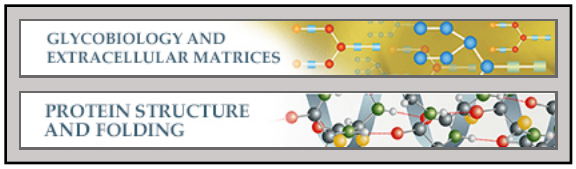
Glycobiology and Extracellular Matrices:

Trimeric Structure of Langerin



Hadar Feinberg, Alex S. Powlesland, Maureen E. Taylor and William I. Weis

J. Biol. Chem. 2010, 285:13285-13293.
doi: 10.1074/jbc.M109.086058 originally published online February 24, 2010

⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 17, pp. 13285–13293, April 23, 2010

Trimeric Structure of Langerin*[S]

Received for publication, November 17, 2009, and in revised form, January 25, 2010 Published, JBC Papers in Press, February 24, 2010, DOI 10.1074/jbc.M109.086058

Hadar Feinberg[‡], Alex S. Powlesland[§], Maureen E. Taylor[§], and William I. Weis[‡1]

From the [‡]Departments of Structural Biology and Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, California 94306 and the [§]Division of Molecular Biosciences, Department of Life Sciences, Imperial College, London SW7 2AZ, United Kingdom

Langerin, an endocytic receptor of Langerhans cells, binds pathogens such as human immunodeficiency virus by recognition of surface glycoconjugates and mediates their internalization into Birbeck granules. Langerin has an extracellular region consisting of a C-type carbohydrate-recognition domain (CRD) and a neck region that stabilizes formation of trimers. As in many other C-type lectins, oligomerization is required for high affinity binding to glycan ligands and is also likely to be important for determining specificity. To facilitate structural analysis of the human langerin trimer, a truncated form of the extracellular region, consisting of part of the neck and the CRD, has been characterized. Like the full-length protein, truncated langerin exists as a stable trimer in solution. Glycan array screening with the trimeric fragment shows that high mannose oligosaccharides are the best ligands for langerin. Structural analysis of the trimeric fragment of langerin confirms that the neck region forms a coiled-coil of α-helices. Multiple interactions between the neck region and the CRDs make the trimer a rigid unit with the three CRDs in fixed positions and the primary sugar-binding sites separated by a distance of 42 Å. The fixed orientation of the sugar-binding sites in the trimer is likely to place constraints on the ligands that can be bound by langerin.

Langerin (CD207), a C-type lectin of Langerhans cells, binds pathogens, including *Candida albicans* and human immunodeficiency virus (HIV),[2] through recognition of surface glycoconjugates containing mannose or related sugars (1–4). Langerin is an endocytic receptor associated with formation of Birbeck granules, subdomains of the endosomal compartment specific to Langerhans cells (3–6). Glycoconjugate ligands internalized via langerin are degraded, and it is likely that langerin plays a role in antigen processing and presentation, the main function of Langerhans cells. Langerin can mediate uptake and processing of antigens for presentation by both major histocompatibility class I and class II molecules and has also been implicated in the processing of mycobacterial nonpeptide antigens for presentation by CD1a (7, 8). Langerin can prevent transmission of HIV from Langerhans cells to T cells by mediating internalization and degradation of the virus (9, 10).

Langerin is a type II transmembrane protein with an extracellular region consisting of a neck and a C-terminal C-type carbohydrate-recognition domain (CRD) (3, 4). Like many other C-type lectins, langerin exists as an oligomer, forming trimers stabilized by a coiled-coil of α-helices in the neck region (11). Trimer formation is essential for binding to oligosaccharide ligands because, as is typical for C-type CRDs, the CRD of langerin has only low affinity for monosaccharides (11, 12). Oligomerization of C-type lectins is also important for determining selectivity for particular oligosaccharide structures. For example, in serum mannose-binding protein, three CRDs in the trimeric unit are held in a fixed position via interactions between the CRDs and an α-helical neck region so that the binding sites are arranged to interact with arrays of sugars in polysaccharides of bacterial cell walls, but not with mammalian high mannose-type oligosaccharides (13). In contrast, CRDs in the tetramer of the dendritic cell receptor DC-SIGN and in the related receptor DC-SIGNR have a more flexible arrangement, which allows movement of the CRDs in the tetramer relative to each other to facilitate engagement of multiple ligands with variable spacing such as the high mannose-type oligosaccharides of gp120 on the surface of HIV (14, 15).

Crystal structures of the CRD of human langerin in complex with mannose or maltose show that it binds monosaccharides by ligation to a bound Ca^{2+} at a site that is conserved in all C-type CRDs (16). Interestingly, the co-crystals also show the presence of a second sugar-binding site that has not been seen in other C-type lectins. Both monosaccharide residues of maltose, or the monosaccharide mannose, are bound in this second site largely via polar interactions with backbone residues in a cleft formed between two of the large loop regions in the top half of the domain (16). This cleft is wider in langerin than in other C-type CRDs, most likely due to the absence of auxiliary Ca^{2+} sites present in many other CRDs, including those of

* This work was supported by Grant 075565 from the Wellcome Trust (to M. E. T.). This work was also supported in part by National Institutes of Health Grant GM62116 (to the Consortium for Functional Glycomics). Analytical ultracentrifugation was performed in a facility supported by the Wellcome Trust and the Biotechnology and Biological Sciences Research Council. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a national user facility operated by Stanford University on behalf of the U.S. Dept. of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Dept. of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

⌘ *Author's Choice*—Final version full access.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table 1.

The atomic coordinates and structure factors (code 3KQG) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed: Dept. of Structural Biology, 299 Campus Dr. West, Stanford, CA 94305. Tel.: 1-650-725-4623; Fax: 1-650-723-8464; E-mail: bill.weis@stanford.edu.

[2] The abbreviations used are: HIV, human immunodeficiency virus; CRD, carbohydrate-recognition domain; DC, dendritic cell; DC-SIGN, dendritic cell-specific intercellular adhesion molecule 1-grabbing nonintegrin; DC-SIGNR, DC-SIGN-related receptor; PEG, polyethylene glycol; PDB, Protein Data Bank.

mannose-binding protein and DC-SIGN (17, 18), and is also more flexible (see below). Modeling studies suggest that high mannose oligosaccharides such as Man_9 could bind to langerin through ligation of the terminal mannose residue of one branch to Ca^{2+} at the primary sugar-binding site and interaction of two mannose residues of another branch at the secondary site (16). Computational docking of linear three mannose fragments of Man_9, αMan1,2αMan1,2αManOMe, and αMan1,2αMan1,3αManOMe correlates well with the two binding sites found in the structure of the CRD complexed with mannose and maltose and suggests that the mannose bound to the Ca^{2+} in the primary binding site is the central mannose with 3-OH and 4-OH coordinating the Ca^{2+} (19).

