

Protein-carbohydrate interactions: basic molecular features

Florante A. Quiocho

Howard Hughes Medical Institute and Department of Biochemistry,
Baylor College of Medicine, Houston, Texas 77030

Abstract — Highly refined X-ray structures (R -factors = 0.15) of the liganded forms of the L-arabinose-binding protein at 1.7 Å resolution and the D-galactose-binding protein at 1.9 Å resolution have revealed the following features of the atomic interactions between proteins and saccharides: 1) hydrogen bonds are the main factors in conferring specificity and affinity to protein-carbohydrate interactions; 2) all the polar groups (hydroxyls and ring oxygens) of the monopyranoside substrates (L-arabinose and D-glucose) are involved in extensive hydrogen bonding; 3) three types of hydrogen bonds are formed — ‘cooperative’ hydrogen bonds, bidentate hydrogen bonds and hydrogen networks; 4) the hydrogen bonds are generally strong and exhibit favorable geometries; 5) the sugar binding sites are almost completely populated by residues with planar polar side chains with at least two functional groups capable of engaging in all three types of hydrogen bonds; 6) carboxylate side chains are especially important in binding anomers and particular epimers; 7) numerous van der Waals contacts are formed, involving all the atoms of the bound saccharides; 8) one or two aromatic residues stack on the sugar ring; 9) the bound monopyranoside substrates have the normal 4C_1 full chair conformation; 10) sugar binding induces protein conformational change; and 11) these features and other factors modulate the stability of protein-carbohydrate complexes. These findings lay the foundation for a thorough and highly detailed understanding of the interactions between proteins and sugars, including those in polysaccharides, glycoproteins, DNA, etc.

INTRODUCTION

In living cells, carbohydrates are the central source for mechanical work and chemical reactions. They are also assembled into polysaccharides for use as fuel storage and structural elements, and are constituents of cofactors (e.g., ATP, NADPH, etc.), glycoproteins and polynucleotides. Thus, carbohydrates play key roles in a wide range of biological processes. As many of these processes require the interactions of proteins with saccharides, a molecular understanding of the features and factors associated with these interactions is of prime importance.

Since this article comes on the heels of two other recent reviews from our laboratory (refs. 1, 2), we provide a succinct, but still comprehensive, summary of the major molecular features and factors associated with protein-carbohydrate interactions and also highlight recent crystallographic data. In this review, as well as in the two previous ones, we are mainly concerned with uncharged sugar substrates.

The first two reviews were based primarily on results of X-ray crystallographic studies, especially the 1.7 Å resolution highly refined X-ray structure of the complex of L-arabinose-binding protein with its sugar substrate (ref. 3). This complex has proven to contain essentially all of the molecular features associated with protein-carbohydrate interactions, thereby laying the foundation for understanding these interactions and serving as a benchmark for analysis of other complexes. This knowledge has been solidified by the recent results of the extensive refinement of the 1.9 Å resolution crystal structure of the D-galactose-binding protein with bound D-glucose, also an excellent substrate (Fig. 1; ref. 4). Moreover, a much clearer picture of the mode of binding of oligosaccharides is beginning to emerge as the structure of the maltose-binding protein, which was determined at 2.3 Å resolution (TABLE 1), is being refined crystallographically (unpublished data, J. C. Spurlino & F. A. Quiocho).

These atomic structures show: 1) that hydrogen bonds are major determinants of specificity and affinity in protein-carbohydrate interactions; 2) that these hydrogen bonds are generally strong and almost always exhibit optimal geometries; 3) that three types of hydrogen bonds are formed — 'cooperative' hydrogen bonds, bidentate hydrogen bonds and hydrogen networks; 4) that many of the residues in the carbohydrate-binding sites have planar polar side chains with at least two functional groups capable of engaging in all three types of hydrogen bonds; 5) that a carboxylate side chain is also responsible in binding anomers or particular epimers; 6) that van der Waals forces also contribute significantly to saccharide affinity and specificity; 7) that aromatic residues stack on the sugar ring; 8) that the bound monopyranoside substrates have the normal 4C_1 full chair conformation; and 9) that many factors contribute to the stability of protein-carbohydrate complexes. These features are further discussed, with the aid of Figs. 2 to 5 which depict many of these features.

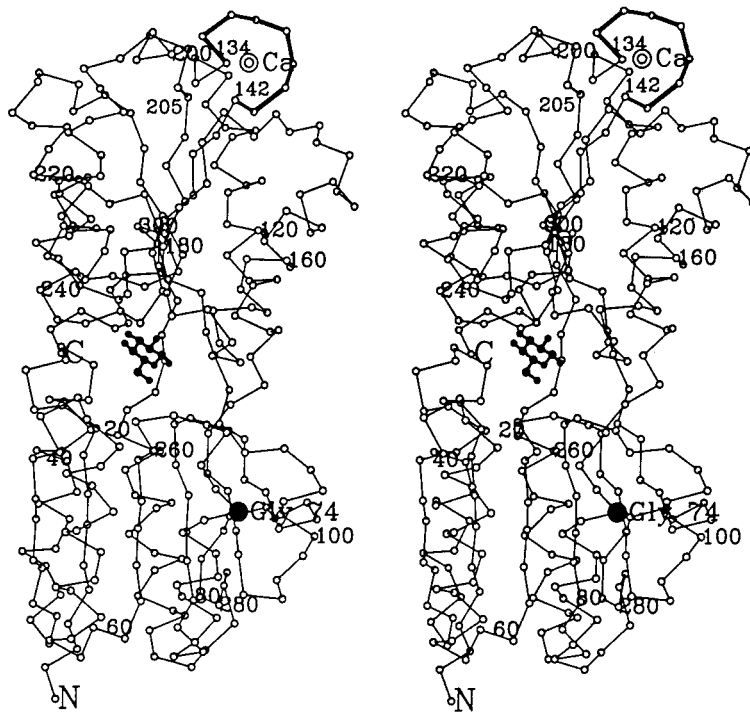


Fig. 1. Stereoscopic view of the α -carbon backbone trace of the structure of the D-galactose-binding protein crystallographically refined at 1.9 Å resolution (ref. 4). Highlighted are 1) the model of the D-glucose substrate (filled circle) bound in the cleft between the two domains; 2) the calcium (depicted as double circles) located at one end of the elongated protein molecule; and 3) Gly 74 (large filled circle) which is part of a site for interacting with the chemotactic transmembrane signal transducer. (Adapted from ref. 3).

To complement our X-ray structure analysis and to gain an understanding of other aspects of sugar binding, we have undertaken a number of other studies of periplasmic sugar-binding proteins in solution. Sugar-induced protein conformational change was assessed by low-angle X-ray scattering measurements (ref. 5) and by ligand binding studies (ref. 7), kinetics of sugar binding were analyzed using stopped-flow rapid-mixing technique (refs. 6, 7), and thermodynamics of complex formation were determined by calorimetric technique (ref. 8).

The L-arabinose-binding protein (ABP)*, D-galactose-binding protein (GBP)*, and D-maltose-binding protein (MBP)* are members of a large group of proteins located in the periplasmic space of gram-negative bacteria. All these monomeric proteins serve as essential components of osmotic-shock-sensitive active transport systems for a variety of carbohydrates, amino acids and ions. Moreover, many of the sugar-binding proteins act as

*Abbreviations used: ABP, L-arabinose-binding protein; GBP, D-galactose-binding protein; MBP, D-maltose-binding protein

initial receptors in the simple behavioral response of bacterial chemotaxis. Nutrient transport and sugar chemotaxis also require distinct components residing in the cytoplasmic membrane. The interaction between a binding protein and the corresponding membrane-bound proteins initiates nutrient translocation or flagellar motion. Because of the variety of interactions that periplasmic binding proteins must undertake (with diverse small molecule nutrients, with membrane-bound transport components, and with transmembrane chemotactic components), these proteins are a gold mine for studies of protein structures and interactions with both large and small ligands (for example, see Fig. 1).

