Bacterial Polysaccharide Specificity of the Pattern Recognition Receptor Langerin Is Highly Species-dependent*§

Received for publication, August 4, 2016, and in revised form, November 29, 2016 Published, JBC Papers in Press, November 30, 2016, DOI 10.1074/jbc.M116.751750

Jonas Hanske§, Jessica Schulze§, Jonas Aretz§, Ryan McBride§, Bernhard Loll†, Henrik Schmidt§, Yuriy Knirel†, Wolfgang Rabsch**, Markus C. Wahl§, James C. Paulson§, and Christoph Rademacher§†

From the §Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces, Potsdam 14424, Germany, the §Department of Biology, Chemistry, and Pharmacy, Freie Universität Berlin, Berlin 14195, Germany; the §Department of Cell and Molecular Biology, Department of Immunology and Microbial Science and Department of Chemical Physiology, Scripps Research Institute, La Jolla, California 92037, the §N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, Moscow 119991, Russia, and the **Robert Koch Institute, Wernigerode Branch, National Reference Centre for Salmonellae and other Bacterial Enteric Pathogens, Wernigerode 38855, Germany

Edited by Luke O’Neill

The recognition of pathogen surface polysaccharides by glycan-binding proteins is a cornerstone of innate host defense. Many members of the C-type lectin receptor family serve as pattern recognition receptors facilitating pathogen uptake, antigen processing, and immunomodulation. Despite the high evolutionary pressure in host-pathogen interactions, it is still widely assumed that genetic homology conveys similar specificities. Here, we investigate the ligand specificities of the human and murine forms of the myeloid C-type lectin receptor langerin for simple and complex ligands augmented by structural insight into murine langerin. Although the two homologs share the same three-dimensional structure and recognize simple ligands identically, a screening of more than 300 bacterial polysaccharides revealed highly diverging avidity and selectivity for larger and more complex glycans. Structural and evolutionary conservation analysis identified a highly variable surface adjacent to the canon binding site, potentially forming a secondary site of interaction for large glycans.

Host-pathogen interactions are substantially facilitated via the recognition of surface glycans, and accordingly, both hosts and pathogens have evolved a large panel of specialized glycan-binding proteins (1–3). The mammalian innate immune system utilizes these pattern recognition receptors (PRRs)2 to detect foreign antigens and combat infections at an early stage (4).

Many members of the calcium-binding myeloid C-type lectin receptor (CLR) family are glycan-binding proteins that are involved in carbohydrate-mediated pathogen recognition and uptake, self/non-self discrimination, and regulation of the inflammatory response (5, 6). The C-type lectin domain fold is the defining structural feature of this protein family, and structural sequence alignments revealed high structural conservation even at low sequence homology among human CLRs (7–9). Binding to carbohydrate monomers is intrinsically weak and promiscuous with dissociation constants in the millimolar range. However, multivalent display of both glycan and receptor enhances binding avidities severalfold, confers sufficient affinity, and fine tunes specificity (10). Whereas other PRRs, such as Toll-like receptors, recognize highly conserved pathogen-associated molecular patterns and are the main initiators of the inflammatory response to pathogens (4), CLRs have a higher specificity and serve as regulators to fine tune the immune response (11). This is paralleled by the findings that expression and ligand specificity of Toll-like receptors are highly conserved in vertebrates (12), whereas some CLRs like DC-SIGN have no true homologs in the mammalian lineage (13).

The myeloid CLR langerin is a PRR expressed by professional antigen-presenting cells (APCs) and is involved in the detection and uptake of a wide set of pathogens, including viruses like HIV (14) and measles (15), fungi (16), and mycobacteria (17). At the same time, langerin recognizes antigens of endogenous origin, such as blood group antigens (18), heparin (19), and hyaluronic acid (20). The trimeric 37.5-kDa type II transmembrane receptor comprises a short intracellular domain, a transmembrane region, an α-helical neck domain, and the characteristic carbohydrate recognition domain (CRD). The CRD exhibits the common CLR fold, with an overall loop structure with N- and C-terminal β-sheets and two flanking α-helices, and harbors one calcium binding site in the long loop region (21). The conserved EPN motif in the primary binding site suggests a mannose-type carbohydrate binding preference (18, 22), but recent studies have revealed that human langerin can also accommo-
date 6-sulfated galactoses (23, 24) and that common polymorphisms shift glycan preference (25).

