Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab

Authors: Lionel Rougé[1], Nancy Chiang[2], Micah Steffek[3], Christine Kugel[4], Tristan Croll[5], Christine Tam[4], Alberto Estevez[1], Christopher P. Arthur[1], Christopher M. Koth[1], Claudio Ciferri[1], Edward Kraft[4], Jian Payandeh[1]*, Gerald Nakamura[2]*, James T. Koerber[2]*, Alexis Rohou[1]*

Affiliations:
[1]Department of Structural Biology, Genentech Inc., South San Francisco, CA, 94080
[2]Department of Antibody Engineering, Genentech Inc., South San Francisco, CA, 94080
[3]Department of Biochemical and Cellular Pharmacology, Genentech Inc., South San Francisco, CA, 94080
[4]Department of Biomolecular Resources, Genentech Inc., South San Francisco, CA, 94080
[5]Cambridge Institute for Medical Research, University of Cambridge, Keith Peters Building, Cambridge CB2 0XY, UK
*Correspondence to: rohou.alexis@gene.com, koerber.james@gene.com, nakamura.gerald@gene.com, payandeh.jian@gene.com

Abstract: Cluster of Differentiation 20 (CD20) is a B-cell membrane protein that is targeted by monoclonal antibodies for the treatment of malignancies and auto-immune disorders, but whose structure and function are unknown. Rituximab (RTX) has been in clinical use for two decades, but how it activates complement to kill B cells remains poorly understood. We obtained a structure of CD20 in complex with RTX, revealing CD20 as a compact double-barrel dimer bound by two RTX antigen-binding fragments (Fabs), each of which engages a composite epitope and an extensive homotypic Fab:Fab interface. Our data suggest that RTX crosslinks CD20 into circular assemblies and lead to a structural model for complement recruitment. Our results further highlight the potential relevance of homotypic Fab:Fab interactions in targeting oligomeric cell-surface markers.

One Sentence Summary: CryoEM structure of CD20 in complex with rituximab reveals membrane protein fold and explains mode of recognition of therapeutic antibody

Main text: The integral membrane protein Cluster of Differentiation 20 (CD20) is a B-cell-specific marker and a clinically-validated therapeutic target for B-cell malignancies and auto-immune conditions (*1*). It is ubiquitously expressed on circulating B cells (*2*) and is predicted to have four transmembrane (TM) helices with two extracellular loops, ECL1 and ECL2, the second of which is much longer and contains a disulfide bond (Figure S1A). These topological features are conserved among a group of membrane proteins called MS4A (membrane-spanning 4-domain family, subfamily A), which includes 18 proteins identified through similarities between their amino acid sequences but whose biological functions are mostly unknown (*3*). Aside from structures of short ECL2 peptide segments from CD20 (*4*, *5*), there exists no high-resolution structural data on any MS4A family member beyond predictions of secondary structure and membrane topology. While CD20 is the best-studied member of the family, even its oligomeric state is poorly understood: available evidence suggests it associates into homo-oligomers and complexes with other proteins (*6*–*8*). In addition, the function of CD20 remains an area of active debate. Early work suggested that CD20 functions as an ion channel because overexpression and knockout of CD20 can increase or decrease Ca^{2+} conductance in B cells, respectively (*6*, *9*). However, more recent work showed that $CD20^{+}$ B cells lacking the B-cell receptor (BCR) are unable to initiate calcium signaling, suggesting that CD20 indirectly regulates calcium release downstream from the BCR (*10*).

CD20-targeted therapies revolutionized the treatment of B-cell malignancies and auto-immune disorders, starting with the monoclonal antibody (mAb) rituximab (RTX; Rituxan®), which was the first approved therapeutic mAb for cancer and continues to be the benchmark for second- and third-generation mAbs (*1*, *11*). Another anti-CD20 mAb, ocrelizumab (OCR; Ocrevus®), is now used in the treatment of multiple sclerosis (*12*). While all anti-CD20 mAbs act by depleting B cells, they employ at least four distinct mechanisms (*13*): direct cell death, FcR effector functions through antibody-dependent cellular cytotoxicity (ADCC) and phagocytosis (ADCP), and complement-dependent cytotoxicity (CDC). Each therapeutic antibody varies in its ability to trigger each pathway and there is no molecular-level understanding of why this is the case, but these distinct functional effects have been useful in categorizing anti-CD20 mAbs into either type I or type II (*1*, *13*). Rituximab is the prototypical type I mAb, characterized by high CDC activity and the ability to cluster CD20 into lipid rafts (*11*, *14*). Other type I mAbs include OCR and ofatumumab (OFA; *15*). Type II mAbs such as obinutuzumab (OBZ; Gazyva®) and tositumomab (B1; Bexxar®) exhibit low CDC activity, lack the ability to localize CD20 into lipid rafts, but induce higher levels of direct cell death (*16*).

These broad categorizations do not explain how CD20:mAb binding and mAb features lead to different modes of action. One hint at a possible molecular underpinning for these differences is that twice as many type I mAbs bind the surfaces of $CD20^{+}$ cells as type II mAbs (*17*), suggesting that CD20:mAb binding stoichiometry plays a role, though it is not clear how such strict stoichiometry might arise. Adding to this mystery, some type II and type I mAbs (*e.g.* OBZ and RTX) have overlapping epitopes centered around the ^{170}ANPSE^{174} motif of ECL2, while at least one type I mAb (OFA) has a completely separate, non-overlapping epitope involving ECL1

and another part of ECL2 (*13*). Thus, the location of the epitope on CD20 cannot be the sole determinant of mAb stoichiometry or of therapeutic mode of action. How do antibodies with virtually identical epitope sequences centered on ^{170}ANPSE^{174} bind with different stoichiometries and trigger remarkably different responses? One hint comes from epitope fine-mapping studies showing that residue Asn176 of ECL2 is involved in OBZ binding, but not RTX binding, and that OBZ binds to ECL2 peptides (but not CD20^{+} cells) with higher affinity than RTX (*5*); another hint comes from X-ray crystallographic structures of peptide-bound antigen-binding fragments (Fabs) of RTX, OCR and OBZ, in which RTX and OCR approach the ECL2 epitope at a similar angle, tilted approximately 70^{o} away from the angle at which OBZ approaches the same peptide (*4*, *5*, *18*, *19*). However, in the absence of a structure of full-length CD20 or of a MS4A homologue, it is difficult to speculate how such differences in binding geometry might impact overall binding stoichiometry and dictate therapeutic mode of action.

Results

CD20 forms dimers bound by two RTX Fabs

To facilitate biophysical and structural analyses, we produced human CD20 recombinantly in insect cells and optimized the construct for increased expression (Figure S1A). Following solubilization and purification in the mild detergent glyco-diosgenin (GDN), CD20 was found to be further stabilized by cholesterol hemisuccinate (CHS; Figure S2). Analyzed in GDN-CHS buffer, size exclusion chromatography with multi-angle light scattering (SEC-MALS) indicated that purified CD20 forms stable complexes in 2:2 stoichiometry with RTX Fabs, and 2:1 stoichiometry with OBZ Fabs (Figure S1C-E; Table S1). Fab binding to purified CD20 was subsequently evaluated using biolayer interferometry (BLI; Figure 1A) and surface plasmon resonance (SPR; Table S2) revealing sensorgrams consistent with a 2-state 2:2 binding of RTX (K_D = 21.4 nM, SPR), and with 2:1 binding of OBZ Fab (K_D = 58.8 nM, SPR). Imaging the resulting CD20:Fab complexes by negative-stain electron microscopy (nsEM) showed that each CD20 particle is bound by either two RTX Fabs or a single OBZ Fab (Figure 1B). Cryogenic electron microscopy (cryoEM) imaging of the RTX:CD20 complex allowed us to determine its structure to a resolution of 3.3 Å (Figure S3), resulting in a near-complete atomic model of the complex (Figure 1C).

The CD20 fold

The RTX:CD20 complex is Y-shaped, with the two RTX Fabs poised on the extracellular face of CD20. CD20 conforms to the predicted 4-transmembrane helix (TM) arrangement, with its 4 TMs arranged anti-parallel and clockwise when viewed from the extracellular side (Figure 2, top view). The core of each CD20 monomer presents a compacted rectangular-fold measuring ~25 × ~20 × ~50 Å (Figure 2). Despite low overall sequence identity across the MS4A family (~30 %), the structure of CD20 reveals that key structural elements are likely shared by all members (Figure S4A). Most notable is a constellation of highly-conserved, small residues that allow for the close inter-helical packing observed in CD20 and are found along TM1 (Gly53, Gly60, Gly67), TM2 (Gly90, Ser97, Gly98), and TM3 (Ser123, Gly130) (Figure S4A). A set of larger residues contributed by TM2 (Tyr86, Tyr94) and TM4 (Leu194, Met197, Ala201, Gln204) forms the bulk of the tightly-packed TM-helical core (Figure 2B) and are conserved as similarly-bulky residues in the MS4A family (Figure S4A). Notwithstanding some conformational heterogeneity observed at the intracellular end of TM1, the close interdigitation of highly-conserved residues over ~30 Å creates a tightly sealed 4-TM bundle within the CD20 monomer that is inconsistent with the formation of a transmembrane permeation pathway.

In contrast to the conserved transmembrane core, the extracellular loops (ECLs) of the MS4A family are extremely diverse in sequence (Figure S4A). In CD20, the first extracellular loop (ECL1) is short and largely shielded by ECL2, leaving only Ile76 and Tyr77 exposed (Figure 2A). The first half of the approximately 35-residue-long ECL2 is an amphipathic sequence that may partially partition into the membrane region and surrounds the perimeter of the 4-helix bundle until it kinks into a single-turn α-helix (Asn153-Arg156), which marks the start of ECL2's extracellular segment (Figure 2A). Following the α-helix, residues His158 to Ile164 form a circumflex-shaped cap above ECL1, and Ile164 and Tyr165 appear to plug a cavity at the center

of the square CD20 fold. The remaining part of ECL2 is stapled by the landmark disulfide bond between Cys167 and Cys183, which is located on an N-terminal extension of helix TM4 (Figure 2, side view). This region of ECL2, akin to a turret, is by far the most solvent-accessible region of CD20 and contains the known peptide epitope (^{170}ANPSE^{174}) for most anti-CD20 antibodies (*13*).

A search of the protein databank for structures similar to CD20 identified Claudin-3 and CD81 as the nearest matches (*20*, *21*). Although both present similar topologies, structural superposition with CD20 demonstrates poor overall correspondence (Figure S4B). For example, the transmembrane cavity and cholesterol-coordinating acidic residue present in CD81 (*20*) are absent in CD20, and while claudins appear to share a similar core 4-TM packing with CD20, they display a different topology and distinct oligomeric assembly interfaces. We conclude that the three-dimensional structure of CD20 represents a distinct membrane protein fold.

CD20 is a dimer

Previous studies have suggested that CD20 exists as a dimer or a tetramer (*6*–*8*). Viewed from the extracellular side, CD20 forms a dimeric double-barrel assembly of approximate dimensions 20 × 50 Å (Figure 2, top view). Two square CD20 subunits abut each other to form a four-helix anti-parallel transmembrane coiled-coil involving the upper halves of TM1 (Leu61, Phe62, Ala65, Leu69) and TM4 (Leu189, Ile193, Val196, Phe200) from each protomer (Figure 2D). The residues at this inter-subunit interface are involved in hydrophobic and van der Waals interactions. In the intracellular half of the transmembrane domain, dimerization is mediated by homotypic contacts between symmetry-related TM1 residues (Thr51, Ala54, Val55, Met58). The close, complementary packing and extensive hydrophobic surface contact at the CD20 dimer interface does not support the existence of an inter-protomer transmembrane conduction pathway. Additionally, structure-based sequence alignment suggests that a CD20-like dimer interface may be shared across the MS4A family (Figure S4A).

