Crystallographic Analysis of the Epimeric and Anomeric Specificity of the Periplasmic Transport/Chemosensory Protein Receptor for D-Glucose and D-Galactose†,‡

Meenakshi N. Vyas,*,§ Nand K. Vyas,§ and Florante A. Quiocho§,∥

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

Received November 30, 1993; Revised Manuscript Received February 16, 1994®

ABSTRACT: The D-glucose/D-galactose-binding protein (M_r = 33 000) found in the periplasm of bacterial cells serves as the primary high-affinity receptor of active transport for and chemotaxis toward both sugar epimers. This protein from *Escherichia coli* binds D-glucose with a K_d of 2×10^{-7} M, which is about 2 times tighter than D-galactose. The 2.0-Å resolution crystal structure of the binding protein complexed with D-galactose has been refined to a crystallographic *R*-factor of 0.167. This structure, combined with that previously refined for the complex with D-glucose [Vyas, N. K., Vyas, M. N., & Quiocho, F. A. (1988) *Science 242*, 1290–1295], provides understanding, in atomic detail, of recognition of sugar epimers and anomers. In the two complex structures, the sugar ring is positioned identically in the binding site, and each hydroxyl group common to both is involved in very similar cooperative hydrogen-bonding interactions with protein residues and ordered water molecules. Only the β-anomer of both monosaccharides is bound, with Asp154 OD1 primarily responsible for accepting a hydrogen bond from the anomeric hydroxyl. Recognition of both sugar epimers is accomplished principally by hydrogen bonding of Asp14 OD1 with the equatorial OH4 of D-glucose and OD2 with the axial OH4 of D-galactose. These results are reconciled with equilibrium and fast kinetics data, which indicate binding of both anomers of the two sugars, and further compared with sugar recognition by other periplasmic sugar-binding proteins with specificities for arabinose/galactose/fucose, maltooligosaccharides, and ribose.

In order to survive, bacteria use a battery of different transport systems. Their versatility in translocating a wide range of nutrients is exemplified by the set of transport systems or permeases that require a group of more than 2 dozen binding proteins found in the periplasmic space of Gram-negative bacteria [reviewed in Furlong (1987)]. All binding proteins serve as initial receptors for nutrients such as amino acids, carbohydrates, oxyanions, peptides, and vitamins. Five of the binding proteins also act in a similar fashion for chemotaxis [reviewed in Macnab (1987)]. This family of proteins is proving to be an excellent system for crystallographic analysis of details of ligand molecular recognition.

In recent years, our laboratory has been engaged in the determination of the high-resolution structures of eight binding proteins with specificities for L-arabinose/D-galactose/D-fucose, D-glucose/D-galactose, linear and cyclic maltodextrins, leucine/isoleucine/valine, leucine, histidine, phosphate, and sulfate [reviewed in Quiocho (1991); several unpublished data]. Despite the lack of sequence homology and uniform size, the structures of these proteins are very similar. They consist of two similar globular domains separated by a cleft or groove in which the ligand is bound.

Focusing on the three sugar-binding proteins underscores the versatility of these proteins in molecular recognition of carbohydrates. The D-glucose/D-galactose-binding protein (GGBP)[1] recognizes both C4 sugar epimers (Anraku, 1968; Miller *et al.*, 1980). How this is accomplished at the atomic level is a major topic of this paper. The L-arabinose-binding protein (ABP) also binds D-galactose and D-fucose with affinities that are 2- and 40-fold lower, respectively, than L-arabinose (Hogg & Englesberg, 1969; Miller *et al.*, 1983). Determination of the structures of complexes of ABP with the three monosaccharides at better than 1.8-Å resolution revealed the key role of bound water molecules in molecular recognition (Quiocho & Vyas, 1984; Quiocho *et al.*, 1989). The maltodextrin- or maltose-binding protein (MBP) binds linear and cyclic dextrins (Ferenci, 1980; Miller *et al.*, 1983). This is reflected in the long binding-site groove seen in the 1.7-Å resolution refined structures of MBP complexed with maltose and β-cyclodextrin (Spurlino *et al.*, 1991, 1992; Sharff *et al.*, 1993), as well as with maltotriose and maltotetraose (L. E. Rodseth, J. C. Spurlino, and F. A. Quiocho, unpublished data). These very high resolution structures have contributed immensely to understanding the atomic interactions between proteins and carbohydrates (monosaccharides, oligosaccharides, and cyclic dextrin) (Vyas, 1991; Quiocho, 1993).

Despite diverse ligand specificities, a remarkable common feature of the sugar-binding proteins, indeed of all periplasmic binding proteins, is the similar high affinity of these proteins for their ligands—K_d values in the micromolar range (Miller *et al.*, 1983; Furlong, 1987). The kinetics of substrate binding to several binding proteins, determined by the stopped-flow rapid-mixing technique, also shows similar reaction rate properties (Miller *et al.*, 1980, 1983).

† This work was supported in part by grants from the NIH, Welch Foundation, and Keck Foundation.

‡ Crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank (reference number 1GLG).

* Author to whom correspondence should be addressed.

§ Department of Biochemistry.

∥ Howard Hughes Medical Institute and Department of Molecular Physiology and Biophysics.

® Abstract published in *Advance ACS Abstracts*, April 1, 1994.

[1] Abbreviations: ABP, L-arabinose-binding protein; GGBP, D-glucose/D-galactose-binding protein; MBP, maltose- or maltodextrin-binding protein; RBP, D-ribose-binding protein; Gal, D-galactose; Glc, D-glucose; Rib, D-ribose.

