An Endogenous *Drosophila* Receptor for Glycans Bearing α1,3-Linked Core Fucose Residues*

Received for publication, March 25, 2002, and in revised form, April 17, 2002
Published, JBC Papers in Press, April 17, 2002, DOI 10.1074/jbc.M202825200

Samuel Bouyain‡§, Nicholas J. Silk‡, Gustáv Fabini¶, and Kurt Drickamer‖

From the §Glycobiology Institute, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom and the ¶Glycobiology Division, Institut für Chemie der Universität für Bodenkultur, Mauthgasse 18, A-1190 Wien, Austria

The genome of *Drosophila melanogaster* encodes several proteins that are predicted to contain Ca\(^2^+\)-dependent, C-type carbohydrate-recognition domains. The *CG2958* gene encodes a protein containing 359 amino acid residues. Analysis of the *CG2958* sequence suggests that it consists of an N-terminal domain found in other *Drosophila* proteins, a middle segment that is unique, and a C-terminal C-type carbohydrate-recognition domain. Expression studies show that the full-length protein is a tetramer formed by noncovalent association of disulfide-linked dimers that are linked through cysteine residues in the N-terminal domain. The expressed protein binds to immobilized yeast invertase through the C-terminal carbohydrate-recognition domain. Competition binding studies using monosaccharides demonstrate that *CG2958* interacts specifically with fucose and mannose. Fucose binds 5-fold better than mannose. Blotting studies reveal that the best glycoprotein ligands are those that contain N-linked glycans bearing α1,3-linked fucose residues. Binding is enhanced by the additional presence of α1,6-linked fucose. It has previously been proposed that labeling of the *Drosophila* neural system by anti-horseradish peroxidase antibodies is a result of the presence of difucosylated N-linked glycans. *CG2958* is a potential endogenous receptor for such neural-specific carbohydrate epitopes.

Animal lectins provide a mechanism for recognition of protein- and lipid-linked glycans. Recognition of endogenous glycoconjugates at the cell surface, in the extracellular matrix, and in serum can lead to intercellular adhesion and signaling as well as uptake and degradation. Animal lectins are diverse in structure, but they usually contain modular carbohydrate-recognition domains (CRDs)\(^1\), (2). The C-type CRDs are the largest and most diverse class of CRDs. These domains share a common fold and show Ca\(^2^+\)-dependent sugar binding activity, although their selectivity for different carbohydrate ligands varies. Binding of sugars involves formation of a ternary complex between the protein, a bound Ca\(^2^+\), and the sugar (3). Selectivity for particular sugar ligands is determined in large part by the disposition of Ca\(^2^+\)-ligating residues in the protein, but can also reflect additional interactions with nearby regions of the protein surface. The C-type CRDs are a subset of a larger family of protein modules, the C-type lectin-like domains (CTLDs) (4). Although the CTLDs share a common fold, many do not bind Ca\(^2^+\) and they often interact with ligands others than sugars.

Profile analysis provides a powerful method for identifying CTLDs in proteins identified by genome sequence analysis. Using information about the structures and sugar binding activities of known C-type CRDs, it is possible to identify CTLDs that are likely to display carbohydrate binding activity. This approach has been applied to the complete genomic sequences of model organisms such as *Caenorhabditis elegans* (5) and *Drosophila melanogaster* (2) as well as to the human genome (see ctdl.glycob.ox.ac.uk for a current update). The screen of the *Drosophila* genome revealed 32 genes that encode proteins containing CTLDs, of which only 6 have appropriate residues at the key positions required for Ca\(^2^+\) and sugar binding in the C-type CRDs. Among these potential CRDs, the domain encoded in the *CG2958* gene (also designated *Lectin-24DB*) shows a particularly strong resemblance to mammalian CRDs that bind mannose, fucose, and N-acetylgalactosamine. The Ca\(^2^+\)-ligating residues are in the same configuration as those in serum mannose-binding protein and other C-type lectins that bind sugars with hydroxyl groups in the orientation corresponding to the 3- and 4-hydroxyl groups of mannose (6). In addition, the CTLD of *CG2958* contains a cluster of basic residues (Arg-Lys-Lys at positions 347–349) that can form a secondary binding site that enhances binding of C-type CRDs to anionic oligosaccharides (7).