An elongated model for the extracellular domain of langerin, constructed using the trimeric structure of mannose-binding protein and the α-helical bundle of the influenza virus hemagglutinin trimer, was found to have good correlation with hydrodynamic measurements (19). This model was used to interpret the organization of langerin in electron micrographs of Birbeck granules. However, the model is limited in its ability to predict details that can influence the selectivity of langerin toward different carbohydrates, such as possible rigidity or flexibility of the hinge between the neck and the CRD, or the precise distance between CRDs in the oligomer. Because the affinity of langerin is enhanced by oligomerization, an experimental structure of the trimer is important for understanding its selectivity for particular oligosaccharides or certain patterns of sugar presentation and is more broadly useful for understanding differences in the binding selectivity among the various C-type lectins that have similar oligosaccharide specificities in the monomeric form.

In this paper, the crystal structure of a trimeric fragment of langerin consisting of the CRDs and part of the neck domain is reported. The structure shows that multiple interactions between the neck region and the CRDs make the trimer a rigid unit with the CRDs in fixed positions and the primary sugar-binding sites separated by a distance of 42 Å. As in mannose-binding protein, the fixed orientation of the sugar-binding sites in the trimer is likely to place constraints on the ligands that can be bound by langerin.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—DNA coding for a fragment of the extracellular domain of human langerin from residue 148 in the neck region to the C terminus (designated "truncated langerin") was cloned into the pT5T expression vector using appropriate restriction sites and a synthetic oligonucleotide designed to bridge the end of the coding sequence and the BamHI site of the vector. The resulting plasmid was transformed into *Escherichia coli* strain BL21/DE3. Growth and induction of bacteria, and extraction and purification of protein on mannose-Sepharose were as described previously for the full-length extracellular domain of langerin (11).

Analytical Ultracentrifugation—Sedimentation equilibrium analysis was carried out in a Beckman Optima XL-A analytical ultracentrifuge using an An60Ti rotor at 20 °C. Truncated langerin at 0.2 mg/ml in 25 mM Tris-HCl, 0.15 M NaCl, 2.5 mM $CaCl_2$ was analyzed at 10,000 rpm and 20,000 rpm. Data were analyzed using UltraSpin software developed by Dmitry Veprintsev.

Glycan Array Analysis—Truncated langerin was labeled with fluorescein by dialysis into 100 mM bicine (pH 9.0), 0.15 M NaCl, 5 mM $CaCl_2$ and incubation with fluorescein isothiocyanate (50 μg/mg protein) overnight. Labeled protein was isolated by affinity chromatography on a 1-ml column of mannose-Sepharose. Fluorescein-labeled truncated langerin was used to probe version 3.2 of the glycan array following the standard procedure of Core H of the Consortium for Functional Glycomics.

Crystallization and Data Collection—Crystals of truncated langerin were obtained by hanging-drop vapor diffusion, with 0.5 μl of protein plus 1 μl of reservoir in a drop. All crystals were grown at 22 °C. Form 1 crystals were grown from a protein solution comprising 10 mg/ml langerin, 9 mM $CaCl_2$, 9 mM Tris (pH 7.8), 22.5 mM NaCl, and 100 mM D-mannose. The reservoir solution contained 8% polyethylene glycol (PEG) 8000, 0.1 M imidazole (pH 8.5), and 0.2 M $Ca(CH_3COO)_2$. The crystals were dehydrated by serial transfer to fresh reservoir solution containing increasing percentages of PEG 8000: 16% for 24 h; 24% for 24 h, then to 32% for 48 h. The crystals were then frozen in liquid nitrogen for data collection.

Form 2 crystals were grown from a protein solution comprising 10 mg/ml langerin, 10 mM Tris (pH 8.0), and 25 mM NaCl. The reservoir solution contained 15% PEG 3350 and 0.2 M NH_4Cl. Crystals were transferred to a fresh reservoir solution containing 35% PEG 3350, and were then frozen in liquid nitrogen for data collection.

Diffraction data were measured at 100 K on beamline 11-1 of the Stanford Synchrotron Radiation Lightsource. Data were processed with MOSFLM and SCALA (20) and are summarized in Table 1.

Structure Determination—The form 1 crystals were solved to a resolution of 4.0 Å by molecular replacement phasing, using coordinates of the langerin CRD (residues 197–326 + Ca^{2+} and mannose; Protein Data Bank (PDB) code 3BC7) as a search model. Self-rotation functions indicated the presence of 4-fold and 3-fold noncrystallographic rotational symmetry. Molecular replacement was performed using the program Phaser (21), searching for one monomer at a time while fixing those previously placed. After eight CRDs were located, it was noted that several pairs of CRDs had a relative rotation of 120°, and their N termini were sufficiently close to one another to likely be part of a trimer. This axis corresponded to the noncrystallographic 3-fold axis observed in self-rotation functions. Assuming a closed 3-fold symmetry for each trimer, the third copy was generated to make a trimer of CRDs. This trimer model was used in a second round of Phaser, this time searching for eight trimers. To save time, this run was manually interrupted after four trimers were placed, and the other four trimers were generated using NCS operations to give eight trimers (24 monomers), arranged in the P_1 unit cell with 432 noncrystallographic symmetry. Rigid body refinement (50–4 Å) gave $r = 44.2\%$, $R_{free} = 44.6\%$ using the trimers as eight rigid groups, and further refinement using the 24 individual CRDs as rigid bodies gave $r = 39.1\%$, $R_{free} = 39.3\%$. No further refinement was done for these data because a higher resolution data set from the form 2 crystals was obtained at that time.

TABLE 1
Crystallographic data statistics

Data collecton SSRL 11-1	Data set 1	Data set 2
Space group	P1	$P2_1$
Unit cell lengths (Å)	a = 119.8	a = 80.8
	b = 120.0	b = 77.5
	c = 120.7	c = 85.9
Angles (°)	α = 90.2	β = 94.4
	β = 106.5	
	γ = 95.2	
Resolution Å (last shell)	119.5-4.0 (4.22-4.0)	85.7-2.3 (2.42-2.3)
R_{sym} (last shell)	4.5 (16.4)	7.9 (34.0)
Mean (I)/S.D.(I)	12.6 (4.6)	11.2 (4.2)
% Completeness (last shell)	94.6 (95.8)	99.6 (99.1)
Average multiplicity (last shell)	2.0 (2.0)	3.7 (3.7)
Refinement	Data set 2	
Resolution for refinement (last shell)	48.0-2.30 (2.35-2.30)	
Residues included in final model	A 159:327 B 167:259 262:325 C 167:325 D 166:259 263:326 E 167:326 F 165:326 6 Ca^{2+} 407 H_2O	
R_{free} [a]	23.8 (29.5)	
R [a]	18.2 (23.0)	
Average B factor	41.9	
Bond length r.m.s.d. [b]	0.005	
Angle r.m.s.d.	0.82	
Ramachandran plot (% in preferred/allowed/ outliers regions) [c]	95.8/4.1/0.1	

[a] $R_{sym} = \Sigma_h\Sigma_i(|I_i(h)| - |<I(h)>|)/\Sigma_h\Sigma_i I_i(h)$ where $I_i(h)$ = observed intensity, and $<I(h)>$ = mean intensity obtained from multiple measurements. R and R_{free} = $\Sigma\ ||F_o|-|F_c||/\Sigma|F_o|$, where $|F_o|$ = observed structure factor amplitude and $|F_c|$ = calculated structure factor amplitude for the working and test sets, respectively.
[b] r.m.s.d, root mean square deviation.
[c] As defined in Coot.