In addition to the L-arabinose- and D-galactose-binding protein, we have also determined and refined the X-ray structures of the sulfate-binding protein and the leucine/isoleucine/valine-binding protein (refs. 9, 10; unpublished data). The structures of all these monomeric proteins are very similar in spite of the lack of significant sequence homology. They are ellipsoidal (axial ratio of 2:1), being composed of two distinct but similar globular domains which are connected by three separate peptide segments (for example, the structure of GBP shown in Fig. 1). Although the two domains of these proteins are each composed of segments from both the amino- and carboxyl-terminal halves of the protein, they have similar supersecondary structure, consisting of a central core of a β -pleated sheet sandwiched by at least a pair of parallel α -helices. Remarkably, while the three separate peptide segments connecting the two domains are widely separated in the amino acid sequence, they are proximal in the tertiary structure and form the base or 'boundary' of the deep cleft between the two domains. The sugar substrate is bound in the cleft and completely engulfed.

Structural studies of the sulfate- and leucine/isoleucine/valine-binding proteins have also led to a greater understanding of electrostatic interactions in protein structures and binding of charged substrates such as sulfate and phosphate dianion and leucine zwitterion (refs. 9, 10, 11, 12). A novel calcium-binding site has also been uncovered (refs. 11, 13).

TABLE 1. Binding protein-carbohydrate complexes

Protein	X-ray Structure Resolution (Å)	Saccharide Binding			
		Substrate	$K_d \times 10^7$ (M)	$k_1 \times 10^{-7}$ ($M^{-1} s^{-1}$)	k_{-1} (s^{-1})
L-Arabinose-binding protein	1.7	L-Arabinose	0.98	2.4	1.5
	1.9	D-Galactose	2.3	0.8	1.8
	1.9	D-Fucose	38	1.2	37
D-Galactose-binding protein	1.9	D-Galactose	4	3.3	3.4
	---	D-Glucose	2	3.5	1.4
D-Maltose-binding protein	2.3	D-Maltose	35	2.3	90
	3.5	Maltotriose	1.6	2.5	8.4
	3.5	Maltotetraose	23	---	---
	---	Maltopentaose	50	---	---
	---	Maltohexaose	34	---	---
	---	Maltoheptaose	16	---	---
	---	Cyclic maltohexaose	40	3.6	110
---	Cyclic maltoheptaose	18	2.2	46	

K_d is the dissociation constant obtained by equilibrium binding measurements. The bimolecular kinetic rates k_1 , or association rate, and k_{-1} , or dissociation rate, were obtained by rapid-mixing, stopped-flow kinetic binding studies (refs. 6, 7; unpublished data).

From the standpoint of diversity and specificity of carbohydrate substrates (TABLE 1), the periplasmic sugar-binding proteins are an excellent system to analyze and study. (TABLE 1 also indicates the considerable progress we have made to determine the structures of a number of these complexes.) ABP binds L-arabinose and D-galactose with similar affinity and D-fucose less tightly. D-galactose and its 4-epimer D-glucose are both excellent substrates of GBP. MBP binds D-maltose but not D-glucose. Remarkably, these sugar-binding proteins bind both anomeric forms of their sugar substrates with identical dissociation constants and kinetic rates (ref. 7).

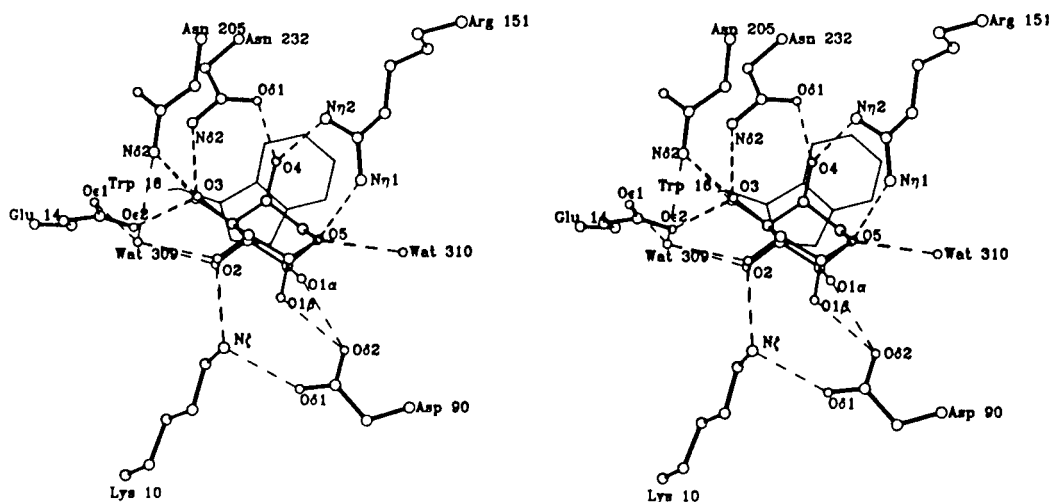


Fig. 2. Stereoscopic drawing of the hydrogen-bonding and stacking interactions between L-arabinose-binding protein and L-arabinose. One domain of the bilobate ABP provides residues 10, 14, 16 and 90 and the other domain donates 151, 205 and 232. (Adapted from ref. 3.)

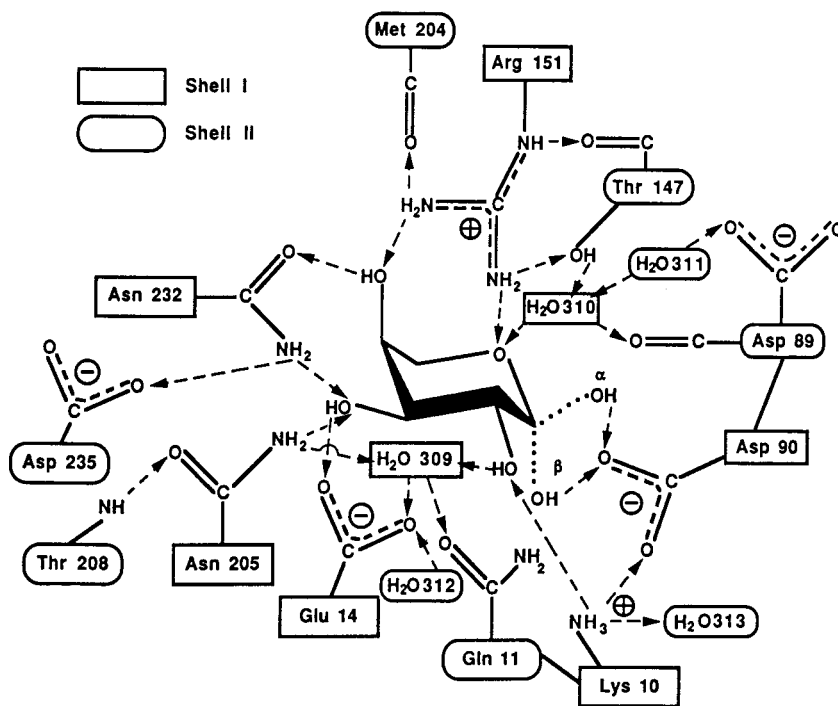


Fig. 3. Schematic diagram of the cooperative hydrogen bonds, bidentate hydrogen bonds, and networks of hydrogen bonds in the complex of L-arabinose-binding protein with L-arabinose. (Adapted from ref. 4.)

The ability of MBP to bind larger oligoglucoses, both linear and cyclic (TABLE 1), provides an unusual opportunity to look at not only the mode of binding but also the conformations of these polysaccharides, many of which are uncrystallizable themselves. What is unique about the geometry of the ligand site of MBP which permits binding of linear and cyclic oligosaccharides?