0006-2960/94/0433-4762$04.50/0

Table 1: Data Collection and Refinement Statistics for the Structures of Complexes of GGBP with D-Glucose and D-Galactose

data	sugar complex		target (σ)
	D-glucose	D-galactose	
diffraction (one crystal used in each case)			
resolution (Å)	1.9	2.0	
total reflections recorded	94000	19700	
R-merge	0.076	0.059	
obsd data in last shell (%)	70 (1.97–1.90 Å)	44 (2.07–2.00 Å)	
PROLSQ refinement			
R-factor[a]	0.166	0.167	
overall *B* value (Å^2)	20.06	17.94	
rms coordinate shift (Å)	0.007	0.004	
rms *B* shift (Å^2)	0.12	0.11	
average ($\|F_o\| - \|F_c\|$)[b]	28.63	32.07	14, −70
bond distances (Å)	0.021	0.014	0.015
angle distances (Å)	0.041	0.038	0.025
planarity (Å)	0.012	0.010	0.010
chiral volume (Å^3)	0.072	0.055	0.050
angle ω (deg)	9.2	8.5	3.0
no. of atoms			
protein	2348	2348	
D-glucose/D-galactose	12	12	
calcium	1	1	
water oxygens	272	215	

[a] *R*-factor ($=\sum\|\|F_o\| - \|F_c\|\|/\sum|F_o|$). [b] The structure factors were weighted in the form $\sigma = A - B(\lambda - {}^1/_6) \sin\theta$.

GGBP has the distinction of being the first binding protein shown to be the primary receptor for both active transport and chemotaxis (Anraku, 1968; Hazelbauer & Adler, 1971). As in all binding protein-dependent processes of active transport and chemotaxis, each process requires other protein components lodged in the cytoplasmic membrane of the Gram-negative bacterium (Furlong, 1987; Macnab, 1987; Ames, 1986).

Examining the sugar reaction rates of GGBP, as well as those of ABP and MBP, Miller *et al.* (1980, 1983) showed that the α- and β-anomers of the respective sugar ligands bind at equal rates. The high-resolution refined structures of complexes of ABP with L-arabinose, D-galactose, and D-fucose revealed a similar mode of binding in the same site of both anomers of each sugar (Quiocho & Vyas, 1984; Quiocho *et al.*, 1989).

Herein, we report the determination of the structure of *Escherichia coli* GGBP complexed with Gal. Comparing this structure with that previously determined for the complex with Glc (Vyas *et al.*, 1988) has enabled us to understand the features associated with sugar recognition.

EXPERIMENTAL PROCEDURES

Protein Purification and Crystallization. GGBP from *E. coli* was purified according to a procedure described by Miller *et al.* (1980). GGBP, like many other purified binding proteins examined subsequently, contains bound endogenous D-glucose (Miller *et al.*, 1980; Quiocho, 1991). This was confirmed following the determination of the structure of GGBP (Vyas *et al.*, 1983, 1988). In order to obtain crystals of the GGBP–Gal complex, we initially attempted to replace the endogenously bound Glc from freshly grown single crystals of the GGBP–Glc complex by soaking them in a solution containing 1 mM Gal. This was completely unsuccessful as evident by the presence of bound Glc in the structure of the soaked crystal.

It was therefore necessary to remove Glc from purified GGBP, form the GGBP–Gal complex, and crystallize the complex. Removal of bound sugar was achieved by treating purified GGBP with 3 M guanidine hydrochloride using a regimen similar to that developed by Miller *et al.* (1980). The sugar-free GGBP was then dialyzed against 1 mM D-galactose (Pfanstiehl Laboratory; sugar used without further purification) in 10 mM sodium citrate, pH 5.1, and 0.02% NaN_3. The GGBP–Gal complex was crystallized using a procedure similar to that described for the GGBP–Glc complex crystals (Quiocho & Pflugrath, 1980; Vyas *et al.*, 1983). Crystals of both complexes have $P2_1$ space group symmetry with almost identical unit cell dimensions. The cell dimensions of GGBP–Gal crystals are a = 65.98 Å, b = 37.09 Å, c = 61.90 Å, and β = 105.83° as compared to those of the GGBP–Glc crystals (a = 66.0 Å, b = 37.07 Å, c = 61.67 Å, and β = 106.67°). The asymmetric unit in both complex crystals contains one protein molecule. The GGBP–Gal crystals exhibit mainly thin-plate morphology with asymmetric hexagonal surfaces whereas the GGBP–Glc crystals show thick-plate morphology.

Diffraction Data Collection. The GGBP–galactose crystals were characterized using diffraction data collected on a San Diego Multiwire two-area detector system (Hamlin, 1985) mounted on an RU-200 rotating anode (Cu Kα) equipped with a monochromator and operated at 110 mA and 40 kV. One crystal was used to collect diffraction data to 2.0-Å resolution, which is 67% complete. The data were reduced with the Howard *et al.* (1985) software package. The statistics of data collection and data reduction are shown in Table 1. As the crystals of the GGBP–Gal complex are thin plates, the percent data in the last annulus are lower than those for the GGBP–Glc complex (Table 1).

Structure Determination and Refinement. As the GGBP–Gal and GGBP–Glc crystals are isomorphous, the structure determination of the former complex was straightforward. The initial calculated phases (α_c) were obtained from the GGBP–Glc structure refined to 1.9-Å resolution (Vyas *et al.*, 1988) excluding the contribution of the bound β-D-glucose and solvent molecules. These phases were then combined with the observed structure factor amplitudes of the GGBP–Gal crystal in order to calculate ($|F_o| - |F_c|$, α_c) and ($2|F_o| - |F_c|$, α_c) Fourier maps. After being established from these maps that the bound sugar was indeed D-galactose, the protein structure, along with the fitted model of the β-D-galactose and water molecules, was refined using PROLSQ (Hendrickson, 1985) with the PROFFT feature (B. C. Finzel, Upjohn Co.). Appropriate restraints for α- and β-D-galactose were incorporated in the restraints dictionary on the basis of

their known small molecule X-ray structures (Sheldrick, 1976). The coordinates of the C3 atom of the crystal structure of β-D-galactose were corrected due to a possible error in the published coordinates (Sheldrick, 1976). In the last few cycles of refinement, contact restraints between sugar and protein were released in order to avoid biasing the interactions between GGBP and the sugar. Despite terminating these restraints, the refinement remained well behaved and the interactions essentially unchanged from those seen in the restrained condition. The coordinates of the GGBP–Gal complex have been deposited in the Protein Data Bank.