The dimeric assembly of CD20 is reinforced by extensive contacts between the extracellular α-helical extension of TM4 and the solvent-exposed region of ECL2. Here, numerous hydrophobic interactions and some polar contacts (Ser179 to Ser179'; Gln181 to backbone amide of Tyr161') contribute to the interface (Figure 2C). In total, the CD20 dimer buries 1,656 Å^2 of surface area and has a shape complementarity score of 0.53, comparable to many established dimeric integral membrane proteins (e.g. *22*, *23*). We further examined Tyr182, located near the symmetry axis in CD20, and found that mutation to cysteine (Tyr182Cys) resulted in purification of a covalent dimeric species in the absence of RTX (Figure S1F), consistent with the dimeric assembly observed in our structure. Overall, we conclude that CD20 forms a tight dimeric assembly that places ECL2 loops in close proximity to each other and presents its main epitope (^{170}ANPSE^{174}) in closely associated pairs, less than 20 Å apart.

ECL2 of CD20 is simultaneously recognized by two RTX Fabs

Previous studies have established that the principal epitope of RTX is centered on the ECL2 sequence ^{170}ANPSE^{174} (*13*). In our structure, this core epitope is simultaneously bound by two Fabs, which we denote RTX and RTX' (Figure 3). The RTX Fab sits atop the CD20 protomer, engaging ECL2 at a shallow angle (~ 22°) relative to the membrane plane (Figure 3, side view) , likely precluding engagement of CD20 by a single IgG. RTX engulfs the core epitope through numerous van der Waals packing contacts, as well as multiple polar interactions including the hydrogen bond pairs HC.Ser58 – Pro169 (backbone), HC.His35 – Asn171, and HC.Asn33 – Ser173 (backbone and side-chain). The second (RTX') Fab extends its heavy chain variable loop 3 (H3) across the dimer interface to present HC.Tyr97', which interacts with Glu174 of CD20 (Figure 3A). At the apex of ECL2, Ser173 organizes an extended network of hydrogen bonds spanning from HC.Asn33, through Glu174, to HC.Tyr97' (Figure 3B). This key interaction network is clamped by both Fabs: Ser173 is stabilized by HC.Tyr52, Glu174 is sandwiched by HC.Trp100b and HC.Gly100', and HC.Tyr97 is stabilized by HC.Trp100b.

The CD20:RTX complex reveals a distinct secondary epitope

Our structure reveals a second CD20 epitope formed by ECL1 and ECL2 and contacted by complementarity-determining region (CDR) loop L1 of RTX (Figure 3D). Completely distinct from the classic ECL2 turret epitope ^{170}ANPSE^{174}, this secondary epitope is recognized primarily by light chain residues LC.Ser28, LC.Ser29 and LC.Ser31. Residues LC.Ser28 and LC.Ser29 are positioned to make van der Waals contacts with Ile76 of ECL1 and Pro160 of ECL2, respectively. The side chain of LC.Ser31 is situated atop ECL2's circumflex cap and interfaces with both CD20 protomers: it stabilizes Tyr161 in a CH2-arene-CH2 sandwich also involving Pro160, while its hydroxyl moiety makes van der Waals contacts with Pro178', which caps the TM4 α-helical extension of the CD20' protomer. Earlier crystallographic studies of the primary ECL2 turret epitope in complex with RTX had measured a buried surface area of only 440 Å^2 (*4*), but these L1 – ECL1/2 interactions increase the contact surface area by almost 50 %, to ~ 640 Å^2 (Figure 4B), suggesting that this secondary epitope likely contributes significantly to RTX's affinity for CD20.

RTX Fabs are engaged in homotypic contacts

The close proximity (~ 20 Å) of the two primary epitopes displayed by the CD20 dimer results in the RTX Fabs accommodating each other along a homotypic interface between their heavy chains (Figure 3). CDR loop 3 (H3) dominates the Fab:Fab interface, engaging with its symmetry mate (H3'), and with the H1' and H2' loops. Residue HC.Tyr97, which is germline-encoded via the D gene segment (Figure S6A), seems essential to the Fab-Fab interaction: its Cβ and Cγ atoms make close van der Waals contacts with their symmetry mates across the dimer axis, while its aromatic ring stabilizes the backbone of the H3 loop, and its hydroxyl hydrogen-bonds with Glu174' of the contralateral CD20' protomer (Figure 3A,B). Two key additional Fab:Fab interactions are mediated by HC.Ser31', whose backbone and side chain directly engage HC.Gly99, while its side chain contacts LC.Tyr49 and HC.Tyr98 (Figure 3C). Overall, this Fab:Fab homotypic interface buries 375 Å^2 of solvent-exposed area (Figure 4B).

Our structure has thus unveiled a composite CD20-Fab' epitope with three components: the primary epitope on ECL2 (^{170}ANPSE^{174}), a secondary ECL1-ECL2 epitope, and a direct homotypic Fab:Fab interface. The total composite buried surface area of ~ 1,000 Å^2 is comparable to traditional prototypical Fab:protein complexes (*24*), and our observations rationalize how RTX achieves nanomolar affinity for full-length CD20 (Table S2) and CD20+ cells (*5*) despite low affinity to the ECL2 epitope peptide (Table S3).

RTX's target recognition and CDC activity require the full composite epitope

To ascertain the functional relevance of previously unobserved structural features of the complete CD20 epitope, we introduced targeted mutations into RTX (Figure 4C). We employed charge reversals and bulky side chains expected to disrupt these molecular interfaces and measured their effects on CDC (Figure 4E), and on IgG binding to purified CD20 (SPR; Table S2), ECL2 epitope peptide (SPR; Table S3), and CD20$^+$ cells (flow cytometry; Figure 4D). For comparison, we included OBZ, which is known to bind cells at levels ~50% lower than RTX and whose Fab can only bind purified CD20 with 1:2 stoichiometry (Figure 1).

RTX variants LC.Ser28Asp and LC.Ser31Asp were generated to probe the importance of the secondary epitope. Mutation LC.Ser28Asp resulted in reduced CD20 affinity (~20 fold), cellular binding (~ 25%), and CDC activity (IC50 > 10x higher) relative to wild-type RTX. LC.Ser31Asp had a notably stronger phenotype, which resulted in a ~100-fold reduction in Fab:CD20 binding, ~50% reduction in IgG:cell binding, and nearly completely abolished CDC activity. These data substantiate the relevance of this secondary epitope in RTX function.

We next evaluated the role of the germline-encoded HC.Tyr97, because it appears central to complex formation (see above; Figure 3A,B). The HC.Tyr97Ser mutation, which we predict may destabilize H3, effectively abolishes target engagement and CDC activity, while mutation HC.Tyr97Phe, which removes only the terminal hydroxyl but maintains the aromaticity of the side chain, reduced affinity to CD20 (~15 fold) as well as cellular binding (~ 25%) and CDC activity (IC50 > 10x higher), suggesting that RTX function is enhanced by the polar interaction between HC.Tyr97 and Glu174'.

To assess the importance of the homotypic Fab:Fab interface, we introduced mutations at HC.Ser31 and HC.Gly99, two positions that are reciprocally involved in interactions at the periphery of the complex. The HC.Ser31Glu mutant had reduced Fab:CD20 affinity (~100-fold), reduced IgG:cell binding and reduced CDC activity (>100-fold). The effect of HC.Gly99Lys was even more marked, with CDC completely abolished. Because these residues are not involved in any interactions with CD20, we conclude that homotypic Fab:Fab interactions potentiate target engagement, cell binding, and CDC activity of RTX.

In summary, we have discovered several RTX mutants that, despite maintaining cell-binding activity comparable to that of OBZ, and largely unaffected binding to the primary ECL2 epitope, are incapable of eliciting CDC. This confirms that the secondary ECL1/2 epitope and Fab:Fab interface contribute to the unique binding properties and high CDC activity of RTX.

Full-length RTX cross-links CD20 dimers into higher-order assemblies

In the RTX Fab:CD20 complex, the distance between the C-termini of the Fab heavy chains (HC.Pro213) is greater than 120 Å, inconsistent with binding of both Fabs from a single IgG to a CD20 dimer (Figure 1C). This suggests two RTX antibodies engage the dimer, each contributing one of its Fab domains. To test this, we formed complexes *in vitro* between full-length IgG and CD20 and examined their structural arrangement. We found that these complexes are stable (Table S2), and nsEM showed that most CD20 particles were bound by two well-resolved Fabs in a similar geometry to that seen in the CD20:RTX Fab structure (Figure S7G). This establishes that 2:2 complex formation is not exclusive to Fab fragments and occurs readily in the context of full-length RTX.

Unlike OBZ or the RTX Fab, RTX IgG cross-links CD20 into cyclical superstructures of 2-to-2 or 3-to-3 IgGs and CD20s, with approximate diameters of 250 and 300 Å respectively (Figure 5A). These closed-ring assemblies feature CD20 dimers and Fc domains splayed outwards, linked by pairs of Fab arms which position the Fab-Fc hinges of RTX on a circle of approximate diameter 100 Å (Figure 5A,B). Because of the striking similarity with our 3D structure of the CD20:RTX Fab complex, we were able to generate a model of these rings as present on the nsEM grids (Figure 5B,C). To understand how these assemblies might relate to RTX function, we endeavored to build a model of an CD20:RTX IgG assembly as it might occur on a cell. This was achieved by rotating each CD20:Fab complex 90 degrees, while keeping the ends of the Fab domains in close proximity to each other (Figure 5D, left). This modeled assembly exhibits precisely the dimensions that would be required for the three Fc domains to “fold in” (Figure 5D, middle) and potentially nucleate assembly of a six-membered Fc platform such as those observed in structures of the complement component C1 in complex with Fc (*26*). The resulting model of a CD20:RTX:C1 complex (Figure 5D, right) provides a structural hypothesis for how Fab:Fab and Fab:CD20 interactions may lay the molecular foundations that promote tight CD20 clustering and complement recruitment, the hallmarks of RTX.

Discussion

CD20 is a clinically-validated target for the treatment of lymphomas and auto-immune diseases, but its structure and function have remained unknown. In contrast to the prevailing view that CD20 is a tetramer (*13*), our structural studies establish CD20 as a compact dimeric double-barrel assembly, with a protein fold that is distinct from any previously determined structure. Electrostatic surface calculations confirm that the transmembrane helices of CD20 are packed predominantly through hydrophobic and van der Waals complementary interactions. Previous reports have suggested the possibility that CD20 may form a plasma membrane ion channel, but our analyses reveal no plausible ion permeation pathway through the monomeric CD20 protomer or along the dimeric packing interface. We conclude that CD20 and other MS4A family members are unlikely to directly function as ion channels.

The dimeric organization of CD20 finally provides a molecular-level explanation for the perplexing observation that twice as many type I as type II mAbs bind $CD20^{+}$ cells (*17*). Our EM and biophysical studies using purified components establish that each CD20 dimer is bound by two type I RTX Fabs, but only one type II OBZ Fab. The two RTX Fabs are brought in close proximity due to CD20's compact symmetrical dimeric arrangement, resulting in an extensive homotypic Fab:Fab interface, which necessitates a shallow angle of approach of the Fabs. This orientation avoids steric clashes between the two RTX Fabs, whereas OBZ's steeper angle of approach (*5*) would be expected to sterically preclude another Fab from binding.

Though it has long been known that RTX promotes CD20 clustering on the cell surface, our observation of circular RTX:CD20 assemblies with a diameter similar to that required for Fc hexamer formation (Figure 5A,B) raises the possibility that RTX-induced cell-surface CD20 clusters may in fact be well-ordered assemblies specifically predisposed to recruit complement (as opposed to loose groupings on the cell surface). Assemblies of this kind could be particularly efficient at complement recruitment by virtue of their biasing of Fc domain positions and orientations towards the formation of the hexameric Fc platforms necessary for complement recruitment (*26*). In a simplistic model, we speculate that dimeric RTX:CD20 building blocks (Figure 1C) can assemble into a 3-to-3 closed-ring configuration that acts as a nucleating scaffold for IgG hexamer formation to ultimately recruit C1q (Figure 5D). One unknown or caveat in this model is that it requires the recruitment of three additional RTX IgG molecules that are not involved in the initial 3-to-3 ring but are needed to achieve Fc hexamerization. Once formed, it seems plausible that 3-to-3 ring assemblies could serve to potentiate Fc-Fc interactions by intermingling with various other RTX:CD20 superstructures (Figure 5A) or assembly intermediates that are likely found and enriched along the cell surface, ultimately leading to efficient Fc-hexamer formation and C1q engagement. Although further experiments are needed to ascertain the precise dynamics and geometrical arrangements of RTX:CD20 complexes on the cell membrane when IgG hexamers are formed, our proposed model for C1q recruitment by RTX (Figure 5D) provides an initial molecular-level hypothesis for why type I mAbs elicit potent CDC. This speculative structure-based model also suggests that CDC functionality may be shared more generally by antibodies which bind oligomeric cell-

surface targets and leave at least one epitope unencumbered and available for further mAb binding.