Whereas most in vitro studies were performed with human langerin, in vivo studies in mice used the native murine homolog that shares 66% of the sequence (26). So far, the two homologs are considered to be functionally interchangeable (8), although there are only limited data on the carbohydrate specificity of murine langerin (27), and no functional studies have been performed directly comparing ligand recognition of the two homologs. Moreover, the expression patterns of the two homologs are not identical; human langerin is expressed on CD1a⁺ APCs of the epithelium, epidermis, and dermis, whereas murine langerin is expressed on CD8⁺ dendritic cells, which do not exist in humans. To gain insight into the level of functional analogy, we solved the crystal structure of the murine langerin CRD and explored human and murine langerin ligand specificities side by side at different levels of ligand complexity, spanning from biophysical studies of simple monosaccharide ligands to microbial interaction with highly complex microbial glycans. Although binding of simple ligands is structurally and thermodynamically nearly identical, we found pronounced differences both in affinity and specificity toward complex bacterial glycans. These differences translated to the microbial level, as could be observed by a competition assay with glycan structures and direct binding of whole bacteria.

Results

Murine and Human Langerin Show Similar Affinity for Simple Mannose Ligands but Differ in Recognition of a Fungal Mannan Polysaccharide—To compare the functional homology of murine and human langerin, the trimeric extracellular domains (ECDs) of both proteins (Fig. 1, A and B) were used to investigate the affinities toward simple mannose-type structures, a commonly reported ligand for langerin (22). Isothermal titration calorimetry (ITC) measurements revealed dissociation constants of 7.0 ± 0.2 mM for the murine and 5.6 ± 1.7 mM for the human langerin (Fig. 1C). These values were identical within the error. To further substantiate the similarity of the mannose recognition, we obtained binding epitope information of a simple disaccharide ligand; α-methyl-(α-1,2)-mannobiose 1 was employed in STD NMR studies to derive epitope information for both homologs (Fig. 2A and supplemental Fig. S1). For both proteins, the binding epitopes were identical; the non-reducing end mannose receives most saturation with the highest effects on protons H2’, H3’, and H4’. These data agree very well with the canonic binding mode of simple glycans to EPN motif-carrying CLRs (28) and the crystal structure of human langerin bound to mannobiose (23). Moreover, the H5 received a strong STD effect when interacting with either of the two homologs, which is in agreement with the crystal structure of the human CRD that revealed a close packing of C5 to Ala-289 if the terminal mannose is coordinated to the Ca²⁺ (23).

Although the two homologs recognize the same epitope, the STD build-up curves of the murine langerin had a higher magnitude of the magnetization transfer, indicating slower off-rates than in the human langerin. In addition, we investigated the binding to the simple monosaccharide ligands N-acetyl glucosamine 2, fucose 3, and 6-sulfated galactose (Fig. 2, B–E, and supplemental Fig. S2), whose binding mode for human langerin has been described before (23). In the case of fucose and 6-sulfated galactose, the binding epitopes were identical for both the human and the murine homolog, whereas in the case of N-acetyl glucosamine, a differential saturation transfer at the 3-hydrogen was observed for the murine form, suggesting a differential binding mode.

Innate immune receptors such as langerin typically recognize substructures of pathogen-associated polysaccharides; hence, we expanded our analysis to mannan, a fungal, mannos-based polysaccharide, and compared binding of the homologs in a plate-binding assay (Fig. 1D). In this assay, the ECD was titrated to the immobilized mannan, resulting in an effective concentration of binding (EC₅₀) of 7.4 ± 0.2 and 28.5 ± 1.4 nM for the murine and human langerin, respectively. Besides this avidity difference, both proteins showed distinct cooperative binding behavior. Whereas the murine langerin association with the plate fits to a dose-response model with a Hill coefficient of 1.0 ± 0.1, suggesting no cooperativity, the human langerin experienced a higher degree of positive cooperativity with a Hill coefficient of 4.1 ± 0.3 (Table 1). To exclude effects resulting from varying lectin activity, we converted the assay to a competition experiment; α-mannose was titrated into a solution of langerin, keeping the protein at a fixed concentration, yielding inhibitory concentrations

![Figure 1. Comparing murine and human langerin extracellular domain binding to simple and complex mannan-type glycans](image)
Species-dependent Polysaccharide Specificity of Langerin

To gain more structural insight, potentially offering an explanation for the differences in mannan binding, we solved the structure of the murine langerin CRD by X-ray crystallography (Fig. 2 and Table 2, PDB entry 5K8Y). The tertiary structure of the domain was almost identical with the structure of the human CRD (PDB entry 3P5F (23)), which aligned with a low Cα-root mean square deviation of 0.4 Å for 111 Cα atoms. In the same affinity with mannose, these data strongly suggest that murine langerin has a higher avidity to the mannose polysaccharide than its human homolog.

To determine the binding specificity of langerin for mannan, we performed a mannan plate assay for human and murine langerin. The results are shown in Table 1.