The higher resolution form 2 crystal structure was solved by molecular replacement, using the program COMO (22). The search model comprised a single trimer of CRDs from the form 1 crystal solution, using data in the resolution range 12–3.5 Å. Two trimers were found in the asymmetric unit. The final correlation coefficient was 39%, and the R value was 41%. Model building and refinement were performed with Coot (23) and Phenix (24). In the initial stages of refinement, noncrystallographic symmetry constraints were imposed but were released in later stages as the electron density maps indicated that the neck of the CRDs does not adopt the same symmetry relationships as the CRDs (see below). Refinement included individual positional and temperature factor refinement with translation-libration-screw modeling. For translation-libration-screw refinement, each chain was split into two domains, the CRD and the oligomerization domain, to give a total of 12 groups. Although no Ca^{2+} was present in the crystallization buffer, it was present during protein purification. Strong electron density (>6σ in $F_o - F_c$ maps) was visible at the known Ca^{2+} site and was therefore modeled as such. All analysis presented here is based on the refined form 2 crystals.

RESULTS AND DISCUSSION

Glycan Binding by a Trimeric Fragment of Langerin Extracellular Domain—To facilitate structural analysis of the trimeric extracellular domain of langerin, a truncated form of human langerin was produced. Truncated langerin corresponds to a proteolytic fragment seen in preparations of the full-length extracellular domain and consists of the last seven heptad repeat sequences (including a 4-residue deletion in the pattern; see below) of the neck region together with the CRD

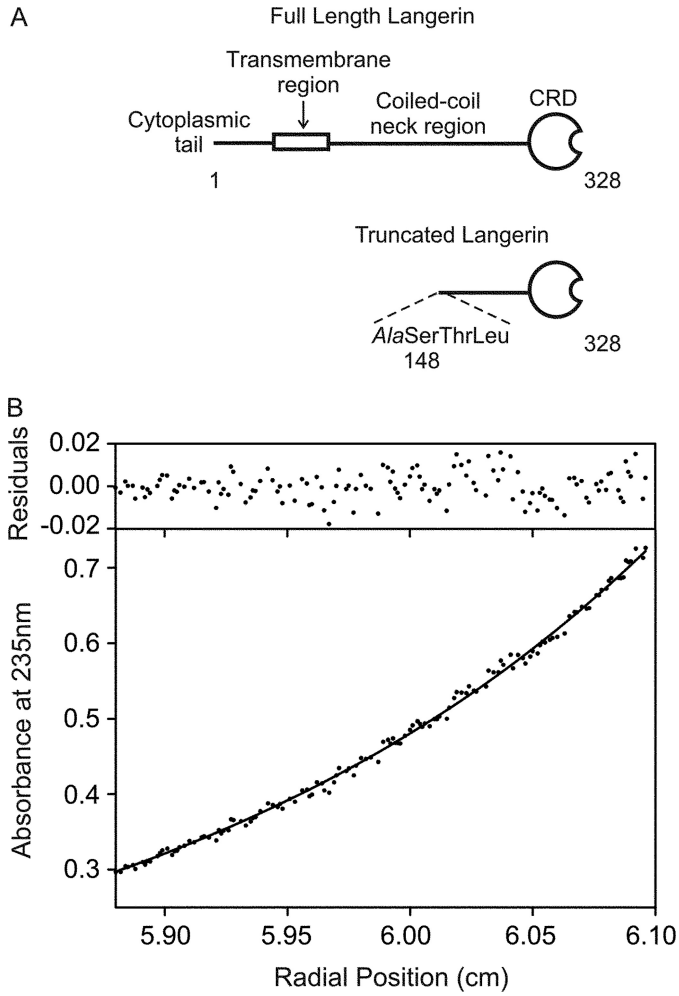


FIGURE 1. **Characterization of a trimeric fragment of human langerin.** *A*, diagram showing domain organization of full-length langerin and the truncated langerin construct. The starting and final amino acid residues are indicated, with numbering based on the residue numbers of full-length langerin. *B*, sedimentation equilibrium analysis of truncated langerin. A scan of a sample after 24 h at 10,000 rpm is shown.

(Fig. 1*A*). Truncated langerin has better solubility properties than the full-length extracellular domain. Sedimentation equilibrium analytical ultracentrifugation was used to confirm that truncated langerin forms a stable trimer, like the full-length extracellular domain, with a measured M_r of 62,530 compared with the value of 61,833 predicted for a trimer (Fig. 1*B*).

The binding specificity of truncated langerin was investigated by screening the glycan array of the Consortium for Functional Glycomics. Oligosaccharide-binding specificity of langerin has been studied previously by screening a smaller array of oligosaccharides and glycoproteins with an IgG-Fc chimera of the extracellular domain of mouse langerin, but analysis of the native trimeric structure has not been reported (25). At a relatively high concentration of fluorescein-labeled truncated langerin, binding to four types of oligosaccharide ligands is observed: high mannose oligosaccharides, structures with terminal GlcNAc residues, structures with terminal fucose residues, and structures with terminal 6-sulfated galactose residues (Fig. 2). Ligands giving the highest signals are Man_9 (oligosaccharide number 311), Galα1–3(Fucα1–2)Gal (blood group B

