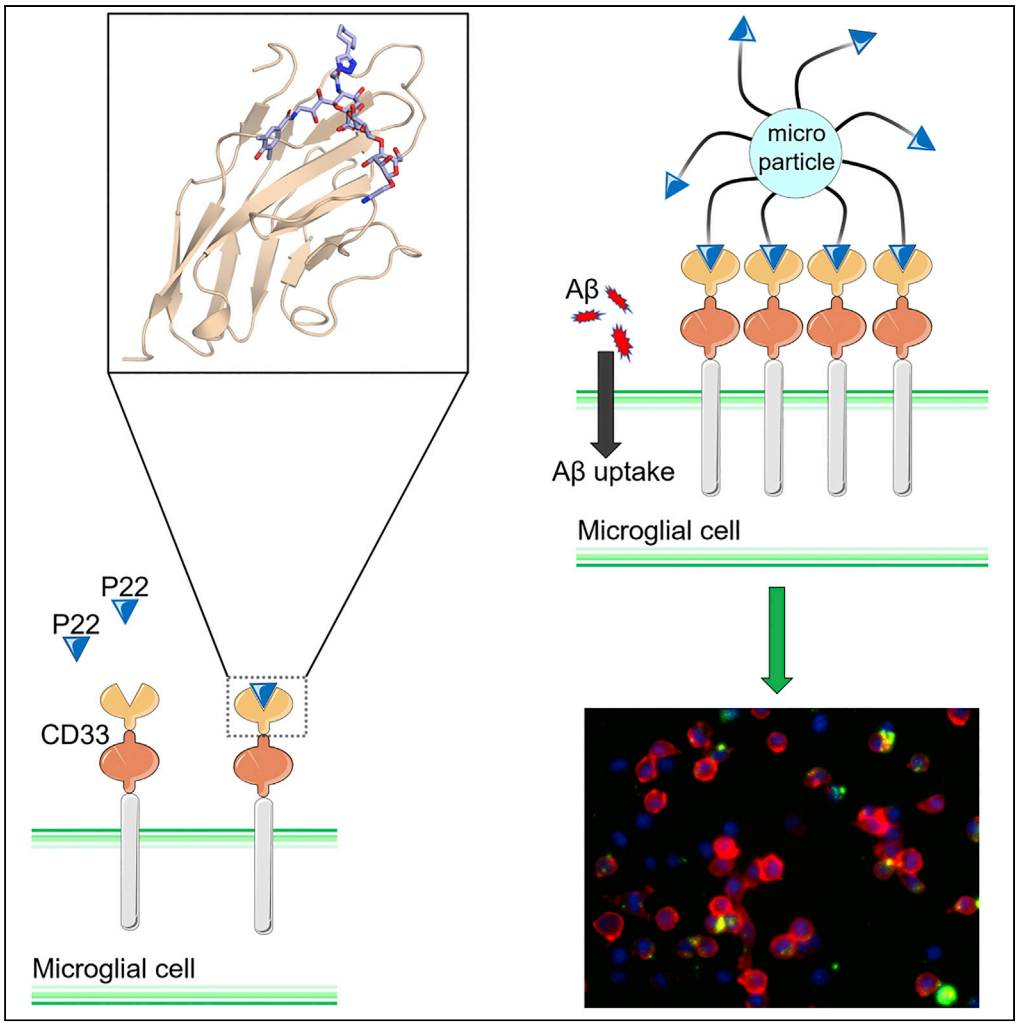
iScience

CellPress

Article

Small Molecule Binding to Alzheimer Risk Factor CD33 Promotes Aβ Phagocytosis



Luke A. Miles, Stefan J. Hermans, Gabriela A.N. Crespi, ..., Ines Royaux, Nancy C. Hancock, Michael W. Parker

lmiles@svi.edu.au (L.A.M.)
mparker@svi.edu.au (M.W.P.)

HIGHLIGHTS

A sialic acid mimetic increases the uptake of Alzheimer peptide into microglia

Crystal structure of cell surface receptor CD33 bound to the sialic acid mimetic

Crystal structure of unliganded CD33 and mutagenesis studies revealed key residues

CD33 carbohydrate-binding site is a pharmacophore for Alzheimer drug development

Miles et al., iScience *19*, 110–118
September 27, 2019
https://doi.org/10.1016/j.isci.2019.07.023

iScience

CellPress

Article

Small Molecule Binding to Alzheimer Risk Factor CD33 Promotes Aβ Phagocytosis

Luke A. Miles,[1,2,*] Stefan J. Hermans,[1,2] Gabriela A.N. Crespi,[1,2] Jonathan H. Gooi,[1,2] Larissa Doughty,[1,2] Tracy L. Nero,[1,2] Jasmina Markulić,[1,2] Andreas Ebneth,[3] Berthold Wroblowski,[3] Daniel Oehlrich,[3] Andrés A. Trabanco,[4] Marie-Laure Rives,[5] Ines Royaux,[3] Nancy C. Hancock,[1,2] and Michael W. Parker[1,2,6,*]

SUMMARY

Polymorphism in the microglial receptor CD33 gene has been linked to late-onset Alzheimer disease (AD), and reduced expression of the CD33 sialic acid-binding domain confers protection. Thus, CD33 inhibition might be an effective therapy against disease progression. Progress toward discovery of selective CD33 inhibitors has been hampered by the absence of an atomic resolution structure. We report here the crystal structures of CD33 alone and bound to a subtype-selective sialic acid mimetic called P22 and use them to identify key binding residues by site-directed mutagenesis and binding assays to reveal the molecular basis for its selectivity toward sialylated glycoproteins and glycolipids. We show that P22, when presented on microparticles, increases uptake of the toxic AD peptide, amyloid-β (Aβ), into microglial cells. Thus, the sialic acid-binding site on CD33 is a promising pharmacophore for developing therapeutics that promote clearance of the Aβ peptide that is thought to cause AD.

INTRODUCTION

Microglia, resident immune cells in the brain, can be activated in response to misfolded proteins found in neurodegenerative diseases leading to neuroinflammation and the release of neurotoxic substances. There is emerging evidence that microglial activation may play an important role in a range of neurodegenerative diseases including Huntington disease, Parkinson disease, amyotrophic lateral sclerosis, frontotemporal dementia, and dementia with Lewy bodies (Heneka et al., 2014; Leyns and Holtzman, 2017; Hickman et al., 2018). In a cell-based model of Alzheimer disease (AD) it has been shown that activated microglia can induce neurite degeneration and cell death, a pathological feature of the disease (Park et al., 2018).

Many of the recently identified genes associated with late-onset AD risk are integral to the innate immune system (Sims et al., 2017; Efthymiou and Goate, 2017). Some of these genes code for microglial proteins, such as the strongest genetic risk factor for AD, namely APOE, and the cell surface receptor CD33 (Krasemann et al., 2017; Pimenova et al., 2017). Genome-wide association studies of late-onset AD demonstrate a link between the CD33 gene and disease susceptibility (Hollingworth et al., 2011; Naj et al., 2011). Tanzi and co-workers showed that knocking out the CD33 gene could mitigate amyloid-β (Aβ42) pathology (Griciuc et al., 2013). They also reported that numbers of CD33-immunoreactive microglia correlated positively with both insoluble Aβ42 levels and amyloid plaque burden in AD brain, that CD33 inhibited clearance of Aβ42 in microglial cell cultures, and that brain levels of insoluble Aβ42/plaque burden were markedly reduced in APPSwe/PS1ΔE9/CD33−/− mice. They also showed that the SNP rs386544, which confers protection from AD, lowered insoluble levels of Aβ42 in the AD brain. The rs386544 allele is in perfect linkage disequilibrium with rs12459419, which is located at a splice site of the sialic acid-binding site containing exon 2 (Raj et al., 2014). Both alleles ultimately lead to lower expression levels of functional CD33 (Malik et al., 2013; Walker et al., 2015), enhanced phagocytic activity of microglial cells, and uptake of Aβ42 (Griciuc et al., 2013; Bradshaw et al., 2013). The widely accepted "amyloid" hypothesis for AD posits that increased production and oligomerization of the Aβ42 peptide initiates a cascade of events leading to neurodegeneration and AD. Thus, small, drug-like inhibitors of CD33 to promote amyloid clearance could represent a novel class of therapeutics for the prevention and treatment of AD.

[1]ACRF Rational Drug Discovery Centre, St. Vincent's Institute of Medical Research, Fitzroy, VIC 3056, Australia

[2]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, VIC 3010, Australia

[3]Janssen Research & Development, a Division of Janssen Pharmaceutica N.V, 2340 Beerse, Belgium

[4]Neuroscience Medicinal Chemistry, Janssen Research & Development, 45007 Toledo, Spain

[5]Molecular and Cellular Pharmacology, Janssen Research & Development, LLC, La Jolla, CA 92121, USA

[6]Lead Contact

*Correspondence: lmiles@svi.edu.au (L.A.M.), mparker@svi.edu.au (M.W.P.)

https://doi.org/10.1016/j.isci.2019.07.023

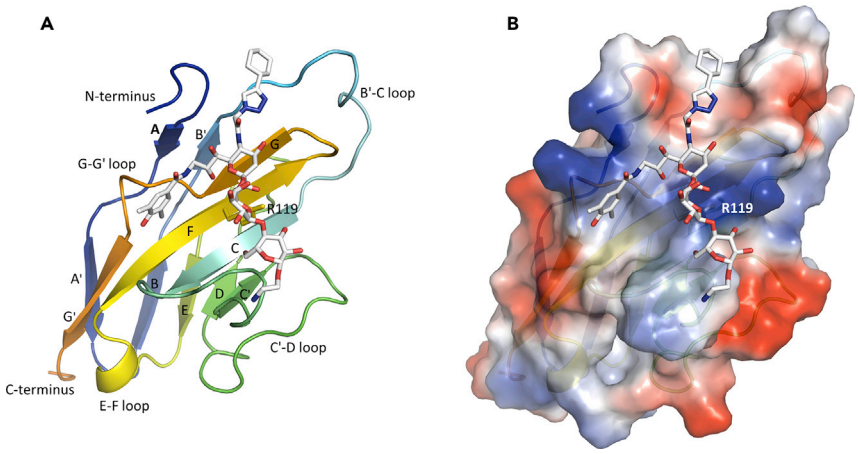


Figure 1. Structure of the P22-CD33 Complex

(A) Cartoon representation of the CD33 N-terminal V-set domain in rainbow coloring from blue (N terminus) to orange (C terminus). The critical sialic acid-binding residue R119 (colored orange) and compound P22 (gray) are shown as sticks.

(B) The electrostatic potential energy mapped onto the molecular surface of CD33 is shown, in the same orientation as (A), with electropositive regions colored blue and electronegative regions red. The location of R119 is indicated.

See also Figure S3 and Table S1.