To probe a possible role of *CG2958* in sugar-mediated recognition, the protein has been expressed and its carbohydrate-binding properties have been examined. *CG2958* appears to be an endogenous receptor for α1,3-linked fucose residues linked to the core GlcNAc residue of N-linked oligosaccharides in *Drosophila*.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes were purchased from New England Biolabs. Both the Advantage 2 polymerase mix and the adult *Drosophila* Matchmaker cDNA library were supplied by CLONTECH. Synthetic oligonucleotides were obtained from Invitrogen. Yeast invertase, glycans, and glycoproteins used in binding assays were obtained from Sigma. Man-BSA, purchased from E-Y Laboratories, had an average density of 30 mol of mannos/mol of protein. Na\(^1,2\)I. Bolton-Hunter reagent, and nitrocellulose were purchased from Amersham Biosciences. Immulon 4 polystyrene wells were from Dynex Technologies. Affinity resins were prepared by the method of Fornstedt and Porath (8). RNase B oligosaccharides were purchased from Oxford GlycoSciences, and oligosaccharides from soybean agglutinin were a kind gift...
of Daniel Mitchell and Brian Matthews (Glycobiology Institute, University of Oxford, Oxford, United Kingdom).

Cloning and Mutagenesis—DNA sequences encoding the CRD of CG2958 were amplified with forward primer 5′-aaggccgccaacetgtctggcgaattgtcagcgc-3′ and reverse primer 5′-ttggcgcgaataacctttgtgacatctggac-3′. The PCR product was sequenced to confirm the alternative forward primer 5′-aaggccgccaacctttgaggtcagcgc-3′ was used. The 5′ end of the cDNA was amplified using forward primer 5′-aaggccgccgagcctattgccttgaggtcagcgc-3′ and reverse primer 5′-ttggcgcgacatcttgacatctggac-3′. The primers include restriction sites for FseI and NotI. Following denaturation at 95 °C for 1 min, 40 cycles of 95 °C for 20 s and 65 °C for 1 min were executed. Fragments were digested with FseI and NotI and inserted into a pIIIINompA2 expression vector modified to contain the restriction sites FseI and NotI downstream of the ompA signal sequence (9). Portions of cDNAs without reverse transcription errors were recombined using convenient restriction sites. The resulting plasmids were used to transform Escherichia coli strain JA221. To avoid possible toxicity of the expressed fragments, the FseI site was introduced in a way that interrupts the reading frame. The correct reading frame was then generated by digesting the vector with FseI followed by trimming of the 3′ extensions with T4 DNA polymerase. Mutagenesis was performed by substituting double-stranded synthetic oligonucleotides for restriction fragments in the expression plasmid.

Expression and Purification of Proteins—Aliquots of an overnight culture of transformed bacteria (20 ml) was used to inoculate 1 liter of Luria-Bertini medium containing 50 mM NaCl, 2.5 mM CaCl2, and 25 mM Tris-Cl, pH 7.8. In most cases, the expressed proteins were further purified on a C4 reverse phase column (4.6 × 25 mm) eluted with an HEPES-buffered saline for 5, 10, and 10 min, radioactivity was detected using a Wallac Wizard counter.

Binding Assays—Aliquots (50 μl) of the peak fraction of CG2958 (0.1 mg/ml) in loading buffer were pipetted into Immulon 4 wells, and the plate was incubated overnight at 4 °C. Protein solution was removed, and wells were filled with 5% BSA in loading buffer. After 2 h at 4 °C, the blocking solution was discarded and wells were washed three times with cold HEPES-buffered saline (136 mM NaCl, 2.7 mM KCl, 10 mM MgCl2, 25 mM CaCl2, 15 mM Tris-Cl, pH 7.5). Portions of 0.5 ml of eluting buffer (150 mM NaCl, 2.5 mM EDTA, and 25 mM Tris-Cl, pH 7.8) were incubated with the purified fragments, and the reaction was allowed to proceed at room temperature for 10 min. After addition of 800 μl of loading buffer, the labeled protein was recovered on invertase-Sepharose as described above. Glycoprotein samples (10 μg) were run on an SDS-polyacrylamide gel under reducing conditions, and the gel was transferred onto a nitrocellulose membrane (12). The membrane was blocked with 2% hemoglobin in HEPES-buffered saline for 1 h at room temperature and incubated for 90 min with a solution of 125I-CG2958 in HEPES-buffered saline in the presence of 2% hemoglobin. Following three washes with cold HEPES-buffered saline for 5, 10, and 10 min, radioactivity was detected using a PhosphorImager (Molecular Dynamics).

