Molecular basis of human CD22 function and therapeutic targeting

June Ereño-Orbea¹, Taylor Sicard¹,², Hong Cui¹, Mohammad T. Mazhab-Jafari¹, Samir Benlekbir¹, Alba Guarné³, John L. Rubinstein¹,²,⁴ & Jean-Philippe Julien¹,²,⁵

CD22 maintains a baseline level of B-cell inhibition to keep humoral immunity in check. As a B-cell-restricted antigen, CD22 is targeted in therapies against dysregulated B cells that cause autoimmune diseases and blood cancers. Here we report the crystal structure of human CD22 at 2.1 Å resolution, which reveals that specificity for α2-6 sialic acid ligands is dictated by a pre-formed β-hairpin as a unique mode of recognition across sialic acid-binding immunoglobulin-type lectins. The CD22 ectodomain adopts an extended conformation that facilitates concomitant CD22 nanocluster formation on B cells and binding to trans ligands to avert autoimmunity in mammals. We structurally delineate the CD22 site targeted by the therapeutic antibody epratuzumab at 3.1 Å resolution and determine a critical role for CD22 N-linked glycosylation in antibody engagement. Our studies provide molecular insights into mechanisms governing B-cell inhibition and valuable clues for the design of immune modulators in B-cell dysfunction.

¹Program in Molecular Medicine, The Hospital for Sick Children Research Institute, Toronto, ON, Canada M5G 0A4. ²Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8. ³Department of Biochemistry and Biomedical Science, McMaster University, Hamilton, ON, Canada L8S 4L8. ⁴Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada M5G 1L7. ⁵Department of Immunology, University of Toronto, Toronto, ON, Canada M5S 1A8. June Ereño-Orbea and Taylor Sicard contributed equally to this work. Correspondence and requests for materials should be addressed to J.-P.J. (email: jean-philippe.julien@sickkids.ca)
ialic acid-binding immunoglobulin-like lectin (Siglec) receptors are a family of 14 cell surface transmembrane proteins that bind specifically to sialic acid (Sia)-containing glycans, facilitating cell adhesion and/or cell signaling. Sigles are found primarily in vertebrates on a wide range of immune cells including granulocytes, neutrophils, monocytes, dendritic cells, eosinophils, mast cells, T cells, and B cells. Their functions are determined by their cellular distribution and ligand specificity. One of the best described Sigles is CD22 (Siglec-2), whose expression is restricted to B cells. CD22 plays a critical role in establishing a baseline level of B-cell inhibition, and thus is a critical determinant of homeostasis in humoral immunity. As a result, CD22 knockout mice have an increased incidence of autoimmune disease and hyperactive B cells.

CD22 is a single-spanning membrane glycoprotein of 140 kDa on the surface of B cells. The extracellular domain (ECD) of CD22 is comprised of seven immunoglobulin (Ig) domains (d1–d7) and 12 putative N-linked glycosylation sites. The most N-terminal domain (d1) is of predicted V-type Ig-like fold and recognizes Sias containing α2,6-linkages. While human CD22 binds preferentially to Sia N-acetyl neuraminic acid (Neu5Ac), murine CD22 has higher specificity to Sia N-glycolyl neuraminic acid (Neu5Gc), highlighting species-dependent specificities for CD22 ligand recognition. Moreover, cell surface sialylated glycans can be modified, typically at the 4, 6, 7, or 9 hydroxyl positions, which can alter their binding specificities to CD22.

Some of these changes are associated with cellular dysregulation. As examples, O-acetylation at the 9 hydroxyl position has been implicated in autoimmunity and in progression of childhood acute lymphoblastic leukemia.

CD22, itself sialylated, forms homo-oligomers in cis on the surface of B cells. CD22 oligomers are located in dynamic nanoclusters and create a signal threshold of antigen binding that must be achieved prior to B-cell activation. CD22 activity is mediated through the intracellular recruitment of phosphatases that facilitate dephosphorylation of stimulatory co-receptors. CD45 has also been implicated as a CD22 ligand in cis.

In addition, trans engagement of Sia-containing ligands on antigen-bearing cells results in the recruitment of CD22 to the immunological synapse and inhibits BCR signaling in response to self-antigens.

The inhibitory function of CD22 and its restricted expression on B cells makes CD22 an attractive target for B-cell depletion in cases of autoimmune diseases and B-cell-derived malignancies. Numerous therapeutic approaches in development harness B-cell inhibition through CD22 to induce tolerance or anergy, or to deplete dysregulated B cells through CD22 targeting by either small molecules or antibody–drug conjugates. Constitutive CD22 clathrin-mediated endocytosis allows for the targeted internalization of CD22 ligands.