GGBP–Glc Complex Structure. The structure of the complex was previously determined and refined at 1.9-Å resolution using intensity data collected from several crystals using a four-circle diffractometer (Vyas *et al.*, 1988). We have since refined the structure against a 1.9-Å data set collected from a single crystal on the area detector using the procedure described above for the refinement of the GGBP–Gal complex structure (Table 1). This structure is extremely similar to that determined originally (Vyas *et al.*, 1988).

RESULTS AND DISCUSSION

The determination of the 2.0-Å structure of the GGBP–Gal complex reported herein, together with that of the 1.9-Å structure of the GGBP–Glc complex, provides a detailed understanding of the atomic features that confer sugar specificity. The structure of the GGBP–Gal complex, like that of the complex with Glc, is well refined (Table 1). Of significance to further discussion on the details of the atomic interactions between GGBP and sugars, it is important to note that the electron densities for the protein residues as well as the bound sugar in the binding site region are well resolved. This observation, together with the very low temperature factors of atoms in this region (about 6 Å^2 relative to 18 Å^2 for the entire structure), shows that this region is very well ordered and has coordinate errors significantly smaller than the overall error of the entire refined structure of about 0.2 Å obtained from a Luzzati plot. Similar features are observed in the refined 1.9-Å structure of the GGBP–Glc complex (Table 1) [see also Vyas *et al.* (1988)]. The averaged *B* values of the bound D-glucose and D-galactose are 6.90 and 6.62 Å^2, respectively.

Fitting of the D-galactose to electron density maps calculated with ($|F_o| - |F_c|$, α_c) and ($2|F_o| - |F_c|$, α_c), first carried out after structure refinement of the protein and bound water was nearly completed (*R*-factor = 0.17), clearly showed binding of only the β-anomer of the monosaccharide. This finding was confirmed in the ($|F_o| - |F_c|$) map contoured at 3.5σ (Figure 1A) and the ($2|F_o| - |F_c|$) map contoured at 1σ (not shown) calculated with structure factors from the final least squares cycle with the contribution of the Gal molecule omitted. The structure of the GGBP–Glc complex refined at 1.9-Å resolution also revealed binding of only the β-anomer of Glc (Figure 1B) (Vyas *et al.*, 1988). Mowbray *et al.* (1990) and Zou *et al.* (1993) also observed only β-anomers bound in the structures of *Salmonella typhimurium* GGBP.

The polypeptide chain conformations of the structures of complexes of GGBP with Gal and Glc are very similar. Superpositioning of the α-carbon backbones of both structures indicated a root-mean-square deviation of 0.2 Å. The binding site regions in both structures, including the identical atoms of the bound sugars, superimposed almost exactly (Figure 2). The hydrogen-bonding interactions associated with sugar binding are shown in Table 2. The bound Glc also makes 57 van der Waals contacts (of less than 4 Å) with a mean distance of 3.71 Å, and Gal forms 47 contacts with a mean distance of 3.55 Å. Several of these contacts are due to the sandwiching of each sugar by Phe16 and Trp183 (Figure 2). This stacking of aromatic residues against the sugar faces is observed in all but two other known structures of complexes of proteins and enzymes with carbohydrates (Vyas, 1991; Quiocho, 1993).

Table 2: Hydrogen Bonds (≤3.5 Å) Associated with Monosaccharide Binding to GGBP

atom		distance	
protein/water	sugar	β-D-glucose	β-D-galactose
Asn91 ND2	O1	3.44	3.25
Asp154 OD1		2.64	2.66
Arg158 NH2		3.22	3.35
Asn256 ND2		3.06	3.13
Arg158 NH1	O2	2.78	2.65
Arg158 NH2		3.22	3.40
Asp236 OD1		3.39	2.70
Asp236 OD2		2.60	2.42
Asn25 ND2		3.43	(3.60)[a]
Asn211 ND2	O3	3.02	2.88
Asp236 OD1		2.64	2.79
Wat313 O		2.91	2.90
Wat314 O		3.38	(3.66)
Asp14 OD1	O4	2.65	2.86
Asp14 OD2		3.50	3.23
Wat313 O		3.50	(3.78)
Asn91 ND2	O5	3.02	2.88
Asn91 OD1	O6	2.62	2.55
His152 NE2		2.77	2.79

[a] Values exceeding the limit of 3.5 Å are in parentheses.

As can be seen in Figures 1 and 2 and Table 2, an identical residue (Asp14) is principally involved in interacting with the OH4 of both epimers. The orientation of the carboxylate of Asp14, which is virtually identical in both complex structures, makes OD2 the principal acceptor of a hydrogen bond from the axial OH4 of Gal and OD1 the major acceptor from the equatorial OH4 of Glc. Gal OH4 and Glc OH4 also make weak hydrogen bonds with OD1 and OD2 of Asp14, respectively. The epimeric hydroxyl of Glc further forms a hydrogen bond with a water molecule (Wat313) which in turn is hydrogen bonded to the main-chain NH of Phe16 (Table 2 and Figure 2). This hydrogen bond if present in the Gal binding is very weak.