CD33 is a pattern recognition receptor belonging to the Siglec (sialic acid-binding Ig-like lectins) receptor family. It is a type-1 membrane protein with an extracellular region consisting of an N-terminal V-set domain that recognizes sialylated ligands and an Ig juxtamembrane (C2-set) domain (Griciuc et al., 2013; Rillahan et al., 2014). Although CD33 recognizes sialylated glycoproteins and gangliosides that extensively coat amyloid plaques, its natural ligand has not been identified and up until recently known ligands had only millimolar affinity. Paulson and co-workers recently identified sialic acid-based ligands with high selectivity and micromolar affinity for human CD33 (Rillahan et al., 2014). To provide a molecular basis for understanding ligand specificity and facilitate the design of drug-like inhibitors of CD33 activity, we have solved high-resolution (1.8-Å) crystal structures of the human CD33 V-set domain, with and without a CD33-selective 2,5,9-trisubstituted sialic acid mimetic called P22 (Figure S1). Importantly, we show that P22 binding to the sialic acid-binding pocket of CD33 can increase uptake of the toxic Aβ42 peptide into microglial cells, and hence CD33 is a promising target for structure-based drug discovery for the treatment of AD.

RESULTS AND DISCUSSION

Crystal Structure of the P22-CD33 Complex

We collected separate X-ray diffraction data from two P22-CD33 complex crystals, which we refer to as Crystal no. 1 (corresponding to P22 binding mode 1) and Crystal no. 2 (corresponding to P22 binding mode 2) in Table S1 and solved their structures. In each crystal, two complexes were found in the asymmetric unit, designated by their chain identifiers A and B, and they overlay very closely with root-mean-square deviation (RMSD) over Cα atoms of 0.27 and 0.29 Å for the two structures (Figure S2).

The crystal structure of P22 bound to the CD33 V-set domain is shown in Figure 1 (see Transparent Methods). The domain forms a typical Siglec V-set domain fold, rich in β-sheet and composed of 11 β-strands (A, A′, B, B′, C, C′, D, E, F, G, and G′). There is an intra-domain disulfide bond between C41 and C101. Comparison with the V-set domains of other human Siglec subtypes available in the Protein DataBank (Figure S3) reveals very similar folds with the B′-C, C-C′, C′-D, and G-G′ loops being the most distinguishing features, exhibiting little or no sequence similarity with the other human Siglecs.

P22 exhibits high specificity for human CD33 over other human Siglecs (Rillahan et al., 2014). We observe two P22 binding modes, with one showing more interactions and straddling the length of one face of CD33 (Figure 1) burying a surface area of 509 $Å^2$. Although the space group was the same and the cell dimensions were very similar between the two crystal forms, the location of the C2-substituent in the CD33 carbohydrate-binding pocket is different. This substituent is highly mobile and shows little

or no engagement with CD33 directly (see Figure S4). In binding mode 1 the surface buried by P22 is 509 Å^2, with the C2-substituent within hydrogen bonding distance of the backbone carbonyl and side-chain hydroxyl of S68. The B-factor of the C2-substituent is over 55 Å^2, whereas the B-factors of the C5- and C9-substituents, and contacting side-chains, are less than 35 Å^2. In binding mode 2 the buried surface area is 370 Å^2, with the C2-substituent making no contact with the protein, but it does form interactions with another CD33 molecule in the crystal lattice. The B-factors of the C2-substituent are over 45 Å^2, whereas the B-factors of the C5- and C9-substituents are less than 30 Å^2 and those of the contacting side-chains are less than 25 Å^2. We chose to perform all our subsequent structural analyses on the binding mode 1 structure (Crystal no. 1).

The P22 molecules themselves within each crystal form superimpose very closely (RMSD over all heavy atoms of 0.32 Å in Crystal form no. 1 and 0.46 Å in Crystal form no. 2). Despite the overall similarity of the two complexes there are a few subtle differences. Several amino acid side-chains are surface exposed and exhibit flexibility between molecules, such as K52, R98, H137, and K130. Several residues are packed against a molecule from an adjacent asymmetric unit resulting in side-chain movement to accommodate the close packing; these are W22, L78, Q81, and R91. In the case of R91, side-chain movement facilitates the formation of additional hydrogen bonds to E27 of another asymmetric unit. Aside from these minor differences, the four molecules of CD33 bound to P22 across our two reported structures are essentially identical.

Interactions between P22 and CD33 observed across all structures reported here include one salt bridge, three hydrogens bonds, and numerous van der Waals interactions with four surrounding residues (Table S2). The central sialic acid moiety is engaged by R119 in a salt bridge (Figure 2C). R119 is a highly conserved residue in Siglecs, located in a highly basic pocket in CD33 (Figures 1B and S3D). Other P22 polar interactions include hydrogen bonds contributed by the main-chain atoms of K126 and S128. A notable van der Waals packing interaction occurs between the C5- and C9-substituents of P22 and an aromatic knob consisting of F21, H45, and Y127 (Figure 2B). The C9-substituent of P22 extends away from the sialic acid-binding site along the G-strand, past the G-G′ loop (Figure 1A). Interestingly, the sequence in this region (TKYSYKSPQ) shows no similarity with other human Siglecs (Figure S3D), suggesting a molecular basis for the affinity and selectivity of CD33 for P22 (Rillahan et al., 2014).

P22 Binds CD33 with Micromolar Affinity

Paulson and co-workers estimated an IC_{50} for P22 of 11 μM using a flow cytometry assay (Rillahan et al., 2014). We next performed surface plasmon resonance (SPR) measurements on P22 binding to CD33 wild-type, and selected mutants, to obtain detailed kinetic data (see Transparent Methods). The CD33 V-set domain has a free cysteine residue (C36), which caused anomalous results in some of our SPR experiments due to disulfide-linked oligomers forming in solution. Hence, we mutated this cysteine to serine for these experiments. The K_D for P22 binding to captured CD33^{C36S} was determined by SPR as 95 ± 39 μM (close to the wild-type value of 118 ± 41 μM) (Figure 3 and Table S3). With biotinylated P22 (Figure 3C) captured on the chip, analysis of CD33^{C36S} binding to P22 yielded a K_D of 67 ± 15 μM (Table S3). In either orientation, no binding was observed for CD33 when the key R119 was mutated to alanine.

Crystal Structure of Unliganded CD33

We determined the crystal structure of the unliganded CD33 V-set domain to answer the question whether CD33 recognizes carbohydrate ligands using an induced fit mechanism. Our crystals of unliganded CD33 contain four copies of the molecule in the asymmetric unit, chains A to D. They are assembled as non-biological dimers, AB and CD, with a free cysteine, C36, forming a disulfide bond between monomers. This cysteine would normally form an inter-domain disulfide with C169 of the C2 domain in the intact receptor. Each CD33 molecule superimposed closely with each other (maximum RMSD of 0.39 Å over 115 Cα atoms) with the largest deviations resulting from side-chain movements to accommodate the packing of the two non-biological dimers in each asymmetric unit. Residues W22, Q24, and F44 show equivalent side-chain positions in chains A and B, with an alternative position in chains C and D, to facilitate packing of the asymmetric unit. Additional differences between each CD33 molecule of the asymmetric unit are observed in surface-exposed and flexible amino acids, such as E27, K52, R69, L78, R98, R111, and K130. As the unliganded structures are so similar all further comparisons in the text focus on chain A as the reference unliganded CD33 molecule. Each of the four unliganded structures superimpose closely on the P22 crystal structures (maximum RMSD of 0.64 Å over 116 Cα atoms) (Figure S2C).

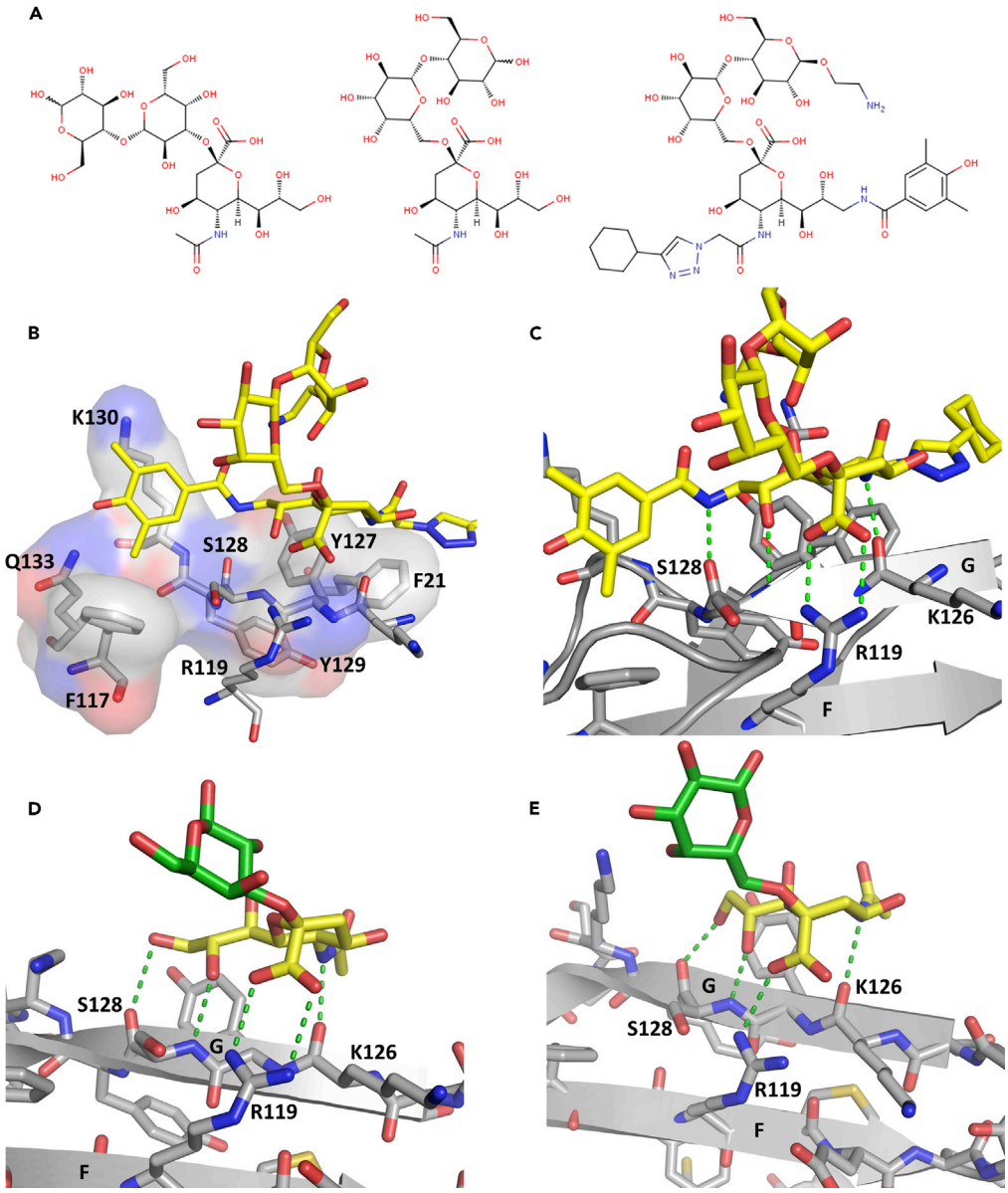


Figure 2. P22 Interactions with CD33

(A) The chemical structures of 3′-sialyllactose, 6′-sialyllactose, and P22 from left to right, respectively. The sialic acid moiety of the three ligands is shown in the same view.

