The asialoglycoprotein receptors and many other C-type (Ca\(^{2+}\)-dependent) animal lectins specifically recognize galactose- or N-acetylgalactosamine-terminated oligosaccharides. Analogous binding specificity can be engineered into the homologous rat mannose-binding protein A by changing three amino acids and inserting a glycine-rich loop (Iobst, S. T., and Drickamer, K. (1994) J. Biol. Chem. 269, 15512–15519). Crystal structures of this mutant complexed with β-methyl galactoside and N-acetylgalactosamine (GalNAc) reveal that as with wild-type mannose-binding proteins, the 3- and 4-OH groups of the sugar directly coordinate Ca\(^{2+}\) and form hydrogen bonds with amino acids that also serve as Ca\(^{2+}\) ligands. The different stereochemistry of the 3- and 4-OH groups in mannose and galactose, combined with a fixed Ca\(^{2+}\) coordination geometry, leads to different pyramidal ring locations in the two cases. The glycine-rich loop provides selectivity against mannose by holding a critical tryptophan in a position optimal for packing with the apolar face of galactose but incompatible with mannose binding. The 2-acetamido substituent of GalNAc is in the vicinity of amino acid positions identified site-directed mutagenesis (Iobst, S. T., and Drickamer, K. (1996) J. Biol. Chem. 271, 6686–6693) as being important for the formation of a GalNAc-selective binding site.

Ca\(^{2+}\)-dependent (C-type) animal lectins are a family of proteins whose members contain one or more homologous carbohydrate-recognition domains (CRDs). 1 The majority of C-type lectins bind to mannose- or D-glucose-related sugars (Man-type ligands), or to β-galactoside and its derivatives (Gal-type ligands). The mammalian hepatocyte asialoglycoprotein receptors, which play a part in serum glycoprotein homeostasis, are the best known of the Gal-binding C-type lectins. C-type lectins with high affinity for glycoconjugates bearing terminal galactose residues have also been identified on the surfaces of peritoneal macrophages and Kupffer cells, and appear to mediate recognition of tumor cells.

C-type CRDs with lower affinity for Gal-type ligands are found in proteoglycan core proteins of cartilage and other tissues and are presumed to contribute to the organization of the extracellular matrix.

Previous crystallographic analyses of rat mannose-binding proteins (MBPs) A and C have shown that Man-binding C-type lectins recognize their sugar ligands by formation of direct coordination bonds between a Ca\(^{2+}\) (designated site 2) and a lone pair of electrons from each of two vicinal hydroxyl groups possessing the same stereochemical arrangement as the equatorial 3- and 4-OH groups of α-mannose (8, 9). The Ca\(^{2+}\) is 8-coordinated in a pentagonal bipyramidal arrangement, with the two sugar hydroxyls bisecting one of the apical positions (8) (see Fig. 1a). In addition, the same OH groups form hydrogen bonds with amino acid side chains that are Ca\(^{2+}\) site 2 ligands, producing an intimately linked ternary complex of protein, Ca\(^{2+}\), and sugar (see Fig. 1a). Only one other contact, an apolar van der Waals contact between a ring carbon and the C\(_{β}\) of residue 189 contributes significantly to binding (8, 10).

Studies with derivatized sugars have shown that free 3- and 4-OH groups are essential for binding to mammalian asialoglycoprotein receptors as well as Man-binding C-type lectins, whereas substitutions at other ring positions have little or no effect on binding (11). However, the 3- and 4-OH groups of galactose have an equatorial/axial arrangement, so the mechanism of Gal- and Man-type ligand recognition must be different. Sequence analysis reveals that of the Ca\(^{2+}\) 2 ligands, positions equivalent to Glu\(^{183}\), Asn\(^{259}\), and Asp\(^{186}\) of MBP-A are highly conserved among C-type lectins regardless of specificity. In contrast, positions 185 and 187 are found to be Glu and Asn in Man-binding family members, whereas Gal-binding C-type lectins have Gin and Asp at these positions. The Glu\(^{183}\) → Gin/Asn\(^{187}\) → Asp mutant of MBP-A, designated “QPD”, binds to galactose in preference to mannose by a factor of 3 but with relatively low affinity for either sugar (12). Position 189 of MBP-A (Fig. 1a) is not conserved among Man-binding C-type lectins but is always either Trp or Phe in Gal-binding family members. Replacement of His\(^{189}\) of MBP-A with Trp in the QPD mutant to make “QPDW” gives a protein with affinity for Gal comparable with natural Gal-binding C-type lectins but that still does not discriminate well between Gal and Man (13).