Table 1 Crystallographic data collection and refinement statistics

| CD2220-330, SA native (SVKJ) | CD2220-330, SA HgCl2 | CD2220-330, SA + α2-6 sialyllactose (5VKM) | Epratuzumab Fab (SVKL) | CD2220-330, 4Q + epratuzumab Fab (SVL3) |
|-------------------------------|----------------------|--------------------------------------------|------------------------|------------------------------------------|
| **Data collection**            |                      |                                            |                        |                                          |
| Space group                   | C2                   | C2                                         | C2                     | P1                                      | P1                                      |
| Cell dimensions               | α, b, c (Å)          | 126.8, 56.6, 49.4                        | 126.8, 56.6, 49.1      | 124.3, 57.9, 48.1                      | 56.7, 61.5, 65.3                      |
|                               | α, β, γ (°)          | 90, 110.7, 90                            | 90, 110.2, 90          | 90, 107.0, 90                          | 71.8, 81.1, 75.9                      |
| **Wavelength**                |                      | 0.9794                                    | 1.0051                 | 1.0033                                  | 0.9794                                  |
| **Resolution (Å)**            | 46.24–212            | 46.24–230                                 | 46.24–230              | 46.24–230                               | 46.24–230                               |
|                               | (2.20–1.21)          | (2.40–2.30)                               | (2.40–2.30)            | (2.40–2.30)                            | (2.40–2.30)                            |
| **Rmerge**                    | 0.063 (0.475)        | 0.049 (0.464)                             | 0.044 (0.459)          | 0.048 (0.464)                           | 0.106 (0.615)                           |
|                               | (0.464)              | (0.464)                                   | (0.464)                | (0.464)                                 | (0.464)                                 |
| **Rfree**                     | 0.037 (0.282)        | 0.022 (0.229)                             | 0.019 (0.229)          | 0.021 (0.229)                           | 0.041 (0.321)                           |
|                               | (0.229)              | (0.229)                                   | (0.229)                | (0.229)                                 | (0.229)                                 |
| **I/σ(I)**                    | 15.8 (2.3)           | 15.8 (1.7)                                | 18.8 (1.8)             | 17.4 (2.0)                             | 9.9 (1.6)                               |
|                               | 99.9 (80.1)          | 99.9 (83.9)                               | 99.9 (85.1)            | 99.9 (86.1)                            | 99.5 (59.7)                            |
| **Completeness (%)**          | 99.8 (99.8)          | 95.8 (74.9)                               | 95.5 (73.4)            | 98.7 (90.1)                            | 99.6 (99.8)                            |
|                               | 3.8 (3.8)            | 3.6 (2.1)                                 | 3.6 (2.1)              | 3.8 (2.8)                              | 3.8 (3.8)                              |
| **Redundancy**                |                      |                                          |                         |                                        |                                        |
|                               | 32.94 (2.40)         | 32.2 (2)                                  | 32.2 (2.40)            | 52.019 (5.050)                         | 65.849 (6.779)                         |
|                               | (2.20)               | (2.40)                                    | (2.40)                 | (2.40)                                 | (2.40)                                 |
| **Resolution (Å)**            | 46.2–2.1             |                                          | 46.2–2.1               | 46.2–2.1                               | 46.2–2.1                               |
| **No. reflections**           | 18,725 (1848)        | 16,769 (1681)                             | 16,769 (1681)          | 52,019 (5.050)                         | 65,849 (6.779)                         |
|                               | (0.227/0.223)        | (0.217/0.256)                             | (0.217/0.256)          | (0.206/0.248)                          | (0.277/0.298)                          |
| **No. atoms**                 | 2678                 | 2604                                      | 2604                   | 7260                                   | 22,916                                  |
|                               | 2436                 | 2436                                      | 2436                   | 6680                                   | 22,598                                  |
| **Protein**                   | 85                   | 99                                        | 99                     | 12                                     | 318                                     |
| **Water**                     | 143                  | 53                                        | 53                     | 514                                    | 0                                       |
| **B factors**                 |                      |                                          |                        |                                        |                                          |
| **Protein**                   | 39.8                 | 49.2                                      | 49.2                   | 36.8                                   | 76.9                                    |
| **Hetero**                    | 54.0                 | 63.3                                      | 63.3                   | 42.8                                   | 102.7                                   |
| **Water**                     | 42.2                 | 46.0                                      | 46.0                   | 40.0                                   | NA                                      |
| **r.m.s. deviations**         |                      |                                          |                        |                                        |                                          |
| Bond lengths (Å)              | 0.002                | 0.005                                     | 0.005                  | 0.007                                  | 0.006                                   |
| Bond angles (°)               | 0.57                 | 0.78                                      | 0.78                   | 0.91                                   | 1.32                                    |

NA not applicable
Values in parentheses are for highest-resolution shell
crystals (Table 1).

solved by multiple anomalous dispersion using mercury-soaked diffracted X-rays to 2.1 Å resolution and the crystal structure was understood from mice studies⁴, 22 and molecular co-localization imaging¹¹, but poorly delineated at the atomic level. Using X-ray crystallography, single-particle electron microscopy (EM) and small-angle X-ray scattering (SAXS) techniques, we solved the molecular structure of the extracellular portion of human CD22 alone and in complex with its ligand α₂-6 siayllactose. Our structural analysis of the full-length extracellular portion of CD22 reveals that the CD22 ECD adopts an extended conformation with low flexibility optimally configured to form nanoclusters and interact with self-ligands at the immune synapse. We also structurally delineate the CD22 site targeted by the therapeutic antibody epratuzumab and determine a critical role for CD22 N-linked glycosylation in therapeutic antibody engagement, with potential implications for CD22 recognition on dysfunctional B cells.

Results
Crystallization of the human CD22 ectodomain. To facilitate crystallization of CD22 ECD, we created a truncated construct that contains the first three Ig domains (residues 20–330) with five of the six N-linked glycosylation sites mutated to alanines (S5A: N67A, N112A, N135A, N164A, N231A) (Supplementary Fig. 1). Alanine mutation at N101 disrupted protein expression, (5A: N67A, N112A, N135A, N164A, N231A) (Supplementary Fig. 1). Crystals of the CD2220 – 330,5A construct as a string and extend ~110 Å (Fig.1). As expected, d1 adopts a V-type fold. Unexpectedly, d2 adopts a C1-type fold, rather than the predicted C2-type fold (Fig. 1). Siglec-4, myelin-associated glycoprotein (MAG), was also reported to have a C1-type fold for d223 (Supplementary Fig. 2a). The d2 of Siglec-5 adopts the predicted C2-type fold (Supplementary Fig. 2a)²⁴. Thus, the CD22 structure further highlights a heterogeneity in the V-type/d2 structural dispositions among Siglec family members, which might be dictated by whether they are members of classic or CD33-related Siglecs sub-classes.

