The human genome sequence has been examined to identify the complete set of proteins related to the human glycans-binding receptor, DC-SIGN. In addition to five SIGNR proteins previously described, a pseudogene, encoding a hypothetical SIGNR6, and a further two expressed proteins, SIGNR7 and SIGNR8, have been identified. The ligand-binding properties of these novel proteins and of the previously described mouse SIGNs have been systematically investigated in order to define the mouse proteins that most resemble human DC-SIGN and DC-SIGNR. Results from screening of a glycan array demonstrate that only mouse SIGNR3 shares with human DC-SIGN the ability to bind both high mannose and fucose-terminated glycans in this format and to mediate endocytosis. The finding that neither SIGNR1 nor SIGNR5 binds with high affinity to specific ligands in a large panel of mammalian glycans is consistent with the suggestion that these receptors bind surface polysaccharides on bacterial and fungal pathogens in a manner analogous to serum mannose-binding protein. The data also reveal that two of the mouse SIGNs have unusual binding specificities that have not been previously described for members of the C-type lectin family; the newly identified SIGNR7 binds preferentially to the 6-sulfo-sialyl Lewis x oligosaccharide, whereas SIGNR2 binds almost exclusively to glycans that bear terminal GlcNAc residues. The results presented demonstrate that the mouse homologs of DC-SIGN have a diverse set of ligand-binding and intracellular trafficking properties, some of which are distinct from the properties of any of the human receptors.

The human receptor designated as the dendritic cell-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin (DC-SIGN) has been identified both as an adhesion molecule that facilitates attachment of T cells to dendritic cells and as a potential pathogen-binding receptor (1, 2). DC-SIGNR or L-SIGN is a closely related receptor found on endothelial cells of liver, lymph nodes, and placenta (3). Both of these receptors have been of considerable interest because of their ability to bind human immunodeficiency virus and present it to CD4-positive T cells, greatly facilitating the efficiency of infection (4). Similar trans activity for other viruses has been reported, and the receptors can also directly mediate infection of cells in cis (3, 5, 6). There is also evidence that these receptors interact with bacterial pathogens and with parasites (7). Although the original names do not encompass all of the functions and expression patterns of DC-SIGN and DC-SIGNR, these acronyms have generally been retained and used as simple designations for the molecules.

Both DC-SIGN and DC-SIGNR bind to surface glycoproteins of viruses by interacting with high mannose oligosaccharides. They bind with highest affinity to larger glycans that contain 8 or 9 mannose residues (8, 9). In addition, DC-SIGN, but not DC-SIGNR, binds to fucose-containing glycans, such as those present on the surfaces of nematode parasites (7, 9). The sugar–binding activity of each protein is conferred by a C-type or Ca2+-dependent carbohydrate recognition domain that is located at the C terminus of the receptor polypeptide. This domain is separated from the membrane anchor by a neck region consisting of multiple 23-amino acid repeats. The neck forms an extended structure that associates to create a tetramer at the cell surface (10). Signals in the N-terminal cytoplasmic domain of DC-SIGN direct internalization of the receptor, which can thus mediate endocytosis and degradation of glycoproteins (9). DC-SIGNR lacks such signals and seems not to be a recycling receptor.

Initial screening of mouse cDNA libraries led to the identification of multiple mouse homologs of DC-SIGN and DC-SIGNR. These have been designated DC-SIGN and SIGNR1 through SIGNR4 (11). Based on the sequences and expression patterns of these molecules, it has been difficult to define which of the cDNAs encode the mouse molecules that perform functions analogous to the two human proteins. Identification of such “functional orthologs” is essential if mice are to be useful models for illuminating the biology of human DC-SIGN and DC-SIGNR. Since it is not clear that the molecule originally designated mouse DC-SIGN is the closest homolog of human DC-SIGN, it is referred to here as SIGNR5.

In the present work, further mouse homologs of the human SIGNs have been identified, and the biochemical and cell biological properties of all the mouse SIGNs have been compared. The results reflect recent, independent divergence of the SIGN

---

*This work was supported by Wellcome Trust Grant 075565 and National Institutes of Health Grant GM62116 (to the Consortium for Functional Glycomics). 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.

1 These authors contributed equally to this work.
2 Recipient of a Vacation Scholarship from the Wellcome Trust.
3 To correspondence should be addressed: Division of Molecular Biosciences, Biochemistry Bldg., Imperial College, London SW7 2AZ, United Kingdom. Tel.: 44-20-7594-5282; E-mail: k.drickamer@imperial.ac.uk.
4 The abbreviations used are: DC-SIGN, dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin; DC-SIGNR, DC-SIGN-related; CRD, carbohydrate-recognition domain; BSA, bovine serum albumin; Bicine, N,N-bis(2-hydroxyethyl)glycine.
family in the human and mouse lineages. The ligand-binding specificity of mouse SIGNR3 is most like human DC-SIGN, whereas several of the other mouse SIGNs have properties that are distinct from either of the human receptors.