In the case of RTX, the simultaneous binding of both CD20 subunits is made possible by an intricate geometrical arrangement involving Fab:CD20 contacts at a secondary epitope and a large Fab:Fab interaction surface. We note that all of the RTX residues involved in this homotypic interaction are germline encoded in mice (Figure S6A). This observation suggests that the homotypic Fab:Fab interaction was inherent to the progenitor RTX B cell prior to somatic hypermutation. These residues are also highly conserved among the RTX-like type I mAbs (Figure S6B), a majority of which are also mouse-derived. Similar Fab:Fab contacts mediate crystal-packing of the isolated RTX Fab-ECL2 peptide complex (*4*) (Figure S5), which, taken together with our findings, indicates that Fab-Fab homotypic interactions are energetically favorable and an essential feature of the RTX-like type I mAbs. This raises the question of whether RTX Fabs may exist as pre-formed dimers prior to CD20 binding. We evaluated this possibility but found only weak Fab:Fab interactions at extremely high concentrations (>100 μM; data not shown), suggesting that the homotypic Fab:Fab interactions are nucleated by CD20 binding. We are aware of two other examples of Fab:Fab homotypic interaction, and in each case, these interfaces are central to antibody function: neutralization of a malaria parasite (*27*, *28*), and cross-linking-independent activation of TRAIL-R2 (*29*). Given the functional relevance of homotypic Fab:Fab interfaces in these three exemplar cases, we propose that these observations can be exploited in the discovery and optimization of next-generation therapeutic antibodies.

References and notes:

1. M. J. E. Marshall, R. J. Stopforth, M. S. Cragg, Therapeutic antibodies: What have we learnt from targeting CD20 and where are we going? *Front. Immunol.* **8** (2017), doi:10.3389/fimmu.2017.01245.
2. P. Stashenko, L. M. E. E. M. Nadler, R. Hardy, S. F. Schlossman, Characterization of a human B lymphocyte-specific antigen. *J. Immunol.* **125**, 1678–85 (1980).
3. L. E. Kuek, M. Leffler, G. A. Mackay, M. D. Hulett, The MS4A family: Counting past 1, 2 and 3. *Immunol. Cell Biol.* **94**, 11–23 (2016).
4. J. Du *et al.*, Structural basis for recognition of CD20 by therapeutic antibody Rituximab. *J. Biol. Chem.* **282**, 15073–80 (2007).
5. G. Niederfellner *et al.*, Epitope characterization and crystal structure of GA101 provide insights into the molecular basis for type I/II distinction of CD20 antibodies. *Blood*. **118**, 358–67 (2011).
6. J. K. Bubien, L. J. Zhou, P. D. Bell, R. A. Frizzell, T. F. Tedder, Transfection of the CD20 cell surface molecule into ectopic cell types generates a Ca2+ conductance found constitutively in B lymphocytes. *J. Cell Biol.* **121**, 1121–1132 (1993).
7. M. J. Polyak, J. P. Deans, Alanine-170 and proline-172 are critical determinants for extracellular CD20 epitopes; heterogeneity in the fine specificity of CD20 monoclonal antibodies is defined by additional requirements imposed by both amino acid sequence and quaternary structure. *Blood*. **99**, 3256–3262 (2002).
8. M. J. Polyak, H. Li, N. Shariat, J. P. Deans, CD20 homo-oligomers physically associate with the B cell antigen receptor: Dissociation upon receptor engagement and recruitment of phosphoproteins and calmodulin-binding proteins. *J. Biol. Chem.* **283**, 18545–18552 (2008).
9. T. W. Kuijpers *et al.*, CD20 deficiency in humans results in impaired T cell-independent antibody responses. *J. Clin. Invest.* **120**, 214–22 (2010).
10. C. A. Walshe *et al.*, Induction of cytosolic calcium flux by CD20 is dependent upon B cell antigen receptor signaling. *J. Biol. Chem.* **283**, 16971–16984 (2008).
11. D. G. Maloney *et al.*, IDEC-C2B8 (Rituximab) anti-CD20 monoclonal antibody therapy in patients with relapsed low-grade non-Hodgkin's lymphoma. *Blood*. **90**, 2188–95 (1997).
12. X. Montalban *et al.*, Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N. Engl. J. Med.* **376**, 209–220 (2017).
13. C. Klein *et al.*, Epitope interactions of monoclonal antibodies targeting CD20 and their relationship to functional properties. *MAbs*. **5**, 22–33 (2013).
14. J. P. Deans, S. M. Robbins, M. J. Polyak, J. A. Savage, Rapid redistribution of CD20 to a low density detergent-insoluble membrane compartment. *J. Biol. Chem.* **273**, 344–348 (1998).
15. W. G. Wierda *et al.*, Ofatumumab as single-agent CD20 immunotherapy in fludarabine-refractory chronic lymphocytic leukemia. *J. Clin. Oncol.* **28**, 1749–1755 (2010).
16. E. Mössner *et al.*, Increasing the efficacy of CD20 antibody therapy through the engineering of a new type II anti-CD20 antibody with enhanced direct and immune effector cell-mediated B-cell cytotoxicity. *Blood*. **115**, 4393–402 (2010).
17. H. T. C. Chan *et al.*, CD20-induced lymphoma cell death is independent of both caspases

and its redistribution into triton X-100 insoluble membrane rafts. *Cancer Res.* **63**, 5480–5489 (2003).

18. J. Du, H. Yang, Y. Guo, J. Ding, Structure of the Fab fragment of therapeutic antibody Ofatumumab provides insights into the recognition mechanism with CD20. *Mol. Immunol.* **46**, 2419–2423 (2009).
19. J. Du *et al.*, Crystal structure of chimeric antibody C2H7 Fab in complex with a CD20 peptide. *Mol. Immunol.* **45**, 2861–2868 (2008).
20. B. Zimmerman *et al.*, Crystal Structure of a Full-Length Human Tetraspanin Reveals a Cholesterol-Binding Pocket. *Cell*. **167**, 1041-1051.e11 (2016).
21. S. Nakamura *et al.*, Morphologic determinant of tight junctions revealed by claudin-3 structures. *Nat. Commun.* **10**, 1–10 (2019).
22. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, R. MacKinnon, X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature*. **415**, 287–94 (2002).
23. F. Li, J. Liu, Y. Zheng, R. M. Garavito, S. Ferguson-Miller, Protein structure. Crystal structures of translocator protein (TSPO) and mutant mimic of a human polymorphism. *Science*. **347**, 555–8 (2015).
24. T. Ramaraj, T. Angel, E. A. Dratz, A. J. Jesaitis, B. Mumey, Antigen-antibody interface properties: Composition, residue interactions, and features of 53 non-redundant structures. *Biochim. Biophys. Acta - Proteins Proteomics*. **1824**, 520–532 (2012).
25. S. Bondza *et al.*, Real-time characterization of antibody binding to receptors on living immune cells. *Front. Immunol.* **8**, 1–11 (2017).
26. D. Ugurlar *et al.*, Structures of C1-IgG1 provide insights into how danger pattern recognition activates complement. *Science*. **359**, 794–797 (2018).
27. K. Imkeller *et al.*, Antihomotypic affinity maturation improves human B cell responses against a repetitive epitope. *Science (80-.).* **360**, 1358–1362 (2018).
28. D. Oyen *et al.*, Cryo-EM structure of P . falciparum circumsporozoite protein with a vaccine-elicited antibody is stabilized by somatically mutated inter-Fab contacts (2018).
29. T. Tamada *et al.*, TRAIL-R2 Superoligomerization Induced by Human Monoclonal Agonistic Antibody KMTR2. *Sci. Rep.* **5**, 17936 (2016).
30. L. C. Simmons *et al.*, Expression of full-length immunoglobulins in Escherichia coli: Rapid and efficient production of aglycosylated antibodies. *J. Immunol. Methods*. **263**, 133–147 (2002).
31. G. B. Fields, R. L. Noble, Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* **35**, 161–214 (1990).
32. T. Grant, A. Rohou, N. Grigorieff, cisTEM, User-friendly software for single-particle image processing. *Elife*. **7** (2018), doi:10.7554/eLife.35383.
33. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
34. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
35. T. I. Croll, ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 519–530 (2018).

36. P. V Afonine *et al.*, Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D, Struct. Biol.* **74**, 531–544 (2018).
37. R. Norel, S. L. Lin, H. J. Wolfson, R. Nussinov, Shape complementarity at protein–protein interfaces. *Biopolymers*. **34**, 933–940 (1994).
38. T. D. Goddard *et al.*, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
39. L. Holm, Benchmarking Fold Detection by DaliLite v.5. *Bioinformatics* (2019), doi:10.1093/bioinformatics/btz536.
40. F. Madeira *et al.*, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
41. G. Cardone, J. B. Heymann, A. C. Steven, One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* **184**, 226–236 (2013).

Acknowledgements

Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH R01-GM129325 and P41-GM103311. Thanks to the entire Biomolecular resources (BMR) department at Genentech for support throughout the project. Thanks to the Peptide Synthesis group in the department of Early Discovery Biochemistry at Genentech. **Funding:** T.C. was supported by a Principal Research Fellowship from the Wellcome Trust to Randy Read (grant number 209407/Z/17/Z) and gratefully acknowledges the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. All other authors are employees of Genentech, Inc. **Author contributions:** Project conceptualization: C.Ko., A.R.; CD20 construct design: C.Ko., L.R.; CD20 protein production: C.Ku., C.T., E.K.; Electron microscopy: A.E., C.A., C.C.; EM image processing: L.R., A.R.; Model building and refinement: A.R., T.C.; Structure analysis: A.R., J.P., J.T.K., L.R.; Antibody protein production, flow cytometry and CDC assays: N.C., G.N.; Antibody mutant design and selection: A.R., J.T.K., G.N., L.R.; Manuscript writing: A.R., J.P., J.T.K.; SPR: M.S.; CD20 purification, EM sample preparation, BLI, DSF, SEC-MALS: L.R.; J.P., G.N., J.T.K. and A.R. are co-senior authors; A.R. supervised the project. **Competing interests:** all authors except T.C. are employees of Genentech, Inc. **Data and materials availability:** Materials will be made available upon request and material transfer agreement with Genentech. Accession numbers of the CD20:RTX Fab model and map are PDB 6VJA and EMD-21212, respectively.

Supplementary Materials:

Materials and Methods
Tables S1-S4
Figures S1-S7
References 30-40

Figure legends

Figure 1. **Characterization of CD20:Fab complexes and cryoEM structure of CD20:RTX Fab.** **A.** Biolayer interferometry (BLI) traces. A serial dilution of RTX (left) or OBZ (right) Fabs was flowed in for the first 1,500 seconds of the experiments, followed by a dissociation step. **B.** Negative-stain EM (nsEM) of expressed, solubilized and purified CD20 in complex with RTX (left) or OBZ (right) Fab. Scale bar: 50 Å. **C.** CryoEM reconstruction of the CD20:RTX Fab complex, at a resolution of 3.3 Å. Left panel: isosurface rendering, with the GDN micelle rendered in transparent grey, the RTX Fab heavy chain in purple and the light chain in pink. Right panel: two orthogonal side views (along the plane of the membrane) of a ribbon rendering of the structure.