TABLE 1
Fitting results for mannan plate assay for human and murine langerin
Shown are values ± S.D.

|              | EC50 (nM) | Hill slope | IC50 (mM) | Hill slope |
|--------------|-----------|------------|-----------|------------|
| Murine langerin | 7.3 ± 1.0 | 1.1 ± 0.1 | 86.0 ± 77.0 | 1.5 ± 2.6 |
| Human langerin   | 29 ± 5.0  | 3.8 ± 0.9 | 5.7 ± 1.8  | 5.4 ± 4.5 |

(IC50) of 85.2 ± 1.3 and 6.5 ± 0.1 mM with Hill coefficients of 1.6 ± 0.1 and 4.5 ± 0.5 for murine and human langerin, respectively. Consequently, because both proteins share the same affinity with mannose, these data strongly suggest that murine langerin has a higher avidity to the mannose polysaccharide than its human homolog.
protein backbone, the only deviations were observed in the short loop between the canonical β2 and β2′ strands. However, this short loop has been shown to adopt different conformations in the crystal structure of the trimeric human langerin (PDB entry 3KQG (29)) and is thus probably flexible. Whereas both the residues and the side chain orientations in the primary binding site are highly conserved, there are two lysine residues in the vicinity of the binding site in the human homolog that are altered in the murine form (Lys-299 → Arg-302 and Lys-313 → Asn-316) (Fig. 3A). Co-crystallization of the murine CRD with glucose revealed that although always coordinated by the 3- and 5-hydroxyl groups, the monosaccharide can bind in two orientations by flipping of the ring (Fig. 3B and Table 2, PDB entry 3KQG (29)). The expansion shows charged residues at the trimerization interface of the CRD. The bottom left corner shows a top view of the trimer. E, rotated surface charge representation of the human and the murine CRD with the trimerization interface highlighted by a green arrow. The murine langerin contains several positively charged residues at the interface, probably resulting in electrostatic repulsion.

| TABLE 2 |
| Crystallographic data for murine langerin CRD with and without glucose |

|                     | Native                   | Glucose-bound              |
|---------------------|--------------------------|----------------------------|
| Data collection     |                          |                            |
| PDB entry           | 5K8Y                     | 5M62                       |
| Space group         | P23                      | P23                        |
| Wavelength (Å)      | 0.91841                  | 0.97630                    |
| Unit cell a = b = c (Å) | 143.1                  | 144.4                      |
| Resolution (Å)      | 50.00–2.40               | 50.00–1.70                 |
| (2.54–2.40)         | (1.74–1.70)              |                            |
| Unique reflections  | 19,343 (3,065)           | 54,929 (4,019)             |
| Completeness (%)    | 100.0 (99.9)             | 100.0 (100.0)              |
| Rwork (%)           | 4.8 (2.1)                | 28.2 (3.3)                 |
| Rfree (%)           | 0.803 (1.528)            | 0.116 (1.783)              |
| CC1/2 (%)           | 95.8 (14.7)              | 100.0 (85.2)               |
| Redundancy (%)      | 20.1 (19.5)              | 40.4 (41.0)                |
| Wilson B-factor (Å) | 28.1                     | 29.1                       |

| Refinement          |                          |                            |
| Resolution (Å)      | 50.00–2.40               | 50.00–1.70                 |
| (2.46–2.40)         | (1.74–1.70)              |                            |
| Reflections (%)     | 18.638 (1.421)           | 52.330 (4.019)             |
| Rwork (%)           | 0.197 (0.254)            | 0.144 (0.354)              |
| Rfree (%)           | 0.194 (0.428)            | 0.164 (0.421)              |
| Non-hydrogen atoms  | 2,366                    | 2,722                      |
| Average B-factor (Å) | 52.1                    | 28.1                       |
| No. of protein chains | 2                      | 2                          |
| Protein atoms (No./B-factor (Å)) | 2,242/52.8 | 234/27.5                  |
| Ca2+ (No./B-factor (Å)) | 2/53.0             | 2/25.0                     |
| Glucose (No./B-factor) | 2/33.4              |                            |
| Ligand molecules (glycerol, PEG) | 4/77.0              | 10/36.5                    |
| Water molecules (No./B-factor (Å)) | 80/42.8            | 276/39.3                   |
| RMSD (%)            | 0.008                    | 0.009                      |
| Bond angles (degrees) | 1.237               | 1.398                      |
| Ramachandran outliers (%) | 0.0              | 0.0                        |
| Ramachandran favored (%) | 95.9              | 98.6                       |
| MolProbity score    | 1.45                     | 0.81                       |

a Values in parentheses refer to the highest resolution shell.

b Rwork = Σ||Fo - Fc||/Σ|Fo|, where Fo is the mean intensity of symmetry-equivalent reflections and n is the redundancy.

c Rfree is the same as Rwork but calculated on 5% of the data excluded from refinement.

d Root mean square deviation from target geometries.