CD22 has elongated D and E strands in d2 that generate a remarkably large interface area with d1 and extensively stabilize the orientation of the ligand binding domain (Fig. 1). As such, d2 of CD22 buries 721 Å² of surface area on d1, which is substantially greater than that for Siglec-4 (684 Å²) and Siglec-5 (461 Å²) (Supplementary Fig. 2a)²⁵. Within the D–E loop of d1, clear electron density is observed for the N-linked glycan at N101, accounting for all monosaccharides in the GlcNAc₂Man₃ core (Fig. 1; Supplementary Fig. 1f). The N101 glycan is positioned in a hydrophobic environment at the d1/d2 junction (Supplementary Fig. 1f), burying 348 Å² of surface area, which helps explain its importance for CD22 expression. Comparison of the human CD22 sequence with Siglec-2 orthologs and Siglecs-4, -5, and -7 reveals that the N101 glycosylation site is well conserved (Fig. 2a). As predicted, d3 adopts a C2-type fold (Fig. 1). CD22 has elongated D and E strands in d2 that generate a remarkably large interface area with d1 and extensively stabilize the orientation of the ligand binding domain (Fig. 1). As such, d2 of CD22 buries 721 Å² of surface area on d1, which is substantially greater than that for Siglec-4 (684 Å²) and Siglec-5 (461 Å²) (Supplementary Fig. 2a)²⁵. Within the D–E loop of d1, clear electron density is observed for the N-linked glycan at N101, accounting for all monosaccharides in the GlcNAc₂Man₃ core (Fig. 1; Supplementary Fig. 1f). The N101 glycan is positioned in a hydrophobic environment at the d1/d2 junction (Supplementary Fig. 1f), burying 348 Å² of surface area, which helps explain its importance for CD22 expression. Comparison of the human CD22 sequence with Siglec-2 orthologs and Siglecs-4, -5, and -7 reveals that the N101 glycosylation site is well conserved (Fig. 2a). In the crystal structures of Siglecs-4 (PDB ID: 5LFR) and -7 (PDB ID: 1OT7), the equivalent N101 glycan also buries a significant surface area on the protein (589 and 317 Å², respectively)²³, 25, 26. As predicted, d3 adopts a C2-type Ig domain topology (Fig. 1). The CD22 d2/d3 interface is smaller (376 Å²) and more hydrophilic than the d1/d2 junction, yet contributes significant inter-Ig domain contacts (including extensive H-bonding networks) that stabilize the CD22 d1–d3 N-terminus in a specific disposition (Fig. 1).

The CD22 V-type d1 domain displays unique features compared to other Siglec family members: (1) the C–C’ strands
Fig. 2 Differences in ligand binding domain d1 among the Siglec family. a Sequence alignment of the V-type domain of mammalian CD22 and other Siglecs. Siglec-1, -4, -5, -7, and -8. hCD22 secondary structure elements are represented atop the alignment. Putative N-glycosylation sites N67, N101, N112, and N135 are highlighted in blue boxes. Conserved R120 involved in binding Sia ligands is marked by a red box. b Comparison of available three-dimensional structures of V-type domains from classic Siglecs (including the CD22 structure reported here, Siglec-1 (PDB ID: 1QFP)67 and Siglec-4 (PDB ID: 5LFR)23) and CD33/Siglec-3-related Siglecs (Siglec-5 (PDB ID: 2ZG2)24, Siglec-7 (PDB ID: 1O7S)26, and Siglec-8 (PDB ID: 2N7A)51). Differences in C-terminus (containing the elongated α-β-hairpin (hereafter named C1 and C2) that shapes the Sia binding site; (2) strand G is continuous, without a loop insertion; and (3) the C-D loop protrudes away from the core V-type structure (Fig. 2b). These unique features result in the CD22 V-type domain being the most structurally distant member (highest root-mean-square-deviation (r.m.s.d.)) across Siglecs of known structure (Fig. 2c).

CD22 specificity for α2-6 sialyllactose ligands. To understand the structural basis of CD22 specificity for α2-6 Sia ligands, we soaked CD22,20,330,5A crystals with α2-6 sialyllactose [Neu5Acα(2-6)Galβ(1-4)Glc] and solved the liganded structure at 2.2 Å resolution (Table 1). The CD22 binding site for α2-6 sialyllactose is formed by strands F and G and loop C–C′ (containing the elongated C1/C2 β-hairpin) (Fig. 3a). No electron density was observed for the glucose moiety of α2-6 sialyllactose [Neu5Acα(2-6)Gal] in our crystal structure. This is likely due to a lack of interactions for this moiety with CD22. Similar to other Siglecs, the majority of CD22 interactions occur through the Sia portion of the ligand (181 Å² of buried surface area for Sia out of a total of 276 Å² for Neu5Acα(2-6)Gal) (Supplementary Table 1). The negatively charged C1 carboxylate of Sia interacts via a salt bridge with the guanidinium group of the highly conserved R120 (Fig. 3a, b). Substitution of R120 to either A or E completely abrogates binding to α2-6 sialyllactose in isothermal titration calorimetry (ITC), compared to a 281 ± 10 μM binding affinity for WT CD22 at 25 °C (Supplementary Fig. 3), as consistent with previous findings27. E126 and W128 make key contacts with Sia (Fig. 3a, b), and these interactions corroborate previous biochemical studies that delineated the CD22 binding site by mutagenesis of residues to lysine and arginine, respectively28. The C1/C2 β-hairpin constrains the binding pocket of CD22. Y64 stacks against the CD22 β-hairpin and the hydrophobic face of galactose of α2-6 Sia ligands (Fig. 3c). Y64 thus largely dictates specificity for the α2-6 glycosidic linkage, and not the α2-3 glycosidic linkage (Fig. 3c). The equivalent position of human CD22 Y64 in mouse is F68, making the aromatic ring a conserved feature across CD22 in different species to participate in stacking interactions with α2-6 Sia ligands (Supplementary Fig. 3i). R131 H-bonds with the C2 galactose hydroxyl, further contributing to this specificity. Substitution of R131 to A, K, or Q only marginally impacted the binding affinity to α2-6 sialyllactose (Supplementary Fig. 3c–f), indicating a peripheral role for this residue in mediating CD22 ligand binding. Modeling of an extended (and biologically relevant) Sia-Gal-GlcNAc-Man glycan in our
Crystal structure reveals that CD22, and particularly the C1/C2 \( \beta \)-hairpin, is optimally configured to extensively interact with branches of complex N-glycans with terminal \( \alpha \)-2-6 glycosidic linkages (Fig. 3d). We also note that the tip of the C1/C2 \( \beta \)-hairpin itself has a conserved putative N-linked glycan at position N67 (Fig. 2a; Supplementary Fig. 1), which potentially contributes carbohydrate–carbohydrate interactions that further stabilize the CD22 ligand contact.