(B) Close-up view of the CD33 residues (gray sticks) directly interacting with P22 (yellow sticks). Residues making hydrophobic interactions with P22 are overlayed with a semi-transparent molecular surface.

(C) The extended polar contact network (green dashes) P22 makes with the F- and G-strands of CD33.

(D and E) (D) 3′-sialyllactose and (E) 6′-sialyllactose binding to the key CD33 R119 residue. The sialic acid moiety of the ligand is shown in yellow and the galactopyranosyl ring in green. Polar contacts between the ligand and CD33 are indicated by the green dashes. In both the 3′- and 6′-sialyllactose complexes with CD33, the glucopyranosyl ring is not visible (PDB: 5J06 and 5J0B).

See also Figures S1, S2, and S4 and Table S2.

The unliganded CD33 superimposes closely on the P22 complex crystal structure. The most significant difference between the structures in the P22-binding pocket is in the aromatic cluster consisting of F21, H45, and Y127 (Figure S2). On binding P22, the H45 side chain rotates down about the Cβ carbon to accommodate F21 in the binding site with the latter forming an aromatic π-stack with Y127.

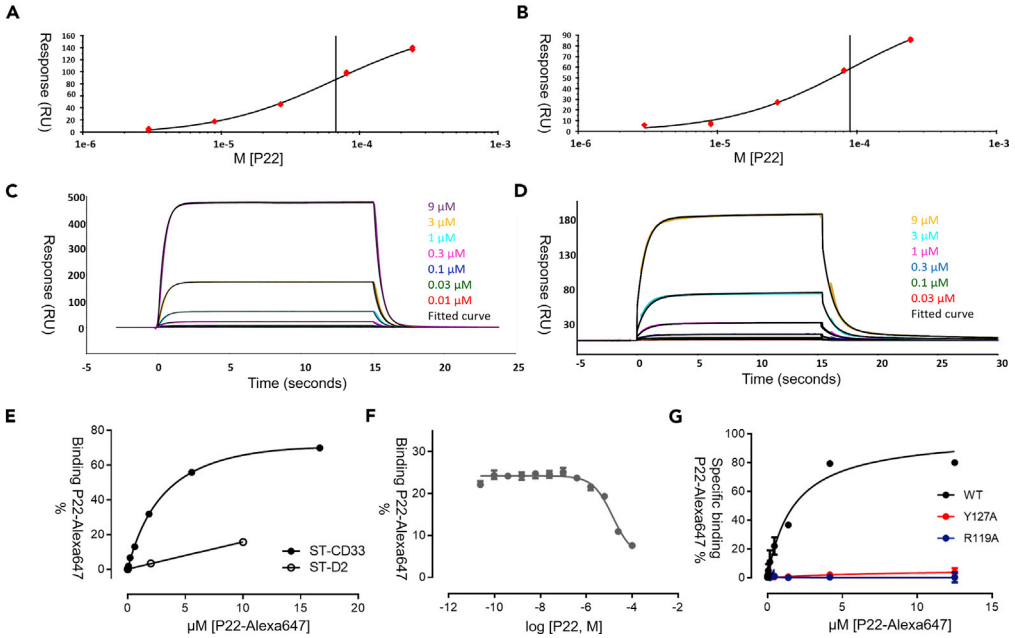


Figure 3. P22 Binding to CD33 Wild-Type and Mutants

(A) An example of affinity analysis for 1–243 µM (3-fold dilutions) of P22 binding to immobilized $CD33^{C36S}$. Data obtained in duplicate using single-cycle analysis and fit to a steady-state binding model to determine the equilibrium binding constant.

(B) An example of affinity analysis of P22 binding to immobilized CD33 wild-type (WT) as per example in (A).

(C) Sensorgram data for 0.01–9 µM (3-fold dilutions) of $CD33^{C36S}$ binding to immobilized P22. Data obtained in triplicate using single concentration cycles and fit to a 1:1 binding model for kinetic analysis.

(D) Sensorgram data for 0.03–9 µM (3-fold dilutions) CD33 WT binding to immobilized P22. Data obtained in triplicate using single concentration cycles and fit to a bivalent analyte model for kinetic analysis. Measurements in triplicate showing mean ± SD.

(E) Saturation total binding experiments on either CD33 (black circles) or the dopamine D2 receptor (open circle) showing that P22-Alexa647 binds specifically to CD33. Data presented are representative of three independent experiments performed in quadruplicate for each compound. Data were normalized on P22-Alexa647 estimated B_{max} on CD33 and are represented as averages ± SD.

(F) Competition binding experiment on CD33 using 2 µM P22-Alexa647. Data presented are representative of three independent experiments performed in quadruplicate for each compound. Data were normalized on P22-Alexa647 estimated B_{max} on CD33 and are represented as averages ± SD.

(G) The specific binding of P22-Alexa647 (total binding – binding in presence of 100 µM non-labeled P22) was evaluated for the WT CD33 and two other single point mutation variants. Binding was completely abolished by the Y127A and R119A variants. Data presented are representative of three independent experiments performed in quadruplicate for each compound. Data were normalized on P22-Alexa647 estimated B_{max} on CD33 and are represented as averages ± SD. See also Tables S3 and S4.

The crystal structures of the entire extracellular region of CD33 bound to two sialic acid derivatives (3′- and 6′-sialyllactoses see Figure 2A) have been deposited in the Protein DataBank (PDB: 5J06, 5J0B) but have not yet been published. The V-set domain of the deposited structures and the ones here superimpose closely (maximum RMSD of 0.41 Å over 112 Cα atoms) (Figure S2). Superposition of the sialic acid-binding sites reveals that CD33 recognizes the sialic acid moieties in very similar ways, with R119 engaging the sialic acid by a salt bridge and main-chain atoms of K126 and S128 forming hydrogen bonds with the carbohydrate in every case. However, the sialyllactoses are much smaller than P22 (Figure 2A) and as a consequence do not form as many interactions with CD33 (Figure 2), consistent with the fact they bind CD33 with only millimolar affinity (Blixt et al., 2003).

Structural Comparison of Siglec Structures

There are 14 other known human Siglecs, ranging in pairwise sequence identity with the CD33 V-set domain from 22% to 60% (Figure S3). The V-set domains from the published crystal structures all adopt the same

fold but show some variation on superposition with CD33: RMSD on Cα atoms of 1.75 Å over 61 residues for Siglec-2 (PDB: 5VKJ) (Ereño-Orbea et al., 2017), 0.85 Å over 113 residues for Siglec-5 (PDB:2ZG2) (Zhuravleva et al., 2008), 0.38 Å over 93 residues for Siglec-7 (PDB: 1O7S) (Dimasi et al., 2004), and 0.73 Å over 106 residues for Siglec-8 (PDB: 2N7A) (Pröpster et al., 2016) (see Figure S3A).

We noted above that there is a movement of aromatic side-chains in the carbohydrate-binding pocket surrounding the C5-substituent when P22 binds to CD33 (and see Figure S2). This side-chain rearrangement is seen in some, but not all, CD33 molecules in the 3′- and 6′-sialyllactose complex structures (PDB: 5J06, 5J0B). Furthermore, it is not observed in unliganded and ligand-bound structures of human Siglec-2 (PDB: 5VKJ, 5VKM) (Ereño-Orbea et al., 2017), Siglec-5 (PDB: 2ZG1, 2ZG2, 2ZG3) (Zhuravleva et al., 2008), Siglec-7 (PDB: 1NKO, 2DF3, 2GFR, 2HRL) (Dimasi et al., 2004; Attrill et al., 2006a, 2006b), or Siglec-8 (PDB: 2N7A, 2N7B) (Pröpster et al., 2016). However, all the ligands in these cases have very short C5-substituents compared with P22. In contrast, side-chain rearrangements are observed in the human Siglec-2, CD33, and Siglec-7 structures, where the C2- (Siglec-2 and 7) and C9- (CD33 and Siglec-7) substituents are reasonably long (Ereño-Orbea et al., 2017; Attrill et al., 2006a, 2006b). Thus, we predict that the side-chain rearrangements we observe in CD33 may only occur with larger carbohydrate and mimetic ligands.

Site-Directed Mutagenesis Reveals Key CD33-Binding Residues

We then used our structural data to design mutations of key residues interacting with P22 to assess which interactions were the most important for binding affinity. The residues were mutated into alanine, and the effect of these mutations on P22 binding affinity was assessed using a SNAP-tag homogeneous time-resolved fluorescence binding assay (see Transparent Methods and Figures 3E–3G). The R119A mutation, which destroys the salt bridge between P22 and CD33, was one of the most detrimental to binding (>30-fold loss) (Figure 2C and Table S4). Y127A also completely abolished P22 binding, which can be explained by its location between the C5- and C9-substituents of P22 (Figure 2B and Table S4). An F21A mutation saw a 6-fold loss in binding affinity, possibly due to a loss of edge-to-face π-π engagement of P22, due to the orientation of Y127 (the tyrosine residue might become more dynamic in the absence of the ring stack), or due to reduced protein stability as the A-strand becomes more flexible. H45 and S131 showed little to no significant impact on P22 binding affinity, suggesting that these residues weakly interact with P22 (Table S4).

P22 Promotes Phagocytosis in a CD33-Dependent Manner

Previous studies have suggested that the CD33 V-set domain is responsible for Aβ uptake (Griciuc et al., 2013; Malik et al., 2013), but it is not known which region of the domain is responsible for this activity. Having shown that P22 binds specifically to the sialic acid-binding site of this domain we now have a chemical tool to answer the question of whether the sialic acid-binding site is the appropriate target for small molecule ligands that might promote Aβ uptake. Siglecs preferentially recognize multivalent presentation of their ligands (Rillahan et al., 2014; Büll et al., 2016), and for this reason we conjugated the biotinylated form of P22 to microparticles labeled with both FITC and streptavidin (P22 microparticles) (see Transparent Methods). More than 60% CD33-transfected BV2 cells were positive for P22 microparticles compared with only 30% cells that were only incubated with control particles (unconjugated to P22) (Figure 4A). Although some P22 microparticles also bind to non-transfected cells, the difference remains significant for the cells transfected with CD33. This led us to believe that the compound bound preferentially to CD33 (Figure 4A). After showing that the P22 microparticles bind to CD33, we studied the effect of P22 on phagocytosis. For this, Aβ42 was labeled with pHrodo and added to CD33-transfected cells that had previously been incubated with P22 and control microparticles. The quantification of Aβ uptake shows that there is an increase of almost 20% in phagocytosis in transfected cells that have been treated with P22 microparticles with respect to those incubated without P22 (Figure 4B). As a control, we tested the core structural unit of P22, i.e., 3′-sialyllactose (Figure 2A), a much weaker binder of CD33 (millimolar affinity) and saw no difference in phagocytosis even when cells were pre-treated with millimolar concentration of the ligand. To further confirm our observations in murine BV2 cells, we examined the ability of P22 to modulate phagocytosis in the human THP-1 cell line in both wild-type and CD33 knockout cells. Differentiated THP-1 cells were either left untreated or were pre-treated with P22 monomers, P22 conjugated to microparticles, or microparticles alone for 30 min before the addition of *E. coli* BioParticles. Treatment with P22 monomers had no effect on the phagocytosis of the *E. coli* BioParticles; however, when P22 was conjugated to microparticles we observed a 35% increase in phagocytosis. Treatment with microparticles alone had no effect. The increase in phagocytosis with P22 conjugated microparticles was not observed in CD33-deficient