The binding of sugar substrates to the various sugar-binding proteins is both a rapid and high affinity process (refs. 1, 7; TABLE 1). The kinetics of ligand binding is a second order. It is also noteworthy that the binding protein-saccharide complexes are some of the tightest known protein-carbohydrate complexes. Variations in the dissociation constants of ligands (e.g., fucose binding) can be attributed primarily to changes in k_{-1} or dissociation rates (refs. 6, 7; TABLE 1). An explanation for the tight sugar binding is offered in a later section. Ligand affinity and the kinetics of substrate binding are fundamentally related to the functions of binding proteins in active transport and chemotaxis (refs. 1, 3, 7, 11). The association rate constant is a measure of the minimum response time for chemotaxis, and the dissociation rate defines the maximum possible velocity for the corresponding transport systems. The ratio of the association and release rates (viz., the affinity) determines the sensitivity of the two physiological functions.

HYDROGEN BONDS

Both the ABP-arabinose and GBP-glucose complexes have set a record for the largest number of hydrogen bonds formed as the result of the binding of monopyranoside substrates (for comparisons, see ref. 1). Indeed, all the polar groups of the bound monopyranosides are involved in the formation of hydrogen bond: ten in the ABP-arabinose complex and twelve in the GBP-glucose complex (Figs. 2, 4).

Hydrogen bonds play three major roles in protein-carbohydrate interactions: they provide stability, confer specificity and control dynamics. These roles are not too surprising for the following general reasons: 1) hydrogen bonds are stable enough to significantly contribute the requisite ligand-binding affinity (stability) but are of sufficiently low strength to allow rapid ligand dissociation (dynamics), features especially important for proteins involved in active transport; 2) compared to the non-directional van der Waals forces, hydrogen bonds are highly directional, which is important in specificity; and 3) in regards to the substrates, hydroxyl groups of carbohydrates not only participate heavily in these bonds but also constitute the major functional groups of carbohydrates, which are stereospecific and highly exposed. The importance of a hydroxyl group is also related to its ability to engage in three hydrogen bonds — as a donor of one hydrogen bond and an acceptor of two through the sp^3 lone pairs. Moreover, as a donor, a hydroxyl group has the added advantage of having rotational freedom about the H-C-O-H torsion angle, thus enabling it to attain the most linear hydrogen bond possible with an acceptor group.

It is perhaps surprising that hydrogen bonds in the various binding protein-sugar complexes take on three forms (Figs. 2 to 5): cooperative hydrogen bonds, bidentate hydrogen bonds, and networks or arrays of hydrogen bonds.

'COOPERATIVE' HYDROGEN BONDS

Cooperative hydrogen bonds are those resulting from the simultaneous participation of a sugar hydroxyl as donor and acceptor of hydrogen bonds (Figs. 2 to 5). With the possible exception of the anomeric hydroxyl of L-arabinose, all of the hydroxyls of the bound L-arabinose and D-glucose substrates are engaged in cooperative hydrogen bonding. Remarkably, essentially all of the cooperative hydrogen bonds seen thus far in our studies can be depicted simply as:



where OH is a sugar hydroxyl, NH and O are hydrogen bond donor and acceptor groups, respectively, of residues in the the binding site, and $n = 1$ or 2.

The hydroxyl groups at positions 2, 3, and 4 of L-arabinose bound to ABP and at all positions of D-glucose bound to GBP are involved in cooperative hydrogen bonds. (Note that it is also possible for OH1 of L-arabinose to accept a hydrogen bond, as described below).

BIDENTATE HYDROGEN BONDS

When two adjacent hydroxyls of a 4C_1 pyranoside each interact with a different atom of the same planar polar side chain residue, they form bidentate hydrogen bonds (Figs. 2 and 4). Bidentate hydrogen bonds have geometrical requirements as evidenced by the following findings. The bidentate hydrogen bonds found thus far are formed when adjacent pairs of hydroxyls are both equatorial (e.g., OH1 and OH2 pair and OH2 and OH3 pair of D-glucose, Fig. 4) or when one is equatorial and the other axial (e.g., OH3 and OH4 of L-arabinose, Fig. 2). As further shown in Figs. 2 and 4, the sugar ring oxygen, when paired with the axial OH4 of L-arabinose or

with D-glucose's OH6 (oriented with χ_5 or torsion angle C4-C5-C6-O6 of 177°), can also participate in bidentate hydrogen bonding. In all four cases, the distance between the pair of sugar oxygen atoms is about 2.8 \AA , ideal for bidentate hydrogen bonding with the planar side chains. An adjacent pair of axial hydroxyls of a pyranoside is less likely to form bidentate hydrogen bonds because the hydroxyls are much more separated (3.6 \AA) and pointing in opposite directions.

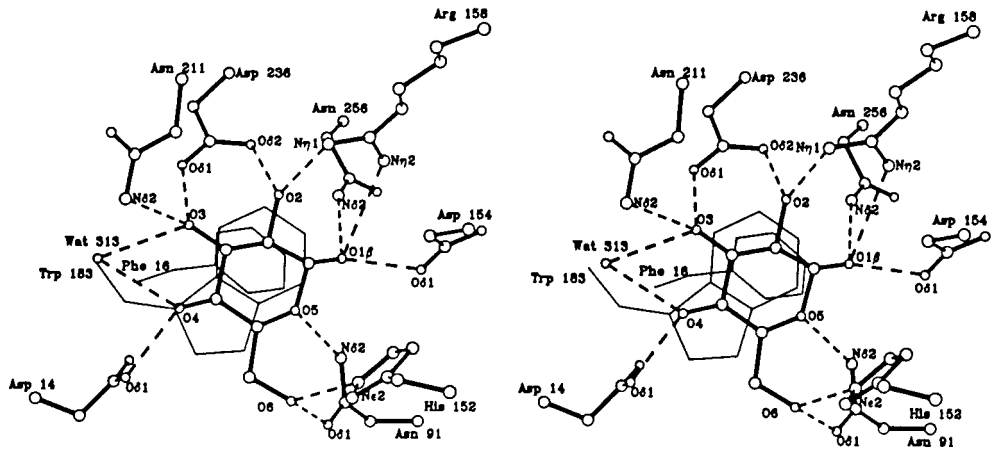


Fig. 4. Stereo drawing of the atomic interactions (both hydrogen-bonding and stacking) between residues of the D-galactose-binding protein and D-glucose. Residues 14, 16, 91, 256 are located in one domain and 152, 154, 158, 183, 211, and 236 are located in the other domain (see Fig. 1). (Adapted from ref. 4.)

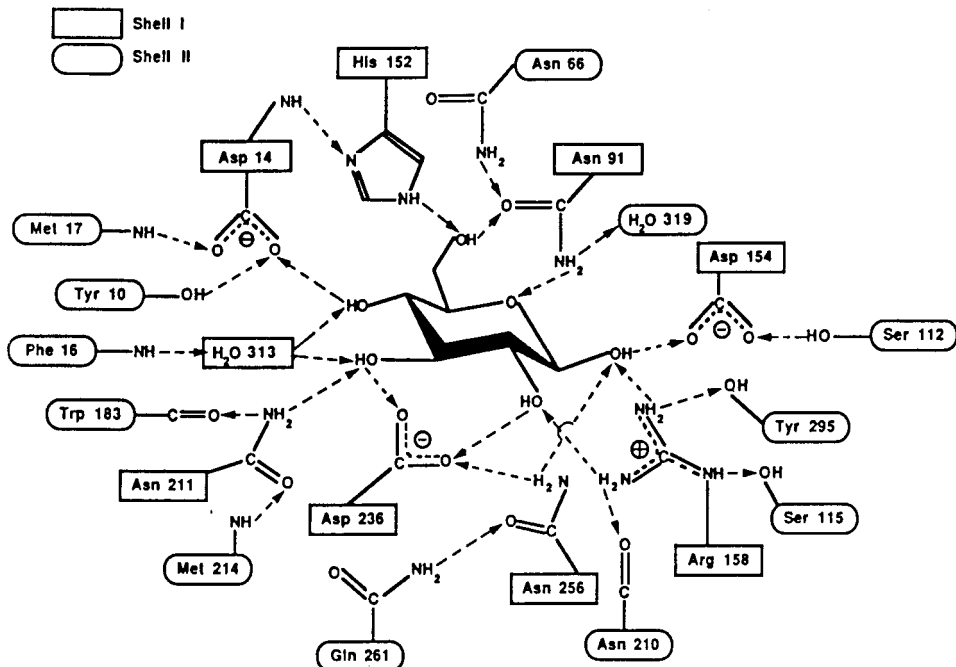


Fig. 5. Schematic diagram of the cooperative hydrogen bonds, bidentate hydrogen bonds, and networks of hydrogen bonds in the complex of D-galactose-binding protein with D-glucose. (Adapted from ref. 4.)