EXPERIMENTAL PROCEDURES

Data Base Screening—The Ensembl annotation of the mouse genome was screened with InterPro profile IPR001304 to identify genes containing potential C-type CRDs. One of the genes (encoding SIGNR6) was identified in earlier screening of the genome annotation but is not currently identified as a gene. Similar searches of the SwissProt and GenBank sequence data bases were performed. DC-SIGN homologs were distinguished from the general pool of C-type lectins based on the presence of a characteristic extra disulfide bond found only in human DC-SIGN, DC-SIGNR, and CD23 as well as overall relatedness of sequences. The complete set of C-type lectin-like proteins is described in the Genomic Resource for Animal Lectins (available on the World Wide Web at www.imperial.ac.uk/research/animallectins). Cluster analysis was performed using the AlignX software from Informax.

cDNA Screening and Cloning—Primers of 30–36 nucleotides were designed from the genomic sequences and cDNAs for SIGN R1, R3, and R5 were amplified from spleen, SIGN R2, R4 and R7 from lung, and SIGN R8 from liver cDNA libraries using the Advantage 2 polymerase (both from Clontech) following the PCR protocol described previously (12) and were cloned using the TOPO cloning kit (Invitrogen). Multiple clones were sequenced, and, where necessary, portions of different cDNAs were combined to create error-free clones. PCR screening of tissues was conducted with a panel of mouse cDNAs (Clontech) using 1 μl of template in the same PCR protocol run for 30 cycles.

Protein Expression and Purification—For expression in bacteria, a fragment of each receptor encoding the extracellular domain was preceded by the nucleotides encoding the sequence Met-Ala in the expression vector pT5T. Expression of the protein and isolation of inclusion bodies followed the procedure used for human DC-SIGN and DC-SIGNR (8). Inclusion bodies from 6 liters of bacteria were solubilized in 100 ml of 6 M guanidine HCl containing 100 mM Tris-Cl, pH 7.0, and renatured following various different schemes. In the direct dialysis protocol, the protein was dialyzed against three changes of 20 volumes of loading buffer (1.25 M NaCl, 25 mM Tris-Cl, pH 7.8, and 25 mM CaCl₂). In the dilution protocol, the sample was diluted slowly into five volumes of loading buffer and dialyzed against three changes of 10 volumes of loading buffer. In some cases, 1% Triton X-100 was added to the sample in guanidine before dialution. Following dialysis against loading buffer, some of the samples were dialyzed further into water, lyophilized, and dissolved in one-twentieth of their original volume of loading buffer.

After each sample was spun at 100,000 × g for 30 min, the supernatant was applied to a 5-ml column of mannosose-Sepharose prepared by the divinyl sulfone method (13). Columns were washed with 20 ml of loading buffer and eluted with 10 ml of elution buffer (1.25 mM NaCl, 25 mM Tris-Cl, pH 7.8, and 2.5 mM EDTA). For proteins that bound less tightly, 2-ml fractions were collected during the wash, and in all cases, 1-ml fractions were collected during elution. Fractions were examined on 17.5% SDS-polyacrylamide gels, which were stained with Coomassie Blue.

Formation of Tetrameric Proteins—Synthetic oligonucleotides were used to append sequences encoding a short glycine linker followed by the biotinylation sequence Gly-Leu-Asn-Asp-Ile-Phe-Glu-Ala-Gln-Lys-Ile-Glu-Trp-His-Glu at the C termini of the extracellular domain fragments (14). Proteins were expressed in Escherichia coli strain BL21/D3 containing plasmid pACYC-184 with the birA gene (Avidity, Denver, CO) and grown in the presence of 20 μg/ml chloramphenicol and 50 μg/ml ampicillin. For most of the proteins, tetramers were formed following protein expression as described above. Fractions from the affinity columns were pooled and adjusted to 25 mM CaCl₂ and streptavidin (Sigma) was added to approximately half the concentration of the tagged protein and incubated overnight at 4 °C. Complexes were isolated by affinity chromatography on 2-ml columns of mannosose-Sepharose that were washed with 10 ml of loading buffer and eluted with 5 ml of elution buffer. In some cases, streptavidin was added directly to the renatured protein before affinity chromatography.