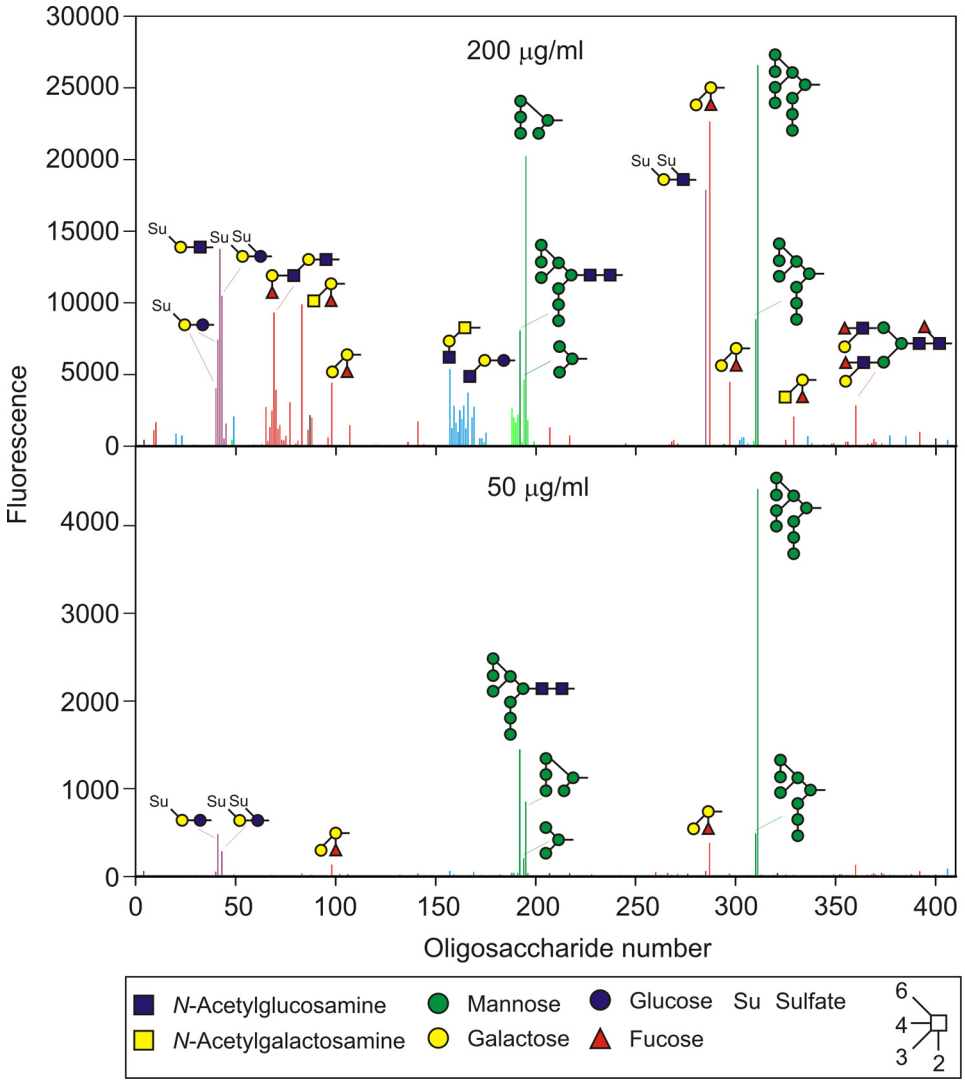


FIGURE 2. **Binding of fluorescently labeled truncated langerin to a glycan array.** The ligands giving the highest signals are identified. *Green bars* indicate high mannose structures. *Red bars* indicate structures with terminal fucose residues. *Blue bars* indicate structures with terminal GlcNAc. *Purple bars* denote structures with terminal 6-sulfated galactose residues. The data shown represent the mean of four replicates. Numerical values with standard deviations as well as a full list of glycans on the array are available in supplemental Table 1.

antigen; oligosaccharide number 287), Man_5 (oligosaccharide number 195), and 6-O-SO_3Galβ1–4(6-O-SO_3)GlcNAc (oligosaccharide number 285).

Binding to oligosaccharides with terminal mannose, fucose, or GlcNAc residues is consistent with the specificity of the CRD of langerin for these monosaccharides (11). There is, however, some selectivity for particular structures bearing these residues. In most of the fucose-terminated glycans on the array that are bound by langerin, fucose is linked α1,2 to galactose, although not all structures with α1,2-linked fucose are bound. Also, unlike DC-SIGN and the scavenger receptor C-type lectin, langerin does not bind to glycans containing the Lewis^X structure: Galβ1–4(Fucα1–3)GlcNAc (26, 27). All of the GlcNAc-terminated structures bound by langerin contain GlcNAc linked to galactose or GalNAc, and no binding is seen to structures terminating in GlcNAcβ1,2Man. Langerin thus differs from the liver endothelial cell C-type lectin LSECtin, which binds with high specificity to structures with terminal GlcNAcβ1,2Man (28). Lack of binding of human langerin to structures containing Lewis^X is consistent with studies showing that transfected cells expressing langerin did not bind beads coated with Lewis^X and that Lewis^X did not inhibit binding of HIV gp120 to these cells (10).

Screening the glycan array at a 4-fold lower concentration of labeled langerin gives an indication of relative affinities of langerin for the different structures. At the lower concentration of langerin the only binding detected is to high mannose oligosaccharides, to a few structures with terminal 6-sulfated galactose residues, and to the blood group B antigen. No binding to other structures with terminal fucose residues or to any GlcNAc-terminated structures is detected, suggesting that these oligosaccharides are only low affinity ligands for langerin. The highest signals are again given by Man_9 and other high mannose oligosaccharides, indicating that these glycans are likely to be the best ligands for langerin.

Binding to oligosaccharides with 6-sulfated galactose residues, as well as to high mannose structures, was demonstrated previously for the IgG-Fc chimera of mouse langerin (25). It is not clear how sulfated galactose is bound by langerin as monosaccharides with an axial 4-hydroxyl group would not be expected to fit into the primary binding site of the mannose-type CRD, which binds sugars with equatorial 3 and 4 hydroxyl groups (11). Langerin does not bind to galactose, so interactions with the sulfate group must contribute significantly to binding of sulfated galactose. There is specificity to these interactions because both mouse and human langerin bind to ligands with terminal 6-sulfated galactose but not to those with galactose sulfated at the 3 position (25). It is possible that 6-sulfated galactose binds to langerin in the secondary binding site identified in the CRD (16).

Structure of Langerin Trimer—To visualize the arrangement of CRDs in the langerin trimer, the trimeric fragment used in the glycan-binding studies was crystallized, and the structure determined at 2.3 Å (Table 1 and Fig. 3). There are two trimers in the asymmetric unit, comprising chains A, B, and C (trimer 1) and D, E, and F (trimer 2). Residues 167–325 are visible in all six crystallographically independent protomers, except that residues 260–261 are absent in copy B and 260–262 in copy D. Additional N-terminal residues are seen in some copies (see

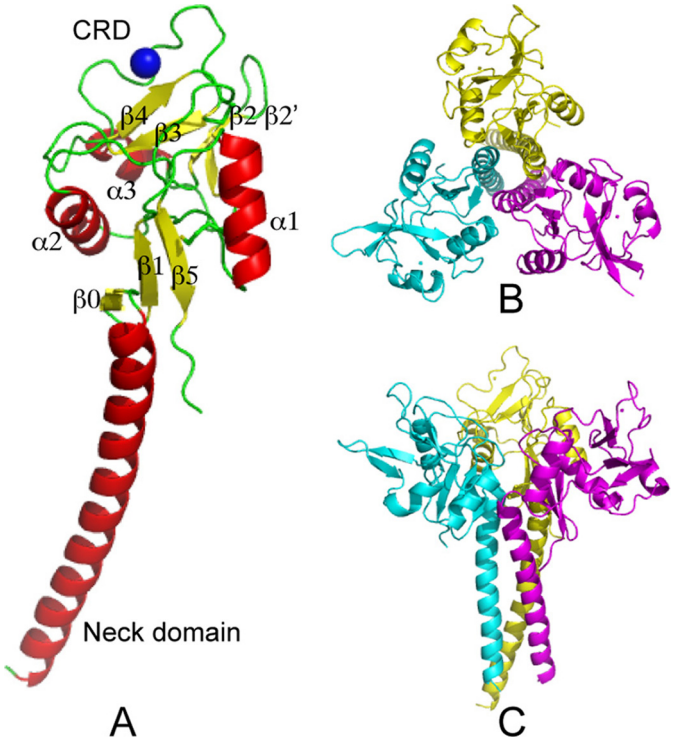


FIGURE 3. **Diagram of the overall structure of truncated langerin.** *A*, secondary structure of protomer A. α-Helices are shown in *red*, β-sheets in *yellow*, loops in *green*. Ca^{2+} is shown as a *blue sphere*. *B*, top view of the trimer structure. Protomer A is shown in *yellow*, B in *magenta*, and C in *cyan*. *C*, side view of the trimer.