Species-dependent Polysaccharide Specificity of Langerin

FIGURE 3. Murine and human langerin share the same primary binding site but differ in surface charge of a secondary site and at the trimerization interface. A, carbohydrate binding site of the human (left, PDB entry 3P5J) (23) and the murine (right) langerin. Whereas residues coordinating the Ca2+ ion are highly conserved, altered residues are located in proximity of the conserved amino acids Asn-297/Asn-300 that have been associated with a secondary binding site. B, binding of glucose and glycerol to murine langerin observed in the crystal structure. The binding mode of glucose is flipped in the two CRD molecules of the unit cell with one of the models shown transparent. The vicinal hydroxyl groups of glycerol are coordinated by Asn-300. Amino acids coordinating Ca2+ are drawn in yellow. C, surface charge representation of the human (top) and the murine (bottom) langerin CRD. Whereas the charge at the Ca2+ binding site (yellow arrow) is conserved, the secondary site (magenta arrow) is less positively charged in the murine homolog compared with the human variant. D, structural model of the murine langerin trimer (colored) generated by superposition to the human langerin trimer (gray, PDB entry 3KQG (29)). The expansion shows charged residues at the trimerization interface of the CRD. The bottom left corner shows a top view of the trimer. E, rotated surface charge representation of the human and the murine CRD with the trimerization interface highlighted by a green arrow. The murine langerin contains several positively charged residues at the interface, probably resulting in electrostatic repulsion.

The construct of the murine langerin CRD used for crystallization contained some residues of the N-terminal neck region that were resolved in the crystal structure adopting a short α-helix. Superposition of the CRD with the trimeric structure of the human ECD (PDB entry 3KQG (29)) resulted in a structural model that suggested that the relative position of the neck and the CRD is conserved in both proteins (Fig. 3D). In this structured model that suggested that the relative position of the neck and the CRD is conserved in both proteins (Fig. 3D). In this
model, however, the trimerization interface of the CRDs contained several positively charged residues that induce electrostatic repulsion (Fig. 3, D and E), suggesting a more opened conformation of the trimer in the murine homolog.

Human and Murine Langerin Differentially Recognize Bacterial Glycans—To investigate whether the differences observed in mannan binding translate to other complex polysaccharides, murine and human langerin were tested against a microbial glycan microarray comprising over 300 bacterial carbohydrates (Fig. 4 and supplemental Table S1) (31). Under saturating amounts of lectin, both proteins show very similar binding to mannan (Fig. 4 C). Remarkably, the recognition of bacterial glycans is highly different between the two homologs. The murine langerin recognized a broad set of polysaccharides with highly heterogeneous structures spread over a large panel of bacterial species (Fig. 4 A). Notably, the highest fluorescence intensities were observed for the binding to Streptococcus pneumoniae serotype 6B, which exceeded those seen for mannan by about 3-fold. Conversely, the human langerin showed only weak signals with some Shigella antigens as well as with structures derived from Escherichia coli O106, E. coli 180/C3, and Shigella boydii B10 were not recognized by human langerin on the array (23). However, an array with limited structures was tested in this previous study, which did not comprise mannan or any of the other binders we identified for human langerin ECD (supplemental Table S1); therefore, the differences probably originate from the different cut-off values that were chosen to define binders.

To validate these target structures from the array, we compiled a list of nine epidemiologically relevant and structurally diverse hits to investigate the differences of the interactions in more detail (Fig. 5 A). The purified lipopolysaccharide (LPS) and O-antigen polysaccharide (OPS) structures were tested in a plate-based, direct binding assay (Fig. 5 B) and in a competition format with mannan-coated plates (Fig. 5 C). Binding of murine langerin was confirmed for six of the structures. Human langerin bound to one of the selected Y. pestis surface glycans, Proteus mirabilis O5, and to E. coli O106 LPS and OPS structures as well as to the positive control laminarin, a branched /H9252 1,3-glucan polymer. Of the remaining three, two were the polysialic acids of E. coli K92 and Neisseria meningitides that were hits in the glycan array but not bound by either protein in the plate assays, a discrepancy that we attribute to a potential false-positive effect on the array, such as the high charge density of the sialic acid moiety. Lack of direct binding to P. mirabilis O5 (Fig. 4 B, top) in the plate assay was most likely due to insufficient immobilization of the glycan on the microtiter plate. However, the polysaccharide competed the binding to mannan...
and was thus validated as binder. Overall, the plate-based assays verified the glycan array results for six glycans.

**Human and Murine Langerin Differentially Bind to Heat-inactivated Bacteria**—Next, to show that our results translate to lectin recognition of polysaccharides presented on microbes, we used heat-killed *E. coli* O106 and performed direct binding assays in flow cytometry (Fig. 6A). Fluorescently labeled murine langerin bound the bacterial cells with high affinity, whereas almost no binding was seen for the human homolog, hence in line with our previous results from the glycan array and plate-based assays. Finally, known inhibitors and competitors of langerin-carbohydrate binding demonstrated the specificity of interactions of the murine ECD with the bacteria (Fig. 6B). Taken together, the murine langerin binding to bacteria provides an additional line of evidence that interspecies differences are profound and can be translated to direct microbial recognition.