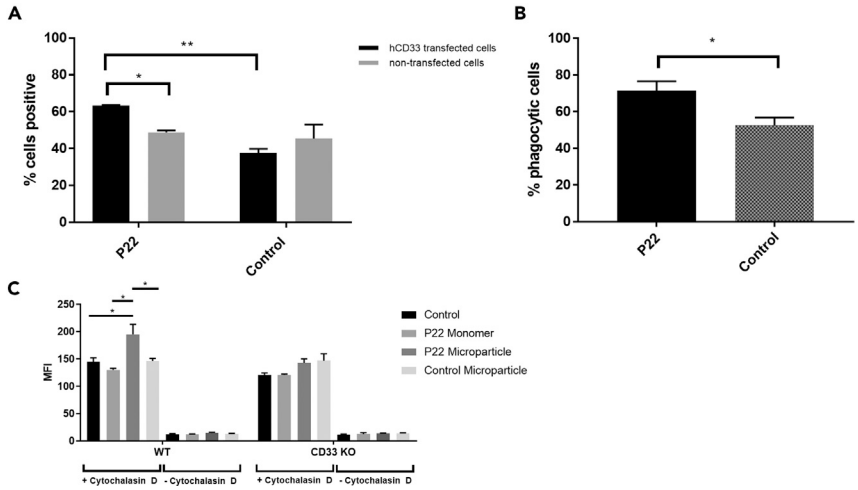


Figure 4. Effect of P22 on Phagocytosis

The cells were treated with P22 conjugated to microparticles labeled with both streptavidin and fluorescein isothiocyanate (FITC), or with microparticles without P22 conjugated as control.

(A) The percentage of positive cells for FITC microparticles was measured by flow cytometry, demonstrating that P22 preferentially binds to human CD33-transfected cells. Data were obtained in triplicate from n = 3 independent experiments and represented as mean ± SEM. *$p < 0.05$, **$p < 0.01$; ANOVA followed by the Tukey test.

(B) Quantification of phagocytic human CD33-transfected cells for Aβ42 pHrodo after being incubated with microparticles reveals an increase of Aβ uptake for cells treated with P22. Data were obtained in triplicate from n = 3 independent experiments and represented as mean ± SEM. *$p < 0.05$; t test.

(C) Quantification of phagocytic differentiated human THP-1 cells after incubation with P22 monomer, P22 conjugated microparticles, or microparticles without P22 conjugation. Differentiated THP-1 cells were either left untreated or were pre-treated with P22 monomers, P22 conjugated to microparticles, or microparticles alone for 30 min before the addition of *E. coli* BioParticles. Treatment with P22 conjugated microparticles resulted in a 35% increase in phagocytosis. This increase was not observed in CD33-deficient cells. Co-treatment with the phagocytosis inhibitor, cytochalasin D, blocked phagocytosis in all treatment groups. Data were obtained in triplicate from n = 3 independent experiments and represented as mean ± SEM. *$p < 0.05$; two-way ANOVA.

THP-1 cells (Figure 4C). Co-treatment with the phagocytosis inhibitor, cytochalasin D, blocked phagocytosis in all treatment groups (Figure 4C). In summary, P22 increases phagocytosis in cells that express CD33.

Concluding Remarks

Structural insights presented in this study reveal the molecular interactions that underpin sialic acid ligand recognition by CD33. We have shown that the sialic acid-based compound P22, when presented on microparticles, can increase phagocytosis of microglial cells and increase Aβ uptake into these cells. A link between CD33-like Siglecs (Siglec-9, Siglec-10, and Siglec-11) and their engagement of sialylated ligands resulting in inhibition of the innate immune response, including phagocytosis, is well established (Carlin et al., 2009; Chen et al., 2009; Wang and Neumann, 2010). The data suggest that these surface receptors can be neuroprotective by dampening immune response and neuronal damage in neuroinflammation. In contrast, the recent genetic data discussed above points to a protective effect of CD33 dysfunction in AD.

The sialic acid-binding site on CD33 is a promising target for developing therapeutics to promote clearance of the toxic Aβ peptide. The results presented here will be of great value in future structure-based drug discovery efforts targeting CD33 in AD. Inhibition of glycoprotein or glycolipid binding by CD33 may also be a useful therapeutic approach for other neurodegenerative diseases. For example, a Parkinson disease risk allele is correlated with increased expression of CD33 (Chan et al., 2016) and CD33 expression is differentially regulated in a murine model of amyotrophic lateral sclerosis (Ferraiuolo et al., 2007).

Limitations of the Study

This study did not investigate the mechanism of action of P22 on CD33-mediated microglial uptake of Aβ ligands. We speculate that the mechanism of action of small molecule CD33 ligands might be to inhibit

engagement of microglia with heavily sialylated ligands on neuritic plaques, thus releasing microglia to clear toxic Aβ species around the plaque. It remains unclear whether P22 microparticles act as an agonist or antagonist of CD33.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.07.023.

ACKNOWLEDGMENTS

This research was partly undertaken on the MX1 and MX2 beamlines at the Australian Synchrotron, ANSTO, Victoria, Australia, and we thank the beamline staff for their assistance. This work on the MX2 beamline made use of the ACRF detector. We thank Susanne Feil for help in early X-ray data collection. This work was supported by grants from the Bethlehem Griffiths Research Foundation, Mason Foundation, Yulgilbar Foundation, the Pratt Family, the Australian Cancer Research Foundation, and the Alzheimer's Drug Discovery Foundation to M.W.P. This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (IMI2 JU) under grant agreement No. 115976. Funding from the Victorian Government Operational Infrastructure Support Scheme to St. Vincent's Institute is acknowledged. J.M. is an Australian Rotary Health/Rotary District 9650 scholar and M.W.P. is a National Health and Medical Research Council of Australia Research Fellow. This research was supported in part by a research collaboration agreement established by Johnson & Johnson Innovation between St. Vincent's Institute and Janssen Pharmaceuticals.

AUTHOR CONTRIBUTIONS

M.W.P. and L.A.M. devised the study, supervised the research component at SVI, and wrote the first draft of the manuscript together. L.A.M. performed the initial bioinformatics for designing domain boundaries and cloning strategies and developed the initial refolding and purification strategies that led to the first unliganded CD33 structure. Refinement of construct design and protein production strategies were developed by J.M. and S.J.H. Protein production was subsequently performed by S.J.H., N.C.H., and G.A.N.C. L.A.M. (initial unliganded structure) and S.J.H. (P22 complex and unliganded structures) performed the crystallographic studies and, together with T.L.N., did the structural analyses. A.A.T. and D.O. synthesized P22 and its biotinylated derivative. M-L.R. developed the SNAP-tag-binding assay and performed the experiments. L.D. designed and conducted the SPR-binding studies and related data analyses. A.E., B.W., D.O., and M-L.R. designed the functional studies and related data analyses. I.R. made the CD33 KO THP-1 cells. G.A.N.C. and J.H.G. performed the phagocytosis assays. L.A.M., L.D., S.J.H., T.L.N., A.E., B.W., D.O., A.A.T., M-L.R., and G.A.N.C. prepared the figures. All authors discussed the results and commented on the manuscript.

DECLARATION OF INTERESTS

A.E., B.W., D.O., A.A.T., M-L.R., and I.R. are employees of Janssen Pharmaceuticals. SVI and Janssen have a research collaboration and license agreement in the area of microglial modulators, and thus all authors have a conflict of interest by virtue of this agreement.

Received: September 8, 2018
Revised: January 16, 2019
Accepted: July 15, 2019
Published: September 27, 2019

REFERENCES

Attrill, H., Takazawa, H., Witt, S., Kelm, S., Isecke, R., Brossmer, R., Ando, T., Ishida, H., Kiso, M., Crocker, P.R., and van Aalten, D.M. (2006a). The structure of siglec-7 in complex with sialosides: leads for rational structure-based inhibitor design. Biochem. J. *397*, 271–278.

Attrill, H., Imamura, A., Sharma, R.S., Kiso, M., Crocker, P.R., and van Aalten, D.M. (2006b). Siglec-7 undergoes a major conformational change when complexed with the alpha(2,8)-disialylganglioside GT1b. J. Biol. Chem. *281*, 32774–32783.

Blixt, O., Collins, B.E., van den Nieuwenhof, I.M., Crocker, P.R., and Paulson, J.C. (2003). Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein. J. Biol. Chem. *278*, 31007–31019.

Bradshaw, E., Chibnik, L.B., Keenan, B.T., Ottoboni, L., Raj, T., Tang, A., Rosenkrantz, L.L., Imboywa, S., Lee, M., Von Korff, A., et al. (2013). CD33 Alzheimer's disease locus: altered monocyte function and amyloid biology. Nat. Neurosci. *16*, 848–850.

Büll, C., Heise, T., Adema, G.J., and Boltje, T.J. (2016). Sialic acid mimetics to target the sialic acid-siglec axis. Trends Biochem. Sci. *41*, 519–531.

Chan, G., White, C.C., Winn, P.A., Cimpean, M., Replogle, J.M., Glick, L.R., Cuerdon, N.E., Ryan, K.J., Johnson, K.A., Schneider, J.A., et al. (2016). Trans-pQTL study identifies immune crosstalk between Parkinson and Alzheimer loci. Neurol. Genet. *2*, e90.

Carlin, A.F., Uchiyama, S., Chang, Y.C., Lewis, A.L., Nizet, V., and Varki, A. (2009). Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. Blood *113*, 1722–1725.

Chen, G.Y., Tang, J., Zheng, P., and Liu, Y. (2009). CD24 and Siglec-10 selectively repress tissue damage-induced immune responses. Science *323*, 1722–1725.

Dimasi, N., Moretta, A., Moretta, L., Biassoni, R., and Mariuzza, R.A. (2004). Structure of the saccharide-binding domain of the human killer cell inhibitory receptor p75/AIRM1. Acta Crystallogr. D Biol. Crystallogr. *60*, 401–403.

Efthymiou, A.G., and Goate, A.M. (2017). Late onset Alzheimer's disease genetics implicates microglial pathways in disease risk. Mol. Neurodegener. *12*, 43.

Ereño-Orbea, J., Sicard, T., Cui, H., Mazhab-Jafari, M.T., Benlekbir, S., Guarné, A., Rubinstein, J.L., and Julien, J.P. (2017). Molecular basis of human CD22 function and therapeutic targeting. Nat. Commun. *8*, 764.

Ferraiuolo, L., Heath, P.R., Holden, H., Kasher, P., Kirby, J., and Shaw, P.J. (2007). Microarray analysis of the cellular pathways involved in the adaption to and progression of the motor neuron injury in the SOD1 G93A mouse model of familial ALS. Neurobiol. Dis. *27*, 9201–9219.

Griciuc, A., Serrano-Pozo, A., Parrado, A.R., Lesinski, A.N., Asselin, C.N., Mullin, K., Hooli, B., Choi, S.H., Hyman, B.T., and Tanzi, R.E. (2013). Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. Neuron *78*, 631–643.