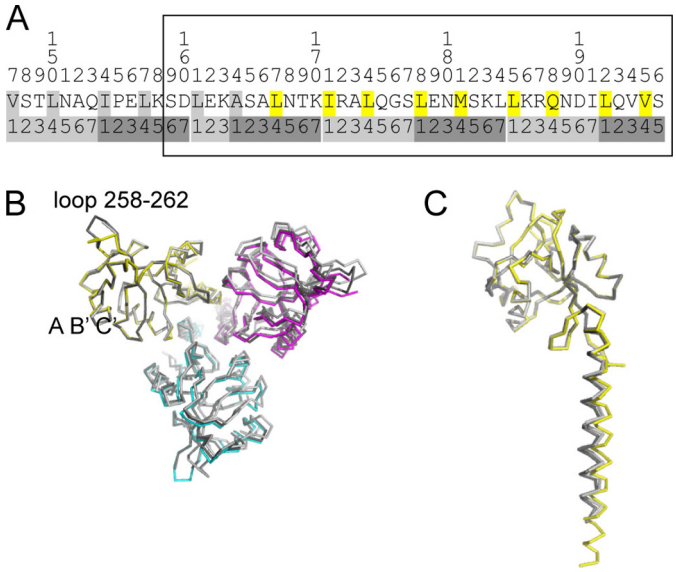


FIGURE 4. **Neck structure and deviation from 3-fold symmetry.** *A*, sequence of the neck domain of truncated langerin. The amino acid sequence numbers are marked on *top*, and the heptad repeat numbering is on the *bottom*. Residues 1 and 4 of each heptad repeat are *highlighted*; those *highlighted* in *yellow* are visible in all crystallographically independent copies. The *box* indicates the visible portion of protomer A. *B*, superposition of protomer B CRD residues 200–257 and 263–325 onto A is shown in *light gray*, and C onto A is in *dark gray*. Protomers A, B, and C are shown in *yellow*, *magenta*, and *cyan*, respectively. Note that when protomer B or C is superimposed on protomer A, the remaining protomer does not superimpose well as a result of the deviation from perfect 3-fold symmetry. *C*, flexibility in the neck region revealed by the same superpositioning as in *panel B* but showing only protomer A in *yellow*, protomer B superimposed on A in *light gray*, and protomer C superimposed on protomer A in *dark gray*. When only the CRDs are superimposed, the N-terminal part of the neck does not superimpose well.

below). The CRDs deviate slightly from perfect 3-fold symmetry, with protomers A and B superimposing at a relative rotation angle of 122.9°, and A and C superimposing at an angle of 119.6°; similar deviations are observed in the second crystallographically independent trimer. With the exception of a loop at residues 258–262, the CRDs are virtually identical in structure (see below).

The neck sequence is composed of series of heptad repeats. As expected, the visible portion of the neck, residues 167–199, is all α-helical through residue Ser^{196}. There is a "stammer" (29) corresponding to a deletion of 4 residues in the normal heptad repeat near residue 167, giving rise to a pattern of hydrophobic residues starting at residue 157 of *…d-a-a-d-a-d…* (*i.e.* hydrophobic residues spaced 4-3-3-4-3 instead of 4-3-4-3-4-3) (Fig. 4*A*). This sequence irregularity correlates with the loss of electron density N-terminal to residue 167, suggesting that it is a point of flexibility. Only protomer A is visible starting at residue 159, but this can be ascribed to contacts with a crystal symmetry mate. Of the five other neck regions in the asymmetric unit, one starts from residue 165, one from residue 166, and three from residue 167. Flexibility or rotation in the region prior to the visible N termini is likely required because modeling of longer helices indicates that residues N-terminal to those observed in the electron density would not contact one another.

The neck also displays a pronounced curvature that causes a deviation from the symmetry of the CRDs. Protomers in the C-terminal part of the neck, close to the beginning of the CRD, superimpose on one another with rotational relationships similar to those of the CRDs, indicating a fixed, rigid relationship between the neck and CRD. However, the neck starts to diverge toward the N-terminal section, where it twists with respect to the C-terminal portion (Fig. 4). This can be seen by dividing the visible portion of the neck into N- and C-terminal segments, spanning residues 167–184 and 185–195. Protomers in the C-terminal region were superimposed, and this transformation was applied to the entire neck. Then, the N-terminal portions of the neck were superimposed. A rotation of the helix of 8.3° was needed to superimpose the N-terminal region of helix B onto the corresponding portion of helix A, and a rotation of 8.2° was needed to superimpose the N-terminal region of helix C onto helix A.

Interactions Stabilizing the Trimer and Positioning the CRDs—Interactions between four regions of the protein contribute to oligomerization of the trimer and setting the position of the CRDs: (i) interhelical contacts within the neck; (ii) contacts between the end of the neck domain of one protomer and the CRD of another; (iii) contacts between adjacent CRDs; and (iv) contacts between the neck domain and the CRD within a protomer. Because the CRD is a monomer without the neck domain (11), it can be concluded that the neck is necessary for oligomerization but that the other interactions seen in the trimer structure, both between CRDs and between the neck domain of one monomer and the CRD of another, contribute to oligomerization and also influence the spacing between the CRDs. The positioning of one CRD relative to another is also influenced by the interactions between the neck and CRD