**Langerin Amino Acid Sequence Is Highly Conserved among Mammalian Species with the Exception of the Extended Binding Site**—Knowing that the overall domain structure of human and murine langerin is highly conserved, we assessed the evolutionary conservation of 28 mammalian langerin homologs (Fig. 7A and supplemental Fig. S3). The residues of the secondary structure elements, with the exception of the β4 strand, are almost identical in all sequences. Surprisingly, low sequence conservation was found adjacent to the conserved primary binding site. In the human langerin, this region harbors the interaction site for the sulfate group (Lys-299 and Lys-313) when 6-sulfated mannose is present on the glycosylation of the antiphagocytic module.

**FIGURE 5. Plate-based assays validate glycan specificities of murine and human langerin.** A, structures of selected bacterial polysaccharide hits for validation of glycan array results. B, direct species-dependent polysaccharide specificity of langerin binding of murine (orange) and human (blue) langerin to immobilized polysaccharides on microtiter plates. C, competition assay with polysaccharides titrated to compete with binding of both proteins to mannan-coated microtiter plates. Polysaccharide concentrations are given in the legend. Data are represented as mean ± S.D. (error bars) (n = 3).

**FIGURE 6. Murine langerin binds preferentially to heat-inactivated E. coli O106.** A, histogram representation of fluorescently labeled murine (left) and human (right) langerin binding to heat-inactivated *E. coli* O106, as measured by flow cytometry. B, murine langerin shows a clear dose-dependent increase in binding, whereas human langerin binds only weakly. C, binding of murine langerin can be inhibited by the addition of known inhibitors EDTA, mannan, and laminarin. Data are represented as mean ± S.D. (error bars) (n = 3).
Species-dependent Polysaccharide Specificity of Langerin

![Figure 7](image)

Langerin sequence is highly conserved among mammalian species with exception of a probable secondary interaction interface. A, residue conservation scores in langerin derived from alignment of 28 mammalian homologs (supplemental Fig. S3) mapped as a color map on the human langerin structure shown as a schematic (left) and surface view (right). The surface view shows a highly variable region in the vicinity of the Ca$^{2+}$ site forming a potential secondary interaction site for complex glycans. Ca$^{2+}$ is shown as a green sphere. B, surface representation of trimeric human langerin color-coded as in A. The highly variable region faces toward the shallow groove in the center of the trimer, potentially forming a secondary interaction site for complex glycans. C, surface representation of ASGPR (PDB entry 1DV8 (44)) color-coded according to residue conservation score derived from sequence alignment of 27 mammalian homologs (supplemental Fig. S4). The surface shows an overall high degree of conservation.

galactose is bound (23) and expands toward the center of the trimer. Notably, these highly variable residues are part of a surface facing the center and providing an extended binding site for polysaccharides (Fig. 7B). Moreover, Ala-289 of the human langerin, which was suggested to provide additional interactions with high mannose glycans, is not conserved, possibly altering specificity for this class of glycans (23). Intriguingly, the residue conservation of the asialoglycoprotein receptor, a heptic CLR that recognizes self-structures, exhibited a high overall conservation (Fig. 7C). Taken together, our evolutionary conservation analysis showed that the region adjacent to the binding site, facing the center of the trimeric ECD, is highly variable, probably influencing the receptor’s specificity for complex glycans.

Discussion

Langerin is a myeloid C-type lectin receptor involved in innate host defense. Whereas other CLR s, such as DC-SIGN, have no true homologs among different species (13), langerin is highly conserved in the mammalian lineage (Fig. 7). To assess whether carbohydrate recognition is also conserved, we compared human and murine langerin side by side. Most strikingly, whereas simple ligands were bound with almost identical affinities (Fig. 1C) and binding modes (Fig. 2), significant differences between the two homologous receptors were observed for complex polysaccharides. Not only did the avidities between the two homologs differ (Fig. 1D), but also the specificities (Fig. 4), suggesting that the interaction with large polysaccharides is not determined by the presence of simple molecular patterns per se, but by their arrangement in the polysaccharide. These differences became evident in our biochemical assays revealing significantly enhanced interdomain cooperativity within the trimeric ECDs (Fig. 1D). Our thermodynamic data strongly suggest that such cooperativity is not triggered by simple glycans but is associated with the extended structure of a polysaccharide. Consequently, these results question the common prediction of polysaccharide specificities based on binding data from simple glycans.

This differential behavior became even more apparent when the two proteins were used to interrogate a microbial glycan array. Whereas the murine homolog binds a large panel of structures giving rise to high signal intensities, the human homolog only recognizes a small set of structurally related glycans (Fig. 4). Notably, several of the repeating units of the polysaccharides recognized by murine langerin have no available vicinal hydroxyl groups that could form the canonic coordination observed in human langerin (23). Hence, the molecular details of this differential behavior cannot be understood by analysis of the monovalent interaction and must be embedded in the overall architecture of the receptor.