For dot blots, nitrocellulose membranes were soaked in gel transfer buffer and clamped into a 96-well dot blot apparatus. Protein samples prepared in 100 μl of gel transfer buffer were filtered through the membrane under vacuum, followed by another 100 μl of gel transfer buffer. The membrane was removed from the apparatus and blocked and incubated with 125I-CG2958. The bound gel blot was washed and resolved by thin layer chromatography following published procedures (13). The chromatograms were blocked and with 125I-CG2958 as described for the gel blots.

Analytical Methods—Polyacrylamide gel electrophoresis was performed by the method of Laemmli, using gels containing 17.5% acrylamide (2% acrylamide–0.8% bisacrylamide) and 5% Triton X-100. Molecular mass standards were used for calibration. The gels were stained and destained according to standard procedures.

RESULTS

Sequence and Expression of CG2958—The deduced amino acid sequence of CG2958 is shown in Fig. 1. The CTLD was previously identified using profile scanning algorithms, and the extent of the signal sequence was predicted using hydrophathy plots combined with the sequence preference of signal peptidase (18, 19). The domain organization of the remainder of the protein was investigated by screening the SwissProt data base of protein sequences with the portion of CG2958 between the signal sequence and the CTLD. The results revealed that several proteins in Drosophila contain sequences similar to residues 21–90 of CG2958. Within this N-terminal domain, two cysteine residues are conserved in most of these proteins. No related sequences in other organisms were detected. The remaining portion of CG2958, between the N-terminal domain and the CTLD, was rescanned against the protein sequence data base, but no significant sequence similarity was detected.

A CG2958 cDNA was created using the polymerase chain reaction. Sequencing of multiple independent clones revealed a consistent discrepancy at the codon for amino acid 134, which encodes lysine in the Drosophila genome sequence (codon AAA) and arginine in the cDNAs (codon AGA). This difference may reflect a strain difference or a polymorphism in the Drosophila population. Based on the sequence analysis, vectors were created from the amplified cDNA to express the entire CG2958 polypeptide, a fragment lacking the N-terminal domain, and a smaller fragment consisting only of the CTLD. In each case, a bacterial signal sequence was fused to the open reading frame to allow export into the periplasm of E. coli. Induction of protein synthesis in the presence of Ca2+ allows correct folding of many CTLDs under such conditions (10). In the case of CG2958, sonication of the bacteria in Ca2+-containing buffer and passage over a column containing immobilized yeast invertase resulted in binding of the expressed fragments, which could be eluted with EDTA (Fig. 2). The intact protein and the CRD fragment were retained more effectively on the column than were the fragments consisting only of the CTLD.

Weak binding of isolated CRDs to affinity columns and stronger binding of larger lectin fragments often reflects multivalent binding by oligomers of the larger fragments. Portions of the lectin polypeptide outside the CRDs are often necessary to stabilize such oligomers (4). The oligomeric states of the CG2958 fragments were investigated by gel electrophoresis
and sedimentation analysis (Fig. 3). When analyzed by SDS-polyacrylamide gel electrophoresis, the intact protein runs as a covalent dimer, whereas the fragments that lack the N-terminal domain run as monomers. Thus, CG2958 consists of covalent dimers of polypeptides that are held together by disulfide bonds between cysteine residues in the N-terminal domain. Sedimentation analysis gives a molecular mass of 156,800 Da, which corresponds to a tetramer with a calculated molecular mass of 156,900 Da. These results demonstrate that CG2958 consists of a dimer of covalent dimers, in which each polypeptide comprises an N-terminal oligomerization domain, an intervening neck, and a C-terminal CRD. Although formation of the covalent oligomer requires the N-terminal domain, the relatively tight binding of the neck + CRD fragment to the invertase-Sepharose column suggests that noncovalent oligomerization can be mediated by the neck region.