Binding Studies—Solid-phase competition assays and pH dependence assays were both performed using CRDs immobilized on polystyrene and probed with ¹²⁵I-Man-BSA (E-Y Laboratories) as described before. Additional binding assays with fluorescein-labeled, glycosylated polyacrylamide polymers (GlycoTech) were performed in the same format and read on a Victor3 plate reader from PerkinElmer Life Sciences. In these studies, the buffer was 150 mM NaCl, 25 mM Tris-Cl, pH 7.4, 2 mM CaCl₂, with 1% BSA present during ligand binding. Following subtraction of background, the results were fitted to a simple binding equation (ligand bound proportional to ligand input/(Kᵦ + ligand input)) using SigmaPlot software. For probing of the glycan array, tetramerized proteins were labeled with fluorescein isothiocyanate (9) following dialysis into 1.25 M NaCl, 25 mM Na-HEPES, pH 8.0, 25 mM CaCl₂ or after repurification on a mannosose-Sepharose column that was rinsed with buffer containing 1.25 M NaCl, 25 mM Bicine-Cl, pH 9.0, and 25 mM CaCl₂ and eluted with 1.25 M NaCl, 25 mM Bicine-Cl, pH 9.0, and 2.5 mM EDTA. The array was screened by Core H of the Consortium for Functional Glycomics following their standard protocol.

Protein Characterization—Cross-linking with bis-sulfosuccinimidyl suberate and analytical ultracentrifugation were performed following published procedures. Gel filtration was performed on a Superose 20 column (Amersham Biosciences) in 100 mM NaCl, 25 mM Tris-Cl, pH 7.8, and 2.5 mM EDTA at a flow rate of 0.5 ml/min.

Assay of Endocytic Activity—Full-length cDNAs inserted into the vector pVCos along with the neomycin resistance gene were transfected into Rat-6 fibroblasts (15). Following selection in 400 μg/ml G418, individual clones were selected and tested for endocytosis and degradation of ¹²⁵I-mannose-BSA (16). Protein expression was verified by Western blotting with antibodies to the bacterially expressed extracellular domain, raised in sheep.

RESULTS

Sequence Comparisons of Mouse SIGNs—Potential mouse orthologs of human DC-SIGN were previously identified by
screening of cDNAs from mouse dendritic cells and other cells of the immune system. In order to complement this approach, the annotated mouse genomic sequence and the SwissProt and GenBank™ data bases were screened for sequences that contain profile matches to the C-type CRD motif. The CRDs from these sequences were excised and subjected to cluster analysis, leading to identification of eight possible genes containing CRDs that are closely related to the human DC-SIGN and DC-SIGNR sequences. In addition to overall sequence similarity, the potential CRDs encoded in these sequences contain an extra disulfide bond that is characteristic of human DC-SIGN and DC-SIGNR as well as CD23, the lymphocyte low affinity IgE Fc receptor, and another recently described sugar-binding receptor designated LSECtin. The genes encoding each of these DC-SIGN homologs are located between 3.0 and 3.5 megabases on mouse chromosome 8, directly adjacent to the genes encoding CD23 and LSECtin (Fig. 1). Five of the genes encode the previously described mouse SIGNR1 to -R5, whereas three novel genes were identified, and the encoded proteins were designated SIGNR6, SIGNR7, and SIGNR8. cDNAs corresponding to SIGNR7 and -R8 have been characterized as part of the RIKEN large scale cDNA sequencing project.

The probable domain organization of the expanded mouse SIGN family is shown in Fig. 2. The sequence for SIGNR6 does not appear to encode a full-length CRD, because the N-terminal part of the deduced sequence does not contain characteristic cysteine residues and hydrophobic amino acids. It has previously been noted that SIGNR2 lacks a membrane anchor. The remaining mouse SIGNs have relatively short neck regions compared with the human proteins. The relationships between the neck regions are summarized in diagrammatic form in Fig. 2. Each of the mouse SIGNs except SIGNR6 contains a sequence immediately preceding the CRD that is somewhat similar to the sequence repeated seven times in the human proteins. This sequence is repeated four times in SIGNR1 and twice in SIGNR2 but with different intermediate spacers (Fig. 3). The remaining SIGNs do not have repeated sequences and thus have very short neck regions.

Based on the predicted sequences of SIGNR6, -R7, and -R8, primers were designed and used to amplify corresponding cDNAs from libraries constructed from various mouse tissues. The cDNAs for mouse SIGNR7 and SIGNR8 were detected in multiple tissues (Fig. 4), but no signal for SIGNR6 was detected using any of three different forward and reverse primers. These results, combined with the truncated 3’ end of the gene, suggest that the SIGNR6 gene is a pseudogene. cDNAs for SIGNR7 and -R8 were successfully cloned, and the cDNA for SIGNR7 exactly matched the genomic sequence as well as the amended cDNA sequence deposited in the GenBank™ data base under accession number XM_284376. The cDNA for SIGNR8 matched the genomic sequence but differed from the cDNA deposited as XM_284386 in GenBank™ because of the absence of a short inserted sequence just before the CRD (Fig. 3). The cDNA in the data base would result from alternative splicing at the 5’ end of exon 6 and corresponds to a slightly
larger messenger RNA detected as a secondary band in the PCRs shown in Fig. 4.