In contrast to this dramatic effect, comparison of the CRD structures of human and murine langerin revealed that the tertiary structures and the primary binding sites are conserved, which is paralleled by the conserved binding epitope for simple carbohydrate ligands (Fig. 2). However, residues adjacent to the binding sites differ (Fig. 3). This observation highlights a second determinant of the altered polysaccharide specificity apart from interdomain cooperativity; subtle changes in or close to the carbohydrate recognition site can yield diverging ligand specificities, an effect described previously for other CLR s, such as DC-SIGN and DC-SIGNR (32). For langerin, Lys-299 and Lys-313 have been attributed as binding determinants for 6-sulfated galactoses (23, 24). Although these changes do not alter the recognition for monovalent mannose, fucose, and 6-sulfated galactose (Fig. 1), the observed subtle changes in binding mode of N-acetyl glucosamine and glucose (Figs. 2 and 3) could potentially have profound effects in the context of larger polysaccharides. In addition, weak additional interactions originating from secondary sites could provide a rationale for the differential recognition of complex glycan structures (33), a notion that is further supported by the promiscuity of the primary binding site, as has been observed for glucose herein (Fig. 3) and has been reported previously for the human langerin (23). Our evolutionary analysis revealed that these residues are highly variable among all mammalian langerin homologs and form part of an extended stretch of residues facing the shallow cavity of the trimer (Fig. 7). Notably, this region has also been suggested before as an extended binding site for heparin in human langerin (30). Overall, our data provide insight into how
changes in secondary binding sites as well as in the interdomain communication combine into significantly altered polysaccharide specificities that cannot be inferred from the analysis of simple glycans.

These altered polysaccharide specificities potentially relate to the high evolutionary pressure under which each of these two PRRs coevolved with the microbial threat. A recent account identified both positive and balancing selection on inter- and intraspecies levels for langerin (34). In particular, CLRs involved in pathogen recognition exhibit higher interspecies diversity with respect to their glycan specificity than CLRs specialized in the recognition of self-antigens (8). Another indicator for an adaptive evolutionary behavior of langerin comparing mice and humans is the divergence in the cellular expression patterns. Human langerin+ APCs are predominantly encountered in the skin and mucosa, where the evolutionary pressure might have driven specificity toward opportunistic fungi, mycobacteria like *Mycobacterium leprae*, and viruses invading via the mucosa similar to HIV. Conversely, murine langerin is also found in APC subsets of the blood, lung, and gut, where it encounters different pathogens, such as *S. pneumoniae*, *Providencia*, and *Proteus* species, respectively. The high interspecies variability of the proposed secondary interaction site also suggests how these receptors rapidly adapt their specificities for complex pathogen glycans while maintaining the same structure and even preserving the primary binding site (Fig. 7). This is further evidenced by the finding that ASGPR, a CLR not involved in pathogen recognition, harbors no regions with elevated interspecies variability (Fig. 7C).

**Experimental Procedures**

All chemicals applied in this study were purchased from Sigma-Aldrich if not otherwise indicated.

**Cloning, Protein Expression, and Purification**—Codon-optimized truncated murine langerin ECD (residues 150–331) was cloned into a pUC19-derived expression vector containing a C-terminal TEV cleavage site and a Strep-tag II. Codon-optimized murine langerin ECD (residues 150–331) was expressed in *E. coli* BL21(DE3), refolded, and purified as described before (22). Murine CRD was expressed in *E. coli* Lemo21(DE3) (New England Biolabs); solubilized in 6 M guanidinium HCl; refolded by rapid dilution into 0.8 M L-arginine in TBS, pH 7.5, containing 2.5 mM reduced and 0.5 mM oxidized glutathione; and purified via a StrepTactin column followed by dialysis against 25 mM MES, 40 mM NaCl, pH 6. Protein concentration was determined by absorbance at 280 nm using the calculated molar extinction coefficients of 56,170 M⁻¹ cm⁻¹ for the human langerin ECD and CRD and 56,170 M⁻¹ cm⁻¹ and 54,680 M⁻¹ cm⁻¹ for the murine ECD and CRD, respectively.

**Chemical Cross-linking**—Chemical cross-linking of human and murine langerin ECD was performed by the addition of 0.1–5 mM amine-reactive BS3 (1 mg/liter, freshly dissolved in DMSO) to 100 μl of protein solution (5 mg/ml) in HEPES-buffered saline, pH 7, and subsequent incubation for 1 h on ice. The reaction was quenched by the addition of 100 mM Tris, pH 8, and samples were taken for SDS-PAGE analysis.