Analysis of Sugar Binding Activity—The observation that CG2958 binds to immobilized yeast invertase in a Ca\textsuperscript{2+}-dependent manner suggested that it interacts with the extensive mannose-containing glycans attached to this protein. To obtain direct evidence for carbohydrate-dependent binding, the fragment representing the head plus neck region of CG2958 was immobilized on polystyrene wells. This fragment was chosen because it binds tightly to the affinity resin and is produced in high yield. Radiolabeled Man-BSA was found to bind to the immobilized protein and was therefore used as a reporter ligand to test the binding of potential sugar ligands in competition assays. These assays were first used to establish relative affinities of CG2958 for various monosaccharides (Table I).

The inhibition data reveal that, among the neutral monosaccharides tested, the most effective inhibitors of Man-BSA binding are L-fucose and the \(\alpha\)- and \(\beta\)-methyl fucosides. Among the D-hexoses tested, only mannose showed inhibition in the low millimolar range. Weak inhibition observed with free galactose is probably caused by interaction with the anomeric hydroxyl group of the free sugar, as has been previously observed (20),

![Fig. 1. Comparison of the sequence of Drosophila protein CG2958 with other proteins sharing related domains.](http://www.jbc.org/)

The N-terminal portion of the deduced amino acid sequence of CG2958 is compared with other Drosophila proteins that contain homologous sequences. The C-terminal CTLD is compared with C-type CRDs from human E-selectin and rat serum mannose-binding protein. Elements of secondary structure are \(\alpha\)-helix (\(\alpha\)), \(\beta\)-strand (\(\beta\)), and loop (L). Ca\textsuperscript{2+}-binding residues are designated by the numbers 1 and 2. Conserved cysteine residues throughout the sequences and conserved amino acids in the N-terminal domain are shaded. The stretches of three basic amino acid residues in CG2958 and E-selectin are underlined, and residues that have been mutated are indicated with asterisks.
because α-methyl-D-galactoside does not inhibit binding. The selective binding of mannose and α-methyl-D-mannoside suggests that, as with other C-type lectins, the relative orientation of the 3- and 4-hydroxy groups is important for binding to CG2958. The assays also reveal that binding is sensitive to the orientation and nature of the 2-substituent, because glucose, GlcNAc, N-acetylmannosamine, and 2-deoxyglucose are all poor inhibitors. Thus, the data suggest that the 2-, 3-, and 4-hydroxy groups of mannose may interact with the binding site of CG2958.

In the light of the selectin-like aspects of the CG2958 sequence, various potential anionic ligands were tested. In the inhibition assay, only sialic acid showed any inhibition at the concentrations tested. The measured $K_I$ indicates that it binds to CG2958 with higher affinity than glucose, but the affinity is still 7-fold lower than the affinity for mannose. Polymers containing glucuronic acid as well as various sulfated sugars were also tested, but none of these have enhanced affinity compared with glucuronic acid. Finally, the result using fucoidan as an inhibitor indicates that sulfated fucose shows decreased rather than increased affinity compared with fucose. Direct incubation with radioiodinated sialyl Lewis$^a$-BSA also failed to demonstrate any binding (data not shown).

**Possible Orientations of Monosaccharides in the Binding Site of CG2958**—The high selectivity for fucose and the submillimolar $K_I$ are unusual in C-type CRDs, so it was of interest to determine what structural features of the CRD might enhance binding to this monosaccharide. In previous studies, the sequence Lys-Lys-Lys has been inserted into the CRD from mannos-binding protein in the positions corresponding to the Arg-Lys-Lys sequence in CG2958 (21). The crystal structure of this modified CRD complexed with the fucose-containing sialyl-Lewis$^a$ tetrasaccharide (22) was used as a starting point to model the binding site of CG2958. There are four orientations in which pairs of hydroxyl groups of fucose can be superimposed on the hydroxyl groups that interact with Ca$^{2+}$ in the binding site (Fig. 4). Two of these structures involve interactions between Ca$^{2+}$ and the 2- and 3-hydroxy groups (orientations A and B) and two involve interaction with the 3- and 4-hydroxy groups (orientations C and D). In orientation D, there are clashes with the modeled protein surface, so this orientation can be dismissed. In orientation C, the fucose residue approaches the side chains of residues that correspond to Asn$^{327}$ and Arg$^{347}$ of CG2958 in the model. Binding of the 2- and 3-hydroxy groups, as in the original crystal structure, projects the sugar mostly away from the protein surface, but in orientation B there is a potential hydrogen bond between O1 of the sugar and Arg$^{347}$.