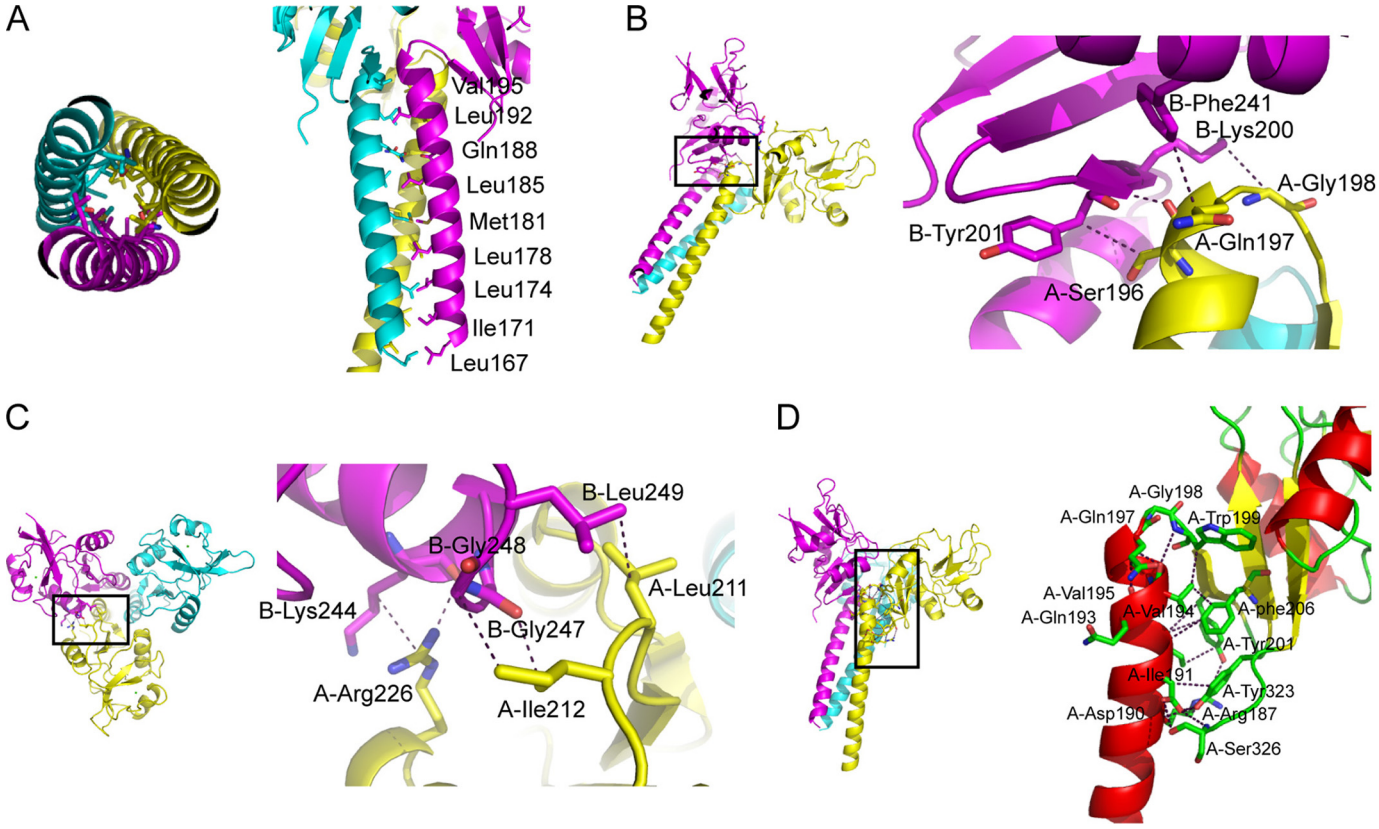


FIGURE 5. **Interactions stabilizing the trimer and positioning the CRDs.** Protomer A is shown in *yellow*, B in *magenta*, and C in *cyan*. *A*, interactions between the neck regions. *Left*, top view. *Right*, side view. *B*, interactions between the end of the neck domain of one monomer and the CRD from another monomer. *Left*, overall view with a *box* showing the area where the interactions take place. *Right*, expanded view of the *box* on the *left*. *C*, interactions between CRDs. *Left*, overall view with a *box* showing the area where the interactions take place. *Right*, expanded view of the *box* on the *left*. *D*, neck to CRD interactions in one monomer. *Left*, overall position of the interactions between the neck domain and the CRD. *Right*, close up view showing only monomer A. Helices are shown in *red*, β-strands in *yellow*, loops and carbon atoms in *green*, oxygen atoms in *red*, and nitrogen atoms in *blue*.

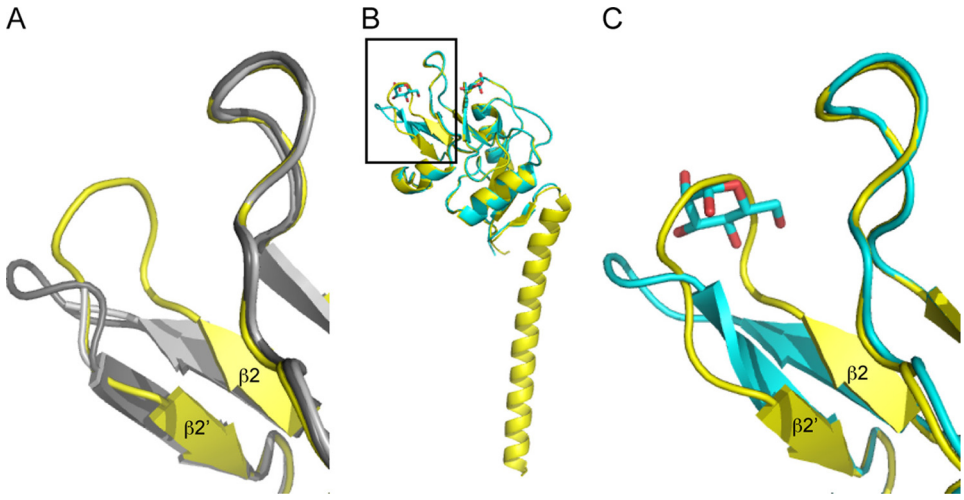


FIGURE 6. **Flexibility of the loop region comprising residues 258–262.** *A*, superposition of the CRD in chain B onto the CRD in chain A, in *light gray*, and chain C onto chain A, in *dark gray*. Chain A of truncated langerin is shown in *yellow*. *B*, superposition of the mannose-CRD complex structure, PDB 3BC7, on the CRD of truncated langerin protomer A. The *box* indicates the largest difference. *C*, close up view of the superposition shown in *B*. Note that in both this superposition as well as those of the different copies of the CRD within the trimer described here, the biggest change is in residues 258–262, which form the loop between β-strand 2 and β-strand 2′ (Fig. 3) in the vicinity of the secondary sugar-binding site.

within one monomer. Collectively, these interactions make the trimer a rigid unit, so it is not surprising that the two crystallographically independent trimers display similar spacings of their CRDs. Moreover, rigid-body refinement of the lower resolution crystal form (see "Experimental Procedures"; data not shown) indicates that the CRDs in the eight crystallographically independent trimers in that crystal lattice also have the same spacing of CRDs. The following analysis was performed on the trimer designated by chains A, B, and C in the coordinates, but similar interactions are present in the other trimer.