The affinity of GGBP for Glc, determined by equilibrium and reaction kinetics techniques, is 2-fold tighter than for Gal (Miller *et al.*, 1980). This is reflected in the results of ligand-binding competition studies (Miller *et al.*, 1980) and structure determination (Vyas *et al.*, 1988) which showed that GGBP, as purified, has bound endogenous Glc exclusively. The data shown in Table 2 provide a likely structural basis for the difference in ligand affinities between the two monosaccharide substrates; the hydrogen bonding, as well as the van der Waals interactions (not shown), associated with Glc binding is more favorable overall than that with Gal binding. Whereas the distance of all the hydrogen bonds to the glucose are within the limit, three hydrogen bonds to galactose exceed the limit (Table 2). Moreover, there are 57 van der Waals contacts (≤4 Å) to the bound Glc, 10 more than to the bound Gal.

As the Asp154 residue makes the shortest hydrogen bond with the β-anomeric hydroxyl, it is the principal residue involved in the recognition of the hydroxyl and potentially of the α-anomeric hydroxyl (discussed below) of both sugar epimers (Figures 1 and 2 and Table 2). (The other hydrogen bonds associated with the anomeric hydroxyls all exceed 3 Å.) In contrast to the epimeric sugar recognition by Asp14, the conformation of Asp154 permits the major involvement

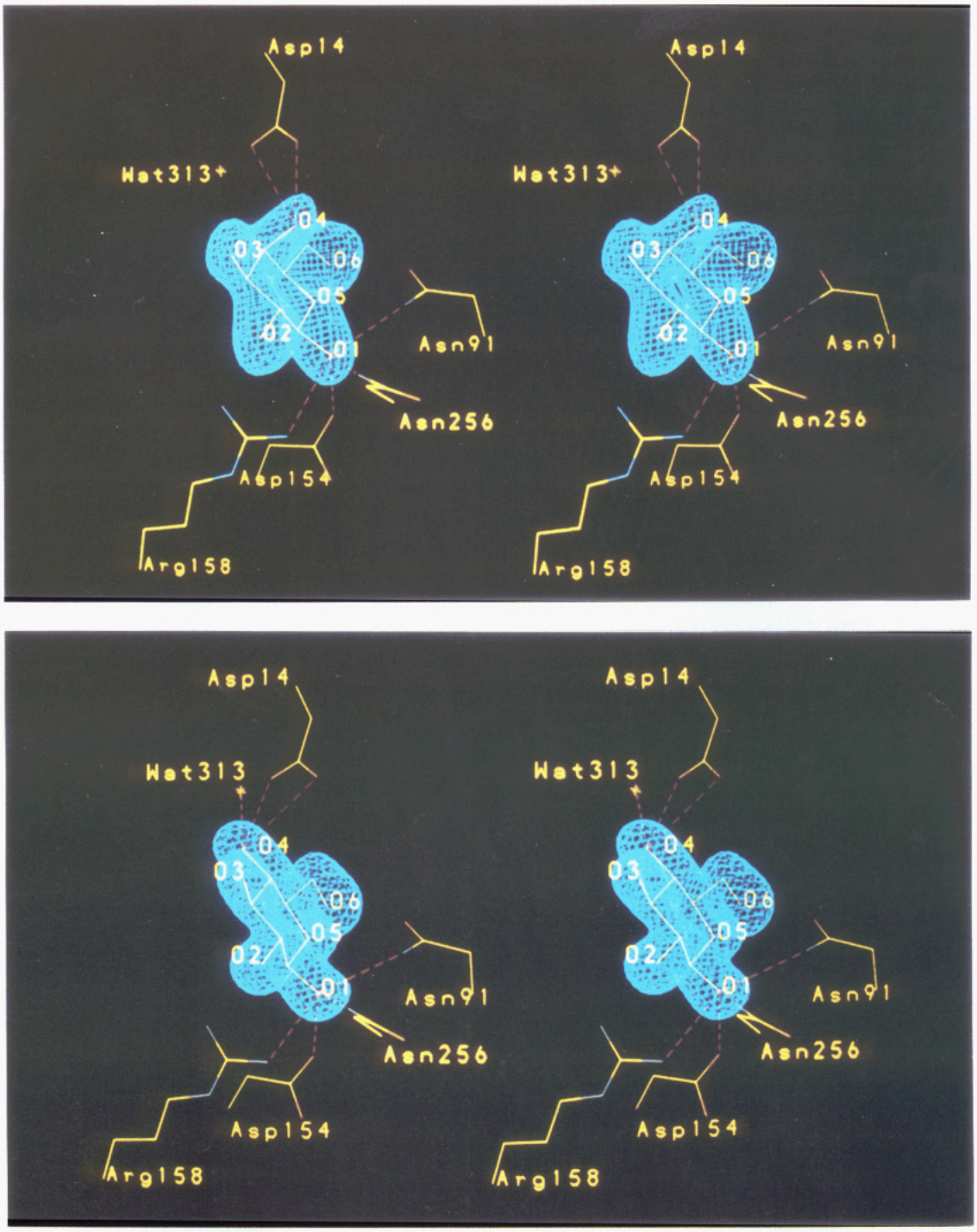

FIGURE 1: Stereoview of the difference Fourier maps of β-D-galactose at 2-Å resolution (A, top) and β-D-glucose at 1.9 Å (B, bottom) bound to GGBP. The maps were calculated using coefficients ($|F_o| - |F_c|$) and α_c calculated phases from the refined 2-Å Gal-bound and 1.9-Å Glc-bound structures. The calculated structure factors (F_c) include contributions from all refined atoms with the exception of the sugar atoms which were excluded. The electron density maps are contoured at a 3.5σ level and a grid spacing of 0.25 Å. The sugar difference density fits the $^4C_1$ conformer of only the β-anomer of both monosaccharides. No other electron density is seen in the binding site. For emphasis, only hydrogen-bonding interactions (red dashed lines) involving O1 and O4 hydroxyls of both sugars with protein residues and a bound water molecule are shown. Atoms are color coded as follows: carbon, yellow; nitrogen, blue; oxygen, red.

of only one carboxylate oxygen atom in the anomeric hydroxyl recognition.