Heneka, M.T., Kummer, M.P., and Latz, E. (2014). Innate immune activation in neurodegenerative disease. Nat. Rev. Immunol. *14*, 463–477.

Hickman, S., Izzy, S., Sen, P., Morsett, L., and El Khoury, J. (2018). Microglia in neurodegeneration. Nat. Neurosci. *21*, 1359–1369.

Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.C., Carrasquillo, M.M., Abraham, R., Hamshere, M.L., Pahwa, J.S., Moskvina, V., et al. (2011). Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat. Genet. *43*, 429–435.

Krasemann, S., Madore, C., Cialic, R., Baufeld, C., Calcagno, N., El Fatimy, R., Beckers, L., O'Loughlin, E., Xu, Y., Fanek, Z., et al. (2017). The TREM2-APOE pathway drives the transcriptional phenotype of dysfunctional microglia in neurodegenerative diseases. Immunity *47*, 566–581.

Leyns, C.E.G., and Holtzman, D.M. (2017). Glial contributions to neurodegeneration in tauopathies. Mol. Neurodegener. *12*, 50.

Malik, M., Simpson, J.F., Parikh, I., Wilfred, B.R., Fardo, D.W., Nelson, P.T., and Estus, S. (2013). CD33 Alzheimer's risk-altering polymorphism, CD33 expression, and exon 2 splicing. J. Neurosci. *33*, 13320–13325.

Naj, A.C., Jun, G., Beecham, G.W., Wang, L.S., Vardarajan, B.N., Buros, J., Gallins, P.J., Buxbaum, J.D., Jarvik, G.P., Crane, P.K., et al. (2011). Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nat. Genet. *43*, 436–441.

Park, J., Wetzel, I., Marriott, I., Dréau, D., D'Avanzo, C., Kim, D.Y., Tanzi, R.E., and Cho, H. (2018). A 3D human triculture system modelling neurodegeneration and neuroinflammation in Alzheimer's disease. Nat. Neurosci. *21*, 941–951.

Pimenova, A.A., Marcora, E., and Goate, A.M. (2017). A tale of two genes: microglial Apoe and Trem2. Immunity *47*, 398–400.

Pröpster, J.M., Yang, F., Rabbani, S., Ernst, B., Allain, F.H., and Schubert, M. (2016). Structural basis for sulfation-dependent self-glycan recognition by the human immune-inhibitory receptor Siglec-8. Proc. Natl. Acad. Sci. U S A *113*, E4170–E4179.

Raj, T., Ryan, K.J., Replogle, J.M., Chibnik, L.B., Rosenkrantz, L., Tang, A., Rothamel, K., Stranger, B.E., Bennett, D.A., Evans, D.A., et al. (2014). CD33: increased inclusion of exon 2 implicates the Ig V-set domain in Alzheimer's disease susceptibility. Hum. Mol. Genet. *23*, 2729–2736.

Rillahan, C.D., Macauley, M.S., Schwartz, E., He, Y., McBride, R., Arlian, B.M., Rangarajan, J., Fokin, V.V., and Paulson, J.C. (2014). Disubstituted sialic acid ligands targeting siglecs CD33 and CD22 associated with myeloid leukaemias and B cell lymphomas. Chem. Sci. *5*, 2398–2406.

Sims, R., van der Lee, S.J., Naj, A.C., Bellenguez, C., Badarinarayan, N., Jakobsdottir, J., Kunkle, B.W., Boland, A., Raybould, R., Bis, J.C., et al. (2017). Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nat. Genet. *49*, 1373–1384.

Walker, D.G., Whetzel, A.M., Serrano, G., Sue, L.I., Beach, T.G., and Lue, L.F. (2015). Association of CD33 polymorphism rs3865444 with Alzheimer's disease pathology and CD33 expression in human cerebral cortex. Neurobiol. Aging *36*, 571–582.

Wang, Y., and Neumann, H. (2010). Alleviation of neurotoxicity by microglial human Siglec-11. J. Neurosci. *30*, 3482–3488.

Zhuravleva, M.A., Trandem, K., and Sun, P.D. (2008). Structural implications of Siglec-5-mediated sialoglycan recognition. J. Mol. Biol. *375*, 437–447.

ISCI, Volume *19*

Supplemental Information

Small Molecule Binding to Alzheimer Risk Factor CD33 Promotes Aβ Phagocytosis

Luke A. Miles, Stefan J. Hermans, Gabriela A.N. Crespi, Jonathan H. Gooi, Larissa Doughty, Tracy L. Nero, Jasmina Markulić, Andreas Ebneth, Berthold Wroblowski, Daniel Oehlrich, Andrés A. Trabanco, Marie-Laure Rives, Ines Royaux, Nancy C. Hancock, and Michael W. Parker

SUPPLEMENTAL INFORMATION

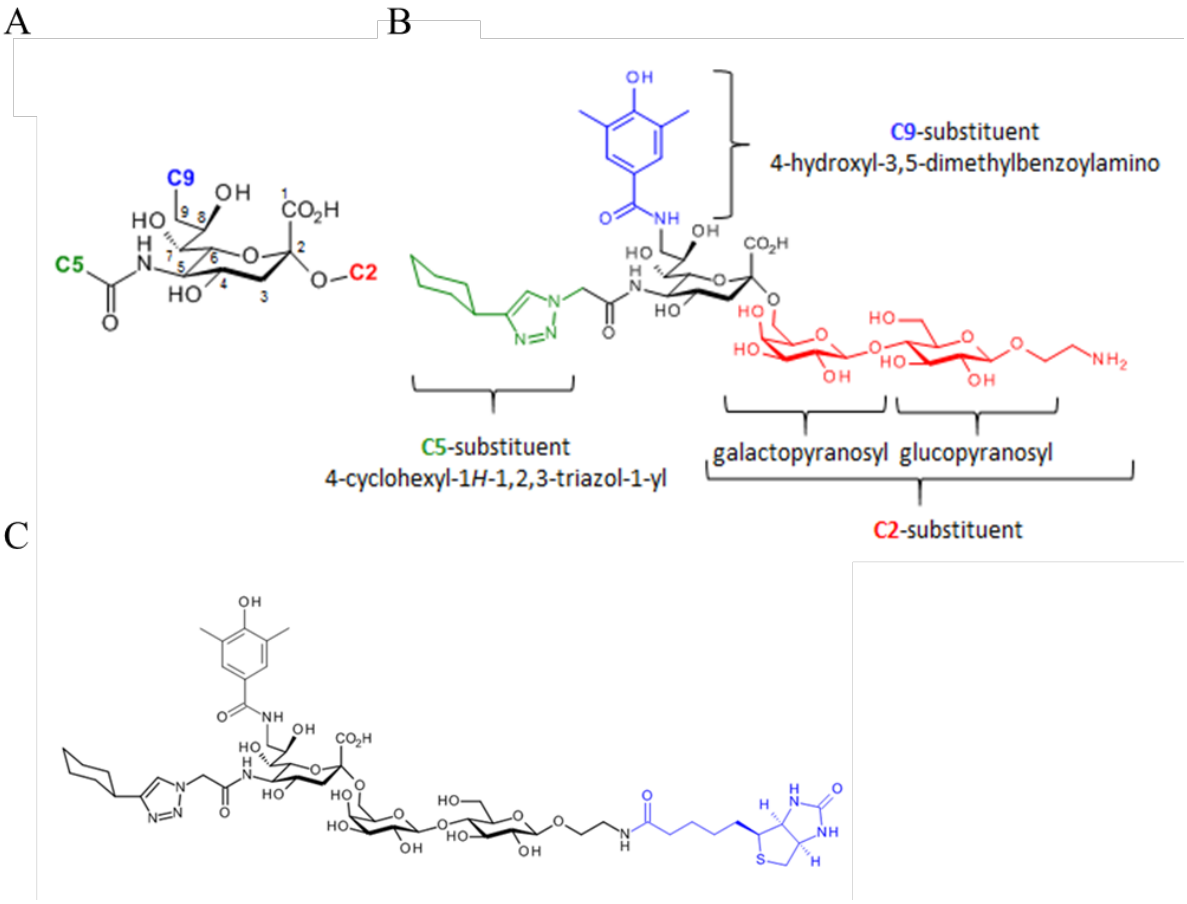


Figure S1. Chemical Structure of P22, related to Figure 2
(A) Numbering convention for the carbon backbone of the P22 sialic acid moiety.
(B) Chemical structure of P22 (β-D-glucopyranoside, 2-aminoethyl *O*-[*N*-[2-(4-cyclohexyl-1*H*-1,2,3-triazol-1-yl)acetyl]-9-deoxy-9-[(4-hydroxy-3,5-dimethylbenzoyl)amino]-α-neuraminosyl]-(2→6)-*O*-β-D-galactopyranosyl-(1→4)-; CAS registry number [1809735-27-0]).
(C) Biotinylated P22 (biotin label in blue).

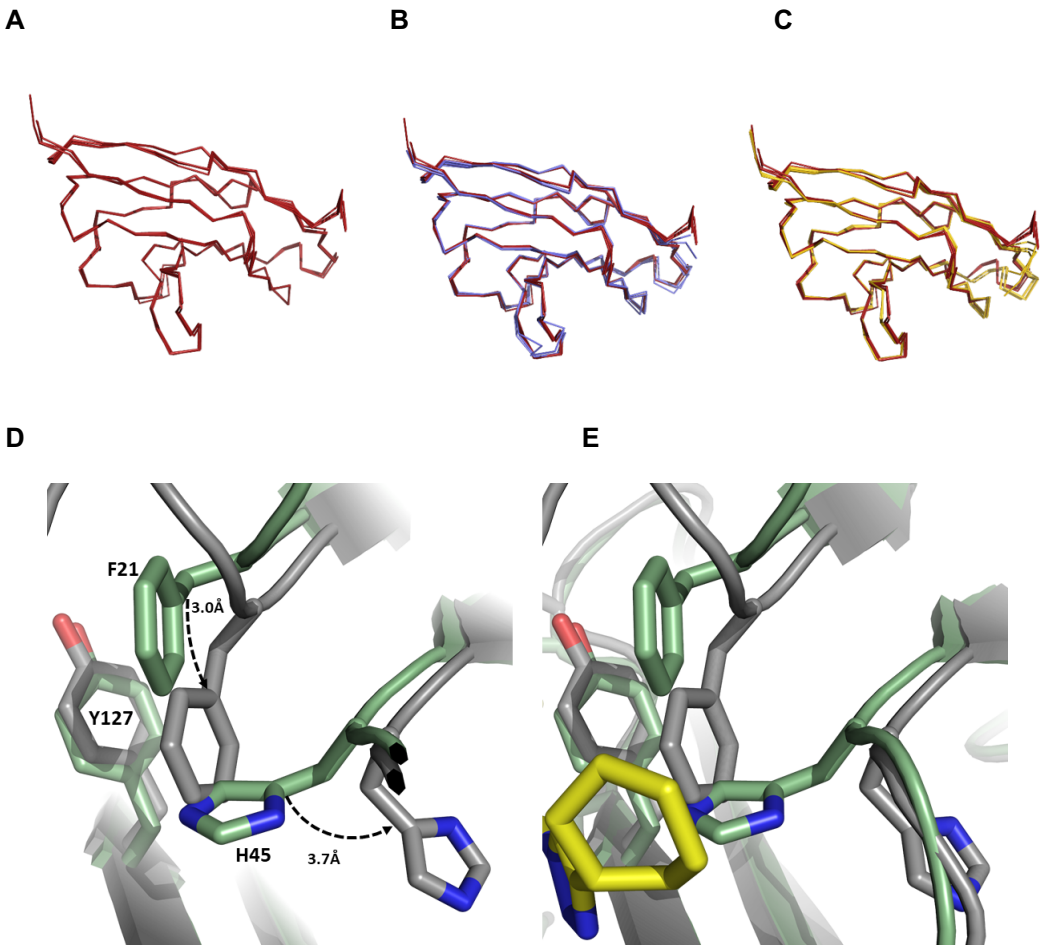


Figure S2. Comparison of Ligand Bound and Unbound CD33 V-set Domain Structures, related to Figure 2

(A) Alignment of the four CD33 chains identified in the unliganded structure, colored red. Maximum RMSD of matched Cα atoms between chains calculated to be 0.39-Å.