The relationships between the sequences of the CRDs from the human and mouse SIGNs are summarized in the dendrogram in Fig. 5. The most striking finding is that the two human proteins are more closely related to each other than they are to any of the mouse SIGNs. Thus, although the human and mouse families have a common ancestor, they have radiated independently in the lineages leading to the two species, and it is therefore not possible to identify individual mouse orthologs of either of the two human proteins. Because these findings make it difficult to predict which mouse proteins share functions with the human proteins based on sequence alone, further biochemical characterization was undertaken.

**Demonstration of Carbohydrate-binding Activity of Mouse SIGNs**—The sequences of the CRDs of all but one of the expressed mouse SIGNs show conservation of the key residues that form Ca\(^{2+}\)/H\(^{11001}\)- and carbohydrate-binding sites in these proteins (Fig. 3). Therefore, it seemed likely that most would bind some combination of mannose- or fucose-containing ligands like the human SIGNs. Fragments corresponding to the extracellular domains of each protein were expressed in a bacterial expression system previously used to generate high levels of the human proteins. Following denaturation of protein from inclusion bodies using guanidine hydrochloride, various methods for renaturation were tested, involving dialysis or rapid dilution followed by dialysis, either in the presence or absence of Triton X-100. The proteins were then passed over affinity columns.
containing high densities of immobilized mannose in the presence of Ca\(^{2+}\) and eluted with EDTA. In some cases, yields were improved by concentrating the renatured proteins by dialysis against water and lyophilization.

Of the seven mouse SIGN extracellular domains that were expressed, five were purified successfully by affinity chromatography on mannose-Sepharose (Table 1 and Fig. 6). However, several of the proteins bounded relatively poorly and could be washed off shorter affinity columns even in the presence of Ca\(^{2+}\). This behavior was previously observed for the CRDs of the human SIGNs, which are monomeric, whereas the tetramers formed by the full extracellular domains bind tightly and can only be eluted with EDTA. Sedimentation and cross-linking analysis of the five mouse proteins that could be purified by affinity chromatography revealed that four are clearly monomeric and show no self-association. Only SIGNR3 associates weakly to form dimers in equilibrium analytical centrifugation experiments (data not shown). Thus, like the human proteins, the CRDs in these mouse SIGNs bind with modest affinity to mannosese-Sepharose. However, in contrast to the human proteins, the neck segments of the mouse SIGNs are insufficient to form stable oligomers.

It was notable that SIGNR1 is one of the two mouse SIGNs that was not successfully purified on mannosese-Sepharose, although this protein has been reported to bind to dextran, mannan, and other bacterial surface polysaccharides (17–20). The results for the other mouse SIGNs suggested that this lack of binding might reflect particularly weak affinity of monomeric CRDs for the immobilized monosaccharide. In order to improve the binding characteristics of the mouse SIGNs, sequences containing a single lysine residue in an appropriate context for biotinylation by biotin ligase were appended to the C-terminal ends of each of them (21). These proteins were expressed in bacteria co-expressing the bacterial birA gene, which encodes biotin ligase. Blotting studies with radioiodinated streptavidin revealed that this procedure resulted in efficient biotinylation of each of the tagged proteins (data not shown). Purification results for the proteins containing biotin tags were identical to those for the CRDs alone, so that four of the tagged proteins could be isolated using mannosese-Sepharose. Following incubation with streptavidin, these proteins could be repurified by affinity chromatography, and in each case, they now bound tightly to the column and required EDTA for elution (Fig. 6B). Analysis by gel filtration chromatography revealed that most of these complexes are homogeneous species with molecular weights corresponding to four extracellular domain fragments complexed with a single streptavidin tetramer. Similar complexes of SIGNR1 were also observed, but there was also evidence of aggregation, probably reflecting the propensity of the extended neck region in this extracellular domain fragment to self-associate.

Based on these results, further attempts were made to purify SIGNR1 and -R4, which were precomplexed with streptavidin following renaturation. This procedure led to the successful isolation of a tetrameric streptavidin-SIGNR1 complex that bound well to mannosese-Sepharose and eluted with EDTA (Fig. 6). However, no binding activity was observed for SIGNR4. This finding is probably not surprising, since SIGNR4 is the only