Overall, the complex and unliganded structures of CD22 are highly similar (Cr r.m.s.d. of 0.35 Å for d1), indicating that carbohydrate recognition by CD22 is largely mediated by a preformed binding site (Supplementary Fig. 2b). Extensive intramolecular H-bonds between C1 and C2 in the \( \beta \)-hairpin and van der Waals interactions between F71 and M129 are major determinants of the preformed binding site (Supplementary Fig. 2c). We note minimal interactions within the crystal lattice that might have artfactually constrained the C1/C2 \( \beta \)-hairpin in our soaking experiments. Conversely, Siglecs-4 and -7 undergo a conformational change from an unordered state to a ligand binding conformation (Supplementary Fig. 2b). For CD22, a preformed binding site results in an enthalpy-driven interaction with \( \alpha \)-2-6 sialylated glycans (Supplementary Fig. 3). Evidently, the C–C′ loop in the Siglec family dictates specificity for terminal moieties of the carbohydrate by adopting one of two different conformations: (i) pointing toward the ligand binding pocket as in CD22, and Siglecs-4, -5, and -8; or (ii) extending away from the ligand binding pocket as in Siglecs-1 and -7 (Supplementary Fig. 2b).

### Molecular Recognition of CD22 by Therapeutic Antibodies

The localization of CD22 in nanoclusters and its extensive N-linked glycosylation likely impact how CD22 can be targeted therapeutically. Consequently, we next characterize the antigenic surface of CD22 recognized by two leading therapeutic antibodies in clinical trials: epratuzumab and pinatuzumab. Binding competition between epratuzumab and pinatuzumab revealed

---

**Fig. 3** CD22 specificity for the \( \alpha \)-2-6 glycosidic linkage. a Sia and Gal moieties (green) of the \( \alpha \)-2-6 sialyllactose ligand extensively H-bond (black dashed lines) with human CD22 (gray). b Schematic representation of the CD22-Neu5Ac(2-6)Gal interaction network. Surface exposed residues are contoured with gray dashed lines. Pink circles with black outline: polar residues, with red outline: acidic residues, with blue outline: basic residues. Green circles represent hydrophobic residues. H-bonds implicating side chains are presented as green dashed lines, and implicating main chain atoms in blue dashed lines. Rendering generated with MOE. c Specificity for the \( \alpha \)-2-6 glycosidic linkage is largely due to residues Y64 and R131 (left). Glycans with \( \alpha \)-2-3 linkages would clash (overlapping dotted spheres) with Y64 (right). d Modeling of Sia-Gal-GlcNAc-Man (cyan surface) reveals that the CD22 C1/C2 \( \beta \)-hairpin is optimally configured to extensively interact with the branch of a complex N-glycan of \( \alpha \)-2-6 glycosidic linkage. The dashed cyan line indicates the direction of the continued glycan branch.
Our structural epratuzumab binds primarily at the base of CD22 d2, with corroborated our low-resolution EM data that showed how slightly acidic pHs (Supplementary Fig. 6e). The structure we con

as previously reported34, 35. All three epratuzumab heavy-chain area (Supplementary Table3) and extends beyond only CD22 d3, shown that the epratuzumab epitope is not located in the ligand

Fig. 4 The CD22 ectodomain adopts a tilted rod-like structure with low flexibility. a The eight primary 2D class average images of CD22 ECD obtained by negative-stain EM (top). The number of particle images in each class is indicated. 2D projections of the ab initio EM reconstruction (middle) and the SAXS 3D volume (bottom) account for nearly all particles classified in 2D classes. Scale bar represents 10 nm. b Views of the ab initio EM reconstruction (blue) and SAXS volume (yellow). The crystal structure of CD2220–330,5A is shown as a surface fitted into the EM map and SAXS volume rendered with UCSF Chimera63

they recognize non-overlapping epitopes on CD22 (Supplemen-

tary Fig. 6a).

To delineate the epratuzumab epitope at high resolution, a glutamine-resurfaced CD2220–330,40Q construct was designed, complexed with epratuzumab Fab, and resulting crystals diffracted to 3.1 Å resolution (Table 1). CD22 d1–d3/epratuzumab Fab crystals were obtained at a relatively low pH (4.6) and we confirmed by bilayer interferometry (BLI) that epratuzumab is capable of binding to CD22 with high affinity in a range of slightly acidic pHs (Supplementary Fig. 6e). The structure corroborated our low-resolution EM data that showed how epratuzumab binds primarily at the base of CD22 d2, with additional interactions with d3 (Fig. 5a; Supplementary Table 3). Our structural findings agree with previous studies33, 34 that have shown that the epratuzumab epitope is not located in the ligand binding site.

The epratuzumab epitope consists of 1308 Å2 of buried surface area (Supplementary Table 3) and extends beyond only CD22 d3, as previously reported34, 35. All three epratuzumab heavy-chain complementarity determining regions (HCDRs), the light-chain CDR1 (LCDR1) and LCDR3 interact with CD22 d2 (Fig. 5b), while HCDR2, LCDR1, and LCDR3 mediate contacts with d3 (Fig. 5c). These interactions were confirmed by mutagenesis in binding experiments (Supplementary Fig. 6b, c). Our 2.0 Å resolution structure of unliganded epratuzumab Fab (Table 1) indicates that its paratope is largely pre-configured for binding its antigenic site (r.m.s.d. of 0.50 Å) (Supplementary Fig. 6d).

Perhaps not surprisingly, the epratuzumab paratope includes an N-linked glycan at position N231, for which clear electron density is observed for the first two GlcNAc residues (Fig. 5d). The CD22 N231Q mutant (a knockout of this glycosylation site) resulted in a 25-fold increase in binding on-rate, and an overall six-fold improvement in binding affinity compared to WT CD22 (Fig. 5e; Supplementary Fig. 7). Consequently, binding kinetics and thermodynamics of epratuzumab Fab to CD22 constructs with different glycan contents revealed an increasing affinity to CD22 with reduced glycan size, with up to a 14-fold improvement in affinity for smaller glycans (327 vs. 24 nM in BLI; 188 vs. 58 nM in ITC) (Fig. 5e; Supplementary Fig. 7). This effect was not as pronounced for pinatuzumab (Supplementary Fig. 7). A tighter affinity for epratuzumab in the presence of smaller N-linked glycans is primarily due to faster rates of association and a sharp decrease in favorable binding entropy (Supplementary Fig. 7). Together, our data indicate that glycosylation on CD22 impacts the ability of epratuzumab to access its epitope.

Discussion

CD22 is a B-cell-restricted co-receptor that plays a critical role in the maintenance of B-cell homeostasis. Our structural delineation of CD22 presented here, combined with extensive literature on the function of human and mouse CD22 orthologs, provides an in-depth molecular understanding of its mode of action.