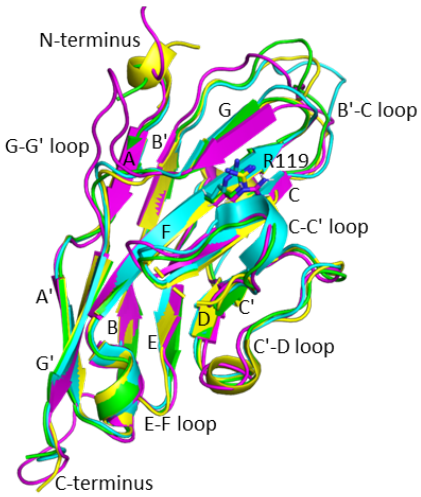
(B) Alignment of all four CD33 molecules in the unliganded structure, colored red, and available CD33 structures deposited in the PDB (PDB: 5IHB, 5J0B, 5J06), colored blue. Maximum rmsd of matched Cα residues calculated to be 0.41-Å to 5IHB chain B.

(C) Alignment of all four CD33 molecules in the unliganded structure, colored red, and all four P22 bound CD33 molecules identified across both co-crystal structures reported, colored yellow. Maximum rmsd of matched Cα residues between chains calculated to be 0.62-Å.

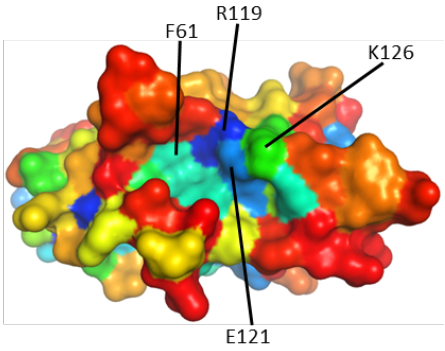
(D) Conformational changes on P22 binding to CD33. The unliganded CD33 structure is shown in green sticks and the P22 complex structure in grey sticks. Magnitude of displacement of aromatic side-chains is shown.

(E) Location of P22 (yellow sticks) in relation to the aromatic cluster.

A



B



C

Human Siglec	1	2	4	5	6	7	8	9	10	11	12	14	15	16
% Sequence identity with CD33	22	25	28	56	60	55	58	58	44	41	53	57	26	41

D

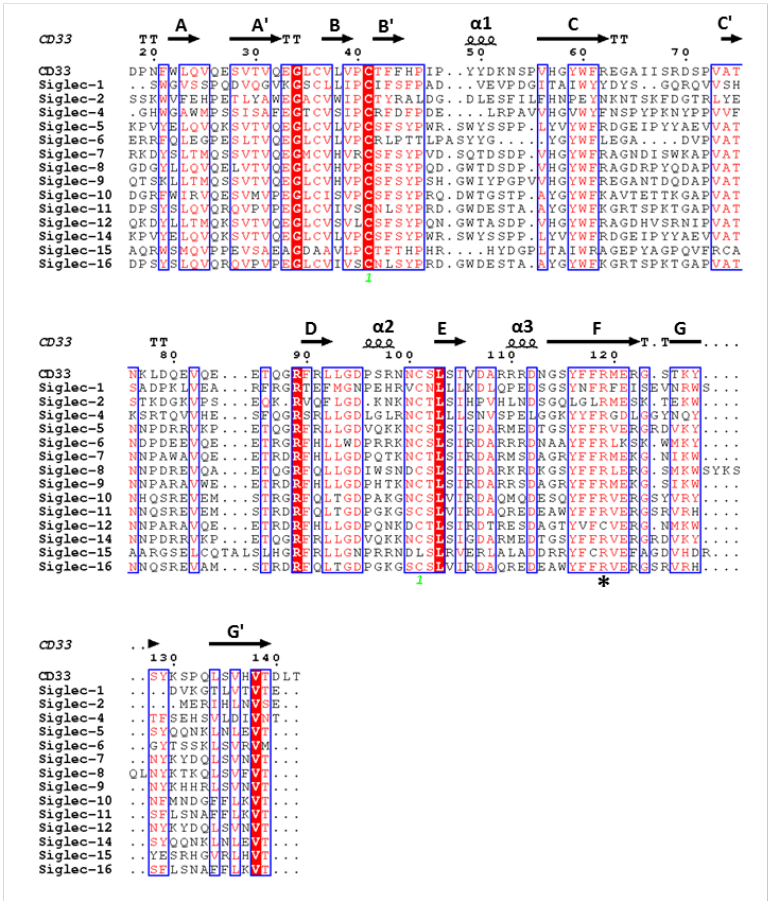


Figure S3. Comparison to Other Human Siglec V-set Domains, related to Figure 1

(A) Superimposition of the unliganded structure of CD33 (Siglec-3; green) with Siglec-5 (PDB: 2ZG2, cyan) (Zhuravleva et al., 2008), Siglec-7 (PDB: 1NKO, yellow) (Dimasi et al., 2004) and Siglec-8 (PDB:2N7A, magenta) (Pröpster et al., 2016). The lectin fold shows highly conserved β-strands labeled from the N-terminus forward as A, A', B, B', C, C', D, E, F, G and G'. The side-chains of the highly conserved R119 (numbering from CD33) are shown as sticks near the end of the central F-strand.

(B) The surface of CD33 with sequence variation between human Siglecs shown as a color spectrum from blue (most conserved) to red (least conserved), based on the alignment shown in panel D.
(C) Amino acid sequence identities between human Siglecs.
(D) Sequence alignment of human Siglecs using Clustalw (Thompson et al., 1994) and ESPript3.0 (Robert and Gouet, 2014). The β-strands and α-helices for CD33 are labeled above the sequence, the location of the C41-C101 disulfide bond is indicated by the green number 1 below the Siglec-16 sequence. The conserved arginine residue is indicated by the asterix. Sequence numbering is for CD33.

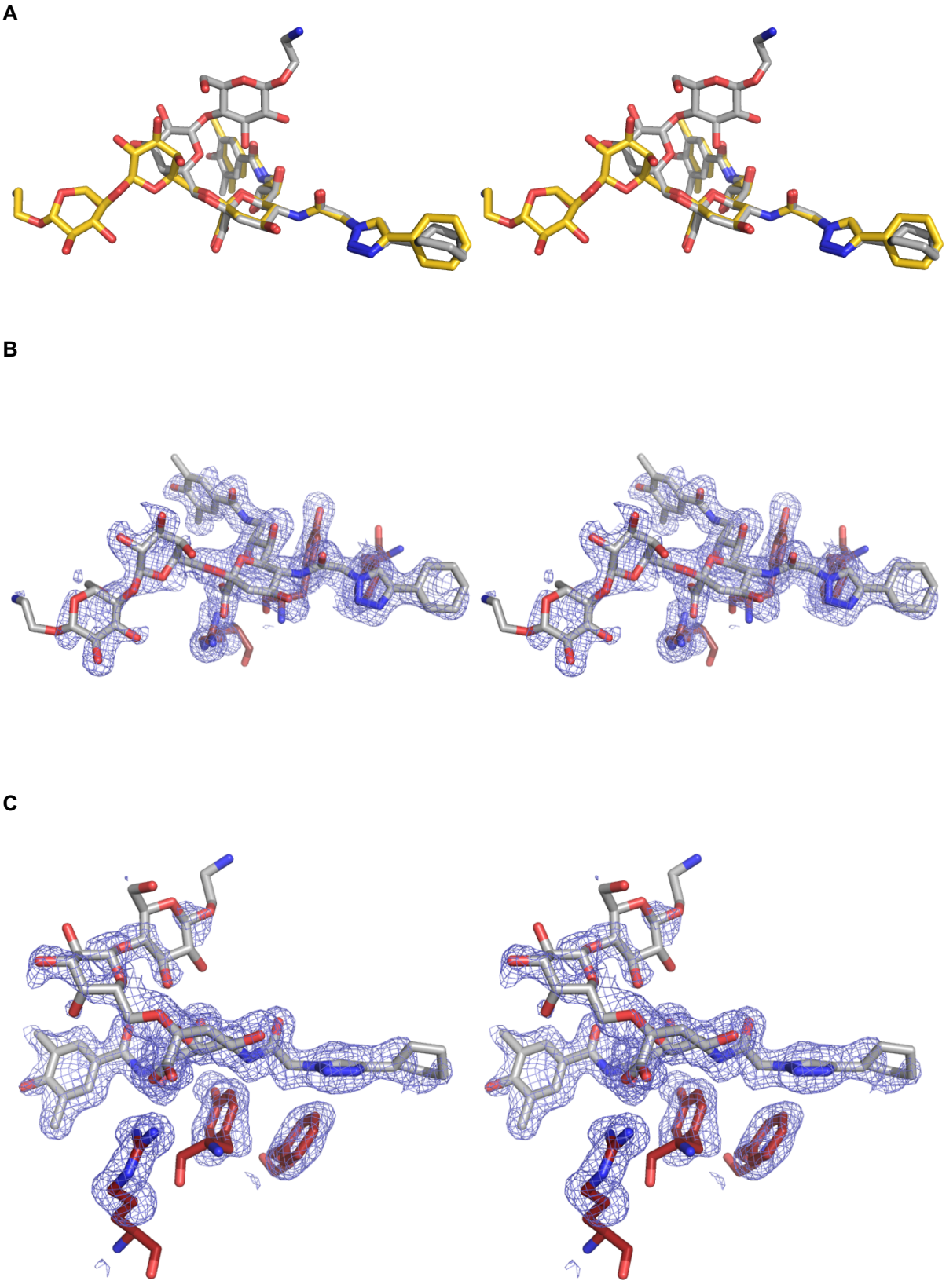


Figure S4. Structures of P22 Bound to CD33, related to Figure 2
(A) Stereo view of the two P22 binding modes (binding mode 1 in yellow and binding mode 2 in grey) observed in the crystal structures of P22 bound to CD33.
(B) Stereo simulating annealing omit map of binding mode 1 (seen in Crystal form no. 1) after the P22 molecules and the side-chains beyond the Cβ of Phe21, Tyr127 and Arg119 were removed before

refinement and calculation of the map. CD33 side-chains are shown in brown stick and the P22 molecule in grey stick, all with element coloring. The F_o-F_c map is shown in blue, contoured to 2.0 σ. (C) As for panel B but showing binding mode 2 (Crystal form no. 2).