TABLE 1

| Protein         | Alternate names | Purification protocol | \(K_{s,Fucose}/K_{s,Mannose}\) | \(pK_a\) |
|-----------------|----------------|-----------------------|--------------------------------|---------|
| Hu-DC-SIGN      | CD209          | Diluted, lyophilized  | 0.6 ± 0.1 a                   | 5.92 ± 0.05 a |
| Hu-DC-SIGNR     | CD209L, L-SIGN | Diluted, lyophilized  | 1.7 ± 0.1 a                   |         |
| Mo-SIGNR1       | CD209B         | Streptavidin complex   | 0.76 ± 0.12 c                 | 5.95 ± 0.05 |
| Mo-SIGNR2       | CD209C         | Diluted               | 1.0 ± 0.1 d                   | 6.08 ± 0.04 |
| Mo-SIGNR3       | CD209D         | Diluted               | 0.62 ± 0.03 c                 | 6.41 ± 0.07 |
| Mo-SIGNR4       | CD209E         | Diluted, lyophilized  | 1.21 ± 0.06 b                 |         |
| Mo-SIGNR5       | CD209A, DC-SIGN| Diluted, lyophilized  | 0.60 ± 0.15 c                 | 5.66 ± 0.06 |
| Mo-SIGNR6       | Undiluted      |                       | 1.9 ± 0.1 c                   | 6.22 ± 0.05 |
| Mo-SIGNR8       | Undiluted      |                       |                               |         |

a Data from Ref. 8.

b Data from Ref. 9.

c Reverse dependence on pH.
mouse SIGN in which some of the residues that ligate the critical Ca\(^{2+}\) that forms the core of the primary sugar-binding site differ from the pattern seen in the human SIGNs and other mannose-binding C-type CRDs (Fig. 3). Although it remains possible that this protein binds to other sugars, these findings indicate that it is unlikely to function as an ortholog of the human SIGNs.

**Carbohydrate-binding Activities of Mouse SIGNs**—Initial characterization of the sugar-binding characteristics of the six mannose-binding mouse SIGNs was undertaken using solid-phase competition binding assays in which \(^{125}\text{I}-\text{mannose-BSA}\) served as a reporter ligand. The purified extracellular domains of SIGNR2, -R3, -R5, -R7, and -R8 were used to coat polystyrene wells, whereas the tetrameric form of SIGNR1 was used. The relative abilities of mannose and fucose to compete for binding were compared, since the ratio of inhibition constants for these two sugars distinguishes the two human SIGNs. The results, which are summarized in Table 1, reveal that mouse SIGNR1, -R3, and -R7 all show preferential binding to fucose over mannose, with a ratio of \(K_I\) values that is similar to the value of 0.6 obtained for human DC-SIGN. Only SIGNR5 and -R8 show preferential binding to mannose, to an extent similar to that observed for human DC-SIGN. Mouse SIGNR2 seems to bind equally well to mannose and fucose. Thus, to a first approximation, the best candidates for functional orthologs of human DC-SIGN are SIGNR1, -R3, and -R7.

Further insights into the binding characteristics of the mouse SIGNs were obtained by screening a glycan array using fluorescently labeled proteins. Preliminary tests with some of the monomeric forms of the extracellular domains indicated that these did not bind with sufficient affinity to yield good results in the array, so the tetrameric complexes of biotin-tagged proteins were used in all cases. The glycan array contains over 170 mono- and oligosaccharides immobilized as biotin glycosides bound to streptavidin-coated wells. Despite the fact that the tetrameric complexes of mouse SIGNR1 and -R5 bind well to mannose-containing affinity columns, they fail to bind selectively to any glycans on the array. Similar results have been obtained for well characterized glycan-binding proteins, such as serum mannose-binding protein, that interact exclusively with terminal sugars.\(^5\) In these cases, although the CRDs are clustered in oligomers, the interaction between individual CRDs and terminal monosaccharides is too weak to retain the protein on the plates during the washing process. The density

---

\(^5\) A. Powlesland and K. Drickamer, unpublished observations.
of terminal sugars on the streptavidin-coated surfaces is apparently significantly lower than the density on the highly substituted resins used for the purification.

In contrast to the results for SIGNR1 and -R5, mouse SIGNR3 binds selectively to specific subsets of sugars on the array (Fig. 7). The best ligands fall into two structural categories consisting of high mannose glycans or fucose-containing oligosaccharides. The latter almost invariably display paired terminal fucose and galactose or GalNAc residues. This dual specificity is very similar to the pattern previously observed for human DC-SIGN, providing strong evidence that, in terms of ligand binding, SIGNR3 is the mouse protein that most closely resembles human DC-SIGN.

Although mouse SIGNR7 fails to bind to most glycans on the array, two targets gave signals substantially above background. The 6'-sulfo-sialyl Lewis^s^ structure, which contains a sulfate group on the galactose residue in the Lewis^s^ trisaccharide, was the best ligand, and binding to an extended sialyl-Lewis^s^ structure was also detected (Fig. 7). In contrast, 6-sulfo-sialyl Lewis^s^, in which the sulfate is attached to the GlcNAc residue (glycan 148 in Fig. 7), failed to show significant binding. These results indicate that mouse SIGNR7 displays an unusual preference for a specific glycan. Further confirmation of this selective interaction was obtained by quantifying binding of glycosylated polyacrylamide derivatives to immobilized SIGNR7 (Fig. 8). When the relative affinities of structurally related ligands were compared using this assay, 6'-sulfo-sialyl Lewis^s^ was found to be the ligand with highest affinity.