NETWORKS OF HYDROGEN BONDS

The formation of cooperative and bidentate hydrogen bonds leads to the creation of networks of hydrogen bonds between the sugar and essential residues (Figs. 3 and 5). The polar groups of the essential planar side chains (Shell I) also hydrogen-bond with other adjacent polar residues (Shell II), further creating a more extensive and elaborate network of hydrogen bonds (Figs. 3 and 5).

GEOMETRY AND PARAMETERS OF HYDROGEN BONDS

Because the structures of the liganded forms of ABP and GBP are well-refined at better than 2 Å resolutions, it has been possible to obtain accurate values of the parameters — distances and angles — of the hydrogen bonds shown in Figs. 2 and 4 (refs. 3, 4). These parameters indicate strong hydrogen bonds; the ten in ABP complexed with either anomer have a mean distance (between H bond donor and acceptor) of 2.82(0.15) Å and a mean angle of 164(9)° (ref. 3). In the GBP-glucose complex, the mean of the hydrogen bond distance is 2.87(0.28) Å and the angle 156(16)° (ref. 4).

It is noteworthy that in both complexes, the hydrogen bonds for which the sugar hydroxyls serve as donors are generally shorter — implying stronger bonds — than the ones for which they serve as acceptor groups (refs. 3 and 4).

Many of the cooperative hydrogen bonds depicted in Scheme 1 have geometries that are close to ideal (see Figs. 2 and 4). For example, the OH3 of L-arabinose and the OH1 and OH3 of D-glucose are fully coordinated ($n = 2$ in Scheme 1) in an arrangement which is essentially tetrahedral, including the sugar C—O bond.

In several cases, where $n = 1$, the geometry is also very favorable. For instance, the atoms C4 and O4 of the L-arabinose, Arg 151 N η 2, and Asn 232 O δ 1 are coplanar (Fig. 2). Additionally, since the hydrogen bond donated to or by OH4 is oriented so as to bisect both lone pairs of electrons (in their predicted locations) on either acceptor group (the pairs on the sugar hydroxyl or the sp^2 pairs on the O δ 1 of Asn 232), it is likely that both lone pairs share in accepting the proton donor. The same conclusion is drawn from the analysis of the hydrogen-bonding system involving OH-2 of L-arabinose (Fig. 2) and OH2 and OH6 of D-glucose (Fig. 4).

The hydrogen-bonding geometries involving the sugar-ring oxygens are also highly regular. Since the sugar ring oxygen of L-arabinose accepts two hydrogen bonds in an almost tetrahedral coordination, each of the lone pair of electrons on the sugar ring oxygen is directed at a hydrogen bond donor group (Fig. 2). On the other hand, both lone pairs of electrons on the O5 of D-glucose share approximately equally in accepting the donatable proton of Asn 91 N δ 2H (Fig. 4).

POLAR PLANAR SIDE CHAINS IN SUGAR BINDING SITES

Since side-chain residues with polar planar groups — Asn, Asp, Glu, Gln, Arg, and His — are the only ones participating in all three forms of hydrogen bonding, they are abundant in sugar-binding sites (Figs. 2 to 5; see also ref. 1). Indeed, these residues, with the exception of Lys 10 in ABP, are the only ones used in hydrogen bonding. In contrast to the sugar hydroxyl groups, all the hydrogen-bond-donating groups of the planar residues are fixed or do not have freedom of torsional rotation. Therefore, the final geometry of the binding site greatly depends on the final folded protein and any attendant ligand-induced conformational change which brings these residues to their correct orientation.

Although Lys 10 of ABP is the only residue in the sugar-binding sites of ABP and GBP that is not polar and planar, like the rest, it is engaged in multiple interactions crucial to ligand binding. Its ammonium side chain is in an excellent position to donate a hydrogen bond to O-2, to fix via a salt-link the alignment of Asp 90 which in turn hydrogen-bonds the anomeric hydroxyls, and to make van der Waals contact or a weak hydrogen bond with each of the anomeric hydroxyls (Figs. 2).

BINDING OF SUGAR ANOMERS AND EPIMERS

As indicated in TABLE 1, the periplasmic sugar-binding proteins have unusual ligand specificities. Understanding of these specificities clearly require knowledge of the geometries of the binding sites.

Prior to obtaining the atomic details of the mode of binding of sugar substrates to ABP and GBP (Figs. 2 to 5),

sugar-binding studies were undertaken in our laboratory, employing equilibrium and rapid stopped-flow kinetic techniques. These studies showed that the L-arabinose-, D-galactose- and D-maltose-binding proteins bind both anomers of their sugar substrates with indistinguishable properties (ref. 7). Moreover, TABLE 1 indicates that the 4-epimers D-glucose and D-galactose are bound with similar affinity. In light of the high resolution structures, we have a clear understanding of the manner in which these unusual stereospecificities are achieved. The structures show that planar polar side chains, especially carboxylate residues, also play a specific role in anomeric and epimeric sugar recognition.

The key to ABP's ability to recognize both anomers of L-arabinose, while using the same essential residues in hydrogen bonding, is the precise alignment of the O δ 1 of Asp 90, which enables this atom to accept a hydrogen bond from either the α (equatorial) or β (axial) anomeric hydroxyl (see Figs. 2 and 3). Even though ABP binds either anomeric form, the orientations of all of the essential residues remain unchanged and the hydrogen bonds to both anomers show essentially the same parameters (ref. 3). As noted previously, Lys 10 is held in place by charge-coupling and hydrogen-bonding with Asp 90.

Asp 154 O δ 1 of GBP undertakes the same role as the corresponding atom of Asp 90 of ABP in accepting, because of its precise location, a hydrogen bond from the hydroxyl of either D-glucose anomer (ref. 4). The Asp 154 of GBP is also coupled with Arg 158.

Whereas the abilities of ABP and GBP to bind both anomers largely depend on one oxygen atom of an Asp residue for accepting a proton from either anomeric hydroxyl, GBP's ability to bind the D-galactose and its 4-epimer D-glucose require two different oxygen atoms of the same carboxylate side chain. As shown in Fig. 4, the equatorial OH4 of the bound D-glucose donates a hydrogen bond to O δ 1 of Asp 14 of GBP. On the other hand, modeling experiments show that the axial OH4 of the D-galactose has to be donated to O δ 2 of Asp 14, instead of O δ 1. Otherwise, both pyranoside substrates are bound identically. Taken together, the data indicate a novel sugar-binding site geometry of GBP, designed to bind both anomers of the D stereoisomers of the galactose and glucose epimers.

The mechanism for recognition of epimers and anomers is relevant to understanding the binding properties of other proteins. For instance, it is known that several of the enzymes of glucose metabolism, such as hexokinase, glucokinase, glucose-6-phosphatase, can also act on both anomeric forms of their respective sugar substrates.