SUPPLEMENTAL TABLES

Table S1. Crystallographic Data Processing and Refinement Statistics, related to Figures 1 and 2

Data Collection	P22 complex Crystal no. 1	P22 complex Crystal no. 2	Unliganded
Temperature (K)	100	100	100
Space group	*P2₁*	*P2₁*	*P2₁*
Unit Cell			
a, b, c (Å)	29.4, 68.8, 73.5	29.4, 68.3, 73.6	60.8, 41.5, 107.5
α, β, γ (°)	90, 97.0, 90	90, 97.5, 90	90, 96.6, 90
Resolution (Å)			
	36.46 - 1.80 (1.84 - 1.80)	34.17 - 1.75 (1.78 - 1.75)	42.48 - 1.78 (1.81 - 1.78)
No. of observations	141033 (8031)	110012 (5510)	171891 (8687)
No. of Unique Reflections	26760 (1534)	28662 (1477)	51251 (2673)
Redundancy	5.3 (5.2)	3.8 (3.7)	3.4 (3.2)
Data Completeness (%)	99.4 (94.8)	97.5 (88.9)	98.5 (90.3)
I/σ_I	12.6 (2.3)	13.4 (2.0)	14.7 (3.1)
R_{merge} (%)	7.3 (59.9)	6.1 (67.6)	4.4 (39.1)
Refinement Statistics			
Resolution (Å)	34.39 - 1.80	34.18 - 1.75	35.60 - 1.78
R_{work}	0.184	0.177	0.182
R_{free}	0.211	0.202	0.220
Non-Hydrogen Atoms			
Protein	2055	2023	4056
Water	253	333	490
r.m.s. deviations from ideality			
Bond angles (°)	1.154	1.001	0.998
Bond lengths (Å)	0.007	0.007	0.007
Average B-factor (Å^2)			
Protein	26.8	20.7	26.6
Water	38.1	33.0	35.8
Ramachandran Plot			
Residues in most Favored Regions (%)	98.4	100	98.2
Outliers (%)	0	0	0

Table S2. P22 Interactions with CD33[a], related to Figures 1 and 2

P22	CD33
Sialic Acid	
Carboxylate	Salt bridge to R119 guanidino moiety
C5-Substituent	
Amide	Hydrogen bond to backbone carbonyl of K126
Amide	vdW with Y127
Triazole ring	π-π edge-to-face interaction with F21
C8-Substituent	
Hydroxyl	Hydrogen bond to backbone amide of S128
C9-Substituent	
Amide	Hydrogen bond to backbone carbonyl of S128
4-hydroxy-3,5-dimethylbenzoylamino	vdWs with F117, K130, Y127, Y129

[a]Interactions only shown if they occur in all P22-CD33 structures.

Table S3. SPR Data for the Binding of CD33 and P22, related to Figure 3

Ligand	Analyte	k_{a1} (1/Ms)	k_{d1} (1/s)	K_{D1} (μM)	k_{a2} (1/RUs)	k_{d2}(1/s)	analysis model
$CD33^{WT}$	P22			118.0±41			1:1 steady state
$CD33^{C36S}$	P22			95.2±39			1:1 steady state
P22-biotin	$CD33^{WT}$	0.7±0.2E+05	1.4±0.4	18.0±01	3.2±2E-05	3.3±0.4E-02	Bivalent analyte
P22-biotin	$CD33^{C36S}$	0.4±0.2E+05	2.5±0.9	67.5±15			1:1 binding

Measurements in triplicate showing mean ± S.D.

Table S4 Effect of CD33 Mutations on P22-Alexa647 Binding Affinity, related to Figure 3

Variant	K_D (μM ± S.D.)	Fold-decrease compared to WT
WT	2.6 ± 1.1	n.a.
R119A	>100[****a]	>30
Y127A	>100[****a]	>30
S131A	3.1 ± 2.1[n.s.]	1.2
H45A	13.1 ± 10.3 [n.s.]	5
F21A	15.0 ± 4.5[*ǂ]	~6

n.a. not applicable

n.s. not significantly different compared to wild-type (WT)

[a]Significantly different from WT (* $p < 0.0332$; ** $p < 0.0021$; *** $p < 0.0002$; **** $p < 0.0001$)

K_D data are representative of n = 3 independent experiments performed in triplicate and presented as the average ± S.D. Statistical significance compared to WT was determined by one-way ANOVA with Dunnett post hoc analysis using GraphPad Prism 7.0 (GraphPad Software, Inc., La Jolla, CA, 92037, USA).

TRANSPARENT METHODS

Synthesis of P22

P22 (β-D-Glucopyranoside, 2-aminoethyl *O*-[*N*-[2-(4-cyclohexyl-1*H*-1,2,3-triazol-1-yl)acetyl]-9-deoxy-9-[(4-hydroxy-3,5-dimethylbenzoyl)amino]-α-neuraminosyl] (2→6)-*O*-β-D-galactopyranosyl-(1→4)-; CAS registry number [1809735-27-0]) was synthesized as described by Paulson and co-workers (Rillahan et al., 2014). Biotinylated P22 (Figure S1C) was prepared by treatment of a solution of P22 (12 mg, 0.012 mmol) in dimethylformamide (1 mL), with DIPEA (8.6 μL, 0.49 mmol) followed by addition of (+)-biotin N-hydroxysuccinimide ester (4.2 mg, 0.012 mmol), the resulting reaction mixture was then stirred for 2 h at RT. After which the reaction mixture was concentrated *in vacuo* and purified by reverse phase, performed via preparative HPLC (Stationary phase: RP XBridge Prep C18 OBD- 5μm,19x150mm, Mobile phase: 0.25% NH_4HCO_3 solution in water, CH_3CN) to afford Biotin-P22 (9 mg, yield 60.9%).

Protein production

Codon optimized (*E. coli*) DNA encoding D18 to H143 of human CD33 was synthesized and sub-cloned into a pET-30a+ vector with an N-terminal octa-histidine tag and a TEV protease cleavage site by Genscript. For SPR studies, a non-cleavable C-terminal AVI-tag was inserted into the clone and the free surface cysteine residue (C36) was mutated to serine by Genscript. The plasmid vector was transformed into *E. coli* (BL21*) and grown in LB broth under kanamycin selection at 37 °C in 2 L batches, in an orbital shaking incubator. Cultures were induced at mid-log phase with the addition of IPTG to 0.2 mM; bacteria were cultured for a further 5 h and harvested by centrifugation. Pelleted bacteria were suspended in lysis buffer (20 mM Tris pH 7.4, 150 mM NaCl, 5 mM EDTA, DNAse 1) and lysed by passage through an EmulsiFlex C-5 high-pressure homogenizer (Avestin). Lysates were subjected to centrifugation (18,000 g for 30 min) to pellet inclusion bodies and cell debris. Insoluble fractions were resuspended using a tissue grinder in wash buffer 1 (20 mM Tris pH 8.0, 150 mM NaCl, 1% Nonidet p 40 substitute, 5 mM EDTA) and pelleted again by centrifugation. This was repeated and insoluble fractions were washed twice with wash buffer 2 (20 mM Tris pH 8.0, 2 M NaCl, 1% Nonidet p 40 substitute, 5 mM EDTA) and twice with wash buffer 3 (20 mM Tris pH 8.0, 150 mM NaCl, 5 mM EDTA). Washed pellets were then resuspended in solubilization buffer (20 mL of 6 M urea, 50 mM Tris pH 8.5, 500 mM NaCl, 5 mM imidazole) and left rotating overnight at 4ºC. The suspension was then diluted 100-fold into refold buffer (400 mM arginine, 100 mM Tris pH 8.5, 200 mM NaCl, 0.5 mM oxidizing glutathione, 5 mM reducing glutathione) and left stirring for at least 60 h at 4 °C. Refold solutions were concentrated, and buffer exchanged into storage buffer (20 mM Tris pH 8.5, 200 mM NaCl) with 10 mM imidazole using a tangential flow system (Sartorius) and loaded onto a Ni-NTA column. Protein was washed, and then eluted with an imidazole concentration gradient up to 500 mM. Pooled protein elution fractions were then dialyzed into storage buffer and incubated for 36 h at 4 °C with 0.1 mg hexa-histidine-tagged TEV protease added at 0 and 18 h. Tag-cleaved samples were passed over a Ni-NTA column (to capture TEV and cleaved tag) and then the eluted, cleaved protein further purified using a HiLoad 26/60 Superdex S200 size exclusion chromatography column equilibrated with 20 mM Tris pH 8.5, 200 mM NaCl. Pure samples of CD33 were concentrated to 3-7 mg/mL and stored in 25 μL aliquots at -80°C until required.

Surface plasmon resonance

All surface plasmon resonance (SPR) experiments were performed in PBS-P+ (KCl 0.2 g/L, KH_2PO_4 0.2 g/L, NaCl 8 g/L, Na_2HPO_4 1.15 g/L, Tween 20 0.05%) with a flow rate of 60 μL/min at 25°C using a Biacore T200. His-tagged CD33 wild-type was captured on a Series S NTA chip (GE) via Ni^{2+}/NTA chelation then covalently immobilized using amine coupling chemistry (GE Healthcare Laboratory Guidelines 29-0057-17 AB). Biotinylated AVI-tagged $CD33^{C36S}$ was immobilized on a SAHC 1500M sensor chip (Xantec) via streptavidin capture, to an average density of 6500 resonance units (RU). In the reverse conformation assay, P22-biotin was captured to an average density of 570 RU.

For affinity analysis of P22 against captured $CD33^{C36S}$, P22 was diluted using a threefold dilution series, each concentration tested twice, from 243 μM to 3 μM and analysis was performed in duplicate using single cycle kinetics. For kinetic analysis of $CD33^{C36S}$ and His-tagged CD33 wild-type against captured P22-biotin, CD33 was diluted using a threefold dilution series, each concentration tested three times, from 9 μM to 0.03 μM and duplicate multiple cycle kinetics was performed. All samples were double referenced against a blank surface and buffer only cycles. All data were analyzed using Biacore T200 Evaluation Software Version 2.0.

SNAP-tag assay

The plasmid encoding human CD33 wild-type, tagged at its N-terminus with a Snap-tag, was generated by sub-cloning human CD33 ORF (NM_001772.3), lacking its signal peptide, into a pT8-SNAP vector (CisBio # PT8SNAPZEO) using EcoRV and XhoI. The Snap-tagged CD33 mutants were generated by site-directed mutagenesis.