The preferential binding of one anomer to GGBP, briefly noted above, is in contrast to previous results, showing the binding of both anomers of L-arabinose, D-galactose, and D-fucose to the L-arabinose-binding protein (Quiocho & Vyas, 1984; Quiocho *et al.*, 1989). As shown in Figure 3 and Table 3, one carboxylate oxygen of Asp90 of ABP makes the best hydrogen bonds to both anomeric hydroxyls. These results, together with the ligand-binding kinetic data which indicated binding of both anomers of sugar substrates to ABP, GGBP, and MBP (Miller *et al.*, 1980, 1983), weighed heavily in our prediction that all periplasmic sugar-binding protein structure will also show binding of both sugar anomers.

On the basis of modeling studies (Table 4), there is nothing in the geometry around the anomeric hydroxyl in both structure of complexes of GGBP with Glc and Gal that would prevent binding of the α-anomeric sugars because of steric reasons. Indeed, the modeled α-hydroxyl of both sugars, based on the β-anomer structures, could be easily accommodated, forming hydrogen bonds with OD1 of Asp154 and other groups albeit at longer distances than those actually found in the refined structures for the bound β-anomers (Table 4).

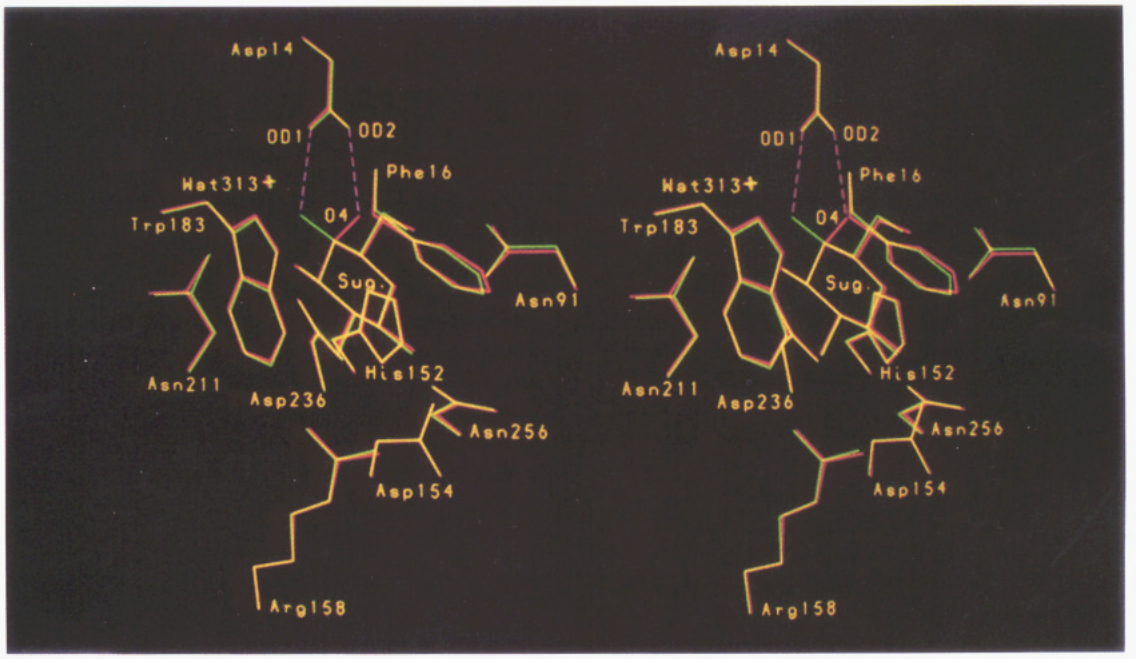

FIGURE 2: Superpositioning of the sugar binding site of the refined structures of complexes of GGBP with the β-anomer of D-galactose (red) and D-glucose (green). Since the unit cell parameters of crystals of both complexes differ very slightly (see Experimental Procedures), the superimposed identical atoms of both sugars served to align the two structures by the least squares method. Yellow-colored bonds are produced when both structures superimposed exactly. All C_α and solvent atoms are labeled with the three-letter amino acid code followed by residue number. The label Sug. refers to the common sugar binding location. Hydrogen-bonding interactions of OD1 and OD2 of Asp14 with Glc and Gal, respectively, are shown as dashes in magenta color. Hydrogen-bonding interactions of the two sugars with the GGBP differ only at the epimeric 4-hydroxyl.