**ITC Measurements**—Isothermal titration calorimetry measurements were performed on a MicroCal ITC 200 (Calimetry) using 100–150 μM protein (monomer concentration) while adding D-mannose (50 mM) as titrant in 18 steps with a 2-μl volume, DP 11, stirring at 750 rpm, at 298 K. The first injection of 0.5 μl was used for equilibration of the system and not included in analysis. Measurements were performed in three replicates. Data were fitted to a one-site of sites model implemented in the supplier’s software with the number of binding sites n set to 1.

**Plate-based Assays**—For mannan plate assays, NUNC Maxisorb 96-well plates (Thermo Fisher Scientific) were coated with 10 μg/ml mannan in 50 mM carbonate buffer, pH 9.6, blocked with 2% BSA in TBS-T for 1 h. Wells were washed three times with TBS-T + 5 mM CaCl₂, incubated with protein in the same buffer for at least 90 min at room temperature, washed, and incubated with StreptTactin-HRP conjugate (IBA) in 2% BSA in TBS-T + 5 mM CaCl₂ for 1 h. Plates were developed by adding TMB solution (tebu-bio). The reaction was quenched by the addition of 0.18 M H₂SO₄. Absorbance was measured at 450 nm on a SpectraMax M5 plate reader (Molecular Devices). For EC₅₀ determination, protein concentrations ranged from 2 to 100 nM. For IC₅₀ determination, mannose was titrated to the protein (50 nM) from 200 to 0.4 mM before incubation on the plate. All measurements were performed in triplicates. Data were fitted to a dose-response model to determine EC₅₀ or IC₅₀, respectively.

**NMR Experiments**—All NMR experiments were performed on an Agilent vnmrs 600-MHz spectrometer equipped with a OneNMR double-resonance gradient probe. All samples were prepared with 50 μM protein concentration, 1 mM 1, and 5 mM 2 and 3, respectively, in 25 mM Tris-α (Eurisotop), 150 mM NaCl, pH 7.0, in D₂O. Spectra were recorded with an STD pulse sequence containing a WATERVER solvent suppression. Selective Gauss-shaped saturation pulses with a −0.5 ppm offset for on-resonance and −80 ppm for off-resonance spectra were applied with saturation transfer times of 0.5, 1, 2, and 4 s with a total of 1,024 scans (2,048 at 0.5 s saturation transfer time). 6-Sulfated galactose was measured under the conditions described above with 4 s of saturation transfer time and 2,048 scans. Difference spectra were processed and analyzed with MestreNova (MestreLab), and the epitope was mapped as described before (35).

**Crystallization and X-ray Structure Determination**—For crystallization experiments, murine langerin CRD was concentrated to 6.0 mg/ml, and crystals were obtained by the sitting-drop vapor diffusion method at 18 °C with a reservoir solution composed of 30% (v/v) polyethylene glycol 600, 5% (w/v) polyethylene glycol 1000, and 10% (v/v) glycerol. Crystals with glucose bound were obtained by co-crystallization with 2.5% D(-)-glucose monohydrate. No further cryo-protection was necessary to flash-freeze the crystals. Synchrotron diffraction data were collected at beamline 14.1 of the BESSY II storage ring (Berlin, Germany) or beamline P14 of PETRA III (Deutsches Elektronen Synchrotron, Hamburg, Germany). X-ray data collection was performed at 100 K. Diffraction data were indexed and processed with XDS (36). Data collection and refinement statistics are given in Table 1. Dif-
fraction data were initially indexed in 1432. Cumulative intensity distribution analysis as well as calculation of the moment of the observed intensity/amplitude distribution performed with PHENIX.XTRIAGE (37) and POINTLESS (38) indicated an unusual intensity distribution, caused by twinning, with I23 as the true space group. The structure was solved by molecular replacement with the program PHASER (39) using the structure coordinates of a human langerin monomer with the trim- erization domain truncated (PDB entry 3KQG (29)) and refined with REFMAC version 5.8.0135. Model building and water picking were performed with COOT (40). Model quality was evaluated with Molprobity (41).

**Protein Labeling**—Both langerin ECD constructs were fluorescently labeled with fluorescein isothiocyanate (Pierce) by slowly adding 100 µl of the dye (1 mg/ml in DMSO) to 2 ml of protein (2 mg/ml) in HEPES-buffered saline, pH 7.2, containing 20 mM D-mannose and 5 mM CaCl₂ and stirred at room temperature for 90 min in the dark. The reaction was quenched by adding 50 mM ethanolamine (1 M, pH 8.5). Labeled protein was purified over a mannann column, and active protein was used for glycan microarray analysis.