The glycan array results for SIGNR2 indicate that it binds preferentially to glycans terminating in GlcNAc (Fig. 7). This result was confirmed in a monosaccharide competition assay, which demonstrated that GlcNAc competes for binding to SIGNR2 dramatically better than mannose does \( \frac{K_d}{K_d} \) (Mannose) = 80 \pm 9). The array results suggest that some additional types of ligands, terminating in galactose, also bind to SIGNR2, but most of the preferred ligands contain multiple terminal GlcNAc residues. Such selectivity for GlcNAc has been observed for the chicken hepatic lectin, which is an endocytic receptor that clears glycoproteins
from chicken serum (22), but GlcNAc-specific binding has not been described for any other human or mouse C-type lectins.

Endocytic Activity of Mouse SIGNs—Studies with human DC-SIGN and DC-SIGNR expressed in fibroblasts were previously used to demonstrate that, in addition to differing in ligand-binding activity, these two proteins have different cellular trafficking properties (9). The ability to release ligands at endosomal pH was shown to be a prerequisite for efficient uptake and degradation of glycoprotein ligands. For this reason, the pH sensitivity of binding was determined for the five membrane-bound mouse SIGNs that bind to sugar ligands (Fig. 9). Two of the proteins, SIGNR1 and -R3, release ligand at pH values commensurate with those found in endosomes. SIGNR7 also releases ligand at low pH, but the midpoint of the pH curve is noticeably lower than the curves for the other two proteins, which are more similar to those of previously characterized endocytic receptors, such as human DC-SIGN and the asialoglycoprotein receptor. In contrast to these three proteins, SIGNR5 does not display pH-dependent ligand release over a physiologically relevant range. In this respect, it appears to be more like human DC-SIGNR than human DC-SIGN. SIGNR8 shows a reverse pH dependence that has not been observed for other C-type lectins.

Although pH dependence of ligand binding is a necessary requirement for endocytic activity, appropriate internalization signals, usually in the cytoplasmic tail of the receptor polypeptide, must also be present. Examination of the cytoplasmic domains of the mouse SIGNs (Fig. 3) reveals that several of the proteins contain paired hydrophobic residues, but only SIGNR3 contains a typical tyrosine-based internalization motif (YSDI). Previous studies with mouse SIGNR1 and -R3 have demonstrated that these receptors have endocytic activity, whereas results for SIGNR5 were ambiguous (19, 23). For comparison, full-length versions of SIGNR3 and SIGNR5 were expressed in fibroblasts and their abilities to take up and degrade mannose-BSA were compared (Fig. 10). As expected from the pH profile for ligand binding to SIGNR5, no evidence of ligand uptake and degradation was observed. Endocytic activity was observed for SIGNR3, providing further evidence that the properties of this receptor closely parallel those of human DC-SIGN.

DISCUSSION

The genomic analysis presented here defines the full set of mouse homologs of human DC-SIGN, but the results shown in Figs. 3 and 5 indicate that sequence comparison alone is not sufficient to suggest that one of these proteins is a specific ortholog. The biochemical properties of the mouse SIGNs summarized in Table 2 reveal that, of the five mouse SIGNs that are sugar-binding membrane receptors, the protein that is most like human DC-SIGN is SIGNR3. Mouse SIGNR3 shares with human DC-SIGN the ability to bind both high mannose and fucose-terminated glycans, and it can mediate endocytosis of glycoprotein ligands. Mouse SIGNR5, which was originally described as the mouse equivalent of human DC-SIGN, is in some respects more like human DC-SIGNR, because it shows preferential binding of mannose over fucose, does not release ligand at endosomal pH, and lacks endocytic activity. However, in glycan array experiments, SIGNR5 does not bind to the set of

FIGURE 8. Binding of fluorescein-labeled, glycosylated derivatives of polyacrylamide to immobilized mouse SIGNR7. Binding was quantified in two independent assays, and the average $K_D$ values measured are indicated.

FIGURE 9. pH dependence of ligand binding to mouse SIGNs. Binding of $^{125}$I-labeled mannosyl-BSA to immobilized SIGN proteins was determined at 1 mM Ca$^{2+}$.
high mannose oligosaccharides that are the primary ligands for human DC-SIGNR. With the full set of mouse SIGNs defined and expressed, it will be possible to develop specific antibodies that can be used to examine the cellular expression patterns of these different proteins, which will provide further insights into the human-mouse comparisons.