However, insertion of a glycine-rich loop found in the major form of the rat asialoglycoprotein receptor, rat hepatic lectin-1 (RHL-1), and other Gal-binding C-type lectins that display strong discrimination against mannose results in a mutant (“QPDWG”) with galactose affinity and selectivity comparable with RHL-1 (13). The affinity for galactose is comparable in QPDW and QPDWG, indicating that the determinants of affinity and selectivity are somewhat distinct.

NMR measurements reveal similar modes of galactose binding by QPDWG and RHL-1, demonstrating that galactose specificity in C-type lectins is determined by a few residues and can be studied in the well characterized MBP-A background. Here we describe the structure of a trimeric fragment of OPD containing the neck and COOH-terminal CRD (14), both alone and complexed with β-methyl galactoside (βMeGal) and N-acetylgalactosamine (GalNAc). The structures reveal the

---

* This work was supported by Grant GM30565 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Pew Scholar in the Biomedical Sciences. To whom correspondence should be addressed: Dept. of Structural Biology, Fairchild Bldg., Stanford University School of Medicine, Stanford, CA 94305-5400. Tel.: 415-725-4623; Fax: 415-723-8464; E-mail: weis@fucose.stanford.edu.

1 The abbreviations used are: CRD, carbohydrate-recognition domain; MBP, mannose-binding protein; βMeGal, β-methyl-α-galactoside; GalNAc, N-acetyl-α-galactosamine; MGR, macrophage galactose receptor; RHL-1, rat hepatic lectin-1; d-, dextrin-pain-4-treated fragment.
molecular basis of selective galactose recognition by C-type lectins. The structure of the QPDWG-GalNAc complex is consistent with results of site-directed mutagenesis experiments that have identified amino acid positions that contribute to the preferential binding of GalNAc over Gal by certain C-type lectins.

**EXPERIMENTAL PROCEDURES**

Materials—Unless otherwise specified, chemicals were obtained from J. T. Baker Inc. LB medium was obtained from Life Technologies, Inc. Guanidinium hydrochloride and isopropylthiogalactopyranoside were obtained from Boehringer Mannheim. Clostripain was obtained from J.T. Baker Inc. LB medium was obtained from Life Technologies, Inc. Preparations were not imposed at any point in the refinement.

Crystallographic Data Collection and Structure Determination—Data collection and structure determination were carried out against data from 5–2.0 Å. Reflections from 10–2.0 Å were then included and refinement continued. This model was subsequently measured from another unliganded cl-QPDWG crystal and used in the final stages of refinement (see Tables I and II).

**TABLE II**

| Crystal  | Resolution range (Å) | Independent reflections | Completeness | Average redundancy | % > 3σ | R_{sym} | Number used (|F| > 2σ(|F|)) | R_{cryst} | R_{free} |
|----------|----------------------|-------------------------|--------------|-------------------|------|--------|-----------------|----------|---------|
| QPDWG    | 10.0–2.0             | 41,049                  | 97.2 (92.9)  | 3.2 (2.7)         | 81.8 (51.7) | 4.9 (21.4) | 39,913          | 23.4     | 28.2    |
| QPDWG + βMeGal | 10.0–2.0            | 46,484                  | 99.0 (97.7)  | 2.9 (2.4)         | 80.4 (49.1) | 5.0 (23.0) | 40,835          | 22.1     | 26.8    |
| QPDWG + GalNAc | 10.0–1.9            | 47,904                  | 97.2 (85.0)  | 3.0 (1.9)         | 82.5 (52.5) | 4.8 (22.3) | 46,236          | 21.7     | 26.1    |

**Table II**

All residues except Lys^{152} of one protomer of the unliganded cl-QPDWG mutant, which lies in a poorly ordered turn, fall within the allowed regions of the Ramachandran plot. No significant differences in structure are observed among the different copies, except in regions of lattice contacts. The average temperature factor of the heavy atom for one protomer is nearly isomorphous with the complex structure than in the sugar-bound structures, but it is not clear if this is a consequence of sugar binding. The side chain of residue His^{99} in protomer 1 of QPDWG was modeled in two conformations. The side chains of residues His^{99} in protomer 1 and Met^{103} and Ser^{129} in protomer 2 were modeled in two conformations in QPDWG + GalNAc. RMSD, root-mean-square deviation.

**RESULTS AND DISCUSSION**

A trimeric fragment of QPDWG containing the neck and COOH-terminal CRD (14) was crystallized, and the structure was solved by molecular replacement, both alone and com-
plexed with βMeGal and GalNAc (Tables I and II). The structures were refined to resolutions of 2.0 Å or better (Tables I and II). Apart from the His189 → Trp change and the glycine-rich insertion at the carbohydrate-binding site, the structures of wild-type MBP-A and the QPDWG mutant are identical to within the coordinate error. In particular, the Ca^{2+} site 2 ligands of the two structures superimpose, with the side chain amide nitrogen of Gln185 and the carbonyl oxygen of Asp187 of QPDWG in the same positions as the carbonyl oxygen of Glu185 and the amide nitrogen of Asn187 in the wild-type protein.