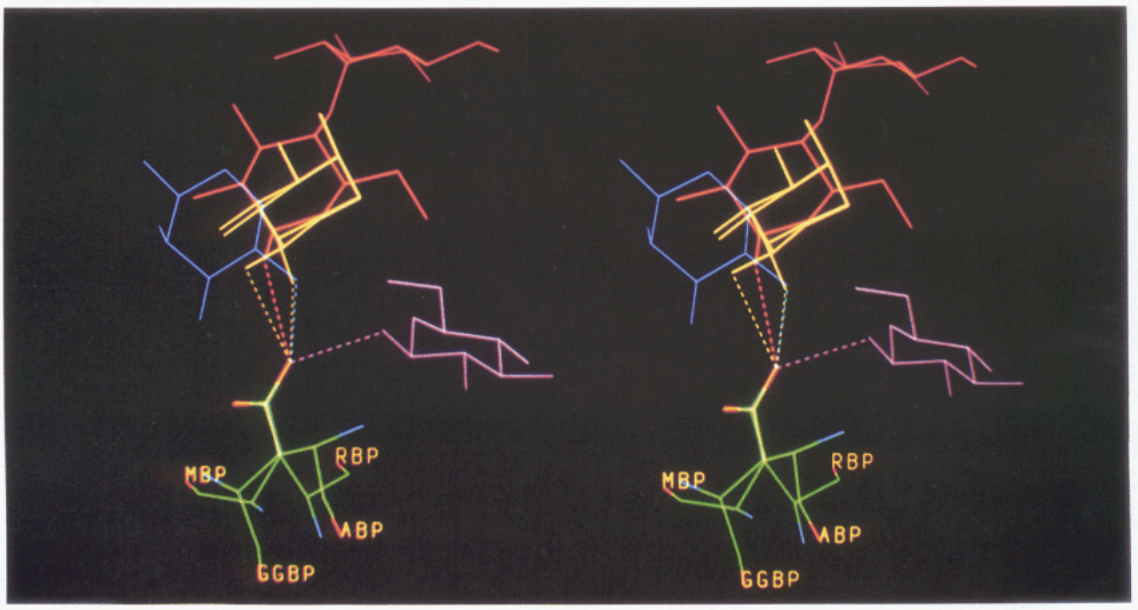

FIGURE 3: Stereoview of the geometry of the hydrogen-bonding interactions between the aspartate residues of ABP, GGBP, MBP, and RBP and the anomeric hydroxyls of the respective sugar substrate. The carboxylate group and CB atom of each Asp residue have been used as common atoms in the superimpositioning. Nitrogen, carbon, and oxygen atoms of Asp residues are colored blue, green, and red, respectively, whereas L-arabinose, D-glucose, D-ribose, and D-maltose are colored yellow, magenta, blue, and orange, respectively. The carbonyl oxygen atom of each Asp residue has been labeled to identify the binding protein. Hydrogen bonds between each Asp and the corresponding sugar's anomeric hydroxyl are shown as dash lines, each colored identical to that of the corresponding sugars. Both α- and β-anomers of Ara are shown in the same color.

Binding of only one sugar anomer has also been observed in the crystal structures of MBP determined at resolutions of 2.3 Å (Spurlino *et al.*, 1991) and 1.7 Å [J. C. Spurlino and F. A. Quiocho, unpublished data; see also Spurlino *et al.* (1992)] and RBP determined at 1.7-Å resolution (Mowbray & Cole, 1992). It is noteworthy that, in these two sugar-binding proteins, Asp residues are also strongly involved in interacting with the anomeric hydroxyls (Figure 3 and Table 3). Thus, in ABP, GGBP, and MBP the geometry of the Asp side chains responsible for binding the anomeric hydroxyl permits the participation mainly of one carboxylate oxygen (e.g., see Figure 3). This geometry is to be contrasted with that of the side chain of Asp14 of GGBP which is responsible for epimer recognition (Figures 1–3).

Superpositioning of the carboxylate side chains of the Asp residues in ABP, GGBP, MBP, and RBP (Figure 3) involved in the interactions with anomeric hydroxyls has further provided insight into the geometry of anomeric recognition. The anomeric hydroxyls of the sugars bound to ABP, MBP, and RBP are located close to each other. This location is different from that of the anomeric hydroxyl of D-glucose bound to GGBP. These two distinct locations of the anomeric

Table 3: Hydrogen-Bonding Interactions of the α- or β-Anomeric Hydroxyl of Sugars Complexed to ABP, GGBP, MBP, and RBP

	atom		hydrogen bond (≤3.5 Å)	
complex	protein	sugar	α-anomer	β-anomer
ABP-L-Ara[a]	Asp90 OD1	O1	2.77	2.74
	Asp90 OD2		3.45	3.48
	Lys10 NZ		3.27	3.14
GGBP-D-Glc[b]	Asn91 ND2	O1		3.44
	Asp154 OD1			2.64
	Arg158 NH2			3.22
	Asn256 ND2			3.06
GGBP-D-Gal[b]	Asp154 OD1	O1		2.66
	Arg158 NH2			3.35
	Asn91 ND2			3.25
	Asn256 ND2			3.13
MBP-Mal[c]	Asp14 OD1	O1	2.79	
	Lys15 NZ		3.09	
RBP-D-Rib[d]	Asp89 OD1	O1		2.70
	Arg90 NE			2.87

[a] Quiocho and Vyas (1984). [b] This work. [c] From the defined 1.7-Å resolution structure of MBP–maltose complex (J. C. Spurlino and F. A. Quiocho, unpublished data). [d] Mowbray and Cole (1992).

hydroxyls are reflected in the stereochemistry of the hydrogen-bonding interactions. Whereas the hydrogen-bonding interactions of the Asp carboxylate side chain in ABP, MBP, and RBP display the *syn* stereochemistry, an *anti* stereochemistry is seen in the GGBP Asp residue interaction. This finding is consistent with the observation that *syn* is the preferred stereochemistry in the interactions involving carboxylate residues in proteins (Gandour, 1981; Carrel *et al.*, 1988). In spite of the close proximity of the locations of the three sugar hydroxyls and the identical stereochemistry, the Asp residues lack a uniform conformation. In fact, all four Asp residues (including the one in GGBP) exhibit different rotamer geometry (Table 5). The χ^1 torsion angles of the Asp residues are close to one of the most probable values of –60° (g^+), 60° (g^-), or 180° (t) (McGregor *et al.*, 1987). However, the χ^2 angles show considerably more deviations from the probable values of 90° or –90°. It should be noted that all the Asp residues are located in loop regions connecting secondary structure elements.