Because interactions of residues Asn$^{327}$ or Arg$^{347}$ with the O1 and O2 substituents of fucose could contribute to selective binding to this sugar, their role in preferential binding to fucose was tested by mutagenesis. The effects of changing these residues to alanine or lysine on the relative binding affinities for mannose and fucose were examined using the competition binding assay (Table II). Changing Asn$^{327}$ to alanine has the most severe effect on the relative affinity for fucose, causing a nearly 3-fold reduction in the $K_I$ for fucose compared with mannose. There is no effect with residue Arg$^{347}$ is modified. The only orientation in which significant interactions with Asn$^{327}$ would be predicted is orientation C, suggesting that this is the most likely way that fucose interacts with CG2958. This orientation is similar to the orientation observed in crystals of E- and P-selectin with bound sialyl-Lewis$^a$ ligand.

**Oligosaccharide Ligands for CG2958**—To screen for potential oligosaccharide ligands for CG2958, the expressed protein
Inhibition constants ($K_i$) for $^{125}$I-Man-BSA binding were obtained from solid-phase competition assays. Only partial inhibition was seen at the indicated highest concentration tested. a Expressed in moles of fucose sulfate. b Expressed in moles of glucuronic acid or glucuronic plus iduronic acid. c Expressed in moles of fucose sulfate.

| Monosaccharide                        | $K_0$ (mM) | $K_i$ | $K_0/K_i$ (M) |
|---------------------------------------|------------|-------|---------------|
| d-Mannose                             | 0.2 ± 0.2  | >200  | >400          |
| o-Methyl-d-mannoside                   | 0.01 ± 0.01| >200  | >200          |
| α-Glucose                             | 0.03 ± 0.03| >200  | >200          |
| N-Acetyl-α-glucosamine                 | 0.04 ± 0.04| >200  | >200          |
| 2-Deoxy-α-glucose                      | >100       | >200  | >200          |
| α-Methyl-α-galactose                   | >200       | >200  | >200          |
| N-Acetylneuraminic acid                | 0.01 ± 0.01| >200  | >200          |
| Mannose 6-phosphate                    | >100       | >200  | >200          |
| α-Glucuronic acid                      | >100       | >200  | >200          |
| GlcNAc 6-phosphate                     | >100       | >200  | >200          |
| GlcNAc 3-phosphate                     | >100       | >200  | >200          |
| Chondroitin sulfate A                  | >100       | >200  | >200          |
| Heparin                               | >100       | >200  | >200          |
| Hyaluronic acid                       | >100       | >200  | >200          |
| (human umbilical cord)                | >100       | >200  | >200          |
| Hyaluronic acid (rooster comb)         | >100       | >200  | >200          |
| Fucoidan                               | >100       | >200  | >200          |

FIG. 4. Possible orientations of fucose in the binding site of a C-type CRD. The model is based on the arrangement of fucose in the binding site of the CRD from mannose-binding protein that has been modified to include three lysine residues analogous to those found in E-selectin (22). Side chains of residues near the binding site have been changed to correspond to the residues found in CG2958: Asn$^{327}$ and Arg$^{347}$. Four possible orientations of fucose, indicated at the bottom of the figure, allow superposition of hydroxyl groups of fucose on O2 and O3 of fucose in the A orientation seen in the crystal structure. Orientation A is shown in white and orientation C in gray. Atoms are shaded black for carbon, gray for nitrogen, and white for oxygen.