The biology of ligands binding to CD22 is complex; for example, it remains unclear how the availability of Sia ligands in their various glycoforms modulate CD22 function, in both health and disease. Two common animal Sias exist: Neu5Ac and Neu5Gc. They differ by a hydroxyl at the 5′-position that is irreversibly added to Neu5Ac by cytidine monophospho-N-acetyl neuraminic acid hydroxylase (CMAH)36. Most mammals, including mouse, mainly express Neu5Gc in their tissues. On the contrary, humans lack the ability to synthesize Neu5Gc because they lack CMAH37. Modeling of mouse CD22 interactions with Neu5Gc based on our human CD22 and α2-6 sialyllactose co-crystal structure suggests that a hydroxyl at the 5′-position in Neu5Gc would lead to favorable interactions with E130 (Supplementary Fig. 3b, i). We also note several differences in the composition of the CD22 binding site between human and mouse: P62 and Y64 located in the C1 strand of human CD22 are tyrosine and phenylalanine in mouse CD22, respectively; and R131 located in stand G of human CD22 is a proline in mouse CD22 (Fig. 2a; Supplementary Fig. 3i). These residues likely play a role in the lower affinity of murine CD22 to Neu5Ac38. Differences in both circulating ligand glycoforms and binding site chemical composition between human and mouse CD22 highlight the complexities that have evolved associated with extrapolating findings about CD22 from mouse models to humans.

The predominant ligand recognized by human CD22, Neu5Ac, itself shows structural diversity that arises from N- and O- substitutions, which are of critical importance for ligand recognition and cellular processes7. For example, 9-O-acetylation is the most commonly observed Sia substitution and has been linked to autoimmunity in human and mouse models6. Modeling suggests that CD22 binding to 9-O-acetylated Sia is sterically impeded by W128 located in strand G of the binding pocket (Supplementary Fig. 3i), providing structural insights into why acetylation of Sia on self-antigens prevents CD22 recognition and increases B-cell-mediated autoimmunity. Future structure/function studies will reveal the spectrum of fine specificities associated with Sia ligand recognition by CD22.

Although the role of CD22 as an inhibitory regulator of BCR signaling is well established, the question of how it binds ligands in both cis and trans is not fully understood. The CD22 preformed binding site and its inter-domain arrangement with low flexibility poses the question of how it binds ligands in both cis and trans. We propose that the binding site in d1 is well positioned to bind flexible N-linked glycans of adjacent CD22 molecules to coordinate clustering of CD22. In this model, the elongated, tilted CD22 structure—and the location of its binding site at the N-terminus—is ideal for inter-molecular interactions with flexible bi-, tri-, and/or tetra-antennary glycans terminated in Sia (Fig. 6).
Modeling the predicted N-glycosylation sites in our CD22 structure reveals that they are predominantly located on one face of the protein (Fig. 6; Supplementary Fig. 1), facilitating inter-molecular interactions in cis and leading to the formation of an interconnected nexus of N-linked glycans atop CD22 nanoclusters (Fig. 6). The ECD of CD22 extends ~300 Å on the surface of B cells, making it optimal in length to also interact selectively with other glycoproteins of similar dimensions, such as protein tyrosine phosphatase optimal in length to also interact selectively with other glycoproteins.

Although cis ligands occupy CD22 on resting B cells, its relatively weak binding affinity for α2-6 Sia-terminated glycans (~250 μM) would allow its predisposed binding site to dynamically exchange interacting partners in the presence of trans ligands on adjacent cells, causing the redistribution of CD22 nanoclusters to the site of cell contact. The length of CD22 may therefore also be optimally evolved to participate in cell–cell recognition through interactions with glycoproteins in trans, allowing CD22 recruitment to the immune synapse to surround BCR clusters and sustain B-cell inhibition in the presence of self-ligands on adjacent cells, causing the redistribution of CD22 to the site of cell contact. The length of CD22 may also be optimally evolved to participate in cell–cell recognition through interactions with glycoproteins in trans, allowing CD22 recruitment to the immune synapse to surround BCR clusters and sustain B-cell inhibition in the presence of self-ligands on adjacent cells, causing the redistribution of CD22 to the site of cell contact.

CD22 undergoes clathrin-mediated endocytosis, and is a recycling receptor that can shuttle cargo between the surface of the B cell and the early/sorting endosomes. This mechanism has been exploited for the delivery of antibody conjugated toxins for B-cell depletion. Epratuzumab functions through an alternative mechanism, whereby it acts as a CD22 agonist causing prolonged B-cell inhibition, in addition to Fc-dependent receptor tropocytosis. Here, we show that epratuzumab is capable of binding to CD22 with high affinity in a range of pH’s, including acidic pH’s corresponding to early/sorting endosomes (pH 5.5–6.5) and late endosomes/lysosomes (pH 4.5–5.5) (Supplementary Fig. 6e). Based on these results, and the very slow off-rates that are observed at the pH of the early/sorting endosomes, it is unlikely that CD22 internalization would lead to dissociation of epratuzumab from CD22, and therefore it is probable that epratuzumab is recycled back to the cell surface alongside CD22, as shown previously for other anti-CD22 antibodies. The recycling of epratuzumab-bound CD22 from the cell surface to the early endosomes may lead to prolonged B-cell inhibition and higher levels of trogocytosis.

Location of the epratuzumab epitope at the d2/d3 interface on a tilted CD22 might ideally promote crosslinking by the IgG, consequently causing accentuated B-cell inhibition and apoptosis as a unique mechanism of action for B-cell depletion through CD22 (Supplementary Fig. 7g). Our biophysical and structural data also indicate that glycosylation on CD22 impacts the ability of this therapeutic antibody to access its epitope, while at the same time favorably contributing binding energy to the epratuzumab–CD22 interaction. This hindrance/dependency relationship for binding to heavily glycosylated proteins has previously been described for broadly neutralizing antibodies recognition of the HIV Envelope trimer. A glycan dependency for epratuzumab binding to CD22 is particularly relevant given that variable glycoforms such as truncations or modifications have been observed on surface glycoproteins in cancer cells due to altered expression of glycosyltransferases. It is currently unclear how the glycosylation patterns of CD22 are altered in B-cell–derived malignancies and autoimmune diseases. Future studies will be required to evaluate whether modifications of the CD22 N231 glycan site on dysfunctional B cells significantly impact on the ability of epratuzumab to engage its protein epitope. Whether the nexus of N-linked glycans atop CD22 nanoclusters offers a yet underappreciated biomarker for dysregulated B cells to be exploited in diagnosis or therapeutic specificities also remains to be determined.
Methods

**CD22 ECD construct design.** Full-length human CD22 ECD (UniprotKB P20273, residues 20–687) was codon-optimized for expression in human cells and synthesized by GeneArt (Life Technologies) (Supplementary Table 4). The construct was subcloned into the pHLsec vector\(^a\) using restriction enzymes AgeI and KpnI, such that a His\(_6\) tag was at the C terminus of the construct to facilitate affinity purification. Truncated constructs 20–330, 20–504, and 20–592 were PCR amplified and cloned into pHLsec as described for the full-length ECD (Supplementary Table 5).

