockefeller group which reports the complete amino-acid sequence of the protein and a definitive description of its molecular structure (Wang *et al.*, *J. biol. Chem.*, **250**, 1490; 1975; Cunningham *et al.*, *ibid.*, **250**, 1503; 1975; Becker *et al.*, *ibid.*, **250**, 1513; 1975; and Reeke *et al.*, *ibid.*, **250**, 1525; 1975).

The X-ray studies have been carried out at 2.0 Å resolution on the tetrameric form of the protein with both types of metal ion bound. The overall shape of the monomer is best described as a flat-based dome with dimensions of 42 Å × 40 Å × 39 Å. Its structure is dominated by two large, entirely antiparallel β sheets. One of these sheets, containing six strands, forms the entire base of the dome and is, unusually, almost lacking the twist normally seen in β sheets. The second β sheet, forming the central slice of the dome, lies at a shallow angle to the first sheet, with its seven strands twisted through 90°. The top of the dome is formed by a number of loops, mainly derived from the immediate N-terminal and C-terminal parts of the polypeptide chain. The quaternary structure of the protein involves two distinct types of subunit interaction, both largely involving the flat β sheet. Interactions of one type cause the monomers to associate in pairs with equivalent strands at the edge of the flat β sheet forming antiparallel hydrogen bonds, to give dimers that contain a 12-stranded β sheet. The second type of interaction involves the face-to-face opposition of these 12-stranded sheets from two dimers, contacts being made through both hydrophobic and ionic side chains that point outward from the sheets. Interestingly the domains of the immunoglobulins, to which the lectins have sometimes been compared, have a similar β -box structure and associate through interactions very like the dimer-dimer interactions in concanavalin A (for references see *Nature*, **253**, 158; 1975). But a more closely related protein is the thyroxine-binding plasma protein, prealbumin, which has both the same type of tertiary structure and subunit interactions (Blake *et al.*, *J. molec. Biol.*, **88**, 1; 1974).

In each subunit the Mn^{2+} and Ca^{2+} ions are located in a surface cavity near

the apex of the dome, only 4.6 Å apart. Each ion is in an octahedral environment composed of four protein ligands, two of which are shared, and two water molecules. The geometry of the binding sites seems to explain the ordered binding of the ions in terms of manganese binding at a proto-binding site, causing conformational changes that complete the site and at the same time form the neighbouring calcium-binding site. About 20 Å from the metal positions is a deep cavity between the two large β sheets in the monomer that seems to be the saccharide binding site. It seems to be composed of two distinct subsites: a deeper one that is predominantly hydrophobic and a more hydrophilic subsite nearer the surface of the molecule. But although this site is occupied by the heavy atom labelled inhibitor, β -(O-iodophenyl)-D-glucopyranoside, other simple sugars that specifically inhibit the protein cause the crystals to dissolve or lose their diffraction patterns. Although this behaviour indicates that the binding behaviour of the protein is altered by crystallisation, Becker *et al.* argue that the alteration is probably a marginal change to the true saccharide binding site in the cavity, rather than the blockage of the true site that is really elsewhere on the molecule. Clearly, this proposal must be tested by either further crystallographic studies or by appropriate solution studies such as NMR, because it is essential to the understanding of the protein's sugar specificity and its interactions with cell surface glycoprotein receptors.

NMR used to study rates of change of conformation

from J. Feeney

A good deal of useful information can be obtained from measurements of the enhanced proton NMR relaxation rates of the water in aqueous solutions of metalloenzymes containing paramagnetic ions. In the past these methods have been used by Cohn, Mildvan, Koenig and their coworkers to determine binding constants for the paramagnetic ion-protein complex and the nature of their ternary complexes with ligands. From studies of the frequency dependence of the relaxation enhancements it is often possible to deduce the number of water molecules in the coordination sphere of the metal, and detailed information about the ligand exchange rates and correlation times of the interactions. When Barber and Carver (*J. biol. Chem.*, **248**, 3353;