23.2. Protein–ligand interactions

BY A. E. HODEL AND F. A. QUIOCHO

23.2.1. Introduction

There are currently over a thousand unique protein-bound ligands described in the Protein Data Bank, illustrating the enormous variety of small molecules with which proteins interact. These ligands serve as cofactors in protein-mediated reactions, substrates in these reactions, and elements that maintain or alter protein structure or macromolecular assembly. The specific binding of small molecules to proteins is a primary means by which living systems interact and exchange information with their environment. The atomic details of protein–ligand interactions are often quite similar to the intramolecular interactions observed within a protein molecule. Examples of all the various non-covalent interactions described in Parts 20 and 22, such as hydrogen bonds, van der Waals forces and other electrostatic phenomena, are observed between proteins and their small-molecule ligands.

Through the diverse interactions observed between proteins and their ligands, a few fundamental patterns of recognition emerge. In general, ligand binding requires that the protein partially or fully sequesters the ligand from the solvent. This demands that the energy of interaction between the protein and the ligand must be strong enough to overcome the interactions between both species and the solvent as well as the translational and rotational entropy which is lost upon fixing the orientation of the ligand relative to the protein. The protein achieves this level of interaction by presenting a binding site that is complementary to the ligand both in shape and electrostatic functionality. Beyond this generalization, each ligand has its own unique and complicated story. Rather than attempt to summarize the enormous subject of protein–ligand interactions in a comprehensive manner, we will instead illustrate several of the unique interactions observed between proteins and other molecules.

23.2.2. Protein–carbohydrate interactions

The interaction between proteins and carbohydrates provides a prototypic example of how proteins specifically recognize small organic ligands. Protein-mediated recognition of carbohydrates is crucial in a diverse array of processes, including the transport, biosynthesis and storage of carbohydrates as an energy source, signal transduction through carbohydrate messengers, and cell–cell recognition and adhesion (Rademacher *et al.*, 1988). Many aspects of protein–carbohydrate recognition are observed in the interactions of proteins with other small-molecule ligands. For example, carbohydrate recognition is largely conferred through a combination of hydrogen bonding and van der Waals interactions with the protein (Quiocho, 1986). These interactions are generally presented in a binding site that is highly complementary to the target ligand in shape and functionality. Carbohydrate-binding proteins also occasionally employ ordered water molecules and bound metal ions to facilitate ligand recognition (Quiocho *et al.*, 1989; Vyas, 1991; Weis & Drickamer, 1996). There are also many examples of 'induced fit' recognition of carbohydrates, where the binding of the ligand induces a conformational change in the protein. This phenomenon of ligand-induced conformational change was observed in one of the first ligand–protein interactions visualized by X-ray crystallography, namely the binding of a carbohydrate substrate to lysozyme (Blake *et al.*, 1967). Since these early studies, a wide variety of carbohydrate–protein structures have been determined, and a number of general themes have emerged from this continually active field (Vyas, 1991). These themes are further applicable in the general study of protein–ligand recognition.

Two general classes of carbohydrate-binding modes have been observed in their complexes with proteins (Quiocho, 1986; Vyas, 1991). The first group of proteins completely sequester the carbohydrate from the surrounding solvent. These proteins, including the periplasmic proteins and the catalytic site of glycogen phosphorylase, tend to have a high affinity for their ligand ($K_d \simeq 10^{-6}$–10^{-7}). The high affinity of these proteins can be attributed to the entropic gain of desolvating the protein and carbohydrate surfaces as well as the exceptionally high degree of functional complementarity in the ligand-binding site. The periplasmic proteins nearly saturate the hydrogen-bonding potential of their carbohydrate ligands (Quiocho, 1986). The second group of proteins bind their ligands in shallow clefts in the solvent-exposed protein surface. This mode of binding, observed in the lectins, lysozyme and the storage site of glycogen phosphorylase, generally has a lower binding constant than the first class of binding proteins ($K_d \simeq 10^{-3}$–10^{-6}). Some members of this second class of proteins, such as the lectins, are able to increase the affinity for their ligands dramatically by clustering a number of low-affinity sites through the oligmerization of the polypeptides (Weis & Drickamer, 1996). The specific arrangement of lectin oligomers allows them to discriminate between large cell-surface arrays of polysaccharides with high affinity and selectivity.

23.2.2.1. *Carbohydrate recognition at the atomic level*

An early review of protein–carbohydrate interactions revealed several atomic level interactions that continue to appear ubiquitously in the structures of protein–carbohydrate complexes (Quiocho, 1986). A generic cyclic sugar in mono- or oligosaccharides appears to be recognized as a disc that displays two flat nonpolar surfaces surrounded by a ring of polar hydroxyls. Proteins recognize these features by hydrogen bonding to the ring of polar hydroxyls while stacking flat aromatic side chains against the nonpolar disc faces.

Cooperative hydrogen bonding, where the hydroxyl group of a carbohydrate participates as both a donor and acceptor of hydrogen bonds, is often observed in the direct interactions between proteins and carbohydrate ligands (Fig. 23.2.2.1). The sp^3-hybridized oxygen atom of a carbohydrate hydroxyl may act as both an acceptor of two hydrogen bonds through the two lone pairs of electrons as well as a donor of a single hydrogen bond. Cooperative

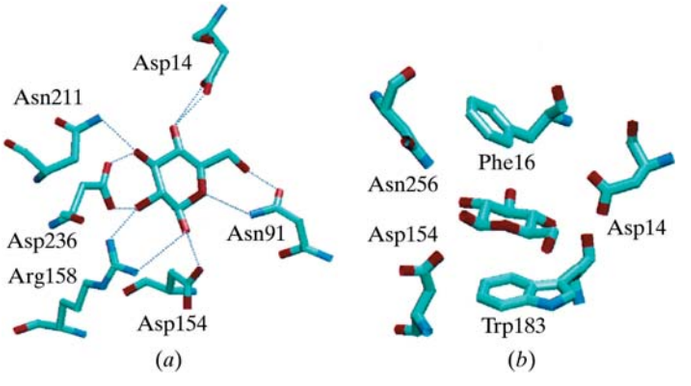

Fig. 23.2.2.1. The atomic recognition of carbohydrates by a protein. The environment of glucose bound to the galactose/glucose binding protein is shown (Vyas *et al.*, 1988). (*a*) The ring of hydrogen bonds around the polar edge of the sugar molecule. Note the 'bidentate' hydrogen bonds with Asp236, Arg158 and Asn91. As the hydroxyl groups 1, 2 and 3 simultaneously donate and accept hydrogen bonds from protein side chains, they are involved in 'cooperative' hydrogen bonds. (*b*) Two aromatic protein side chains, Phe16 and Trp183, 'stack' on the nonpolar faces of the carbohydrate.

references