**Expression and purification of CD22.** CD22 constructs were transiently transfected into HEK293F (Thermo Fisher Scientific) or HEK293 Gnt I (HEK293S) (ATCC CRL-3022) suspension cells to produce CD22 glycoforms displaying either glycans with mature carbohydrates (HEK293F), or of high mannose type (HEK293S). Cells were split in 200 ml cultures at 0.8 \(\times\) 10⁶ cells per ml. About 50 μg of DNA was filtered and mixed in a 1:1 ratio with transfection reagent FectoPRO (Polyplus Transfactions), and incubated at room temperature for 10 min. The DNA:FectoPRO solution was then added directly to the cells, and cells were incubated at 37 °C, 125 rpm, 8% CO₂ in a Multitron Pro shaker (Infors HT) for 6–7 days.

Cells were harvested by centrifugation at 6371×\(g\) for 20 min, and supernatants were retained and filtered using a 0.22 \(\mu\)m filter (EMD Millipore). Supernatants were passed through a HisTrap Ni-NTA column (GE Healthcare) using an AKTA Start chromatography system (GE Healthcare) eluted with 100 mM NaCl, pH 8.0. Eluted fractions were immediately neutralized with 1 M Tris-HCl pH 9.0. Fractions containing protein were pooled and run through a desalting column to change the sample buffer into 20 mM sodium acetate, pH 5.6. Ion exchange chromatography was performed using a MonoS column (GE Healthcare) and eluted with a potassium chloride gradient. Fractions were pooled and eluted on a Superdex 200 Increase gel filtration column (GE Healthcare) to obtain purified samples. Peaks were pooled for crystallization trials and binding studies.

**Cryocrystallization and X-ray data collection.** Purified CD22\(_{20–330}\)A protein was concentrated to 10 mg ml\(^{-1}\) in a buffer containing 20 mM Tris pH 9.0 and 150 mM NaCl. Crystals were obtained by sitting drop vapor diffusion in 30% PEG 4000, 0.2 M lithium chloride, and 0.1 M Tris pH 8.5 in 96-well plates after mixing 0.1 and 0.1 M protein and solution according to an Orius crystallization robot (Douglas Instruments). Crystals were cryo-protected by soaking them in mother liquor solution containing 20% glycerol and flash cooled in liquid nitrogen. X-ray diffraction data were collected at the 08ID and 08BM synchrotron beamlines at the Canadian Light Source (CLS). Initial attempts to solve the structure of CD22\(_{20–330}\)A by molecular replacement did not yield any solution using the V- and C2-type Ig domains of homologous Siglec family members as search models.\(^b, c, d\) To acquire phasing information, we soaked native crystals with 7 mM of mercuric chloride (Analar) for 30 min. At CLS beamline 08BM, a fluorescence scan was performed near the Hg L-III absorption edge (energy range 11,560–11,570 keV) to determine the appropriate wavelengths for collection of multi-wavelength anomalous dispersion datasets. The three wavelengths selected for multi-wavelength anomalous dispersion were 1.0051 Å at the absorption peak, 1.0083 Å at the inflection point, and 1.0033 Å at remote wavelength. Full multi-wavelength anomalous dispersion datasets were collected on a single crystal (Table 1).

Data for CD22\(_{20–330}\)A derived with mercury were processed using XDS\(^a\). Based on the C2 space group, Matthews volume calculations\(^b\) and predicted binding sites of mercury, we estimated one molecule in the asymmetric unit, and expected one anomalous scatterer. Initial phases obtained using AutoSolve\(^c\) were useful for automatic building of the structure by AutoBuild\(^d\). The heavy atom that allowed phased by multi-wavelength anomalous dispersion was bound to C308 (an unpaired CD22 cysteine) in d\(_3\). Iterative rounds of manual model building in Coot\(^e\) and refinement with Phenix\(^f\) followed, with statistics reported in Table 1. Representative electron density is shown in Supplementary Fig. 5a.

Data for CD22\(_{20–330}\)A were collected at 1.2 Å resolution. The crystal structure of the complex CD22\(_{20–330}\)A, with n2-6 sialyllactose was solved by molecular replacement using CD22\(_{20–330}\)A, as a search model in Phaser\(^g\). Representative electron density is shown in Supplementary Fig. 8b. To determine the structure of unliganded epratuzumab Fab, purified sample was concentrated and crystals were obtained at 7 mg ml\(^{-1}\) by sitting drop diffusion.
Epratuzumab Fab crystallized in a condition containing 85 mM Tris, pH 8.5, 25.5% PEG 4000, 170 mM sodium acetate, and 15% glycerol at 20 °C. Crystals were flash frozen in liquid nitrogen. X-ray diffraction data were collected at the 08ID beamline atCLS and processed in space group P1 using XDS. The structure was solved by molecular replacement using Phaser58 with unliganded epratuzumab Fab and diffraction data were collected using synchrotron radiation at the 08ID beamline at APS. The complex was concentrated to 5 mg ml\(^{-1}\) and crystals were grown by vapor diffusion at 20 °C in ambiguous scoring of 0.30 and 0, respectively, and deeming the ab initio reconstruction of the molecular envelope of CD22+ was calculated using cryoSPARC31. 2D projections were calculated using script genproj_evenlyspaced.i90 (https://sites.google.com/site/rubinsteingroup/). Fitting of the CD22 D1–D3 crystal structure in the CD22EM volume using UCSF Chimera32 was unambiguous.