VAN DER WAALS FORCES

Dispersion forces also contribute significantly to protein-carbohydrate interactions. The well-refined structures of the ABP-arabinose and GBP-glucose complexes has enabled us to more clearly identify and describe some aspects of these interactions. All the heavy atoms of the L-arabinose bound to ABP and D-glucose bound to GBP make van der Waals contacts with protein atoms (maximum contact distance of 4 Å) — about 54 contacts in the ABP-arabinose complex and about 60 in the GBP-glucose complex (refs. 1, 3). As is the case with the hydrogen bonds, these many contacts are unusual, given that the sugars are only monopyranosides. This number of contacts is the consequence of the numerous hydrogen-bonding interactions, enabling many more of the other atoms of the polar residues to come within van der Waals distance of the bound sugar substrates, and the complete enclosure of the sugar substrate within the binding site.

STACKING BETWEEN SUGAR AND AROMATIC RESIDUES

Unexpectedly and remarkably, some of the van der Waals contacts in the two complexes are brought on by stacking of one or both sides of the monopyranoside ring with aromatic residues (Figs. 2 and 4). The discovery of these stacking interactions further reveals another unusual facet of protein-sugar interactions.

In the descriptions of the stacking interactions which follows, we have adopted the suggestion by Rose *et al.* (ref. 14) of a simple α/β -face convention for identifying or distinguishing the two faces or sides of ring compounds. In this paper this concerns the faces of the pyranose ring of the more common and biochemically important 4C_1 conformation and the indole ring of tryptophan.

Whereas only the β -face of the L-arabinopyranoside bound to ABP is stacked with the β -face of Trp 16 (Fig. 2), both faces of D-glucopyranoside bound in GBP are stacked — the α -face with Phe 16, located in one domain, and the β -face with the α -face of Trp 183, situated in the other domain (Fig. 4). It is important to emphasize that the complete pairing of all polar groups of the residues and sugar substrates in the solvent-inaccessible cleft enables the stacking interactions to occur.

The locations of the aromatic residues are consistent with the presence of hydrophobic patches especially in pyranosides. However, the sizes of these patches represent — in terms of accessible area — a small portion compared to the total area occupied by the polar hydroxyl groups, varying according to the orientations of the hydroxyls relative to the sugar ring. For instance, D-glucose has a hydrophobic patch on the β -face composed of C3, C5 and C6 and a minor patch on the α -face composed of C2 and C4. On the other hand, L-arabinose has a hydrophobic patch, consisting of C3, C4 and C5, located on the β -face.

The aromatic residues further confer specificity to the site by disallowing binding of particular sugar epimers through steric hindrance and/or an unfavorable non-polar environment (refs. 1, 3, 4). The lack of binding of these epimers could be also attributed to the absence of hydrogen-bonding residues at the right orientations.

SUGAR CONFORMATION

Figs. 2 and 4 clearly show that both the protein-bound monopyranosides have the normal 4C_1 full chair conformation. In our crystallographic refinements, this conformation of the sugars fitted the electron density extremely well; there is no obvious distortion of the conformation. It is also important to note that the binding site regions of ABP and GBP, with the bound substrates, have exceptionally well-resolved electron density and highly accurate molecular structure. These regions are also some of the least mobile regions of the protein molecules, due in part to the complex formation. In spite of the very extensive atomic interactions through an unusually large number of highly optimized hydrogen bonds and van der Waals contacts, the stacking interactions, and the high packing efficiency in the substrate-occupied binding sites, it appears that the ring structure of the bound monopyranosides is not easily distorted.

The C4–C5–C6–O6 of the D-glucose bound to GBP has a χ_5 torsion angle (177°) which places the OH6 hydroxyl away from the equatorial OH4. However, this position is still not close enough to hydrogen-bond to the sugar ring oxygen. This orientation of the $-CH_2OH$ of the bound D-glucose conforms more with those observed in sugar crystal structures rather than with NMR data which show hydrogen bonding of the OH6 with the ring oxygen.

FACTORS THAT MODULATE PROTEIN-CARBOHYDRATE AFFINITIES

Hydrogen bonds make a considerable contribution to the absolute energy or stability of protein-carbohydrate interactions, modulated by several factors (refs. 1, 4): 1) number, 2) type (neutral-neutral or neutral-charged), 3) distance and angle, 4) 'cooperativity', 5) geometry, 7) networking, 8) solvent accessibility or dielectric constant, and 9) ligand-induced protein conformational change. As is clearly evident in the X-ray structures of the complexes of both ABP and GBP with sugars, not only are all these factors contributing to the stability of these complexes, but many are achieved optimally. Figs. 2 to 5 are probably among the best in portraying how an extensive and precise complementarity between proteins and sugars is achieved through hydrogen bonding.

All the hydrogen-bonding groups of the pyranoside substrates are used, resulting in an unusually large number of hydrogen bonds which are almost equally distributed between neutral-neutral and neutral-charged types. The means of the hydrogen bond distances and angles indicate strong bonds. The large number of strong hydrogen bonds mitigates any need for aqueous solvation.

Essentially all the sugar hydroxyls are involved in cooperative hydrogen bondings, which are remarkably simple (Scheme 1), and which exhibit nearly optimal geometries. Cooperative hydrogen bonding enhances charge delocalization and thus effectively strengthens hydrogen bonds. The 'cooperative strengthening' of hydrogen bonds has been noted previously in studies of hydrogen-bonded systems of water molecules (ref. 15) and of crystal structures of simple sugars (ref. 16).

There is pairing of all functional groups of the binding-site polar residues and total involvement of every one of the potential hydrogen bond donor groups of these residues and of the sugars as well, giving rise to extensive networks of hydrogen bonds and to precisely and stably oriented ligand-binding site residues. Remarkably, the hydrogen-bonding schemes shown in Figs. 3 and 4, including the arrows indicating the directions of the hydrogen bonds, are completely dictated by the tertiary structures of the complexes, and leave no ambiguity.

A sugar-induced conformational change brings the binding site residues, polar and non-polar together, located in both lobes of the bilobate binding proteins, to their correct orientations for binding. As a further consequence, the bound sugar substrates and all of the groups directly associated with sugar binding are buried deep in the cleft and are rendered inaccessible to the bulk solvent. The conformational change is described as a

bending motion about a hinge between the two domains (refs. 5, 10, 17). Kinetic studies indicate that the conformational change is not the rate-limiting step in complex formation (refs. 6, 7). These studies portray the periplasmic binding proteins as 'Pac-mans', but with the condition that closure of the cleft is preceded by the ligand binding first to one domain (ref. 10).

Complete enclosing of the sugar-occupied binding site has several consequences favorable to the stability of protein-sugar complexes. Water molecules hydrogen-bonded to the polar groups of both the sugar and the binding site residues are released to the bulk solvent, thus resulting in a favorable positive entropic contribution of hydrophilic origin to the free energy of the complex (further discussed below). The lower dielectric constant, relative to the bulk solvent, within the solvent-excluding cleft hosting the sugar is likely to strengthen hydrogen bonds and van der Waals contacts.

Also, in order to achieve a tightly bound ligand in totally enclosed binding sites, such as the ones of ABP and GBP, the hydrogen bonding capabilities of polar groups of the ligand and residues must be satisfied and the efficiency of hydrogen bonds and van der Waals contacts optimized as much as possible. In such an environment, an unpaired polar group would lead to weaker complex; as much as 4 kcal/mol loss of free energy has been estimated for an unpaired buried polar. It is also unlikely that buried groups with the capability to donate and accept a proton (e.g., hydroxyl groups of sugars and of side chain residues) would serve solely as a hydrogen bond acceptor, thus leaving free the 'active' or donatable proton.

Dispersion forces also provide stability to protein-sugar complexes. Owing to the extensive hydrogen bonding and the complete enclosure of the sugars in the binding sites of the L-arabinose- and D-galactose-binding proteins, there is an unusually large number of van der Waals contacts and a high packing efficiency between protein and sugar atoms. Consequently, the contribution of these forces in the overall stability of the complexes is also significant.