Figure 2. **The CD20 dimer is a compact double square-barrel structure.** Ribbon diagrams of the CD20 structure, with RTX omitted and one of the CD20 protomers transparent, for clarity. **A.** The short loop ECL1 (red), between TM1 and TM2, is almost entirely surrounded by the first half of ECL2 (green). **B.** The core of each protomer is marked by a number of highly-conserved small (mostly glycine) residues (not shown) and a complementary set of bulkier residues shown here in space-filling representation. **C, D.** The extensive dimeric interface of CD20 involves the extracellular domain (C) as well as TM helices 1 and 4 (D).

Figure 3. **Key molecular interactions between CD20 and RTX.** Ribbon diagrams of the CD20:RTX Fab structure, with key amino acid side chains involved in CD20:RTX or RTX:RTX shown in stick representation. **A.** Top view of the center of the complex, where HC.Tyr97 mediates Fab:Fab and Fab:CD20' contacts. **B.** The canonical RTX epitope ^{170}ANPSE^{174}, in addition to being recognized by RTX's heavy chain (left, purple), is also involved in a hydrogen bond network with Tyr97' from the distal RTX Fab (right, grey). **C.** Additional Fab:Fab contacts between heavy chain (HC) loops H3 and H1 and light chain (LC) loop L2. **D.** A secondary epitope consisting of ECL1 and ECL2 is contacted by RTX's LC loop 1.

Figure 4. **Multiple CD20:RTX and RTX:RTX interactions enable cell binding and CDC. A.** Surface representation of the CD20:RTX Fab complex. The surfaces buried by complex formation are colored in yellow (CD20:RTX), orange (CD20':RTX'), or red (RTX:RTX'). **B.** Open-book representation of the same surfaces, with surface area measurements for each buried surface indicated. Residues mutated as part of this study are labeled in the top left panel and are not involved in the primary paratope. The primary epitope (dashed line) only accounts for less than half (440 Å^2 out of 1,015 Å^2) of the total Fab binding area. **C.** Surface representation of the CDR face of RTX Fab, with the ECL2 turret epitope shown (green), as well as the positions of the point mutations under study. **D,E.** Complement-dependent cytotoxicity (**D**) and cell binding (**E**) of RTX mutants are plotted as a function of antibody concentration, and compared to wild-type RTX and OBZ.

Figure 5. **RTX cross-links CD20 into circular super-assemblies. A.** Average nsEM images of CD20 incubated with full-length RTX show cyclical higher-order structures of involving 2-to-2 (top row; diameter of 250 Å) or 3-to-3 (bottom row; diameter of 300 Å) CD20-to-RTX

complexes. The RTX Fc domains appear disordered, presumably because of IgG hinge flexibility. **B,C.** Interpretation of an nsEM class average of a 3-to-3 assembly. **D.** Proposed model for CD20:RTX super-complex formation and complement recruitment. During nsEM experiments, the IgGs and solubilized CD20s are co-planar (C). Modeling these high-order assemblies as they might occur at the surface of $CD20^+$ cells requires rotating the CD20:Fab complexes 90 degrees (D, left). Given the flexibility provided by the IgG hinges, it is then possible to position Fc domains (pink) in a common plane (D, middle). The addition of three further Fc domains possibly contributed by neighboring CD20:IgG assemblies (grey) would complete the Fc hexamer formation and enable recruitment of C1q (D, right). Dashed lines outline IgG molecules. Models used: structure from present work (RTX Fab:CD20 complex), EMDB-4232 (EM map of C1:Fc complex) and PDB 6FCZ (Fc domains and C1q head domains) (*26*).

Supplementary Materials for

Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab

Lionel Rougé[1], Nancy Chiang[2], Micah Steffek[3], Christine Kugel[4], Tristan Croll[5], Christine Tam[4], Alberto Estevez[1], Christopher Arthur[1], Christopher Koth[1], Claudio Ciferri[1], Edward Kraft[4], Jian Payandeh[1]*, Gerald Nakamura[2]*, JT Koerber[2]*, Alexis Rohou[1]*

*Correspondence to: rohou.alexis@gene.com, koerber.james@gene.com, nakamura.gerald@gene.com, payandeh.jian@gene.com

This PDF file includes:

Materials and Methods
Tables S1-S4
Figs. S1-S7
References 30-40

Materials and Methods

Recombinant expression of CD20 constructs

All CD20 human constructs were synthesized (Genescript) and cloned into a modified pAcGP67A vector downstream of the polyhedron promoter. All mutations were performed using Quickchange mutagenesis kit (Qiagen). Recombinant baculovirus was generated using the Baculogold system (BD Biosciences) following standard protocols. *Trichoplusia ni* cells were infected for protein production and harvested 48 hrs post-infection.

Sequence of the construct for structural studies is (the underlined sequence is the 6-His affinity tag and TEV protease cleavage site):

MGSTQSFFMRESKTLGAVQIMNGLFHIALGGLLMIPAGIYAPICVTVWYPLWGGIMYIISGSLLAATEKNSRK
CLVKGKMIMNSLSLFAAISGMILSIMDILNIKISHFLKMESLNFIRAHTPYINIYNCEPANPSEKNSPSTQYCYSI
QSLFLGILSVMLIFAFFQELVIAGIVENEWKRTCSRPKSNIVLLSAEEKKEQTIEIKEEVVGLTETSSQPKNEEDIEI
IPIQEEEEEETETNFPEPPQDQESSPIENDSSPGNS<u>ENLYFQGHHHHHHHH</u>

Sequence of the construct for binding studies is (the underlined sequence is the Flag affinity tag while the Avitag is in bold):

M<u>DYKDDDDK</u>GENLYFQS**GLNDIFEAQKIEWHE**GSMTTPRNSVNGTFPAEPMKGPIAMQSGPKPLFRRMS
SLVGPTQSFFMRESKTLGAVQIMNGLFHIALGGLLMIPAGIYAPICVTVWYPLWGGIMYIISGSLLAATEKNSR
KCLVKGKMIMNSLSLFAAISGMILSIMDILNIKISHFLKMESLNFIRAHTPYINIYNCEPANPSEKNSPSTQYCYSI
QSLFLGILSVMLIFAFFQELVIAGIVENEWKRTCSRPKSNIVLLSAEEKKEQTIEIKEEVVGLTETSSQPKNEEDIEI
IPIQEEEEEETETNFPEPPQDQESSPIENDSSPGNS

Purification of CD20 for structural studies

Trichoplusia ni cells expressing CD20 were resuspended in 25 mM Tris pH 7.5, 150 mM NaCl (Buffer A) supplemented with cOmplete EDTA-free protease-inhibitor cocktail tablets (Roche). Cell suspension was passed through a microfluidizer once at a pressure setting of 15,000 psi. Following cell lysis, suspension was spun down at 8,000 rpm for 20 min. Supernatant was then spun down 40,000 rpm at 4 °C for 1 hour. Isolated membranes were resuspended into Buffer A supplemented with 1% (w/v) GDN (Anatrace) and 0.2% (w/v) cholesterol hemisuccinate and solubilization was carried out with gentle agitation for 2 hrs at 4 °C. After ultracentrifugation at 40,000 rpm at 4 °C for 30 min, clarified supernatant was mixed gently with Ni-NTA resin (Sigma) pre-equilibrated with Buffer B (25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN) for 1 hour at 4 °C. Ni-NTA resin was collected by gravity flow and washed with 5 column volumes of Buffer B supplemented with 45 mM imidazole. His-tagged CD20 was eluted with Buffer C (25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN and 250 mM imidazole). Eluate was passed over a Superdex S200 10/300 GL column (GE Healthcare) equilibrated in Buffer B and peak fractions were collected and incubated overnight with His-tag TEV protease (prepared in-house). Cleaved mixture was then passed over a fresh Ni-NTA column equilibrated with Buffer B and column was washed and eluted as described above. Flow through and wash fractions (containing untagged CD20) were collected, concentrated using an Amicon ULtra-15 Centrifugal Filter Units (100K MWCO, Millipore Sigma) and loaded onto a Superose 6 10/300 GL column (GE Healthcare) pre-equilibrated with Buffer B.

Purification of CD20 for binding studies

Trichoplusia ni cells were co-infected with virus expressing full length CD20 (containing a C-terminal Flag followed by an Avi-tag) along with a virus encoding for the protein BirA. Cells were resuspended in 25 mM Tris pH 7.5, 150 mM NaCl (Buffer A) supplemented with cOmplete EDTA-free protease-inhibitor cocktail tablets (Roche). Cell suspension was passed through a microfluidizer once at a pressure setting of 15,000 psi. Following cell lysis, suspension was spun down at 8,000 rpm for 20 min. Supernatant was then spun down 40,000 rpm at 4° for 1 hour. Isolated membranes were resuspended into Buffer A supplemented with 1% (w/v) GDN (Anatrace) and 0.2% (w/v) cholesterol hemisuccinate and solubilization was carried out with gentle agitation for 2 hrs at 4°C. After ultracentrifugation at 40,000 rpm at 4°C for 30 min, clarified supernatant was mixed gently with anti-M2 FLAG resin (Sigma) pre-equilibrated with Buffer B (25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN) for 1 hr at 4°C. Flag resin was collected by gravity flow and washed with 5 column volumes of Buffer B. Flag-tagged CD20 was eluted with Buffer C (25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN and 0.15 mg/mL Flag peptide. Eluate was passed over a Superose 10/300 GL column (GE Healthcare) equilibrated in Buffer B and peak fractions were collected.

Expression and purification of recombinant antibodies and fragments

Constructs suitable for periplasmic expression of Fab in *Escherichia coli* and containing sequences coding for either Fab fragments of RTX or OBZ were cloned; they were transformed into 34B8 *E. coli* cells and expressed at 30°C under control of the phoA promoter in CRAP phosphate-limiting autoinduction medium (*30*) supplemented with carbenicillin (50 µg/mL). After 24 hr, cells were harvested and resuspended in PBS supplemented with one complete EDTA-free Protease Inhibitor Cocktail tablet (Roche) per 50 mL of lysis buffer, lysozyme (0.1 mg/mL), and benzonase (0.01 mg/mL). The prepared suspension was microfluidized at 15,000 psi and clarified at 50,000 x g for 30 min at 4°C. The supernatant was then resolved on protein G Sepharose beads equilibrated with PBS, using 2 mL packed resin volume per original gram of cell paste. The column was washed extensively with PBS and Fabs were eluted under mildly acidic conditions (0.56 % glacial acetic acid pH 3.6). Eluted Fabs were immediately dialyzed overnight at 4°C against buffer containing 500 mM NaCl, 10% glycerol, and 100 mM Tris (pH 8.0). Fabs was further resolved on an S75 16/60 gel filtration column (GE Healthcare) using PBS (pH 7.2) as the running buffer.

Constructs for mammalian expression of IgGs were generated by gene synthesis. Plasmids encoding for the LC and HC were co-transfected into 293 cells and purified with affinity chromatography followed by SEC using standard methods (MabSelect SuRe; GE Healthcare, Piscataway, NJ, USA).

CDC assay

A complement dependent cytotoxicity (CDC) assay was set up to measure the degree of antibody dependent complement lysis of target cells. Human serum complement protein C1q, that binds to the Fc domain of an antibody bound to a target cell, was used to trigger the initiation of the complement cascade. This action eventually culminates in the formation of the complement protein membrane attack complex resulting in target cell lysis. The assay is

performed in a 96 microwell plate format and in duplicate/triplicate as follows. 50 µL of serially diluted (1:3) anti-CD20 mAbs and controls starting at 900 nM were incubated with 50 µL of B-cell lymphoma WILS-2 (ATCC CRL-8885) cells (50,000) and with 50 µL of a 1:2 dilution of normal human serum complement (Quidel, Santa Clara, CA). After a 2-hour incubation at 37°C, 15 µL of Alamar Blue (Biosource International, Camarillo, CA) was added and incubated for an additional 18 hours at 37 C. The plates were briefly shaken for 15 minutes and then read on a fluorescent plate reader (Ext. 535 nm, Emt 590 nm) to determine the relative fluorescent units (RFU). The RFU value observed was plotted relative to concentration of mAb in KaleidaGraph (Synergy Software, Reading, PA). Curves are plotted using a 4-parameter fit.