was radiiodinated and used to probe a glycoprotein blot containing a range of glycan structures (Fig. 5). As a positive control, yeast invertase was included on the blot. The strong binding observed confirms that the radiiodinated material retains binding activity. No interaction with glycoproteins bearing complex N-linked glycans was detected. Binding to high mannose oligosaccharides is suggested by the interaction with RNAse B, which contains a range of oligosaccharides from Man$_6$GlcNAc$_2$ to Man$_9$GlcNAc$_2$ (23). In contrast, no binding to soybean agglutinin was detected, although this glycoprotein bears Man$_6$GlcNAc$_2$ oligosaccharides (24). To determine whether CG2958 binds preferentially to the smaller oligosaccharides that are more abundantly expressed on RNAse B, the oligosaccharides from RNAse B and soybean agglutinin were released and presented to CG2958 in the form of neoglycolipids on a thin layer chromatogram (Fig. 5). In this format, binding to all of the mannose-containing structures, including Man$_6$GlcNAc$_2$, was observed at levels proportional to their abundance (25). These results suggest that binding of CG2958 to oligosaccharides may be influenced by the proteins to which they are conjugated.

The other glycoprotein ligand may be detected in the blotting procedure by horseradish peroxidase. As a plant glycoprotein, horseradish peroxidase bears glycans containing sugars, such as xylose, and linkages, including fucose residues α,β, linked to the inner core GlcNAc residue, that are not found in mammalian glycoproteins (26, 27). Xylose is not found in N-linked glycans on Drosophila glycoproteins, but the core α,β-fucose residue is present (28). The most striking finding from the comparison of monosaccharide affinities is the relatively high selectivity for fucose, suggesting that the binding to horseradish peroxidase might be a result of the presence of core fucose.

The possibility that CG2958 binds to core α,β-fucose residues was investigated using various modified human serum transferrin preparations, in an approach previously used for studying binding of anti-horseradish peroxidase antibody...
Fig. 6. Binding of CG2958 to glycoproteins detected by dot blot assay. Samples consisted of various concentrations of horseradish peroxidase (A) or 1 μg of each of the modified transferrin preparations (B). Following incubation with radiiodinated CG2958, binding was quantified using a PhosphorImager. Averages of duplicate samples, corrected for background binding, are shown. In all cases, the results for duplicates were within 5% of each other.

(28). After removal of the terminal sialic acid and galactose residues, fucose was added enzymatically in vitro to the inner GlcNAc residue in either α1,3- or α1,6-linkage. The nonreducing terminal GlcNAc residues were then removed to expose the modified Manα1GlcNAc2 core. In preliminary studies, these proteins were run on SDS-polyacrylamide gels, blotted onto nitrocellulose, and probed with radioiodinated CG2958. Only the proteins bearing structures modified with α1,3-linked or α1,3- and α1,6-linked core fucose were bound by the labeled protein (data not shown). Similar modified proteins in which the nonreducing terminal GlcNAc residues were retained were also tested, but no binding was observed. Binding to the fucosylated Manα1GlcNAc2 core structures was quantified by performing a dot blot assay and comparing the binding to that obtained with dilutions of horseradish peroxidase (Fig. 6). Binding to 1-μg aliquots of the modified transferrin samples exceeds binding to comparable amounts of horseradish peroxidase, even though here they are only two N-glycosylation sites in transferrin and seven in horseradish peroxidase and the in vitro fucosylation of transferrin is incomplete. The results confirm that binding is dependent on the presence of α1,3-linked fucose, and they suggest that addition of α1,6-linked fucose enhances the binding. The enhancement of CG2958 binding to modified transferrin by α1,6-fucosylation is consistent with the findings that all α1,3-fucosylated N-glycans found in adult flies are difucosylated and that α1,6-fucosylated N-glycans are the preferred substrates for the α1,3-fucosyltransferase.

**DISCUSSION**

The results presented here demonstrate that *Drosophila* protein CG2958 is a fucose-binding lectin that interacts specifically with α1,3-linked fucose attached to the core of N-linked glycans. This binding can occur with oligosaccharides that are attached to proteins, so CG2958 is a potential endogenous receptor for *Drosophila* glycoproteins that bear core α1,3-fucosylated glycans. CG2958 is predicted to be secreted, and the full-length, expressed protein is soluble in the absence of detergents. Its tetrameric structure indicates that it would have the ability to interact with multiple N-linked glycans simultaneously. Depending on the geometry of the tetramer, such multivalent binding could lead to high affinity attachment to cell surfaces or to cross-linking of glycoproteins.