As expected, nonpolar packing interactions occur between hydrophobic residues at the *a* and *d* positions of the heptad repeats (Fig. 5*A*). The only exception is Gln^{188}, which is a *d* position. The side chain from each copy points toward the 3-fold axis, and there is electron density, presumably a water molecule but possibly Na^+, on the 3-fold axis within hydrogen-bonding distance of the side chains. If this is a water molecule, then there would be a statistical average of

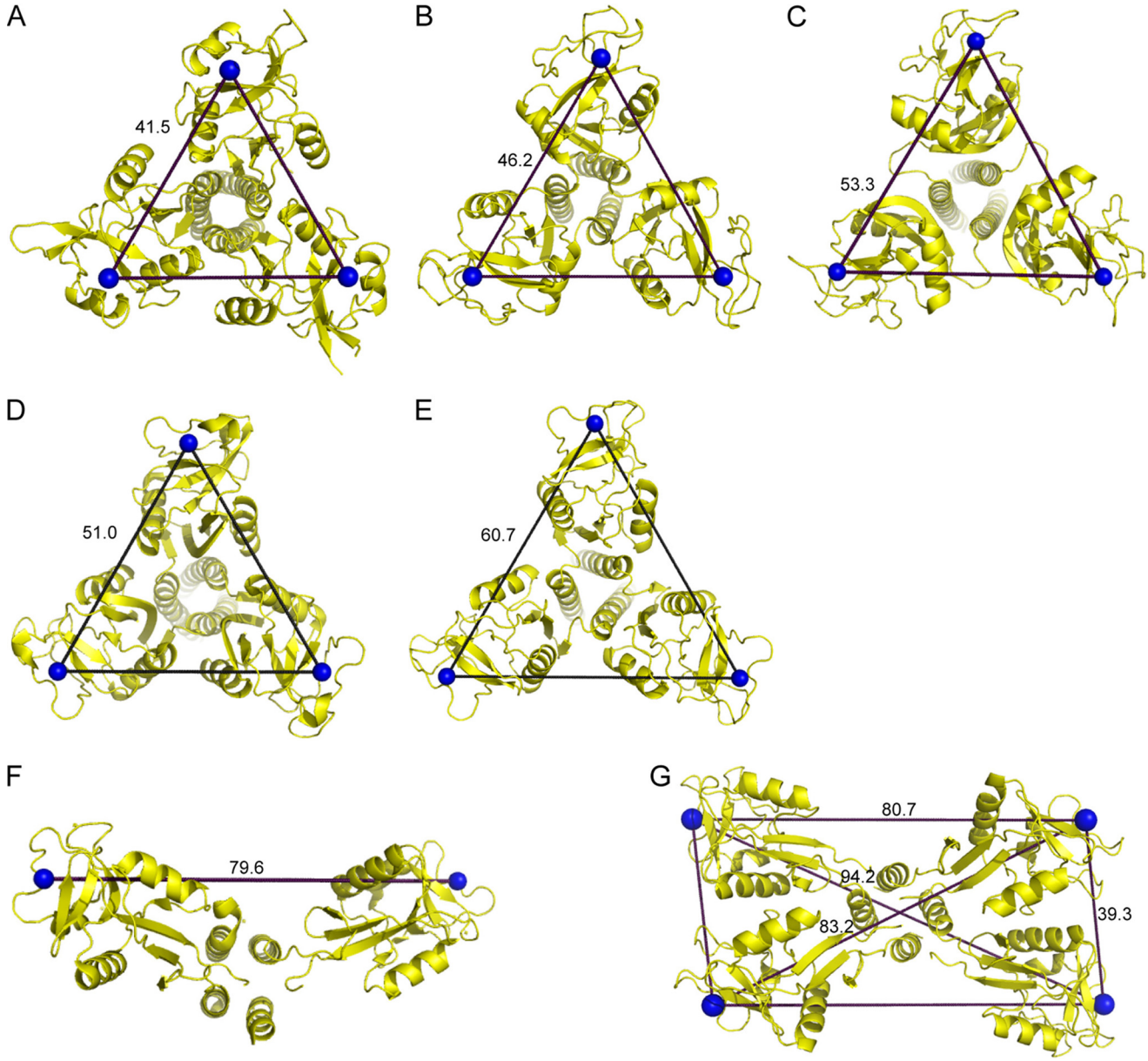


FIGURE 7. **Oligomers of C-type lectins.** *A*, truncated langerin. The spacing shown is the average of the distances. *B*, human mannose-binding protein, PDB 1HUP. The trimer was generated by applying crystallographic symmetry. *C*, rat mannose-binding protein, PDB 1RTM. The spacing shown is the average of the distances. *D*, human lung surfactant protein D, PDB 1PW9. *E*, surfactant protein A, PDB 1R13. The trimer was generated by applying crystallographic symmetry. *F*, DC-SIGNR, dimer structure, PDB 1SL6. The dimer was completed by applying crystallographic symmetry. *G*, DC-SIGNR tetramer structure, PDB 1XAR. The tetramer was generated by applying crystallographic symmetry. The spacing for this model was by calculated by measuring the distance between Na^+ occupying the positions of the Ca^{2+}.

two-thirds of the side chains with either the carbonyl oxygen or amide nitrogen facing the water molecule. In addition, there are several interactions between the loop that connects the neck helix to the CRD with the β-strand 0 and α-helix 2 of a neighboring CRD (Fig. 5*B*). Gly^{198} (chain A) packs against the side chain of Lys^{200} (chain B), Ser^{196} (A) packs against Tyr^{201} (B), and Gln^{197} (A) packs with Phe^{241} (B). There are also hydrogen bonds formed between the main chain of Ser^{196} (A) and Tyr^{201} (B) (Fig. 5*B*). The trimer is further stabilized by inter-CRD interactions: α-helix 1 and the preceding loop of one protomer contact α-helix 2 and the loop that follows it in the neighboring CRD. Specifically, Leu^{211} and Ile^{212} (chain A) pack against Leu^{249}, Gly^{247}, and Gly^{248} (chain B), and Arg^{226} (A) packs against the side chain and also forms a hydrogen bond with the main-chain carbonyl hydrogen of residue Lys^{244} (B) (Fig. 5*C*).

The interactions between the neck and CRD within one protomer involve residues from the α-helical neck and residues from the loop prior to β-strand 0, β-strand 1, and the C-terminal part of the CRD that follows β-strand 5 (Fig. 5*D*). The loop before β-strand 0, comprising residues 197–199, is involved in both hydrogen bond and packing interactions with residues from the C-terminal part of the neck helix (residues 193–195). The side chain of Tyr^{201} in β-strand 0 packs against the side chain of Ile^{191} and also hydrogen bonds to Arg^{187}. The side chain of Phe^{206}, which is part of β-strand 1, is packed against

the side chain of Val^{194} and the Cα of Ile^{191}. Tyr^{323}, positioned at the C-terminal part of the CRD, packs Cβ of Asp^{190} and Cα of Arg^{187} and forms a hydrogen bond with the side chain of Asp^{190}. Ser^{326}, near the C terminus of the CRD, further sets the position of the CRD by a hydrogen bond to Asp^{190}. The side chains of Tyr^{323}, Phe^{206}, Tyr^{201}, and Trp^{199} from the CRD together with the side chains of Ile^{191} and Val^{194} of the neck pack to form a hydrophobic pocket in between the two domains. This pocket is buried between the two domains, locked by hydrogen bond interactions between Asp^{190} and Arg^{187} from the neck and Tyr^{201}, Tyr^{323}, and Ser^{326} from the CRD (Fig. 5*D*).

The langerin trimer structure differs in significant ways from the mannose-binding protein trimer used to generate computational models of langerin (19). This is due principally to the differences imparted to the neck-CRD interactions by the extension at the N terminus of the CRD provided by β0 and surrounding residues. These differences result in different CRD-CRD interfaces within the trimer, although the rigidity of the trimer observed in the different crystallographic copies is reminiscent of mannose-binding protein.