Flow cytometry assay

The relative degree of anti-CD20 mAbs binding to CD20 expressing B-cell lymphoma cell lines was measured by flow cytometry. Antibodies were serially diluted 1:2 in phosphate buffered saline containing 1% of heat inactivated fetal bovine serum (VWR Int. Radnor, PA). 50 µL of the diluted mAbs was dispensed into U-bottom 96-well plates. The concentrations ranged from 200 nM to 0.82 nM. Next, each well received 100,000 WIL2-S (ATCC CRL-8885) B-cell lymphoma cells in a 50 ul volume. The plates were gently mixed and incubated on ice. Following a 1-hr incubation, the cells were pelleted and washed 5 times in cold PBS. 1 ug of a FITC-conjugated goat anti-human Fc specific $F(ab')^2$ (Jackson Immunoresearch, West Grove, PA) was added to each well. After a 30-minute incubation on ice, the cells were pelleted, washed in cold PBS 3-times and then fixed in PBS containing 4% paraformaldehyde (ThermoScientific, Rockford, IL). Cells were analyzed using a BD Biosciences LSRFortessa (San Jose, CA) to determine the geometric MFI (median, fluorescence intensity) of antibody binding to cells.

BLI binding studies

Biotinylated CD20 was captured on streptavidin SA biosensors (FortéBio / Molecular Devices) at 10 µg/mL for 180 seconds and washed to remove any unbound material before conducting measurements in Buffer B supplemented with 0.5 mg/mL BSA. The kinetic data were fit to a simple 1:1 binding model to determine the affinity constant (K_D) using the association (K_{on}) and dissociation (K_{off}) rates. Binding assays were performed in triplicate and average KD values reported with SDs. Buffer containing 0.02% GDN supplemented with 0.002% CHS was used for all washes and dilutions. Sensorgrams were normalized to a reference well containing only buffer.

SPR binding studies

Binding studies were conducted using on a 8K Biacore instrument by immobilizing onto a Sensor Chip C1 (GE LifeSciences), which had been immobilized with 500-800 RU of neutravidin. The N-terminal avi-tagged CD20 was immobilized in the presence of 50 mM HEPES pH 7.5, 150 mM NaCl, 0.2% GDN to a density ~40 RU. A amino-terminally biotinylated (PEGx6) linear peptide with the CD20 epitope (^{163}NIYNCEPANPSEKNSPSTQYCYSIQ^{187}), which had been synthesized using standard solid-phase Fmoc chemistry and purified with reverse phase HPLC (*31*), was immobilized under similar conditions to a density of ~15 RU. The full length IgG and Fab fragments were added using single cycle kinetic methods with 100 second associations at

20 °C and 500 s dissociation phase at 50 µL/min depending. The data was referenced to an un-immobilized neutravidin surface and a blank injection. Non-specific binding to the chip surface was subtracted from each curve. Data was analyzed using Biacore Evaluation software and curves were fit using 1:1 and 2-state kinetic fits.

Complex formation

Recombinant CD20 was incubated with a molar excess of IgG or Fab and incubated on ice for 30 min. Reaction was passed over a Superose 6 10/300 GL column (GE Healthcare) equilibrated in 25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN and peak fractions were collected.

Negative stain analysis

Complexes were diluted to 0.01 mg/mL in buffer B and 4 µL of the solution was placed onto a CF2/1 300 grid from Electron Microscopy Sciences (EMS) previously cleaned using a GloQube glow discharge (Quorum Technologies) system. After 30 sec incubation, sample was blotted with Whatman paper, dipped three times into 30 µL of distilled water and stained twice into 30 µL of a filtered solution of 2% uranyl acetate. Grids were imaged using a 200 KV Talos F200c with a CETA camera (ThermoFischer Scientific) and 2D class averages were computed using cisTEM (*32*).

Differential scanning fluorimetry

Melting experiments were conducted on a Prometheus NT48 (NanoTemper technologies) by measuring the tryptophan fluorescence 330/350 nm ratio of protein samples concentrated at 0.3 mg/mL in a standard capillary.

SEC-MALS analysis

Purified proteins and complexes were run on a Waters XBridge BEH S200 A gel filtration column equilibrated in 25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN in line with a Dawn HELEOS II (Wyatt Technologies) light scattering detector connected to a Wyatt OptiLab rEX refractive index detector. Wyatt Technologies software (ASTRA) was used to determine the corresponding peaks' molecular weight based on the refractive index. In order to determine the unbiased molecular weight of the apo CD20 protein, the protein conjugate analysis within the ASTRA software was performed. Because the dn/dc value for this detergent is unknown, a value of 0.143 mL/gm was arbitrary used to calculate the molecular weight of the GDN micelle (dn/dc values for most detergents is usually comprised between 0.1 and 0.15 ml/gm).

Cryo-EM sample preparation and data acquisition

Recombinant CD20 was incubated with Rituxan Fab at 1:1.2 molar ratio and incubated on ice for 30 min. Mixture was injected over a Superose 6 Increase 3.2/300 column (GE Healthcare) equilibrated in 25 mM Tris pH 7.5, 150 mM NaCl, 0.01% GDN. 3.5 µL of the peak fraction of CD20/Rituxan Fab complex at a concentration of 2 mg/mL was applied to a glow-discharged C-flat holey carbon grid (CF-1.2/1.3-2C, Electron Microscopy Sciences) coated with a thin layer of gold. Grids were blotted in Vitrobot Mark IV (ThermoFisher Scientific) using 5 s blotting time with 100% humidity, and then plunge-frozen in liquid ethane cooled by liquid nitrogen. A total

of 24,743 movie stacks were collected with SerialEM (*33*) on a Titan Krios (Thermo Fisher Scientific) operated at 300 kV and equipped with a BioQuantum energy filter operated with a 20eV energy slit with a K2 Summit direct electron detector camera (Gatan). Images were recorded at a nominal magnification of 165000 x, corresponding to a pixel size of 0.849 Å per pixel. Each image stack contains 40 frames recorded every 0.25 s giving an accumulated dose of 53 electrons/Å² and a total exposure time of 10 s. Images were recorded with a set defocus range of 0.5 to 1.5 µm.

Structure determination

All image processing was performed with cisTEM (*32*). For the 3D structure of CD20 bound by RTX Fab, the processing scheme was as depicted in Figure S3C. For atomic model building, PDB 2OSL was used to place the ECL2 peptide epitope and an initial model of CD20 was built by extending this manually in Coot (*34*). The model for the RTX Fab was started from PDB 4KAQ. This complete model was then rebuilt using interactive molecular dynamics (*35*), and refined in real space (*36*).

Structure analysis

Protein-protein shape complementarity was calculated using the program SC (*37*). Figure preparation and computation of buried solvent-accessible surface area was done with ChimeraX (*38*). 3D homology structural alignments were performed using the DALI server (*39*) and sequence alignments was using EMBL-EBI's Clustal Omega Multiple Sequence Alignment tool (*40*) followed by structure-guided manual curation.

Table S1: MALS estimates of CD20:Fab molecular weights

	Estimated molecular weight of complex (kDa)	Relative to apo CD20 (kDa)
CD20	308.7 (± 3.7 %)	0.0
CD20 + RTX Fab	386.7 (± 3.8 %)	78.0
CD20 + OBZ Fab	348.2 (± 3.6 %)	39.5

Estimated MW includes CD20's detergent GDN micelle and does not correspond to the protein molecular weight.

Table S2: Affinity of mAbs to immobilized full-length CD20

mAb	**K_D (nM)**	**Average SE (%)**	**Kinetics Chi^2 (RU^2)**
RTX	1.65	3.44	0.36
OBZ	12.33	1.01	0.10
RTX (Fab)	21.4	0.55	0.22
OBZ (Fab)	58.8	0.55	0.47
RTX HC.Y97F	24.75	3.77	1.21
RTX HC.Y97S	> 1500	N/A	N/A
RTX HC.S31E	100.35	1.82	1.01
RTX HC.G99K	468.5	0.59	0.24
RTX LC.S28D	37.95	2.93	1.21
RTX LC.S31D	136	2.79	1.17

Surface plasmon resonance measurements of mAb affinity for CD20. For each parameter, the average value (n=2) is shown. Data were fit using a two-state reaction model for all molecules tested. A 1:1 binding model did not fit well to either RTX or OBZ sensorgrams (data not shown).

Table S3: Affinity of mAbs to immobilized CD20 epitope peptide

mAb	**K_D (µM)**	**Average SE (%)**	**Kinetics Chi^2 (RU^2)**
RTX	0.642	18	0.797
OBZ	0.347	16	4
RTX HC.Y97F	1.21	9.7	0.06
RTX HC.Y97S	No binding	N/A	N/A
RTX LC.S28D	2.31	12.8	0.013
RTX LC.S31D	3.75	15.9	0.0128

Surface plasmon resonance measurements of IgG affinity for CD20 peptide were fit using a 1:1 steady state model.

Table S4: cryoEM

Data collection	
Microscope	ThermoFisher Titan Krios
Acceleration voltage	300 kV
Detector	Gatan K2
Energy filter	Gatan BioQuantum
Energy filter slit width	20 eV
Image pixel size	0.849 Å
Electron exposure	48.5 $e^-/Å^2$
Number of frames	40
Image processing	
Number of micrographs	24,743
Number of picks (“soft disc” template)	4,114,800
Number of particles used	466,362
Pixel size for processing & reconstruction	1.0 Å
Box size	310 pixels
Symmetry	C2
Number of particles for final reconstruction	155,719
Highest resolution used in refinement	5.0 Å
Global resolution (FSC = 0.143)	3.3 Å
Model building & refinement	
Software used (de novo build)	coot
Software used (refinement)	ISOLDE, Phenix
Number of amino acid residues	1,196
Model validation (Molprobity)	
MolProbity score	1.54
All-atom clashscore	4.71
Ramachandran plot	
favored	95.72 %
allowed	4.19 %
outliers	0.08 %
Rotamer outliers	0.49 %
Cβ deviations	0.72 %
Peptide plane	
cis-proline	15.25 %
cis-general	0.00 %
twisted proline	0.00 %
twisted general	0.00 %
CaBLAM outliers	1.10 %

Supplementary Figure legends

Figure S1. CD20 constructs, expression, purification and characterization. A. Topological map of CD20 constructs used in this study. The construct used for EM studies was amino-terminally truncated at residue Thr41, and carboxy-terminally His-tagged. The full-length construct was N-terminal FLAG-Avi-tagged for binding studies. **B.** Primary sequence of CD20, annotated with secondary-structure elements. Amino acid residues involved in interactions with RTX Fabs are labeled in pink (light chain) or purple (heavy chain). **C.** SEC-MALS analysis of detergent-solubilized, purified CD20 reveals a protein molecular mass of approximately 67 kDa, consistent with a CD20 dimer. Estimated molecular weight of CD20 embedded within a detergent micelle was determined by subtracting the estimated size of the GDN micelle using the "protein conjugate" feature of the ASTRA software (Wyatt Technology). An arbitrary dn/dc value of 0.143 mL/g was attributed to GDN. **D.** SEC elution profiles of CD20 alone (blue), CD20 with OBZ Fab (red), or CD20 with RTX Fab (black) show that RTX and OBZ Fabs form stable complexes with purified CD20, and that the CD20:RTX complex is of larger dimensions than the CD20:OBZ complex. **E.** BLI equilibrium values from Figure 1A are plotted as a function of Fab concentration, indicating that approximately 2 times more RTX than OBZ Fabs bind to immobilized CD20. Fitting of these data with a 1:1 binding model suggests apparent K_D values of 49 nM for RTX and 70 nM for OBZ. **F.** SDS-PAGE analysis of recombinant wild type full length CD20 and Tyr182Cys mutant under non-reducing and reducing conditions (10 mM TCEP, 5 minutes). Under native conditions, Tyr182Cys mutant runs at a molecular weight consistent with a CD20 dimer; upon addition of reducing agent, band collapses down to the equivalent of a CD20 monomer suggesting mutant was expressed as a disulfide linked dimer. Migration of wild type protein is unchanged regardless of the reducing conditions.