CG2958 was initially considered a potential selectin ortholog, largely based on the presence of a cluster of three basic residues adjacent to the predicted primary sugar-binding site in both the selectins and CG2958. In previous studies, it has been shown that such a cluster of residues can form a secondary, ionic strength-sensitive subsite in C-type CRDs, facilitating the binding of sialyl-Lewis^a^ and other anionic oligosaccharide ligands (7). Additionally, like the selectin CRDs, the CRD in CG2958 appears to contain a single Ca^2+^-binding site, because several of the acidic amino acid residues that usually form the secondary site are not conserved in CG2958 (Fig. 1). The present studies suggest that CG2958 and the selectins share the ability to bind fucose-containing ligands, but their oligosaccharide-binding characteristics are very different. CG2958 does not bind with high affinity to any of the anionic oligosaccharide or polysaccharide ligands tested. Thus, there seems to be little basis for suggesting that CG2958 might function in a selectin-like fashion in *Drosophila*.

Despite the differences between the CRDs in the selectins and CG2958, it is interesting that these proteins do share the ability to bind to fucose attached to N-acetylgalactosamine in α1,3 linkage. The modeling and mutagenesis results suggest a possible orientation of fucose-containing ligands in the binding site of the CRD in CG2958, in which the anomeric hydroxyl group would be near to one of the loops of the polypeptide adjacent to the conserved Ca^2+^ that forms the nucleus of the sugar-binding site (Fig. 4). The portions of the E- and P-selectin CRDs that correspond to the loop containing Asn^327^ in CG2958 form part of the extended sialyl-Lewis^a^-binding sites (30). Differences in sequence between these loops in the selectins and CG2958 preclude exactly analogous contacts, but the results suggest that the loop may be a common determinant of ligand-binding specificity beyond the simple monosaccharide-binding site. Although the modeling studies are suggestive, further structural analysis will clearly be necessary.

The pattern of staining of *Drosophila* embryos with anti-horseradish peroxidase antibodies suggests a role for core α1,3-linked fucose in the nervous system (31–33). The ability of CG2958 to serve as an endogenous receptor for glycoproteins bearing core α1,3-linked fucose further suggests that this protein might also be expressed in the nervous system. The Berkeley *Drosophila* Genome Project (34) has reported expressed sequence tags for the mRNA encoding CG2958 in cDNA libraries from head, brain, and sensory organs in the larval-early pupal stage suggestive of expression throughout development (identifiers HL05328 and LP02926).

Structural data for glycoconjugates in invertebrates, including *Drosophila*, are limited (35, 36). However, the information that is available suggests that there are substantial differences between the glycans in invertebrates and vertebrates. Particularly striking differences are observed in the terminal sugars present on glycans. Although sialic acid and galactose are common terminal elaborations of vertebrate glycans, such sugars are rare, if present at all, in invertebrates. In contrast, fucose appears to be much more abundant in invertebrates. These data suggest that cell surface carbohydrates have evolved somewhat differently in different animal lineages. The results reported here suggest that endogenous receptors for cell surface glycans have evolved in parallel to recognize a distinct complement of sugar structures.

**Acknowledgments**—We thank Dawn Torgersen for help with protein production and binding assays, Russell Wallis for assistance with the analytical ultracentrifugation, and Iain Wilson and Maureen Taylor for comments on the manuscript.

**REFERENCES**

1. Drickamer, K., and Taylor, M. E. (1993) *Annu. Rev. Cell Biol.* 9, 237–264
2. Dodd, R. B., and Drickamer, K. (2000) *Glycobiology* 11, 71R–79R
3. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem.* 65, 441–473
4. Drickamer, K. (1999) *Curr. Opin. Struct. Biol.* 9, 585–590
5. Drickamer, K. and Dodd, R. B. (1999) *Glycobiology* 9, 1587–1599
6. Drickamer, K. (1992) *Nature* 360, 183–186
7. Bouyain, S., Rashton, S., and Drickamer, K. (2001) *Glycobiology* 11, 989–996
8. Fornstedt, N., and Porath, J. (1975) *FEBS Lett.* 77, 187–191
Drosophila Receptor for α1,3-Linked Fucose

9. Grayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y., and Inouye, M. (1984) *EMBO J.* 3, 2437–2442
10. Quesenberry, M. S., and Drickamer, K. (1992) *J. Biol. Chem.* 267, 10831–10841
11. Bolton, A. E., and Hunter, W. M. (1973) *Biochem. J.* 133, 529–539
12. Burnette, W. N. (1981) *Anal. Biochem.* 112, 195–203
13. Mizuochi, T., Loveless, R. W., Lawson, A. M., Chai, W., Lachmann, P. J., Childs, R. A., Thiel, S., and Feizi, T. (1989) *J. Biol. Chem.* 264, 13834–13839
14. Laemmli, U. K. (1970) *Nature* 227, 680–685
15. Wallis, R., and Drickamer, K. (1997) *Biochem. J.* 325, 391–400
16. Cohn, E. J., and Edsall, J. T. (1943) in *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, pp. 370–381, Reinhold, New York
17. Bairoch, A., and Apweiler, R. (2000) *Nucleic Acids Res.* 28, 45–48
18. Engelmann, D. M., Steitz, T. A., and Goldman, A. (1986) *Annu. Rev. Biophys. Biophys. Chem.* 15, 321–353
19. von Heijne, G. (1985) *Eur. J. Biochem.* 133, 17–21
20. Ng, K. K. S., Drickamer, K., and Weis, W. I. (1996) *J. Biol. Chem.* 271, 663–674
21. Blance, O., Jobst, S. T., Gabel, C., and Drickamer, K. (1996) *J. Biol. Chem.* 271, 7289–7292
22. Ng, K. K. S., and Weis, W. I. (1997) *Biochemistry* 36, 979–988
23. Carr, S. A., and Roberts, G. D. (1986) *Anal. Biochem.* 157, 396–406
24. Dorland, L., van Halbeek, H., Vliegenthart, J. F. G., Lis, H., and Sharon, N. (1981) *J. Biol. Chem.* 256, 7708–7711
25. Childs, R. A., Feizi, T., Yuen, C. T., Drickamer, K., and Quesenberry, M. S. (1990) *J. Biol. Chem.* 265, 20770–20777
26. Kurosaka, A., Yano, A., Ish, N., Kuroda, Y., Nakagawa, T., and Kawasaki, T. (1991) *J. Biol. Chem.* 266, 4168–4172
27. Wilson, I. B. H., Harthill, J. E., Mullin, N. P., Ashford, D. A., and Altman, F. (1998) *Glycobiology* 8, 651–661
28. Fabini, G., Freiling, A., Altman, F., and Wilson, I. B. H. (2001) *J. Biol. Chem.* 276, 28058–28067
29. Raju, T. S., Briggs, J. B., Borge, S. M., and Jones, A. J. S. (2000) *Glycobiology* 10, 477–486
30. Somers, W. S., Tang, J., Shaw, G. D., and Camphausen, R. T. (2000) *Cell* 103, 467–479
31. Jan, L. Y., and Jan, Y. N. (1982) *Proc. Natl. Acad. Sci. U. S. A.* 79, 2700–2704
32. Snow, P. M., Patel, N. H., Harrelson, A. L., and Goodman, C. S. (1987) *J. Neurosci.* 7, 4137–4144
33. Katz, F., Muats, W., and Jan, Y. N. (1988) *EMBO J.* 7, 3471–3477
34. Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M., and Harvey, D. A. (2000) *Science* 287, 2222–2224
35. Seppo, A., and Tiemeyer, M. (2000) *Glycobiology* 10, 751–760
36. Altman, F., Fabini, G., Ahorn, H., and Wilson, I. B. H. (2001) *Biochemie* 83, 703–712
An Endogenous *Drosophila* Receptor for Glycans Bearing α1,3-Linked Core Fucose Residues

Samuel Bouyain, Nicholas J. Silk, Gustáv Fabini and Kurt Drickamer

*J. Biol. Chem.* 2002, 277:22566-22572.
doi: 10.1074/jbc.M202825200 originally published online April 17, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M202825200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 35 references, 13 of which can be accessed free at http://www.jbc.org/content/277/25/22566.full.html#ref-list-1