The finding of only the β-anomer of D-ribose bound to the RBP structure has been explained on the basis of the β-anomer being three times more abundant than the α-anomer in solution (Table 6) (Mowbray & Cole, 1992). While this explanation could also account for the pattern observed in the crystal structures of GGBP (Table 6), it is inconsistent with the ABP and MBP X-ray results. Nearly equal amounts of both anomers of the three monosaccharide substrates are bound to ABP in the crystal structure (Quiocho & Vyas, 1984; Quiocho *et al.*, 1989) despite the fact that the equatorial anomeric form (which corresponds to the β form of the D-galactose or α form of L-arabinose) exists in a much greater proportion than the axial anomeric form (Table 6). Moreover, the refined structures of the maltose-bound MBP have bound α-anomer only (Spurlino *et al.*, 1991, 1992) even though the α:β ratio of the disaccharide is 2:3 (Gehring *et al.*, 1991) (Table 6).

Is there one possible unified explanation for why the crystal structures of ABP show the binding of both anomers of Ara, Gal, and Fuc in roughly equal proportion, while the structures of GGBP, MBP, and RBP contain only one bound anomeric sugar? Differences in the binding affinities between anomers could be offered as the best explanation. ABP presumably binds both anomers with almost equal affinity, and regardless of the ratio of the anomers in solution (Table 6), the long period of crystallization (often several weeks and sometimes months) and the high concentration of sugars used in the cocrystallization led to the formation and crystallization of approximately 50% of the proteins with bound α-anomer and the other 50% with bound β-anomer. On the other hand, the affinity of GGBP, MBP, and RBP for one anomer is presumably somewhat greater than that for the other anomer. Consequently, at high sugar concentrations, combined with the long period of crystallization, the tightest complex predominates and is crystallized preferentially. The idea of different sugar affinity is borne out by the data on hydrogen-bonding interactions shown in Tables 3 and 4 for ABP and GGBP. The hydrogen bond distances to both α- and β-anomer hydroxyls of arabinose are very similar (Table 3). On the other hand, the hydrogen bonds involved with the β-anomers of Glc and Gal bound to GGBP appear shorter and hence stronger than those obtained by modeling their α-anomer counterparts (Table 4). The contribution of van der Waals contacts to sugar affinity is difficult to evaluate, but as shown in Table 4, the X-ray data indicate one less favorable contact than the model results.

The above explanation is further supported by the work of Gehring *et al.* (1991), who used tritiated maltooligosaccharides in NMR measurements of oligosaccharide binding to MBP in solution. At a low ratio (about 1:1) of D-maltose to MBP, the proportion of anomers initially bound to MBP reflected more closely the ratio of 2α:3β observed in solution in the absence of the protein (Table 6). However, further equilibration of the MBP–maltose solution for 2–3 days resulted in a significantly higher proportion of the α-sugar bound than the β-sugar, reflecting a tighter affinity for and a shift in equilibrium in favor of the α-anomer. At a higher ratio of maltose to MBP (a condition which approaches that of the crystallization condition) the proportion of bound α-anomer further increased. The NMR study further indicated a 2-fold

Table 4: Interactions of the Anomeric Hydroxyl of D-Glc and D-Gal with GGBP Based on X-ray Structures and Modeling

	atom		distance (Å)			
type	protein/water	sugar	α-D-Glc (model)	β-D-Glc (X-ray)	α-D-Gal (model)	β-D-Gal (X-ray)
hydrogen bonds	Asn91 ND2		(4.82)[a]	3.44	(4.66)	3.25
	Asp154 OD1		3.04	2.64	3.03	2.66
	Arg158 NH1	O1	3.16	(4.11)	3.37	(4.15)
	Arg158 NH2		(3.69)	3.22	(3.79)	3.35
	Asn256 ND2		(4.95)	3.06	(5.03)	3.13
van der Waals	Asp154 CB		3.34	(4.23)	3.38	(4.33)
	Asp154 CG		3.53	3.63	3.59	3.67
	Trp183 CH2	O1	3.29	(5.40)	3.23	(5.50)
	Trp183 CZ2		3.56	(5.60)	3.42	(5.69)
	Wat319 O		(5.24)	3.70	(5.82)	3.64
	Wat401 O		(5.15)	3.77	(5.04)	3.66

[a] Values in parentheses exceed the limits of 3.5 Å for hydrogen bonds and 4 Å for van der Waals contacts.

Table 5: Torsional Angles of Asp Residues Involved in Anomeric Hydroxyl Recognition[a]

protein	Asp residue	χ^1	χ^2	stereochemistry
ABP	Asp90	−65 (g^+)	13	syn
GGBP	Asp154	−76 (g^+)	149	anti
RBP	Asp90	173 (t)	−107	syn
MBP	Asp90	64 (g^+)	−173	syn

[a] Stereochemistry of the interaction between the Asp and sugar hydroxyl (see text also). The χ^1 and χ^2 torsion angles are defined as N–CA–CB–CG and CA–CB–CG–OD. The oxygen atom of the carboxylate side chains making hydrogen bonds to the sugar anomeric hydroxyls is designated as OD (Figure 3).

Table 6: Equilibrium Composition of Sugars

	% composition			
sugar[a]	α-pyranose	β-pyranose	α-furanose	β-furanose
L-arabinose[b]	60.0	35.5	2.5	2
D-galactose[b]	~29	64	~3	~4
D-glucose[b]	38	62		
D-ribose[b]	21.5	58.5	6.5	13.5
D-maltose[c]	40	60		

[a] The α- and β-anomeric hydroxyls of the D sugar, in $^4C_1$ full chair conformation, correspond to axial and equatorial orientations, respectively. For L-arabinose, the α- and β-hydroxyls correspond to equatorial and axial orientations, respectively. [b] Determined by NMR in deuterium oxide (Pigman & Anet, 1972). [c] Determined by tritium NMR spectroscopy (Gehring *et al.*, 1991).

greater affinity of α-maltose than β-maltose. In light of the NMR data, it is not surprising to find that the crystal structures of MBP contained only the bound α-anomer of maltose (Spurlino *et al.*, 1991, 1992) as well as that of maltotriose and maltotetraose (L. E. Rodseth, J. C. Spurlino, and F. A. Quiocho, unpublished data).