Despite the different stereochemistry of the 3- and 4-OH groups, the mechanism of βMeGal and GalNAc binding to QPDWG is similar to that of Man-type ligands to wild-type MBPs, with the full noncovalent bonding potential of 3- and 4-OH groups used for Ca^{2+} coordination and hydrogen bond formation with Ca^{2+} ligands (8, 9) (Fig. 1, a and c). However, maintenance of the pentagonal bipyramidal Ca^{2+} coordination geometry forces the pyranose ring into a very different orientation from that observed in mannose binding to wild-type MBPs (8, 9). The apolar patch formed by the 3, 4, 5, and 6 carbons of βMeGal and GalNAc packs against the side chain of Trp189, an interaction observed in all galactose-lectin interac-

**Fig. 1. Galactose binding to QPDWG.** a, mannose binding to wild-type MBP-A as observed in a Man6GlcNAc2-Asn-MBP-A complex (8). Carbon, nitrogen, and oxygen atoms are shown as white, gray, and black spheres, respectively; Ca^{2+} 2 is shown as a larger white sphere. Coordination and hydrogen bonds are represented by long and short dashed lines, respectively. Carbon atoms of the sugars are numbered. b, stereo pair of the final 2Fo−Fc electron density map in the binding site of the βMeGal-QPDWG complex, contoured at 1.2σ. c, galactose binding to QPDWG. Symbols are as described for a. Parts a and c were made with MOLSCRIPT (28).
as a rigid unit that restricts the conformation of Trp189 rather than providing specific interactions with the sugar or other residues of the protein. Superposition of mannose bound as a guest in the crystallographically independent copies (protomers 1 and 3) of the QPDWG structures explain binding, mutagenesis, and spectroscopic data obtained from several galactose-binding mutants of MBP-A. Proton NMR spectra of βMeGal in the presence of QPDWG show upfield shifts of the H5, H6, and H6’ protons of Gal consistent with their interaction with the delocalized π electron system of the Trp ring observed in the crystal structure (13). The linewidths of the aromatic protons of Trp189 are broadened upon Gal binding to QPDWG, whereas they are only in the absence or the presence of Gal in QPDWG, consistent with the notion that the Gly-rich loop immobilizes Trp189 in a position optimal for interaction with Gal (13). The proteoglycan core protein CRDs have Phe instead of Trp at position 189 and exhibit relatively poor selectivity against Man-type ligands. The corresponding MBP-A mutant QPFDG, which includes Phe189, binds to Gal-type ligands only 6-fold more strongly than Man-type ligands, as opposed to the 40-fold selectivity for Gal-type ligands shown by QPDWG (13). These properties are explained by exclusion of Man by the 6-membered portion of the Trp189 ring (Fig. 2c), which extends farther out than the side chain of Phe.

Several Gal-binding C-type lectins, including RHL-1, display strong preference for GalNAc over Gal, whereas others do not discriminate between these two sugars. An example of the latter is the macrophage galactose receptor (MGR), and the QPDWG mutant of MBP-A mimics MGR in this respect. Site-directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.

The present structures leave unclear how the Glu185 directed mutagenesis of MGR based on sequence comparisons with the asialoglycoprotein receptors has identified residues in four regions of the sequence that provide selectivity for GalNAc over Gal (24). Of these regions, the residue equivalent to Ser154 of MBP-A provides 20-fold of the observed 60-fold selectivity for GalNAc over Gal by RHL-1 (24). Moreover, a histidine equivalent to Thr202 of QPDWG is found in both RHL-1 and MGR.
with side chain carbonyl oxygen and amide nitrogen atoms that conform approximately to this symmetry in the wild-type site but not in the QPD site (Fig. 4). Although the mechanism is not obvious, this difference in symmetry may be related to the weaker affinity of QPD for either Gal- or Man-type ligands (12). The absolute affinity of wild-type MBP-A for Man is similar to that of QPDW or QPDWG for Gal, which implies that the binding energy of Man to the wild-type Ca$^{2+}$ site is greater than that of Gal to the QPD mutant site. Thus the favorable interaction with the aromatic residue at position 189 can be viewed as compensating for the loss of symmetry in the mutant site to provide affinity for Gal comparable with that of wild-type MBP-A for Man. In the absence of the glycine-rich loop, mannose is not excluded from the site but interacts with lower affinity due to the asymmetric arrangement of its hydrogen-bonding partners.