In light of the above exposé, it is not surprising that the binding of a variety of substrates (from monopyranosides to oligosaccharides) to periplasmic sugar-binding proteins results in some of the tightest known protein-carbohydrate complexes (Table 1; refs. 1, 6, 7). It should also be underscored that many of the factors which contribute to the affinity between proteins and saccharides are nonexistent or less effective in the interactions between water molecules and sugars in aqueous solution. Therefore, the formation of hydrogen bonds leads to a considerable net contribution to the absolute stability of protein-sugar interactions, despite the compensatory displacements of water of hydration of the substrate and the binding site residues.

VARIATIONS ON THE THEME AND OTHER RELATED MATTERS

Sugar-binding sites in other proteins and enzymes

Undoubtedly, the features and factors (or combinations thereof) described above also govern all other protein-sugar interactions. Indeed, many of these features are present in other structures of complexes of proteins or enzymes with carbohydrates, although these complexes have not been analyzed at very high resolutions and/or refined as well as those of the periplasmic binding proteins (ref. 1).

There are, however, conditions that would cause variations of these features or factors in other complexes. 1) The number of hydrogen bonds and van der Waals contacts could be fewer. 2) More water molecules may be hydrogen-bonded to the bound saccharide. 3) Cooperative, bidentate, and networks of hydrogen bonds are likely to be less numerous. 4) It may very well be that $n = 0$ in Scheme 1, i.e., the sugar hydroxyl may be solely a proton donor. However, it would be rather unlikely, for the reasons discussed above, for a buried sugar hydroxyl to be a proton acceptor without at the same time being a proton donor, either to a protein group or ordered water molecules. And 5), the dielectric constant of the site could be higher. These variations could very well all occur in a sugar-binding site close to the surface of the protein, with the bound substrate partially accessible to the solvent.

The NH groups of side chains need not always be the ones donating H bonds to sugar hydroxyls (see Scheme 1); indeed, in a few instances, donors have been found to be OH groups of serine, threonine, tyrosine, or water and NH groups of peptide units and tryptophan residues (ref. 1). However, these residues are incapable of participating in bidentate hydrogen bonding and less able to further engage in extensive networks of hydrogen bonds.

Aliphatic residues could very well be adjacent to the hydrophobic patches of sugars, thus achieving a somewhat similar effect as the stacking interactions between aromatic residues and faces of pyranosides as shown in Figs. 2 and 4. Indeed, the methyl group of a methionine residue is close to the e -face of the L-arabinose (unpublished data, N. K. Vyas & F. A. Quicho), whereas the β -face, with a larger hydrophobic patch than the one on the α -face, is stacked with a Trp residue (Fig. 2).

All of the above variations or modifications are likely to create low-affinity protein-carbohydrate complexes.

Polysaccharide substrates

Although large polysaccharides serve as substrates, the binding sites of enzymes with these ligands have all been shown to be of very limited size, capable of interacting with only very short oligosaccharide segments that are often many orders of magnitude smaller than the polysaccharide substrates themselves (ref. 1). Moreover, none of these sites is capable of completely burying its ligand. For example, the binding site of lysozyme, which catalyzes the hydrolysis of bacterial cell walls, has six overlapping subsites and taka-amylase, which hydrolyzes amylose, is able to bind only a hexasaccharide unit (refs. 18, 19). On the other hand, the glycogen storage site of phosphorylase has four subsites (ref. 20). Although MBP binds oligosaccharides, as large as linear and cyclic maltoheptaose (TABLE 1), the binding site is capable of recognizing only three, possibly four, glucosyl units, some parts of which are exposed (unpublished data, J. C. Spurlino and F. A. Quiocho). Notwithstanding these different sizes of binding sites, the germane features of the interactions between the enzymes and MBP and their polysaccharide substrates are all essentially the same, and do not differ from those firmly established from studies of protein-monosaccharide complexes (ref. 1).

Entropic effects

Favorable entropic contributions to the stability of protein-carbohydrate complexes are derived from both hydrophilic and hydrophobic effects. Both effects, generating positive entropies, are due to the release to the bulk solvent of water molecules either hydrogen-bonded to the accessible polar groups or ordered at the interface of apolar groups consequent upon complex formation. Since saccharides may be considered as amphipathic substances (ref. 2), both effects are likely to have influence upon binding.

As already indicated above, the exposed polar groups of the free sugar and of the binding-site residues in the uncomplexed protein will have hydrogen-bonded water molecules. These constrained water molecules may be wholly or almost totally expelled to the bulk solvent on formation of the hydrogen bonds and van der Waals contacts between these groups in the protein-substrate complex, thus providing a significant favorable hydrophilic contribution.

In addition, non-polar groups of the carbohydrate substrates have at their interface more locally organized water structures. There is a disordering of the water structures at these interfaces when the non-polar groups come into proximity and engage in van der Waals interactions with protein atoms, resulting in favorable hydrophobic contribution.

The contribution of hydrophobic effects is proportional to accessible areas (ref. 21). Since non-polar areas usually constitute a smaller proportion of the total solvent accessible areas of saccharides and since sugar-binding sites are lined by many polar residues, hydrophobic effects make less contribution to the free energy change of protein-sugar complexes. Especially in simple sugars, the hydroxyls and ring oxygen are highly exposed, constituting better than 70% of the accessible area of the sugar in the free state (refs. 2, 3). The fact that saccharides are very soluble in water attests to the highly polar nature of these ligands.

Complex formation through hydrogen bonds and van der Waals contacts is further accompanied by an unfavorable (negative) entropic term from loss in translational and rotational motions.

The net contribution of entropic effects to the overall stability of protein-sugar complexes is difficult to pin down (see ref. 1.). However, thermodynamic studies of the reaction between proteins and carbohydrate substrates (including the binding of L-arabinose and D-galactose to ABP (ref. 8)) show negative entropy contributions (ref. 1), thus indicating hydrogen bonds and van der Waals contacts are the major forces in the stability of protein-carbohydrate complexes (refs. 1, 22). The positive entropic contributions, derived from hydrophilic and hydrophobic effects, are likely to be crucial in the initial phase of ligand binding, when the ligand first encounters the protein surface (ref. 22) and desolvation ensues.

Taken together, the results of the X-ray structure and thermodynamic studies indicate that electrostatic interactions, primarily through hydrogen bonds and van der Waals forces, are the main source for the absolute stability of protein-sugar complexes (see also ref. 1), not hydrophobic effects as suggested by Lemieux and co-workers (refs. 23, 24). Hydrophobic effects are but one component of the entire process of protein-carbohydrate complex formation.

Specificity, recognition and conformational change

Although opinions differ on the contributions of the various factors to the absolute stability of protein-ligand complexes, there is no doubt that hydrogen bonds and electrostatic interactions are crucial for ligand specificity of proteins in general (ref. 25). Carbohydrate-binding proteins or enzymes are no exceptions. Indeed, results of structural analyses of complexes of proteins or enzymes with their carbohydrate substrates — from monosaccharides to oligosaccharides — clearly show that hydrogen bonds are mainly responsible for conferring sugar specificity and ensuring correctness of fit of the substrate (ref. 1). This conclusion is based on the understanding that hydrogen bonds are more highly directional as compared to dispersion forces, and also have electrostatic properties which could assist in short range attraction of a polar substrate. It is worth indicating that the hydroxyl groups are also the most accessible and highly stereospecific functional groups of saccharides.

Figs. 2 to 5 show the extensive and precise hydrogen bonding between proteins and sugars. As we have further discovered, hydrogen-bonding interactions provide the only means whereby the periplasmic binding proteins are able, in a simple but precise way, to bind both anomeric sugars and specific epimers (see above). It is inconceivable how hydrophobic or non-polar interactions could play a role in these subtle binding specificities.

The importance of hydrogen bonds for specificity was made clear soon after the complexes of lysozyme with saccharides (ref. 30) were analyzed at high resolution twenty years ago (ref. 18). This and other structural data and the detailed picture presented and discussed here and elsewhere (ref. 1) provide further strong support that hydrogen-bonding interactions are mainly responsible for specificity.