Previous studies in which some of the mouse SIGNs were tested against selected oligosaccharides have produced results that are generally consistent with the findings reported here (23). For SIGNR3, the more comprehensive panel of glycans used in the present studies makes it possible to see a pattern in the types of ligands bound, which also makes the comparison with human DC-SIGN more evident. In the case of SIGNR1, binding to a number of glycans has been previously detected in assays in which densely spaced neoglycolipid ligands on nitrocellulose were probed with highly multivalent complexes of the proteins (23). The array used in the present studies presents glycans at relatively sparse density defined by the spacing between the binding sites in the streptavidin tetramer and is therefore most useful for detecting and comparing relatively strong interactions between carbohydrate recognition domains and glycans. Consequently, serum mannos-binding protein, which interacts only with terminal monosaccharides and has low intrinsic affinity for individual glycans, does not bind effectively to this array, whereas selective binding is readily observed for human DC-SIGN and other proteins with extended binding sites that accommodate oligosaccharides and thus have higher affinity for the individual glycans (9). Therefore, the failure of SIGNR1 to bind to glycans in this format is consistent with the published finding that it interacts preferentially with dense arrays of sugars characteristic of polysaccharides on the surface of microorganisms and mediates protection against bacterial pathogens (17, 18, 20, 24, 25).

The sequence relationships summarized in Fig. 5 suggest that the DC-SIGN family of proteins has undergone substantially different patterns of divergence in the human and mice lineages. The properties of mouse SIGNR3 are sufficiently similar to those of human DC-SIGN that they may fulfill similar functions that have been largely conserved, whereas the gene has been independently duplicated in mice and humans, with the duplicated genes evolving to perform distinct functions. The roles of human DC-SIGN and DC-SIGNR as receptors that facilitate or mediate viral infection of cells suggest that the different spectrum of viral or other pathogens confronted by different species have created evolutionary pressure leading to rapid divergence of this family.

Analysis of SIGNR2 shows that this protein has diverged from the other mouse SIGNs both in its ligand-binding properties and in the biological functions that it mediates. Its capacity for binding GlcNAc suggests that it might bind pathogens in a manner similar to ficolins (26), but it lacks obvious accessory domains that might activate complement or initiate other responses. In fact, it is not clear how the functional protein would be expressed, because all of the cDNAs that have been isolated lack signal sequences to direct the polypeptide to the secretory pathway, which would be required for disulfide bond formation. Thus, it is possible that the mouse SIGNR2 gene is an expressed pseudogene that produces an mRNA but no functional protein.

Preferential binding of SIGNR7 to the 6'-sulfo-sialyl Lewisx oligosaccharide is reminiscent of the specificity observed for some members of the siglec family of sialic acid-binding receptors, including mouse siglec-F and human siglec-8 (27, 28). In humans, the 6'-sulfo-sialyl Lewisx structure is expressed on the mucin-like protein GlyCAM-1, but the sites of expression have not been fully explored. This oligosaccharide has also been

---

![FIGURE 10. Endocytic activity of mouse SIGNs.](image)

**TABLE 2**

Summary of the properties of the mouse SIGNs

| Protein    | mRNA expressed | Membrane anchor | Sugar binding | Endosomal release | Endocytosis | Binding to Lewisx |
|------------|----------------|-----------------|---------------|--------------------|-------------|------------------|
| Mo-SIGNR1  | +              | +               | +             | +/−                | +           | −                |
| Mo-SIGNR2  | +/−            | +               | +             | +                  | ND          | +/−              |
| Mo-SIGNR3  | +              | +               | +             | ND                 | ND          | ND               |
| Mo-SIGNR4  | +              | +               | +             | ND                 | ND          | ND               |
| Mo-SIGNR5  | +              | +               | +             | ND                 | ND          | ND               |
| Mo-SIGNR6  | −              | ND              | ND            | ND                 | ND          | ND               |
| Mo-SIGNR7  | +/−            | +               | +             | +                  | ND          | −                |
| Mo-SIGNR8  | +/−            | +               | +             | +/−                | ND          | ND               |

* ND, not determined.
reported to be a ligand for mouse langerin (23). The siglecs contain CRDs that are based on the immunoglobulin fold and have primary binding sites that selectively recognize sialic acids. Although a similar preference for the 6'-sulfo-sialyl Lewis\textsuperscript{x} ligand is observed for mouse SIGNR7, it is expected that the primary binding site in this case would involve interaction with the terminal fucose residue. This convergent evolution of multiple receptors toward binding of a relatively uncommon oligosaccharide suggests that 6'-sulfo-sialyl Lewis\textsuperscript{x} may be a target for specialized recognition processes.

Acknowledgments—We thank Angela Lee and Rick Alvarez of the Consortium for Functional Glycomics for running the glycan array experiments and Dawn Torgersen for help with protein purification.