HEK293T cells were cultured in DMEM (ThermoFisher Scientific # 11965092) containing 10% FBS (Hyclone # SH30070.03) and 1% Penicillin-Streptomycin (ThermoFisher Scientific # 15140122) and transfected using Fugene HD (Promega # E2312) in 10 cm plates. For each plate, the DNA (µg):Fugene HD (µL) mix ratio was 1:3 and 5 million cells were added on top of 1 mL mix and incubated at 37°C in a CO_2 incubator for 48 h. Forty-eight hours after transfection, the Snap-tag was labeled by directly incubating the cells in 10 cm plates with 50 nM BG-K Terbium (CisBio # SSNPTBX) in complete media for 1 h at 37ºC. After labeling, the cells were washed twice with 1X labeling media (CisBio # LABMED). The cells were then resuspended using Detachin (Amsbio # T100106). The cells were counted, centrifuged and resuspended in PBS at a concentration of 1,000 cells/µL and 10 µL were plated into each well. Increasing concentrations of P22-Alexa647 were incubated on top of the cells for 1 h at 4°C. Non-specific binding was assessed by competition with 100 µM of the non-labeled P22, as well as in control (Snap-tagged dopamine D2R receptor)-transfected cells.

Fluorescence was read with a PHERAstar plate reader using an excitation of 337 nm and emissions of 620 and 665 nm. Data presented are representative of three independent experiments performed in triplicates for each compound. Data are represented as the average ± S.D. Statistical significance was determined by one-way ANOVA with Dunnett post hoc analysis using GraphPad Prism 7.0 (GraphPad Software, Inc., La Jolla, CA, 92037, USA).

Aβ42 phagocytosis assay

A BV2 immortalized microglial cell line was cultured in DMEM low glucose with 10% (vol/vol) heat-inactivated FBS. Cells were seeded in a 12 well plate at 3×10^5 cells/mL and transfected using 0.9 µg human CD33 mRNA (hCD33) and 0.75 µL Lipofectamine® MessengerMAX (Invitrogen) per well.

Biotinylated P22 was dissolved in Hanks' Balanced Salt Solution (HBSS). 50 µg of P22 was added to 1 mg of FITC streptavidin particles (PC-SAFY-0.5, Kisker Biotech) that were previously equilibrated with HBSS buffer. After overnight incubation at 4°C, particles with and without compound were centrifuged (2,000 xg) and washed one time with HBSS. Microparticles were resuspended in 1 mL HBSS and were stored at 4 °C and protected from the light.

Aβ42 lyophilized stock was dissolved in 60 mM NaOH:ddH2O:10XPBS, in that order, in the ratio of 2:7:1. (The high pH buffer was used to prevent aggregation). 4 µL of dye (Invitrogen pHrodo™ Red Microscale Labelling Kit) was used to label 100 µL (1 mg/mL) of Aβ42. Excess dye was removed using 900 µL of resin (Bio-Gel p-6 fine resin suspended in PBS). Peptide was checked on an SDS-gel for the presence of aggregates.

Twenty four hours after transfection, cells were incubated with either 150 µL/mL FITC streptavidin particles alone or conjugated to biotinylated P22 in serum free media at 37°C in a CO_2 incubator for 30 min. After P22 incubation, 2 µg/mL of Aβ42 pHrodo was added for 1 h at 37°C in a CO_2 incubator. Cells were extensively washed with cold PBS and harvested using 200 µL PBS-EDTA for further analyses using flow cytometry. Ten thousand counts were acquired on an LRS Fortessa Flow Cytometer (Beckman). Samples were analyzed using FlowJo version 10. Data were collected in triplicate from n = 3 independent experiments and represented as the mean ± SEM. Statistical analysis was performed using GraphPad Prism 7 software. Comparisons involving more than two groups used one-way ANOVA followed by Tuckey post hoc test. All differences were considered significantly different when $p<0.05$.

Generation of CD33 knockout THP-1 cells

Guide RNA sequences for CRISPR/Cas9 were designed online using the Broad Institute sgRNA design tool at https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design. The gRNA selected for the KO of CD33 targets exon 3 of the human CD33 gene (target sequence: 5'-GGCCGGGTTCTAGAGTGCCA *GGG*-3'; PAM in italics). crRNA and tracrRNA oligos as well as recombinant Cas9 protein were ordered from Integrated DNA Technologies (IDT). The Cas9RNP complexes were made by assembling the 3 components, crRNA:tracrRNA:Cas9, at a 1.25:1.25:1 ratio. Cas9RNP were electroporated at a final concentration of 3.3 µM into 300,000 THP-1 cells using the Neon electroporation device (ThermoFisher Scientific) programmed at 1600 volts, 10 milliseconds and 3 pulses. Eight days after electroporation, cells were stained with PE conjugated mouse anti-human CD33 (Pharmigen cat #555450) and the CD33 negative population was collected by fluorescence-activated cell sorting (FACS) with a FACSCanto™ II cell sorter (BD BioSciences).

THP-1 phagocytosis assay
THP-1 cells (control and CD33KO) were differentiated in growth media (RPMI 1640, 10% FCS) supplemented with 100 ng/mL PMA for 3 days. Following this, cells were returned to normal growth media for an additional 3 days. Differentiated cells were then lifted, counted and 75,000 cells/well were plated in a 96 well plate. Cells were left for 1 h to settle then transferred to serum free media for 2 h. Following this, cells were incubated with P22 monomer, P22 conjugated microparticles or microparticles alone for 30 min prior to incubation with pHrodo *E. coli* Bioparticles for 2 hours per the manufacturer's instruction. All treatments were present for the duration of the experiment. Mean fluorescent intensity (MFI) was determined using a Perkin Elmer Enspire Plate Reader.

Crystallization
Thawed CD33 at 3.6 mg/mL was supplemented with 1 mM P22 and screened against the MCSG1 Suite (Microlytic) in 200 x 200 nL sitting drops, dispensed by a Gryphon LCP (Art Robbins Instruments) crystallization robot. CD33 crystals grew with a well solution of 30% w/v PEG MME 550, 100 mM HEPES pH 7.5, 50 mM $MgCl_2$ at 22°C (Crystal form no.1) and presented in a range of cubic morphologies from 40 – 80 µm in length. Initial hits were reproduced in 2 x 2 µL hanging drops appearing over a week. Similar shaped crystals also grew from a well solution of 18% w/v PEG 4000, 100 mM MES pH 6.5, 600 mM NaCl at 22°C (Crystal form no. 2). Crystals were briefly transferred to a cryo-solution of the mother liquor supplemented with 20% (v/v) glycerol before being flash frozen directly into liquid nitrogen.

Unliganded CD33 crystals were likewise grown from a protein batch concentrated to 15 mg/mL and screened against The PEGs Suite (Qiagen). Initial hits were reproduced in 2 x 2 µL hanging drops appearing overnight. These crystals grew using well solution composed of 25% w/v PEG 4000, 100 mM Tris-HCl pH 8.5 at 22°C but were small and excessive nucleation was apparent despite being well-shaped. The best crystals were achieved with the same well solution made up with 10% (v/v) glycerol; cubic-shaped crystals appeared overnight or soon after and increased in size over a week or so. Crystals were flash frozen directly from the mother liquor.

Structure determination
X-ray diffraction data were acquired on the MX beamlines at the Australian Synchrotron (Clayton, Victoria, Australia) (Cowieson et al., 2015; Aragão et al., 2018) using Blue-Ice software (McPhillips et al., 2002). Data sets were acquired at a single wavelength of 0.9537-Å, in a nitrogen cryostream (100 K). Diffraction images were integrated using XDS (Kabsch, 2010) and scaled with Aimless of the CCP4 program suite (Winn et al., 2011), setting aside 5% of the reflections for the R_{free} set. Data collection statistics are shown in Table S1.

The initial unliganded structure was determined by molecular replacement with Phaser from the Phenix software suite (Adams et al., 2010) in the $P2_1$ space group. A successful molecular replacement solution was achieved with a probe model derived from the Siglec-7 structure (PDB:1NKO (Dimasi et al., 2004)), identifying four CD33 molecules in the asymmetric unit. All other CD33 structures were determined by molecular replacement with Phaser in the $P2_1$ space group using the unliganded structure as the search model. Refinement and rebuilding were performed with Coot (Emsley and Cowtan, 2004) and Phenix Refine (Adams et al., 2010). Structure validation was monitored with MolProbity (Chen et al., 2010). The quality of the final electron density maps for the P22 complex structures is shown in Figure S4. Refinement statistics are shown in Table S1. The final models have better than 98.2% of residues in favored regions of the Ramachandran plot with no outliers. Structural figures were produced using PyMOL version 1.8.2.2 (http://www.pymol.org), chemical structures were drawn using MarvinSketch 6.0.2 (http://www.chemaxon.com), and interactions between CD33 and P22 were identified and visualized using LigPlus (Laskowski and Swindells, 2011) version 1.3.5 (www.ebi.ac.uk/thornton-srv/software/LigPlus).

QUANTIFICATION AND STATISTICAL ANALYSIS
All results are presented as either average ± S.D. or mean ± SEM. Microsoft Excel and GraphPad Prism 7 software were used for statistical analysis. Data were analyzed using one-way ANOVA, or repeated-measures one-way ANOVA, followed by Tukey's or Dunnett's multiple comparisons test, as specified in the Figure legends. Only two-sided tests were used, and all data analyzed met the assumption for the specific statistical test that was performed. Probability levels of $p < 0.05$ were considered statistically significant.

DATA AND SOFTWARE AVAILABILITY

Atomic coordinates and structure factors of the unliganded CD33 V-set domain and the two P22 complex models have been deposited in the Protein Data Bank under the accession numbers 6D48, 6D49 and 6D4A, respectively.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines

Chemically competent *E. coli* BL21cells used to produce recombinant hCD33 proteins were obtained from Thermo Fisher Scientific. BV2 cells used for the phagocytosis assay were a gift from Prof Kevin Barnham, Florey Institute of Neuroscience and Mental Health (Parkville, Australia). HEK293T cells used for the SNAP-tag assay were obtained from and authenticated by ATCC. THP-1 cells used for the phagocytosis assay were obtained from and authenticated by ATCC and by Janssen using STR analysis. Sex information of HEK293T cell lines is not available. THP-1 cell lines are derived from a male donor.

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Michael Parker (mparker@svi.edu.au). All modified cell lines and plasmids of mutants will require completion of a materials transfer agreement with Janssen Pharmaceuticals. There are no fresh supplies available of the tool compound, P22, but the interested reader is directed to the corresponding author of the original paper describing its synthesis (Rillahan et al., 2014).

Supplemental References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213-221.

Aragão, D., Aishima, J., Cherukuvada, H., Clarken, R., Clift, M., Cowieson, N.P., Ericsson, D.E., Gee, C.L., Macedo, S., Mudie, N. et al. (2018) MX2: a high-flux undulator microfocus beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron". J. Synchrotron Rad. *25*, 885-891.

Chen, V.B., Arendall, W.B. 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12-21.

Cowieson, N.P., Aragão, D., Clift, M., Ericsson, D.J., Gee, C., Harrop,S.J., Mudie, N., Panjikar, S., Price, J.R., Riboldi-Tunnicliffe, A. et al. (2015) MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. J. Synchrotron Rad. *22*, 187-190.

Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126-2132.

Kabsch, W. (2010) XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125-132.

Laskowski, R.A., and Swindells, M. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. *51*, 2778-2786.

McPhillips, T.M.., McPhillips, E., Chiu, H.J., Cohen, A.E., Deacon, A.M., Ellis, P.J., Garman, E., Gonzalez, A., Sauter, N.K., Phizackerley, R.P., et al. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Radiat. *9*, 401-406.

Robert, X., and Gouet P. (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*, W320-W324.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673-4680.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235-242.