As a very simple view, the essence of Fischer's 'lock and key' view of protein-substrate interactions, enunciated 104 years ago, remains essentially intact (ref. 26). However, in light of the knowledge that has accumulated in the past two decades, this view, which was based coincidentally on studies of the selectivity of a certain 'carbohydrase' toward certain glucosides, has taken an infinitely clearer definition. Through X-ray crystallographic analysis, we are now able to actually 'see' the atomic interactions between proteins and ligands. Moreover, the 'induced fit' model proposed by Koshland and co-workers (ref. 27) has emerged as a more general and appropriate description of protein-ligand interactions in some proteins and enzymes. While conveying some elements of the 'lock and key', many aspects of the complexes between the periplasmic sugar-binding proteins and their substrates are more consistent with the 'induced fit' idea. In these complexes, there is exact complementarity between a monopyranoside and the binding site (which is the essence of a lock and key), despite the involvement of essential residues poised in both domains and the requirement for a large ligand-induced conformational change in order to bring these residues simultaneously to their correct positions for maximum interaction (the essence of induced fit) and in order to sequester and dehydrate the substrate and many of the residues in the cleft between the two domains.

There has also been several studies which relied heavily on the use of synthetic analogues of oligosaccharides and glycoconjugates as probes of sugar-binding sites and the features associated with ligand binding (e.g., see refs. 23, 28, 29). Some conclusions and ideas have emerged from these studies (for examples, see ref. 24). The long-standing idea of anomeric effect (ref. 24) has contributed considerably to understanding the properties of sugars and of enzyme catalyzed reactions (e.g., phosphorylase activity).

Some of these ideas, having neglected to take into account the wealth of data on protein structures and function, provide little, if any, understanding of protein-carbohydrate interactions. Serious flaws in arriving at these ideas are that the models of the structures of the oligosaccharide analogues (refs. 23, 28, 29) may not be entirely correct and/or that these analogues have different, yet unknown structures when bound to the protein sites. In this conjunction, it has been pointed out (ref. 1) that, although the substrates of taka-amylase and phosphorylase are oligoglucoses which are predicted to bind in a ribbonlike configuration, the bound substrates exhibit different configurations, each following the natural curvature of the binding-site groove in each enzyme. The oligosaccharide bound to taka-amylase has the OH3 of one glucose unit hydrogen-bonded to O5 of the next sugar (ref. 19), whereas the oligosaccharide bound in the storage site of phosphorylase has been shown crystallographically to be in the preferred left-handed helical conformation with the OH2 of one glucose unit hydrogen-bonded to the OH3 of the adjacent sugar (ref. 20).

There is also little, if any, validity to the idea of 'hydrated polar group gate effect' in the formation of protein-carbohydrate complex and the implication that 'oligosaccharides may play roles as biological messengers through the provision of a specific driving force for the organization of a protein into a specific conformation' (ref. 24). (The 'gate effect' is to be contrasted with the 'induced fit model', which has received considerable evidence and support from a variety of studies on different protein or enzyme systems.) There is just simply no polar gate (meaning, a distinct separate site on the protein which is totally polar) to be unlocked by the key (the saccharide substrate) through interactions solely between polar groups, nor completely distinct hydrophobic site or sites to be finally drawn (via a conformational change) to the nonpolar part of the initially bound saccharide. (Also sugars are not known to be factors or to assist in protein folding.) Structural studies, such as those of the periplasmic sugar-binding proteins, show that both polar and non-polar residues participate simultaneously in binding of amphipathic carbohydrates. Moreover, each type of residues is not completely segregated or grouped into one or several discrete segments along the polypeptide chain. Thus, in disagreement with the gate effect, there is no possibility for each type of groups of residues to interact with the same type of the sugar, through conformational rearrangements, in a sequential order — polar interactions first, followed finally by a succession of interactions involving each separate groups of non-polar residues. Indeed, what is typical of ligand-binding sites is that the polar and non-polar residues confined in these sites are not close together in the sequence nor completely segregated from each other in the folded protein (e.g., see Figs. 2 to 5). In the periplasmic sugar-binding protein, these residues are concentrated in both domains, near the interface of the cleft and poised for binding. The absence of several or even any structural intermediates in the kinetic analysis of saccharide binding to proteins (e.g., see TABLE 1 and refs. 1, 7) is inconsistent with the idea of a polar gate effect, with its implied sequential participations of distinct groups of residue types in sugar binding.

The polar gate effect is also not in accord with many observations that binding of ligands to proteins is often not accompanied by protein conformational change. Moreover, in those cases where ligand-induced conformational changes have been demonstrated, these changes have been shown to serve one dominant role: to bring into position and/or optimize the orientation of functional groups (polar and non-polar together) for binding and/or catalysis. This role is the essence of the induced fit model. Conformational changes can be broadly divided into categories — local minor perturbations which affect only residues in the binding site and large changes such as the relative movement between two domains (refs. 1, 5, 17). It is noteworthy that, although the two lobes of the periplasmic sugar-binding proteins participate in binding through hydrogen bonds and van der Waals contacts, the stereospecificity is confined first and foremost to one lobe (refs. 4, 10).

There are data indicating an increase in binding affinity after removal of hydroxyl groups of sugar substrates (for example see refs. 24, 28, 29). In light of our analysis as discussed above, this enhance binding of deoxy-sugar analogues could very well be due to the removal of a hydroxyl group which in the normal sugar is unpaired in the bound state, rather than, as suggested (ref. 24), to an increase in non-polar interactions.

The sugar (and also phosphate) groups in DNA

Because of its obvious, vital importance, the interaction between protein and DNA or RNA has received considerable attention. This interest has been heightened recently by the determination of the tertiary structures of several DNA-binding proteins and, in some cases, of the proteins complexed with DNA fragments.

Two types of interactions occur in protein-DNA complex: base sequence-dependent or specific and non-specific interactions. Hydrogen bonding is an important factor in both interactions and in contributing the requisite affinity of the complex. Specific interaction is accomplished by hydrogen-bonding and stacking interactions between residues and specific bases.

In both specific and non-specific types of interactions, hydrogen bonding through the sugar and phosphate moieties is the common mode of interaction between polynucleotides and proteins. An important inherent feature of this mode of interaction is that it does not necessarily depend on any structural motifs (e.g., 'helix-loop-helix', 'zinc fingers', etc.) of the DNA binding proteins. Hydrogen bonding with the polynucleotide backbone's sugar and phosphate groups could very well also play a major role in non-specific binding. The hydrogen bonds with the sugar groups will undoubtedly have features similar to those described in this paper. Furthermore, we have also described the major factors associated with hydrogen bonding between proteins and charged ligands (either anionic, cationic or zwitterionic), and on the basis of these factors, proposed a general mechanism by which uncompensated charges on these ligands are stabilized (ref. 12). Interestingly, hydrogen bonding with uncompensated, negatively charged groups (e.g., phosphate, sulfate, etc.) also utilizes primarily NH moieties of proteins. However, the NH groups used in hydrogen bonding sugars almost always come from side chain residues (see above), but the NH groups used in hydrogen bonding uncompensated charged ligands come principally from backbone peptide units, which are not necessarily located at the N-termini of helices. These backbone peptide units have also been implicated to play a major factor in stabilizing charges on uncompensated, sequestered charged groups (ref. 12). We have also noted that hydroxyl groups of serine and threonine residues are frequently hydrogen-bonded to charged groups, especially to negative groups such as phosphates, sulfates, etc. (ref. 12).

A good example of the importance of hydrogen bonds, and the pattern of involvement of the various groups of DNA in hydrogen bonding is seen in the well-refined 2 Å structure of a complex of DNase with a nicked DNA octanucleotide (ref. 31). In this complex, there are fourteen hydrogen bonds and, surprisingly, only one salt-link and one stacking interaction. Of the fourteen hydrogen bonds, four are with four bases, six with five phosphate groups, and four with three deoxyribose groups. The groups forming hydrogen bonds with phosphates include three backbone peptide NH groups, two side chain hydroxyls, and one NH of an Asn side chain; none of these protein groups are close to the N-termini of helices.