Another potential source of the different specificities of wild-type and QPD sites is the displacement of ordered water molecules upon sugar binding. High resolution structures of MBP-C show that the 3- and 4-OH of Man-type ligands replace two water molecules that form the same set of hydrogen and Ca$^{2+}$ coordination bonds (9). Unfortunately, the amount of visible, ordered water structure in the uncomplexed QPDWG site varies among the three crystallographically independent
copies, making it difficult to draw firm conclusions. In the best ordered site, two water molecules that form hydrogen bonds with the Ca\(^{2+}\) ligands at 185, 187, 198, and 210 equivalent to those formed by Gal can be discerned. These water molecules are in approximately the same position as the 3- and 4-OH groups of Gal but only one of them appears to be close enough to Ca\(^{2+}\) to form a coordination bond. Only one water molecule is observed in another copy, and no water molecules can be placed with confidence in the third site. Higher resolution structures of the uncomplexed QPDWG site will be required to assess whether or not there is a change in Ca\(^{2+}\) coordination number upon ligand binding.

The different locations of the bound pyranose ring seen in the present structures and the structures of wild-type MBPs complexed with Man-type ligands are a consequence of Ca\(^{2+}\) coordination geometry. This observation and the fact that few other contacts are made with the protein demonstrate the dominant role that Ca\(^{2+}\) coordination plays in sugar recognition by C-type lectins. The different pyranose ring locations dictated by Ca\(^{2+}\) coordination geometry forms the basis of selective recognition of galactose by steric exclusion of Man-type ligands provided by Trp189 and the glycine-rich loop.

Acknowledgments—We thank Shaun Park-Snyder for technical assistance and Kurt Drickamer for providing expression constructs and comments on the manuscript.

REFERENCES
1. Spiess, M. (1990) Biochemistry 29, 10009–10018
2. Drickamer, K. (1991) Cell 67, 1029–1032
3. Ii, M., Kurata, H., Itch, N., Yamashina, I., and Kawasaki, T. (1990) J. Biol. Chem. 265, 11295–11298
4. Hoyle, G. W., and Hill, R. L. (1988) J. Biol. Chem. 263, 7487–7492
5. Vavasseur, F., Perrado, A., Heuze, F., Joret, E., and Meflah, K. (1990) Int. J. Cancer 48, 744–751
6. Sato, M., Kawakami, K., Osawa, T., and Toyoshima, M. (1992) J. Biol. Chem. 267, 331–336
7. Drickamer, K., and Taylor, M. E. (1993) Annu. Rev. Cell Biol. 9, 237–264
8. Wels, W. I., Drickamer, K., and Hendrickson, W. A. (1992) Nature 360, 127–134
Galactose Selectivity in C-type Lectins

9. Ng, K. K.-S., Drickamer, K., and Weis, W. I. (1996) J. Biol. Chem. 271, 663–674
10. Iobst, S. T., Wormald, M. R., Weis, W. I., Dwek, R. A., and Drickamer, K. (1994) J. Biol. Chem. 269, 15509–15511
11. Lee, R. T., Ichikawa, Y., Fay, M., Drickamer, K., Shao, M.-C., and Lee, Y. C. (1991) J. Biol. Chem. 266, 4810–4815
12. Drickamer, K. (1992) Nature 360, 123–128
13. Iobst, S. T., and Drickamer, K. (1994) J. Biol. Chem. 269, 15512–15519
14. Weis, W. I., and Drickamer, K. (1994) Structure 2, 1227–1240
15. Fornstedt, N., and Porath, J. (1975) FEBS Lett. 57, 187–191
16. Ohwino, Z. (1993) in Proceedings of the CCP4 Study Weekend: Data Collection and Processing, 29–30 January 1993 (Sawyer, L., Isaacs, N., and Bailey, S., and Daresbury, U. K., eds) pp. 56–62, Science and Engineering Research Council, Daresbury Laboratory, Warrington, UK
17. Brünger, A. T. (1992) X-PLOR Manual, Version 3.1, Yale University, New Haven, CT
18. Brünger, A. T., Krukowski, A., and Erickson, J. W. (1990) Acta Crystallogr. Sect. A 46, 585–593
19. Brünger, A. T. (1992) Nature 355, 472–475
20. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
21. Angyal, S. J., and Pickles, V. A. (1972) Aust. J. Chem. 25, 1695–1710
22. Sheriff, S., and Hendrickson, W. A. (1987) Acta Crystallogr. Sect. A 43, 118–121
23. Weis, W. I., and Drickamer, K. (1996) Annu. Rev. Biochem. 65, in press
24. Iobst, S. T., and Drickamer, K. (1996) J. Biol. Chem. 271, 6686–6693
25. Bernstein, F. C., Koetzel, T. F., Williams, G. J. B., Meyer, E. T., Brice, M. D., Rose, A. R., Kannard, O., Shimamushi, T., and Tatsui, M. J. (1977) J. Mol. Biol. 112, 535–542
26. Kraulis, P. J. (1991) J. Appl. Cryst. 24, 946–950