SAXS data collection and processing. Samples of purified CD22\(^{20-687}\) were concentrated to 1.25, 2.5, or 5.0 mg ml\(^{-1}\) for analysis by SAXS at the Argonne Advanced Photon Source SAXS beamlines 12-ID-D and 18-ID-D. For each concentration in the refolding and presence of CD22+, 15 molar excess of 2-6 sialyllactose, 15 exposures of 2 s each were collected and the scattering curves were calculated by subtracting the contribution from the buffer. Analysis of the scattering curves using PRIMUS34 showed no signs of aggregation, long-range interactions, or radiation damage, but the 1.25 mg ml\(^{-1}\) concentration had weak signal (as seen in the Kratky plot [I(q)q\(^2\) vs q, Supplementary Fig. 5b). Association between radius of gyration (Rg) and (f(r)) values determined from the Guinier plot and the pair distribution function (P(r)) further confirmed the good quality of the data. As such, determination of the Dmax overall shape, excluded volume, and molar mass were performed with high confidence via analysis of the P(r) function, Kratky plots, and Porod invariant. The highest quality scattering curve (5.0 mg ml\(^{-1}\)) as determined by the AutoRg and AutoGnom functions in PRIMUS34, was selected for ab initio modeling using scattering data up to q = 0.2 Å\(^{-1}\). Since scattering curves are inherently ambiguous, we used the program Ambit35 to assess whether ab initio modeling would yield reliable representations of the CD22+ (α-2,6) sialyltransferase structures in solution. We found that a single shape category (cylinder) was compatible with the scattering curves of CD22\(^{20-687}\) and CD22\(^{20-687}\) + α-2,6 sialyllactose, resulting in ambiguity scores of 0.30 and 0, respectively, and deeming the ab initio reconstructions as potentially unique. Ten ab initio models were then independently generated using the program DAMMIF35 with imposing symmetry. Ab initio models were aligned and averaged by using programs DAMAVER66 and DAMFILT66 to yield a final low-resolution model. Given the quality of the z2-values (Supplementary Table 3) for all DAMMIF models and the fact that the scattering curves could only be modeled with one shape category, we concluded that the CD22\(^{20-687}\) + α-2,6 sialyllactose samples have limited flexibility in solution and, therefore, the ab initio models are a good representation of their structures in solution.

Isothermal titration calorimetry. ITC experiments were performed using an AutoITC200 (Malvern Instruments). ITC measurements of CD22\(^20-330\) WT and CD22\(^{20-330}\) (2,6) sialyllactose were collected using 70–110 μM of CD22+ in the cell and 0.65–1.18 mM of α-2,6 sialylactose (Sigma-Aldrich) in the syringe. A total of 16 injections were performed with an injection volume of 2.5 μl, injection duration of 5 s, and 180 s spacing between injections. The cell temperature was set to 25 °C, with a stirring speed of 750 rpm, for a total period of 5 s. All experiments were repeated at least in duplicates, and values were averaged and standard errors were calculated, based on a near 1:1 restricted binding stoichiometry. For Fab binding to CD22+, 5 μM of CD22+ was placed in the cell and 50 μM Fab was present in the syringe. A total of 16 injections were performed with an injection volume of 2.5 μl, injection duration of 5 s, and 180 s spacing between injections. The cell temperature was set to 35 °C, with a stirring speed of 750 rpm and a filter period of 5 s. All experiments were repeated in triplicate, and values were averaged and standard errors were calculated.

Bilayer interferometry. The binding affinities of epratuzumab and pinatuzumab Fabs to CD22 were measured by BLI using the Octet RED96 BLI system (Pall ForteBio). Ni-NTA biosensors were hydrated in 1x kinetic buffer (1x PBS, pH 7.4, 0.002% Tween, 0.01% BSA) and loaded with 25 ng μl\(^{-1}\) CD22\(^{20-687}\), CD22\(^{20-687}\), CD22\(^{20-687}\), or CD22\(^{20-687}\)–glucan mutants) for 60 s at 1000 rpm. Biosensors were then transferred into wells containing 1x kinetic buffer to baseline for 60 s before being transferred into wells containing a serial dilution of Fabs starting at 500 nM and decreasing to 6.25 nM. The 180 s association phase was subsequently followed by a 180 s dissociation step in 1x kinetic. Measurements were performed using the Octet software, with a 1:1 fit model. All experiments were repeated in triplicate, values were averaged, and standard errors were calculated.

Data availability. The crystal structures, EM reconstructions, and SAXS envelopes reported in this manuscript have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID: 5VKJ, 5VKK, 5VM, and 5VL3), the Electron Microscopy Data Bank, https://www.ebi.ac.uk/pdbe/emdb/ (EMDB ID: EMD-8704 and EMD-8705), and the Small Angle Scattering Biological Data Bank, www.sasdb.org (SASDB ID: SASDC76 and SASDC86). Other data are available from the corresponding author upon reasonable request.