Figure S2. Cholesterol hemisuccinate stabilizes CD20. A. Experimental workflow of sample preparation. Full length CD20 was solubilized in 1% GDN in absence or presence of 0.2% CHS and purified over Ni-NTA column. Concentrated eluates were then diluted 100 fold into CHS free buffer (blue or orange traces, respectively) or into buffer supplemented with 0.002% CHS (red and black traces). Standard error was calculated from three technical replicates. **B.** Differential scanning fluorimetry of CD20 purified in presence or absence of cholesterol measured on NanoDSF instrument. Curves are average of three replicates. For cryoEM studies, CHS was added at the time of solubilization.

Figure S3. **CryoEM structure determination of CD20:RTX Fab. A.** Representative micrograph. **B.** Representative 2D class averages. **C.** Single-particle image processing workflow. **D.** Fourier Shell Correlation (FSC) between two half datasets yields a global resolution estimate of ~ 3.3 Å. No spatial frequencies beyond 1/5 $Å^{-1}$ were used at any point during refinement. E. Heat map representation of the distribution of assigned particle orientations shows a bias towards side views (near the periphery of the plot). **F.** Isosurface rendering of the 3D map with surface coloring according to the local resolution estimated by windowed FSCs (*41*). **G.** Map details showing side chains in stick representation and a mesh representation of

the 3D Coulomb potential map, contoured at the indicated σ level (where σ is one standard deviation above the mean of the map values).

Figure S4. The MS4A family and nearest structural homologues of CD20. A. Multi-sequence alignment of CD20 (MS4A1) and selected members of the MS4A family. Clustal conservation is denoted by asterisk (*; fully conserved residue), colon (:; strongly similar properties), period (.; weakly similar properties). **B.** Cartoon renderings of CD20 and the two closest structural homologues identified, claudin-3 and CD81. In the case of claudin-3, sequence or topology-based alignments do not give good structural overlap ("topological alignment"); a slightly better 3D alignment can be obtained by ignoring the sequence and connectivity of the TM helices ("3D TM alignment").

Figure S5. The CDR surface of RTX is also involved in mediating Fab:Fab contacts in crystallography experiments. A. Bottom and side views of RTX Fab in complex with CD20 or ECL2 peptide, as visualized with cryoEM (this work) or X-ray crystallography (*4*) respectively. For clarity and ease of comparison, the non-ECL2-turret parts of CD20 are hidden from the cryoEM structure. The structures were aligned by superposing CD20 residues 170 to 175, which constitute the core of the main RTX epitope. **B.** Heavy-chain loop 3 adopts significantly different conformations in the cryoEM and X-ray crystallography structures.

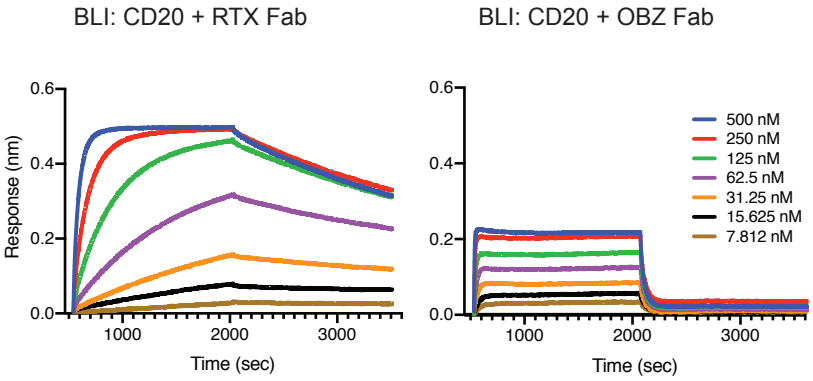
Figure S6. Sequence analysis of RTX and other anti-CD20 antibodies. A. Sequences of RTX and mouse germline. B. Multi-sequence alignment of notable anti-CD20 antibodies. Type I antibodies with overlapping ECL2 turret epitopes are shown in black. Ofatumumab, which is also a type I antibody, but binds a composite ECL1/ECL2 epitope is shown in blue. Type II mAbs are shown in purple.

Figure S7. Full-length IgG RTX and OBZ complex formation with CD20. A. Size exclusion chromatography profiles of CD20 complexed with an excess of OBZ IgG (blue trace) or RTX IgG (red). CD20:RTX elution profile suggests the presence of multiple distinct species whereas CD20:OBZ co-elution peak suggests a unique species, although slightly asymmetrical. **B.** SDS-PAGE associated to SEC profiles shown in panel A. Fractions underlined corresponds to co-elution fractions. **C**. SPR sensorgram. RTX and OBZ (red and blue respectively) were injected in increasing concentrations over surface immobilized CD20. Under saturating conditions the RTX shows a 2:1 stoichiometry compared to the OBZ binding to CD20. Injection spikes were removed for clarity. **D**. Biolayer interferometry (BLI) traces representing binding of a dilution series of Rituxan IgG to immobilized biotinylated CD20. Sensograms were normalized to a reference well containing only buffer. **E**. Biolayer interferometry (BLI) traces representing binding of a dilution series of Gazyva IgG to immobilized biotinylated CD20. Sensograms were normalized to a reference well containing only buffer. **F**. BLI equilibrium values from panels D and E are plotted as a function of IgG concentration. Fitting of these data with a 1:1 binding model suggests

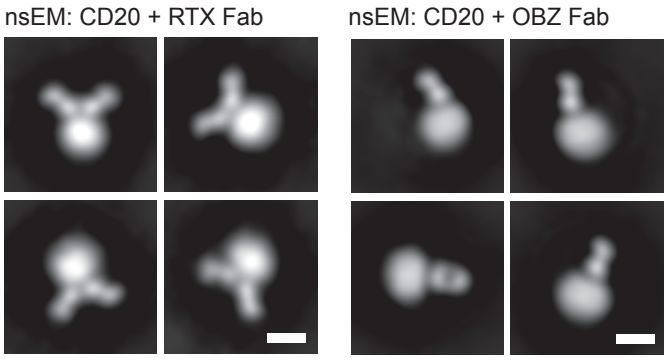
apparent KD values of 1.7 nM for RTX (red trace) and 3.3 nM for OBZ (blue traces) consistent with SPR measurements of the same (Table S2). **G**. Representative micrograph of negative-stain EM analysis of purified CD20 in complex with RTX (left) or OBZ (right) IgG (see Figure S1, panel A and B). A representative cyclical superstructure is highlighted by a white box.

Figure 1

A


BLI: CD20 + RTX Fab
BLI: CD20 + OBZ Fab
Response (nm)
0.6
0.4
0.2
0.0
1000
2000
3000
Time (sec)
500 nM
250 nM
125 nM
62.5 nM
31.25 nM
15.625 nM
7.812 nM

B


nsEM: CD20 + RTX Fab
nsEM: CD20 + OBZ Fab

C

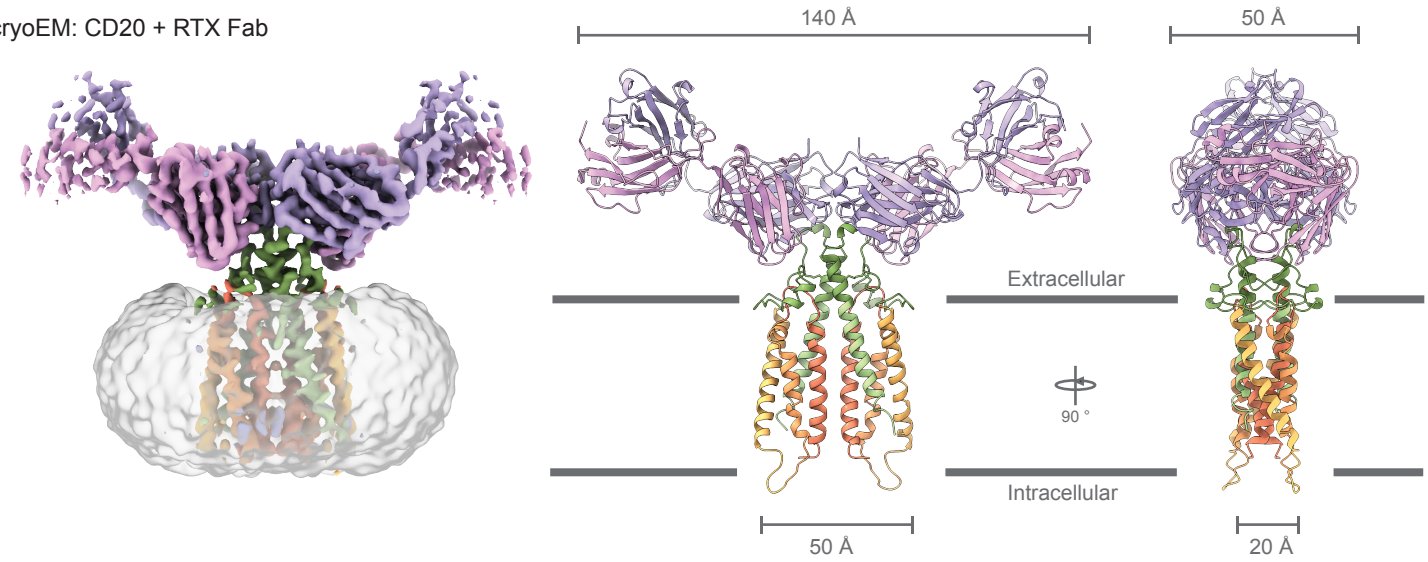

cryoEM: CD20 + RTX Fab
140 Å
50 Å
Extracellular
90 °
Intracellular
50 Å
20 Å

Figure 2

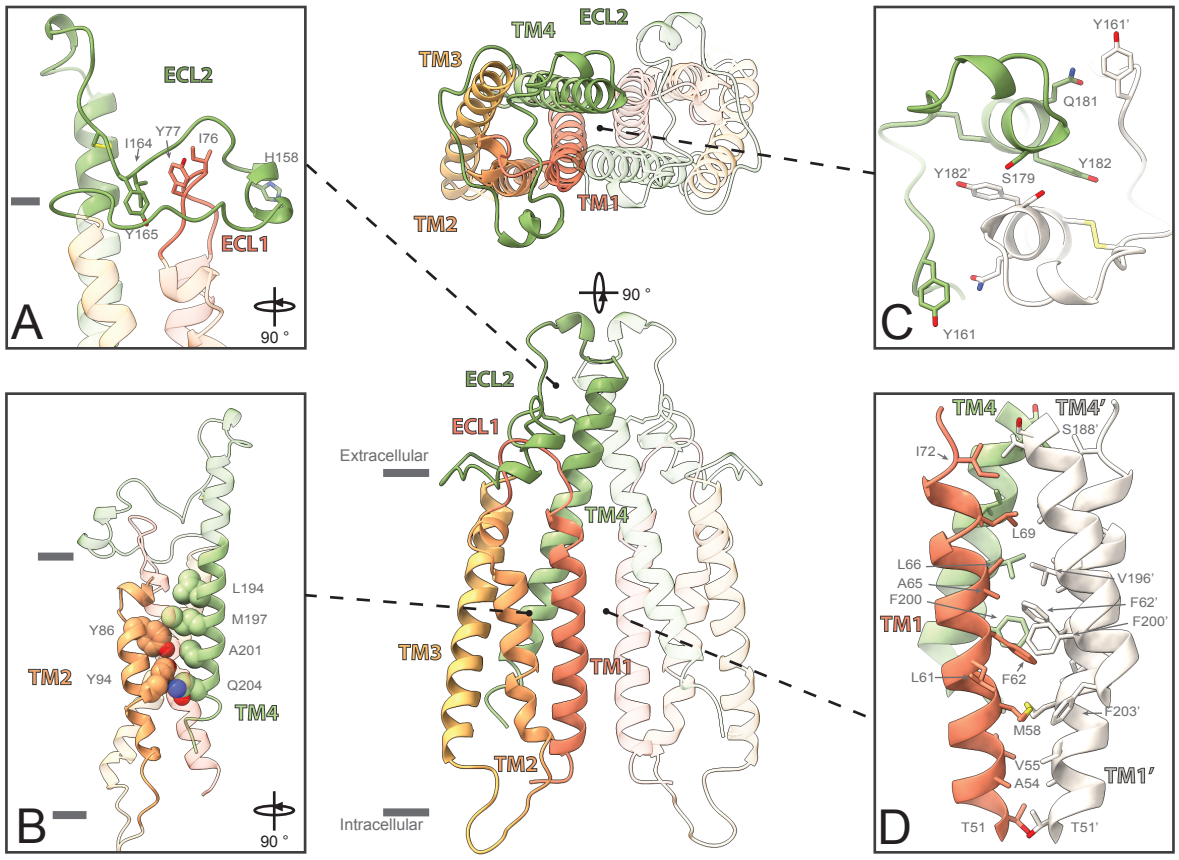

A
ECL2
Y77
I164
I76
H158
Y165
ECL1
90 °
TM4
ECL2
TM3
TM2
TM1
C
Y161'
Q181
Y182'
S179
Y182
Y161
B
L194
M197
Y86
A201
TM2
Y94
Q204
TM4
90 °
ECL2
ECL1
Extracellular
TM4
TM3
TM1
TM2
Intracellular
D
TM4
TM4'
S188'
I72
L69
L66
A65
F200
V196'
F62'
TM1
F200'
L61
F62
F203'
M58
V55
A54
TM1'
T51
T51'