**Glycan Array Analysis**—Production of the microbial glycan microarray has been described previously (31). Soluble, FITC-labeled, trimeric human and murine langerin ECD were diluted to 50 µg/ml in TSM buffer (20 mM Tris-HCl, pH 7.4, 150 mM sodium chloride, 2 mM calcium chloride, 2 mM magnesium chloride, 0.05% Tween 20, 1% BSA) and applied directly to the array surface. Following a 1-h incubation, slides were washed three times by dipping into TSM buffer, then TSM without Tween and BSA, and, finally, deionized water. Washed slides were then dried by centrifugation and immediately scanned on an Innoscan1100AL (Innopsys) using the 488-nm laser at low power (5 milliwatts) and 50% gain. Image intensities were collected using Mapix software (Innopsys). Fluorescence intensity values were calculated using the mean signal minus mean background values of six replicate spots with the highest and lowest signals removed.

**Bacterial Polysaccharide Plate Assays**—Bacterial polysaccharides were isolated from bacteria as described before (31). For direct binding, polysaccharides were immobilized on plates as described above with concentrations of 100, 10, and 1 µg/ml, and protein binding was probed at 50 nM. A competition binding assay was performed analogously to mannose competition on mannan-coated plates with glycan concentrations of 100, 10, and 1 µg/ml.

**Flow Cytometry**—Fluorescently labeled murine and human langerin were incubated in increasing concentrations with about 64 million bacteria in 200 µl of Hanks’ balanced salt solution with 2 mM CaCl₂, and measured in triplicates using an Attune Nxt flow cytometer (Thermo Fisher Scientific) with 12.5 µl min⁻¹ flow rate and 10,000 events. Gates were set in FSC-H/SSC-H with double discrimination to ensure analysis of single bacteria.

**Residue Conservation Analysis**—All mammalian langerin and asialoglycoprotein receptor sequences deposited in UniProtKB were chosen for multiple sequence alignment using Clustal o (42) and subsequent scoring of residue conservation (43).

**Data Processing and Presentation**—All structural representations were prepared in PyMOL version 1.7.2.1 (Schrödinger). For the calculation of the electrostatic surface potentials, structures were prepared in MOE (version 2015.10). Ca²⁺ counterions were removed, and hydrogens were added at 300 K and pH 7. Charges were calculated using the Amber99 force field. All graphs were prepared, and all data fitting was performed in OriginPro version 2015 (OriginLab).

**Author Contributions**—J. H. and C. R. conceived the study and wrote the paper. J. H. designed, performed, and analyzed experiments shown in Fig. 1. J. A. designed and produced recombinant murine langerin and performed the crystallization screen; B. L. collected X-ray data and solved the structure shown in Fig. 2. J. H. provided fluorescently labeled proteins, and R. M. designed, performed, and analyzed the glycan array in Fig. 3. J. H. designed and H. S. performed and analyzed the experiments shown in Fig. 4. J. S. designed, performed, and analyzed the experiments shown in Fig. 5. J. H. and C. R. conducted the analysis shown in Fig. 6. J. H. prepared all figures. Y. K. provided bacterial polysaccharides, and W. R. heat-killed bacteria. J. C. P. and M. C. W. revised the manuscript critically. All authors reviewed the results and approved the final version of the manuscript.

**Acknowledgments**—We thank Daniel Varon Silva for generously providing α-methyl-(α-1,2)-mannobiose and Willie Vann for granting access to polysaccharide acid structures. J. H. thanks Daniel Tesolin for help with optimizing protein expression and labeling. We acknowledge Claudia Alings for excellent technical support. C. R. and J. H. thank Peter Seeberger for providing the research environment and fruitful discussions. We accessed beamlines of the BESSY II (Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung II) storage ring (Berlin, Germany) via the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie. We acknowledge beamtime and support at the EMBL beamline P14 of PETRA III (Deutsches Elektronen Synchrotron, Hamburg, Germany).

**References**

1. Gagneux, P., and Varki, A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology 9, 747–755
2. Varki, A. (2006) Nothing in glycobiology makes sense, except in the light of evolution. Cell 126, 841–845
3. Lebeer, S., Vanderleyden, J., and De Keersmaecker, S. C. (2010) Host interactions of probiotic bacterial surface molecules: comparison with commensals and pathogens. Nat. Rev. Microbiol. 8, 171–184
4. Akira, S., Uematsu, S., and Takeuchi, O. (2006) Pathogen recognition and innate immunity. Cell 124, 783–801
5. Drickamer, K. (1995) Increasing diversity of animal lectin structures. Curr. Opin. Struct. Biol. 5, 612–616
6. van Kooyk, Y., and Rabinovich, G. A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses. Nat. Immunol. 9, 593–601
7. Aretz, J., Wamhoff, E.-C., Hanske, J., Heymann, D., and Rademacher, C. (2014) Computational and experimental prediction of human C-type lectin receptor druggability. Front. Immunol. 5, 323
8. Drickamer, K., and Taylor, M. E. (2015) Recent insights into structures and functions of C-type lectins in the immune system. Curr. Opin. Struct. Biol. 34, 26–34
9. Zelensky, A. N., and Gready, J. E. (2005) The C-type lectin-like domain superfamily. FEBS J. 272, 6179–6217