The X-ray and NMR results demonstrating the preferential binding of only one sugar anomer are not inconsistent with the results obtained by rapid-reaction kinetic measurements of ABP, GGBP, and MBP, indicating binding of both sugar anomers (Miller *et al.*, 1980, 1983). The time scales of the reaction rates are infinitely faster than those required to carry out crystallization and X-ray and NMR structure studies. The association rate (k_{on}) constants for sugar binding to the three sugar-binding proteins in solution are all about (1–3) $\times 10^7$ M^{-1} s^{-1}, and the dissociation rate (k_{off}) constants range from 1 to 100 s^{-1}, which are reflected in the observed variation in the affinity constants. The fast time scale of the reaction rates may be more relevant to the role of the binding proteins in both active transport and chemotaxis. Indeed, as discussed by Miller *et al.* (1980), the large association rate constants allow a fast transport or rapid response and confer, in part, the high sensitivity of the corresponding chemotaxis and transport systems.

In summary, the determination of the crystal structures of GGBP with bound sugars (D-gluocse or D-galactose) has enabled us to take a close look at the mode of binding of epimeric sugars in the same site and orientation without alteration of the conformation of any of the residues in the site and to account for the several facets of molecular recognition of sugars by several periplasmic binding proteins. Asp residues are heavily involved in molecular recognition of sugar epimers and anomers. The geometry of the side chain of Asp14 of GGBP, acting as a bidentate, enables both carboxylate oxygens to participate in binding both epimers, with each oxygen consigned primarily to one epimeric hydroxyl. This geometry is different from that of the Asp residue in ABP, GGBP, MBP, and RBP, which is principally responsible in the recognition of sugar anomers by way of one (or monodentate) carboxylate oxygen. Our results are likely to have some bearing on other proteins or enzymes that bind carbohydrates. Binding of both sugar anomers is not confined to the periplasmic sugar-binding proteins. Several of the enzymes of glucose metabolism (e.g., hexokinase, glucokinase, glucose-6-phosphatase) are able to act on both anomers of their respective sugar substrates.

REFERENCES

Ames, G. F.-L. (1986) *Annu. Rev. Biochem. 55*, 397–425.

Anraku, Y. (1968) *J. Biol. Chem. 243*, 3116–3122.

Carrel, C. J., Carrel, H. L., Erlebacher, J., & Glusker, J. P. (1988) *J. Am. Chem. Soc. 110*, 8651–8656.

Ferenci, T. (1980) *Eur. J. Biochem. 108*, 631–636.

Furlong, C. E. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaecher, M., & Umbarger, H. E., Eds.) pp 786–796, American Society for Microbiology, Washington, DC.

Gandour, R. D. (1981) *Bioorg. Chem. 10*, 169–176.

Gehring, K., Williams, P. G., Pelton, J. G., Morimoto, H., & Wemmer, D. E. (1991) *Biochemistry 30*, 5524–5531.

Hamlin, R. (1985) *Methods Enzymol. 114*, 416–452.

Hazelbauer, G. L., & Adler, J. (1971) *Nature (London) 230*, 101–104.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Hogg, R. W., & Englesberg, E. (1969) *J. Bacteriol. 100*, 423–432.

Howard, A. J., Neilsen, C., & Xuong, N. H. (1985) *Methods Enzymol. 114*, 452–472.

Macnab, R. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaecher, M., & Umbarger, H. E., Eds.) pp 732–759, American Society for Microbiology, Washington, DC.

McGregor, M. J., Islam, S. A., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 198*, 295–310.

Miller, D. M., III, Olson, J. S., & Quiocho, F. A. (1980) *J. Biol. Chem. 255*, 2465–2471.

Miller, D. M., III, Olson, J. S., Pflugrath, J. W., & Quiocho, F. A. (1983) *J. Biol. Chem. 258*, 13665–13672.

Mowbray, S. L., & Cole, L. B. (1992) *J. Mol. Biol. 225*, 155–175.

Mowbray, S. L., Smith, R. D., & Cole, L. B. (1990) *Receptor 1*, 41–45.

Pigman, W., & Anet, E. F. L. J. (1972) in *Carbohydrate Chemistry and Biochemistry* (Pigman, W., & Horton, D., Eds.) pp 165–194, Academic Press, New York.

Quiocho, F. A. (1991) *Curr. Opin. Struct. Biol. 1*, 922–933.

Quiocho, F. A. (1993) *Biochem. Soc. Trans. 21*, 442–448.

Quiocho, F. A., & Pflugrath, J. W. (1980) *J. Biol. Chem. 255*, 6559–6561.

Quiocho, F. A., & Vyas, N. K. (1984) *Nature (London) 310*, 381–386.

Quiocho, F. A., Wilson, D. K., & Vyas, N. K. (1989) *Nature (London) 340*, 404–407.

Sharff, A. J., Rodseth, L. E., & Quiocho, F. A. (1993) *Biochemistry 32*, 10553–10559.

Sheldrick, B. (1976) *Acta Crystallogr. 32*, 1016–1020.

Spurlino, J. C., Lu, G.-Y., & Quiocho, F. A. (1991) *J. Biol. Chem. 266*, 5202–5219.

Spurlino, J. C., Rodseth, L. E., & Quiocho, F. A. (1992) *J. Mol. Biol. 226*, 15–22.

Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732–740.

Vyas, N. K., Vyas, M. N., & Quiocho, F. A. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1792–1796.

Vyas, N. K., Vyas, M. N., & Quiocho, F. A. (1988) *Science 242*, 1290–1295.

Zou, J., Flocco, M. M., & Mowbray, S. L. (1993) *J. Mol. Biol. 233*, 739–752.