Figure 3

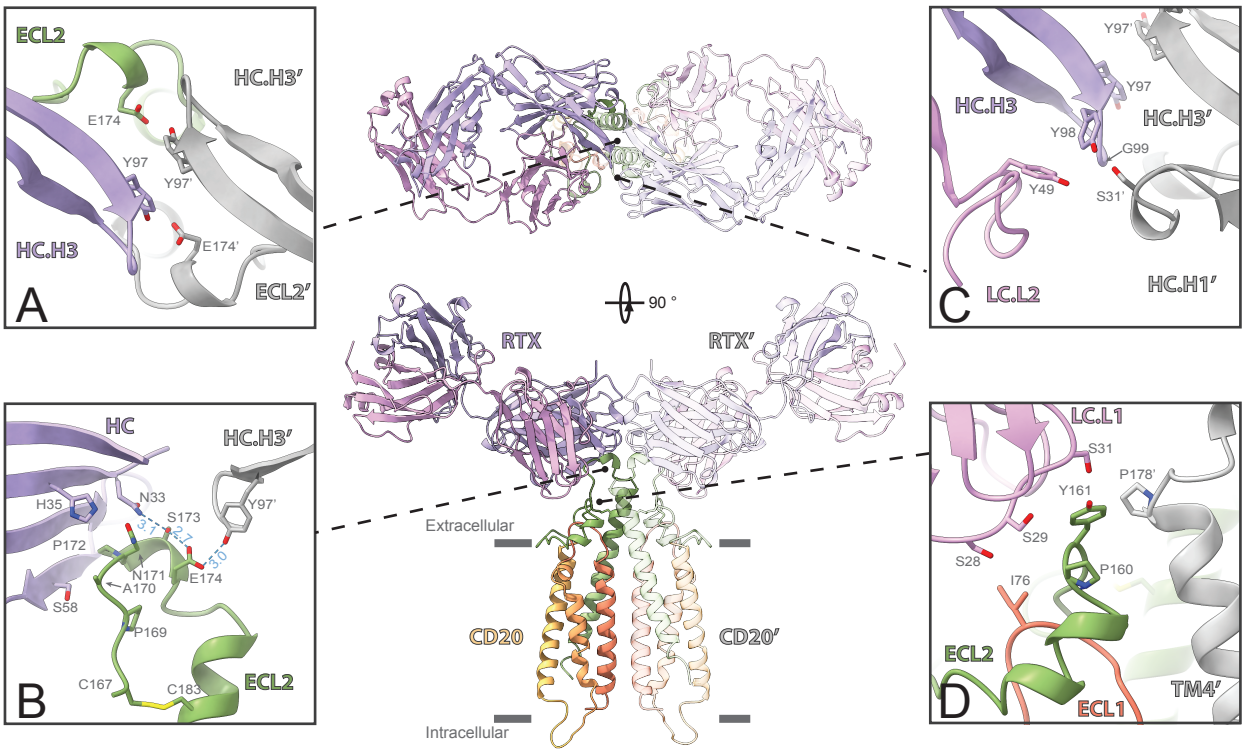

ECL2
HC.H3′
E174
Y97
Y97′
HC.H3
E174′
ECL2′
A
RTX
RTX′
90 °
Extracellular
CD20
CD20′
Intracellular
HC
HC.H3′
H35
N33
S173
Y97′
P172
N171
A170
E174
S58
P169
C167
C183
ECL2
B
Y97′
Y97
HC.H3
Y98
HC.H3′
G99
Y49
S31′
LC.L2
HC.H1′
C
LC.L1
S31
P178′
Y161
S29
S28
P160
I76
ECL2
ECL1
TM4′
D

Figure 4

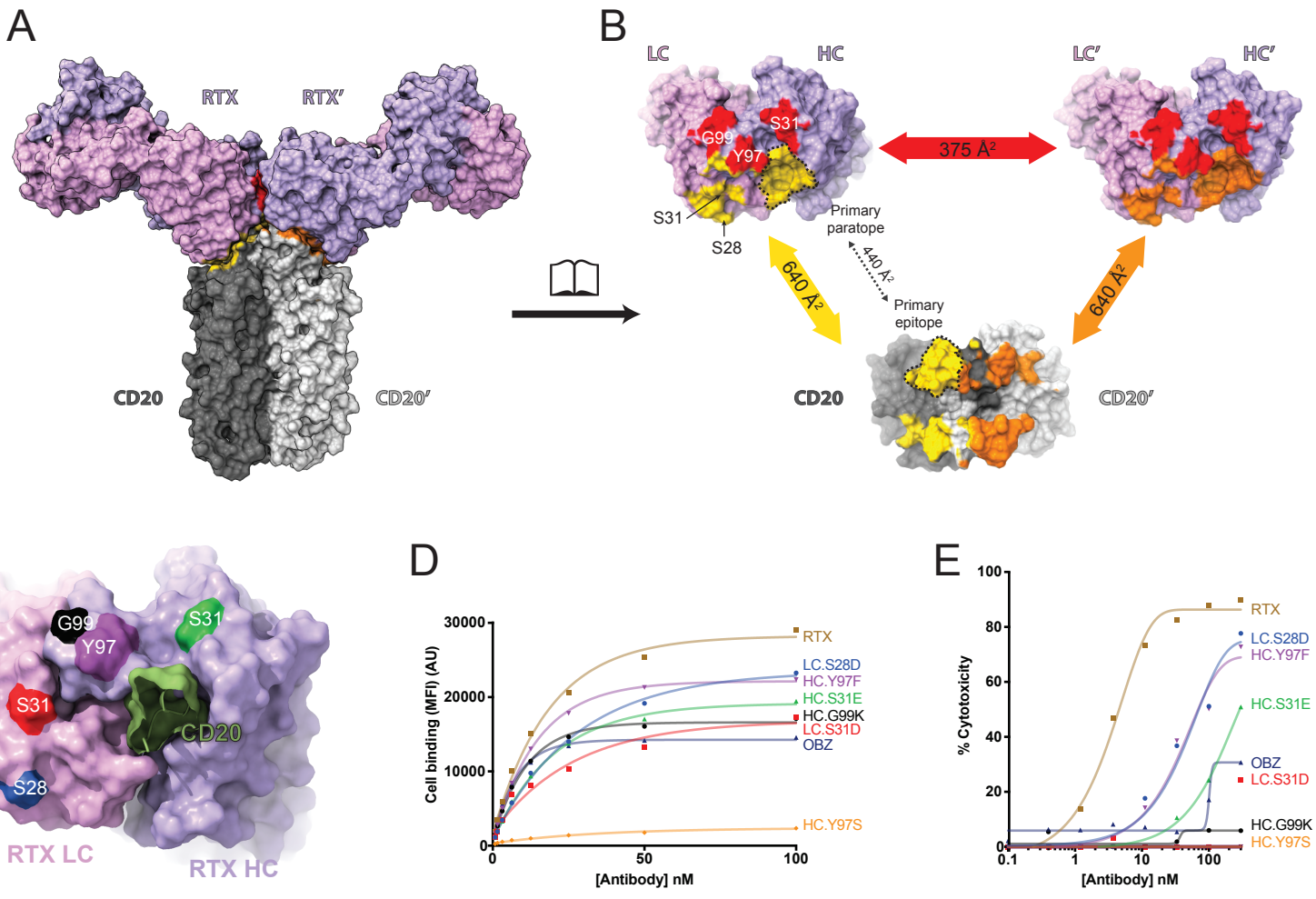

A
RTX
RTX′
CD20
CD20′
B
LC
HC
LC′
HC′
G99
Y97
S31
S31
S28
375 Å²
Primary paratope
440 Å²
Primary epitope
640 Å²
640 Å²
CD20
CD20′
C
G99
Y97
S31
S31
CD20
S28
RTX LC
RTX HC
D
Cell binding (MFI) (AU)
30000
20000
10000
0
0
50
100
[Antibody] nM
RTX
LC.S28D
HC.Y97F
HC.S31E
HC.G99K
LC.S31D
OBZ
HC.Y97S
E
% Cytotoxicity
100
80
60
40
20
0
0.1
1
10
100
[Antibody] nM
RTX
LC.S28D
HC.Y97F
HC.S31E
OBZ
LC.S31D
HC.G99K
HC.Y97S