Received: 8 May 2017 Accepted: 28 July 2017

References
1. Macauley, M. S., Crocker, P. R. & Paulson, J. C. Siglec-mediated regulation of immune cell function in disease. Nat. Rev. Immunol. 14, 653–666 (2014).
2. Pillai, S., Netravali, I. A., Cariappa, A. & Mattoo, H. Siglecs and immune regulation. Annu. Rev. Immunol. 30, 357–392 (2012).
3. Walker, J. A. & Smith, K. G. C. CD22: an inhibitory enigma. Nat. Rev. Immunol. 14, 701–719 (2014).
4. O’Keefe, T. L., Williams, G. T., Davies, S. L. & Neuberger, M. S. Hyperresponsive B cells in CD22-deficient mice. Science 274, 798–801 (1996).
12. Nitschke, L. The role of CD22 and other inhibitory co-receptors in B-cell activation. *Curr. Opin. Immunol.* 17, 290–297 (2005).
13. Green, S. F. & Stavrinou, P. L. B. CD45 regulates tyrosine phosphorylation of CD22 and its association with the protein tyrosine phosphatase SHP-1. *J. Immunol.* 162, 5278–5286 (1999).
14. Coughlin, S. et al. An extracellular function of CD45 in B cells is mediated by CD22. *Proc. Natl Acad. Sci. USA* 112, E6513–E6524 (2015).
15. Jellusova, J. & Nitschke, L. Regulation of B cell functions by the sialic acid-binding receptors Siglec-G and CD22. *Front. Immunol.* 2, 96 (2012).
16. Macauley, M. S. et al. Antigenic liposomes displaying CD22 ligands induce antigen-specific B cell apoptosis. *J. Clin. Invest.* 123, 3074–3083 (2013).
17. Rillahan, C. D. et al. Disubstituted sialic acid ligands targeting Siglecs CD33 and CD22 associated with myeloid leukemias and B cell lymphomas. *Chem. Sci.* 5, 2398–2406 (2014).
18. Mesch, S. et al. From a library of MAG antagonists to nanomolar CD22 ligands. *ChemMedChem* 7, 134–143 (2012).
19. Sullivan-Chang, L., O’Neill, R. T. & Tuscano, J. M. Targeting CD22 in B-cell malignancies: current status and clinical outlook. *BioDrugs* 27, 293–304 (2013).
20. O’Reilly, M. K., Tian, H. & Paulson, J. C. CD22 is a recycling receptor that can shuttle cargo between the cell surface and endosomal compartments of B cells. *J. Immunol.* 186, 1554–1563 (2011).
21. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* 10, 317–327 (2010).
22. Nitschke, L., Carnessi, R., Ocker, B., Köhler, G. & Lamers, M. C. D. CD22 is a negative regulator of B-cell receptor signalling. *Curr. Biol.* 7, 133–143 (1997).
23. Pronker, M. F. et al. Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nat. Commun.* 7, 1–13 (2016).
24. Zhuravleva, M. A., Trandem, K. & Sun, P. D. Structural implications of Siglec–CD22 engagement with epratuzumab on B cell function: implications for the treatment of systemic lupus erythematosus. *Autoimmun. Rev.* 14, 1079–1086 (2015).
25. Brosi, E. A. et al. Tropocytosis of multiple B-cell surface markers by CD22 antagonizing with epratuzumab. *Blood* 122, 3020–3029 (2013).
26. Maxfield, F. & McGraw, T. Endocytic recycling. *Nat. Rev. Mol. Cell Biol.* 5, 121–132 (2004).
27. Jardine, J. G. et al. Minimally mutated HIV-1 broadly neutralizing antibodies to guide reductionist vaccine design. *PLoS Pathog.* 12, e1005815 (2016).
28. Wang, P. Altered glycosylation in cancer: sialic acids and sialyltransferases. *Cancer Biol. Med.* 1, 73–81 (2015).
29. Ohtsubo, K. & Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* 126, 855–867 (2006).
30. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1243–1250 (2006).
31. Pröpster, J. M. et al. Structural basis for sialuron-dependent self-lycan recognition by the human immune-inhibitory receptor Siglec-8. *Proc. Natl Acad. Sci. USA* 113, E4170–E4179 (2016).
32. Kabsch, W. *Acta Crystallogr. D Biol. Crystallogr.* 65, 124–130 (2012).
33. Weichenberger, C. X. & Rupp, B. Ten years of probabilistic estimates of biocrystal solvent content: new insights via nonparametric kernel density estimate. *Acta Crystallogr. B Biol. Crystallogr.* 70, 1579–1588 (2014).
34. Moyo, W. T. M. et al. Automated protein–ligand crystallography for structure-based drug design. *ChemMedChem* 1, 827–838 (2006).
35. Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuilder. *Acta Crystallogr. D Biol. Crystallogr.* 64, 61–69 (2018).
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. B Biol. Crystallogr.* 67, 9280 (1996).
37. Morin, A. et al. Collaboration gets the most out of software. *Elife* 2, e01456 (2013).
38. Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180, 519–530 (2012).
39. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1650–1657 (2004).
40. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36, 1277–1282 (2003).
41. Petoukhov, M. V. & Svergun, D. I. Ambiguity assessment of small-angle scattering curves from monodisperse systems. *Acta Crystallogr. D Biol. Crystallogr.* 71, 1051–1058 (2015).
42. Volkov, V. V. & Svergun, D. I. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* 36, 860–864 (2003).
43. May, A. P., Robinson, R. C., Vinson, M., Crocker, P. R. & Jones, E. Y. Crystal structure of the N-terminal domain of sialoadhesin in complex with 3′-sialyllactose at 1.85 Å resolution. *Mol. Cell* 1, 719–728 (1998).
44. Saphire, E. O. et al. Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. *Science* 293, 1155–1159 (2001).
45. Raadev, S. et al. Structural and functional studies of Igαβ and its assembly with the B cell antigen receptor. *Structure* 18, 934–943 (2010).
46. Croll, T. L. et al. Higher-resolution structure of the human insulin receptor ectodomain: multi-modal inclusion of the insert domain. *Structure* 24, 469–476 (2016).

**Acknowledgements**

We thank X. Zuo and W. Shang for help with SAXS data collection and interpretation, A. Bosch for help with figure preparation, and J. Paulson, B. Treanor, and I. Wilson for discussions and critical comments on the manuscript. X-ray diffraction experiments described in this paper were performed using beamlines 08ID and 08BM at the Canadian Synchrotron Radiation Centre. This research was supported by the National Institute of Allergy and Infectious Diseases (AI090528, AI083623), National Science Foundation (DMS-1312478), and National Institute of General Medical Sciences (GM103461).
Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research. SAXS experiments were performed at beamlines 12-ID and 18-ID of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We would like to acknowledge the Structural & Biophysical Core Facility, The Hospital for Sick Children, for access to the ITC and Octet RED96 BLI instruments. J.E.-O. was supported by Banting Postdoctoral Fellowship BPF-144483 from the Canadian Institutes of Health Research. T.S. is a recipient of a Canada Graduate Scholarship Master’s Award from the Canadian Institutes of Health Research. This work was supported by operating grant PJT-148811 (J.-P.J.) from the Canadian Institutes of Health Research.

**Author contributions**
The following outlines author contributions. Experimental conception and design: J.E.-O., T.S., J.-P.J.; data acquisition: J.E.-O., T.S., H.C.; analysis of data: J.E.-O., T.S., M.T.M.-J., S.B., A.G., J.L.R., J.-P.J.; drafting the article or revising it critically for important intellectual content: J.E.-O., T.S., J.-P.J.

**Additional information**
Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00836-6.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at http://npg.nature.com/reprintsandpermissions/

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.