REFERENCES
1. Geijtenbeek, T. B. H., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Adema, G. J., van Kooyk, Y., and Figdor, C. G. (2000) Cell 100, 575–585
2. Geijtenbeek, T. B. H., Kwon, S. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Middel, J., Cornelissen, I. L. M. H. A., Nottet, H. S. L. M., KewalRamani, V. N., Littman, D. R., Figdor, C. G., and van Kooyk, Y. (2000) Cell 100, 587–597
3. Soilleux, E. J. (2003) Curr. Sci. 104, 437–446
4. Pohlmann, S., Baribaud, F., and Doms, R. W. (2001) Trends Immunol. 22, 643–646
5. Alvarez, C. P., Lasala, F., Carrillo, J., Muñoz, O., Corbí, A. L., and Delgado, R. (2002) J. Virol. 76, 6841–6844
6. Gardner, J. P., Arrigale, R. R., Donovan, G. P., Maddon, P. J., Drickamer, K., and Olson, W. C. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 4498–4503
7. Appelmelk, B. J., van Die, I., van Vliet, S. J., Vandenbroucke-Grauls, M. J. E., Geijtenbeek, T. B. H., and van Kooyk, Y. (2003) J. Immunol. 170, 1635–1639
8. Mitchell, D. A., Fadden, A. J., and Drickamer, K. (2001) J. Biol. Chem. 276, 28939–28945
9. Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I., and Drickamer, K. (2004) Nat. Struct. Mol. Biol. 11, 591–598
10. Feinberg, H., Guo, Y., Mitchell, D. A., Drickamer, K., and Weis, W. I. (2005) J. Biol. Chem. 280, 1327–1335
11. Park, C. G., Takahara, K., Umemoto, E., Yashima, Y., Matsubara, K., Matsuda, Y., Clausen, B. E., Inaba, K., and Steinman, R. M. (2001) Int. Immunol. 13, 1283–1290
12. Fadden, A. J., Holt, O. J., and Drickamer, K. (2003) Glycobiology 13, 529–537
13. Fornstedt, N., and Porath, J. (1975) FEBS Lett. 57, 187–191
14. Schatz, P. J. (1993) Biotechnology 11, 1138–1143
15. Stambach, N. S., and Taylor, M. E. (2003) Glycobiology 13, 401–410
16. Mellow, T. E., Halberg, D., and Drickamer, K. (1988) J. Biol. Chem. 263, 5468–5473
17. Geijtenbeek, T. B. H., Groot, P. C., Nolte, M. A., van Vliet, S. J., Nottet, H. S. L. M., KwaiRamani, V. N., Litman, D. R., Figdor, C. G., and van Kooyk, Y. (2002) Blood 100, 2908–2916
18. Kang, Y.-S., Yamazaki, S., Iyoda, T., Pack, M., Bruening, S. A., Kim, J. Y., Takahara, K., Inaba, K., Steinman, R. M., and Park, C. G. (2003) J. Exp. Med. 197, 177–186
19. Takahara, K., Yamashita, Y., Omatsu, Y., Yoshida, H., Kimura, Y., Kang, Y.-S., Steinman, R. M., Park, C. G., and Inaba, K. (2004) Internat. Immunol. 16, 819–829
20. Kang, Y.-S., Kim, J. Y., Bruening, S. A., Pack, M., Charalambous, A., Pritsker, A., Moran, T. M., Loeffler, J. M., Steinman, R. M., and Park, C. G. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 215–220
21. Chang, E. J., and Cronan, J. E., Jr. (1999) Trends Biochem. Sci. 24, 359–363
22. Kawasaki, T., and Ashwell, G. (1977) J. Biol. Chem. 252, 6536–6543
23. Galustian, C., Park, C. G., Chai, W., Kiso, M., Bruening, S. A., Kang, Y.-S., Steinman, R. M., and Feizi, T. (2004) Internat. Immunol. 16, 853–866
24. Lanoue, A., Clatworthy, M. R., Smith, P., Green, S., Townsend, M. J., Jolin, H. E., Smith, K. G. C., Fallon, P. G., and McKenzie, A. N. J. (2004) J. Exp. Med. 200, 1383–1393
25. Koppel, E. A., Wieland, C. W., van den Berg, V. C. M., Litjens, M., Florquin, S., van Kooyk, Y., van der Poll, T., and Geijtenbeek, T. B. H. (2005) Eur. J. Immunol. 35, 2962–2969
26. Matsushita, M., and Fujita, T. (2001) Immunol. Rev. 180, 78–95
27. Tatem, A. S., Kachroo, P. R., and Paulson, J. C. (2005) Glycobiology 15, 1125–1135
28. Bochner, B. S., Alvarez, R. A., Mehta, P., Bovin, N. V., Blixt, O., White, J. R., and Schnaar, R. L. (2005) J. Biol. Chem. 280, 4307–4312