Flexibility of the Loop Region Forming a Secondary Sugar-binding Site—In the previously reported structure of the monomeric langerin CRD bound to mannose or maltose, sugars were bound to the conserved principal Ca^{2+}-binding site in a mode similar to other mannose-type ligands bound to the C-type CRD. A surprising feature, however, was a second, Ca^{2+}-independent binding site formed by the loop at residues 258–262 (16). In the present structure, the only significant deviation among the CRDs occurs in this loop: if this loop is omitted from the superposition, the root mean square deviation is 0.26 Å for the Cα positions of CRDs of chains A and B and 0.30 Å between chains A and C. The 258–262 loop adopts different conformations in different protomers (Fig. 6), and the various conformers are stabilized by contacts with other trimers in the crystal lattice. These observations indicate that the loop is flexible, and formation of the crystal lattice selects various low energy conformations from the ensemble in solution. One of these conformations is similar to that observed bound to mannose, whereas the other would occlude mannose from the site. Presumably, the bound sugar stabilizes this otherwise flexible loop.

Fixed Orientation of the Primary Sugar-binding Sites in the Trimer—Oligomerization confers high avidity binding to C-type lectins. In some cases, for example, mannose-binding protein, the CRDs are held in fixed positions, which allows them to bind avidly only to dense, repetitive arrays of sugars on target cells while binding only weakly to the more closely spaced terminal sugars on host glycans (13) (Fig. 7). In other cases, for example, DC-SIGN, the CRDs are linked flexibly to the oligomerization domain (14) (Fig. 7), allowing them to adapt and bind avidly to glycans displaying a broader range of spacings. The structural data presented here indicate that the langerin CRDs are held in a fixed orientation, with their principal carbohydrate-binding sites (defined as the position of the principal Ca^{2+}) spaced ~42 Å from one another. This is similar to the fixed spacing of CRDs in human mannose-binding protein. Unlike mannose-binding protein, langerin can bind preferentially to clusters of mannose in high mannose *N*-linked oligosaccharides, but this is a property intrinsic to the CRD rather than the oligomeric structure. The wide, fixed spacing of the langerin CRDs indicates that avidity enhancements due to binding of multiple CRDs might occur only for certain classes of multiply glycosylated glycoprotein or cell surface ligands, which likely contributes to the specificity for particular pathogens.

Acknowledgments—We thank David Smith of the Consortium for Functional Glycomics for performing glycan array analysis and Kurt Drickamer for helpful discussions and comments on the manuscript.

REFERENCES

1. Takahara, K., Yashima, Y., Omatsu, Y., Yoshida, H., Kimura, Y., Kang, Y. S., Steinman, R. M., Park, C. G., and Inaba, K. (2004) *Int. Immunol.* **16,** 819–829
2. Turville, S. G., Cameron, P. U., Handley, A., Lin, G., Pöhlmann, S., Doms, R. W., and Cunningham, A. L. (2002) *Nat. Immunol.* **3,** 975–983
3. Valladeau, J., Duvert-Frances, V., Pin, J. J., Dezutter-Dambuyant, C., Vincent, C., Massacrier, C., Vincent, J., Yoneda, K., Banchereau, J., Caux, C., Davoust, J., and Saeland, S. (1999) *Eur. J. Immunol.* **29,** 2695–2704
4. Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., Caux, C., Lebecque, S., and Saeland, S. (2000) *Immunity* **12,** 71–81
5. Kissenpfennig, A., Aït-Yahia, S., Clair-Moninot, V., Stössel, H., Badell, E., Bordat, Y., Pooley, J. L., Lang, T., Prina, E., Coste, I., Gresser, O., Renno, T., Winter, N., Milon, G., Shortman, K., Romani, N., Lebecque, S., Malissen, B., Saeland, S., and Douillard, P. (2005) *Mol. Cell Biol.* **25,** 88–99
6. McDermott, R., Bausinger, H., Fricker, D., Spehner, D., Proamer, F., Lipsker, D., Cazenave, J. P., Goud, B., De La Salle, H., Salamero, J., and Hanau, D. (2004) *J. Invest. Dermatol.* **123,** 72–77
7. Hunger, R. E., Sieling, P. A., Ochoa, M. T., Sugaya, M., Burdick, A. E., Rea, T. H., Brennan, P. J., Belisle, J. T., Blauvelt, A., Porcelli, S. A., and Modlin, R. L. (2004) *J. Clin. Invest.* **113,** 701–708
8. Idoyaga, J., Cheong, C., Suda, K., Suda, N., Kim, J. Y., Lee, H., Park, C. G., and Steinman, R. M. (2008) *J. Immunol.* **180,** 3647–3650
9. de Witte, L., Nabatov, A., and Geijtenbeek, T. B. (2008) *Trends Mol. Med.* **14,** 12–19
10. de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. A., de Gruijl, T., Piguet, V., van Kooyk, Y., and Geijtenbeek, T. B. (2007) *Nat. Med.* **13,** 367–371
11. Stambach, N. S., and Taylor, M. E. (2003) *Glycobiology* **13,** 401–410
12. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem.* **65,** 441–473
13. Weis, W. I., and Drickamer, K. (1994) *Structure* **2,** 1227–1240
14. Feinberg, H., Guo, Y., Mitchell, D. A., Drickamer, K., and Weis, W. I. (2005) *J. Biol. Chem.* **280,** 1327–1335
15. Menon, S., Rosenberg, K., Graham, S. A., Ward, E. M., Taylor, M. E., Drickamer, K., and Leckband, D. E. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 11524–11529
16. Chatwell, L., Holla, A., Kaufer, B. B., and Skerra, A. (2008) *Mol. Immunol.* **45,** 1981–1994
17. Feinberg, H., Mitchell, D. A., Drickamer, K., and Weis, W. I. (2001) *Science* **294,** 2163–2166
18. Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992) *Nature* **360,** 127–134
19. Thépaut, M., Valladeau, J., Nurisso, A., Kahn, R., Arnou, B., Vivès, C., Saeland, S., Ebel, C., Monnier, C., Dezutter-Dambuyant, C., Imberty, A., and Fieschi, F. (2009) *Biochemistry* **48,** 2684–2698
20. Collaborative Computational Project, N. (1994) *Acta Crystallogr. Sect. D* **50,** 760–763
21. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* **40,** 658–674
22. Tong, L. (1996) *Acta Crystallogr. Sect. A* **52,** 782–784
23. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132

24. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
25. Galustian, C., Park, C. G., Chai, W., Kiso, M., Bruening, S. A., Kang, Y. S., Steinman, R. M., and Feizi, T. (2004) *Int. Immunol.* **16,** 853–866
26. Feinberg, H., Taylor, M. E., and Weis, W. I. (2007) *J. Biol. Chem.* **282,** 17250–17258
27. 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
28. Powlesland, A. S., Fisch, T., Taylor, M. E., Smith, D. F., Tissot, B., Dell, A., Pöhlmann, S., and Drickamer, K. (2008) *J. Biol. Chem.* **283,** 593–602
29. Brown, J. H., Cohen, C., and Parry, D. A. (1996) *Proteins* **26,** 134–145