The 3 Å crystal structure of the complex between Eco RI endonuclease and the cognate oligonucleotide provides a detailed example of the molecular basis of sequence-specific DNA-protein interactions (ref. 32). This complex, unlike in the DNase-oligonucleotide complex, shows more hydrogen bonding between side chain residues and bases — at least 12 hydrogen bonds, of which many are bidentates. There are also a number of hydrogen bonds with the backbone. Note that the DNase and Eco RI do not have the same structural motif, yet there is similarity in the mode of interaction of these proteins with the DNA backbones. The pattern of interactions in the DNase-DNA complex — dominated by hydrogen bonding with the ribose-phosphate polynucleotide backbone — might very well be a good example of non-specific interactions.

In light of this brief consideration, one might ask: "Is it possible that the primordial DNA-protein interaction is mediated primarily through hydrogen bonds between the bi-functional, dipolar peptide backbone groups and the base, phosphate, and ribose groups of DNA? Could the DNA-binding protein have been as simple as a polymer of glycine residues or a copolymer of simpler amino acids capable of only forming hydrogen bonds and van der Waals contacts with DNA?"

EPILOGUE

Twenty three years ago, the determination of the tertiary structures of lysozyme and its complexes with a variety of saccharides by Phillips and co-workers (refs. 30, 33) not only provided the first glimpse of the binding modes of substrates, but also ushered in X-ray crystallography as the only method for obtaining a detailed picture of the atomic interactions in protein-ligand complexes. Structures of other complexes of proteins or enzymes with carbohydrates have also been determined, providing a much clearer view (ref. 1). The highly refined 1.7 Å resolution structure of the L-arabinose-binding protein-sugar complex, which appeared four years ago (ref. 3), remain unsurpassed in terms of detail and accuracy, and in revealing many, if not all, of the molecular features of protein-carbohydrate interactions. These attributes of the complex, as well as those of other recently refined structures of GBP-glucose and MBP-maltose complexes, have paved the way for a thorough understanding of these interactions and the design and analysis of other structure-function studies (e.g., site-directed mutagenesis, ligand binding kinetics, molecular dynamics, etc.) currently underway in our laboratory. We have also obtained equally detailed results on the role of one or two ordered water molecules (e.g., see Figs. 2 and 4) and local conformational change in conferring substrate specificity. Notwithstanding these recent developments, the knowledge that has accrued and discussed in this and two previous reviews (refs. 1, 2) provides a solid base for the analysis of other protein-saccharide complexes (for example, the influenza virus hemagglutinin complexed with its receptor sialic acid (ref. 34)).

Acknowledgements

I am indebted to past and current co-workers, especially Dave Miller, Gary Gilliland, Marcia Newcomer, Nand & Mina Vyas, John Spurlino, and Dave Wilson for their immense contributions. I thank Jack Sack, Denise Taylor, John Spurlino, Nand Vyas, Tim Reynolds and Connie Wallace for their help in preparing this manuscript. Work from this laboratory was supported by the Howard Hughes Medical Institute and grants from the National Institutes of Health and the Welch Foundation. This paper is dedicated to Dr. Frank. T. Wilson.

REFERENCES

1. F. A. Quicho, *Annu. Rev. Biochem.* **55**, 287-315 (1986).
2. F. A. Quicho, *Current Topics in Microbiology and Immunology* **139**, 135-148 (1988).
3. F. A. Quicho and N. K. Vyas, *Nature* **310**, 381-386 (1984).
4. N. K. Nand, M. N. Vyas and F. A. Quicho, *Science*, in press.
5. M. E. Newcomer, B. A. Lewis and F. A. Quicho, *J. Biol. Chem.* **254**, 13218-13222 (1981).
6. D. M. Miller, III, J. S. Olson and F. A. Quicho, *J. Biol. Chem.* **255**, 2465-2471 (1980).
7. D. M. Miller, III, J. S. Olson, J. W. Pflugrath and F. A. Quicho, *J. Biol. Chem.* **258**, 13665-13672 (1983).
8. H. Fukada, J. M. Sturtevant and F. A. Quicho, *J. Biol. Chem.* **258**, 13193-13198 (1983).
9. J. W. Pflugrath and F. A. Quicho, *J. Mol. Biol.* **200**, 163-180 (1988).
10. J. S. Sack, M. A. Saper and F. A. Quicho, *J. Mol. Biol.*, in press.
11. F. A. Quicho, N. K. Vyas, J. S. Sack and N. N. Vyas, *Cold Spring Harbor Symp. Quant. Biol.* **LII**, 453-463 (1987).
12. F. A. Quicho, J. S. Sack and N. K. Vyas, *Nature* **329**, 561-564 (1987).
13. N. K. Vyas, M. N. Vyas, and F. A. Quicho, *Nature* **327**, 635-638 (1987).
14. I. A. Rose, K. R. Hanson, K. D. Wilkinson and M. J. Wimmer, *Proc. Natl. Acad. Sci. USA* **77**, 2439-2441 (1980).
15. J. E. Del Bene and J. A. Pople, *J. Chem. Phys.* **58**, 3605-3608 (1973).
16. G. A. Jeffrey and L. Lewis, *Carbohydr. Res.* **60**, 179-182 (1978).
17. B. Mao, M. R. Pear, J. A. McCammon and F. A. Quicho, *J. Biol. Chem.* **257**, 1131-1133 (1982).
18. C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, *Proc. R. Soc. London Ser. B* **167**, 378-88 (1967).
19. Y. Matsuura, M. Kusunoki, W. Harada and M. Kakudo, *J. Biochem.* **95**, 697-702 (1984).
20. E. Goldsmith and R. J. Fletterick, *Pure Appl. Chem.* **140**, 577-588 (1983).
21. C. H. Chotia, *Nature* **248**, 338-339 (1974).
22. P. D. Ross and S. Subramanian, *Biochemistry* **20**, 3096-3102 (1981).
23. E. A. Kabat, J. Liao, M. H. Burzynska, T. C. Wong, J. Thorgersen and R. U. Lemieux, *Mol. Immunol.* **18**, 873-881 (1981).
24. Lemieux, R. U., In "Proceedings of the Seventh International Symposium on Medicinal Chemistry", August 27-31, 1984, Uppsala, Sweden, Swedish Pharmaceutical Society, Vol. 1, pp. 329-351.
25. W. P. Jencks, *Catalysis in Chemistry and Enzymology*. McGraw-Hill Book Co., pp 352, 399 (1969).
26. E. Fischer, *Ber. Dtsch. Chem. Ges.* **27**, 2985-2993 (1884).
27. D. E. Koshland, Jr., G. Nemethy and D. Filmer, *Biochemistry* **15**, 365-385 (1966).
28. D. R. Bundle, In "Protein-Carbohydrate Interactions in Biological Systems", D. L. Lark, ed. pp. 165-171, Academic Press, London (1986).
29. K. Bock and H. Pedersen, In "Protein-Carbohydrate Interactions in Biological Systems", D. L. Lark, ed., pp. 173-182, Academic Press, London (1986).
30. L. N. Johnson and D. C. Phillips, *Nature* **206**, 761-763 (1965).
31. D. Suck, A. Lahm and C. Oefner, *Nature* **332**, 464-468 (1988).
32. J. A. McClarin, C. A. Frederick, B-C. Wang, P. Greene, H. W. Boyer, J. Grable and J. M. Rosenberg, *Science* **234**, 1526-1541 (1986).
33. C. C. F. Blake, D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, *Nature* **206**, 757-761 (1965).
34. W. Weis, J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel and D. C. Wiley, *Nature* **333**, 426-431 (1988).