Figure 5

Scale bars: 50 Å

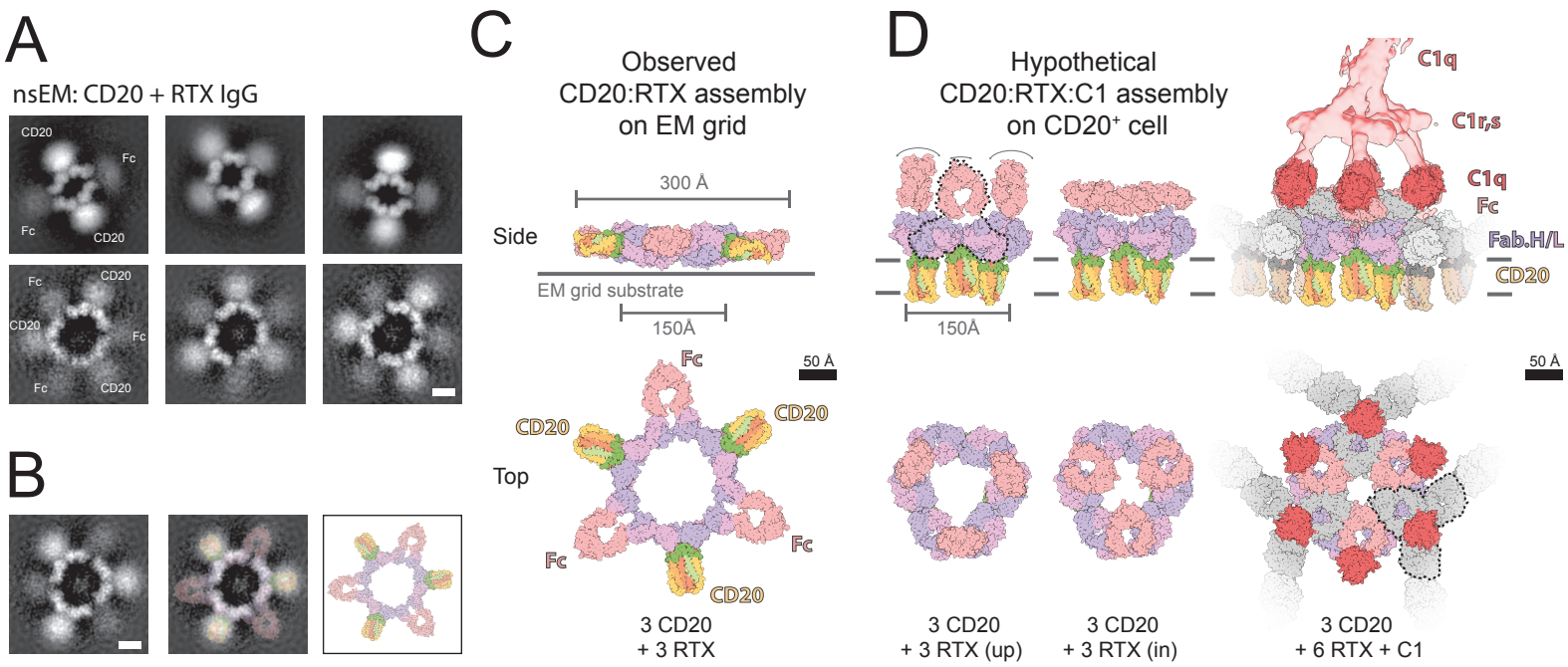


Figure S1

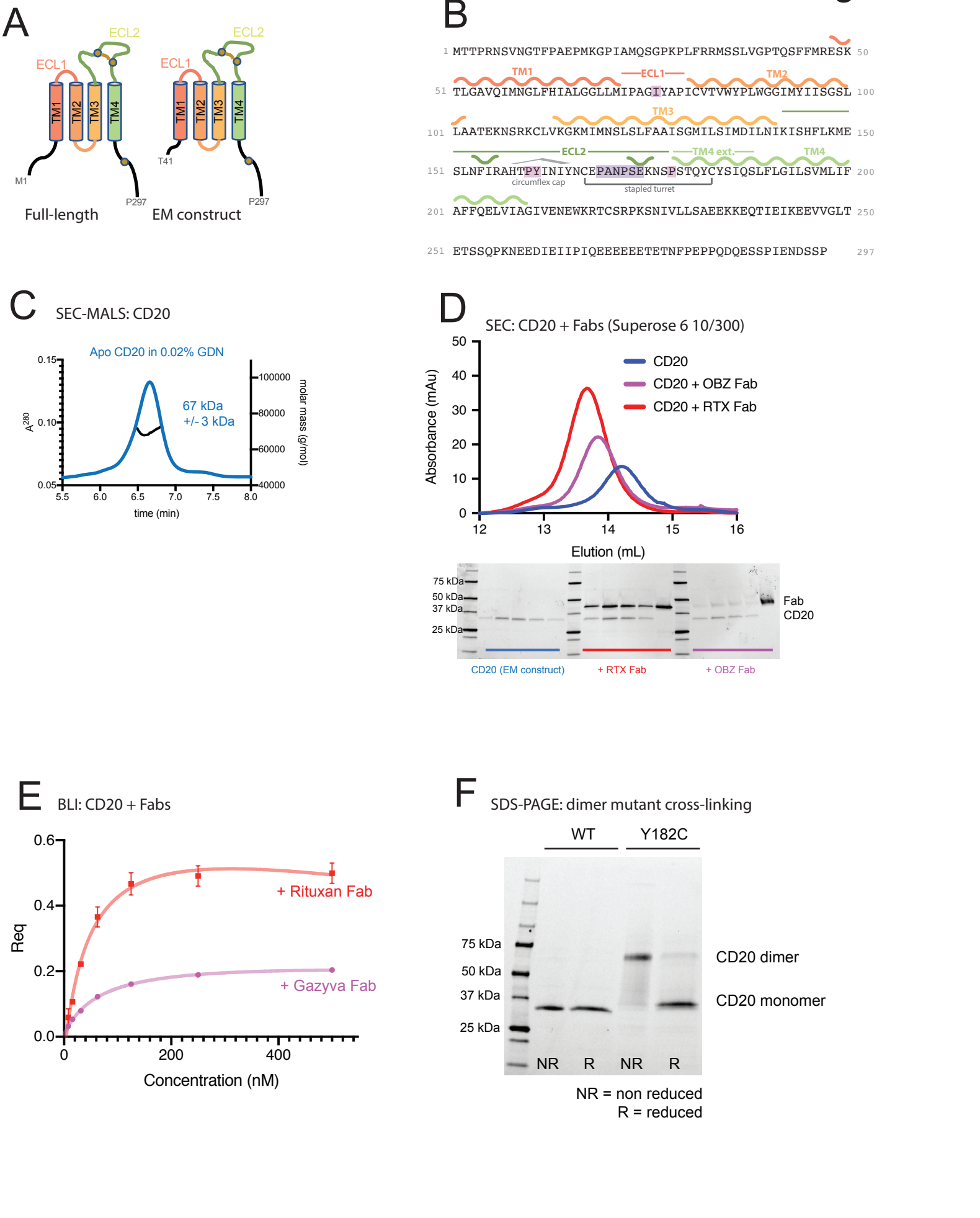
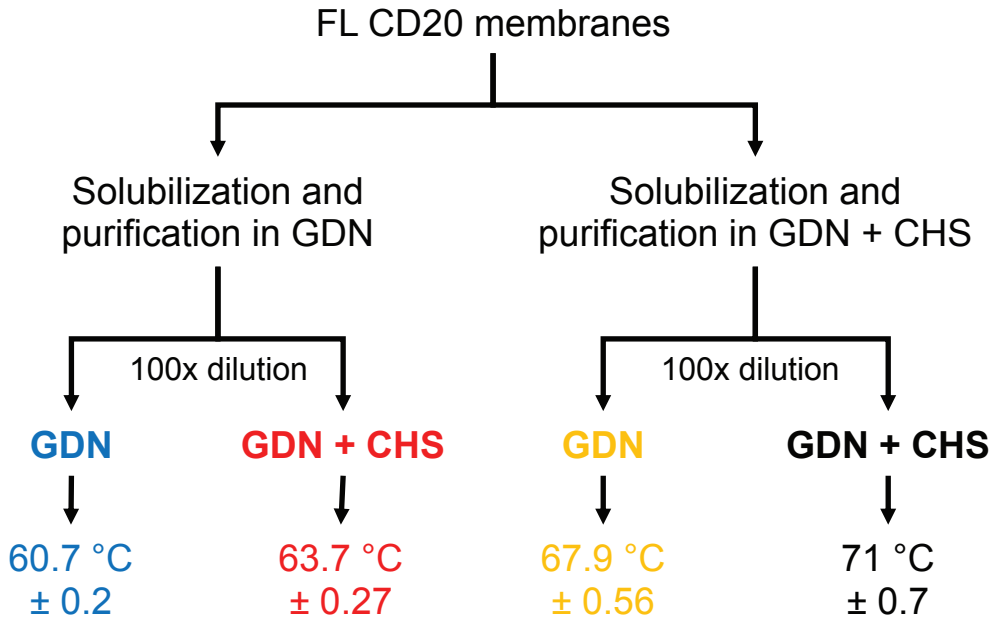

A
ECL2
ECL1
TM1
TM2
TM3
TM4
M1
P297
Full-length
T41
EM construct
B
1 MTTPRNSVNGTFPAEPMKGPIAMQSGPKPLFRRMSSLVGPTQSFFMRESK 50
TM1
ECL1
TM2
51 TLGAVQIMNGLFHIALGGLLMIPAGIYAPICVTVWYPLWGGIMYIISGSL 100
TM3
101 LAATEKNSRKCLVKGKMIMNSLSLFAAISGMILSIMDILNIKISHFLKME 150
ECL2
TM4 ext.
TM4
151 SLNFIRAHTPYINIYNCEPANPSEKNSPSTQYCYSIQSLFLGILSVMLIF 200
circumflex cap
stapled turret
201 AFFQELVIAGIVENEWKRTCSRPKSNIVLLSAEEKKEQTIEIKEEVVGLT 250
251 ETSSQPKNEEDIEIIPIQEEEEEETETNFPEPPQDQESSPIENDSSP 297
C
SEC-MALS: CD20
Apo CD20 in 0.02% GDN
67 kDa
+/- 3 kDa
A280
molar mass (g/mol)
time (min)
D
SEC: CD20 + Fabs (Superose 6 10/300)
CD20
CD20 + OBZ Fab
CD20 + RTX Fab
Absorbance (mAu)
Elution (mL)
75 kDa
50 kDa
37 kDa
25 kDa
Fab
CD20
CD20 (EM construct)
+ RTX Fab
+ OBZ Fab
E
BLI: CD20 + Fabs
+ Rituxan Fab
+ Gazyva Fab
Req
Concentration (nM)
F
SDS-PAGE: dimer mutant cross-linking
WT
Y182C
75 kDa
50 kDa
37 kDa
25 kDa
CD20 dimer
CD20 monomer
NR
R
NR
R
NR = non reduced
R = reduced

Figure S2

A


FL CD20 membranes
Solubilization and purification in GDN
Solubilization and purification in GDN + CHS
100x dilution
100x dilution
GDN
GDN + CHS
GDN
GDN + CHS
60.7 °C ± 0.2
63.7 °C ± 0.27
67.9 °C ± 0.56
71 °C ± 0.7

B

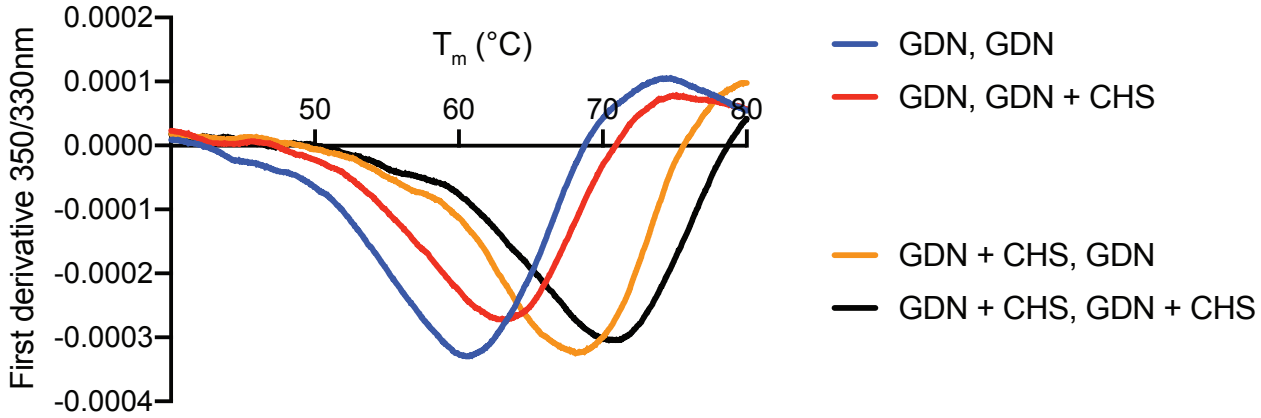
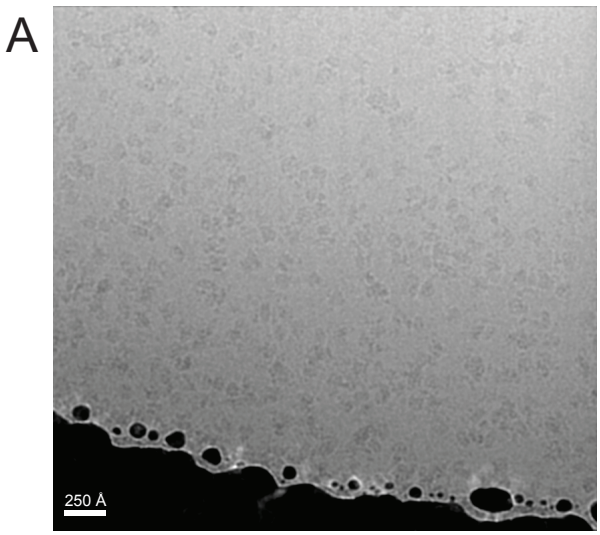
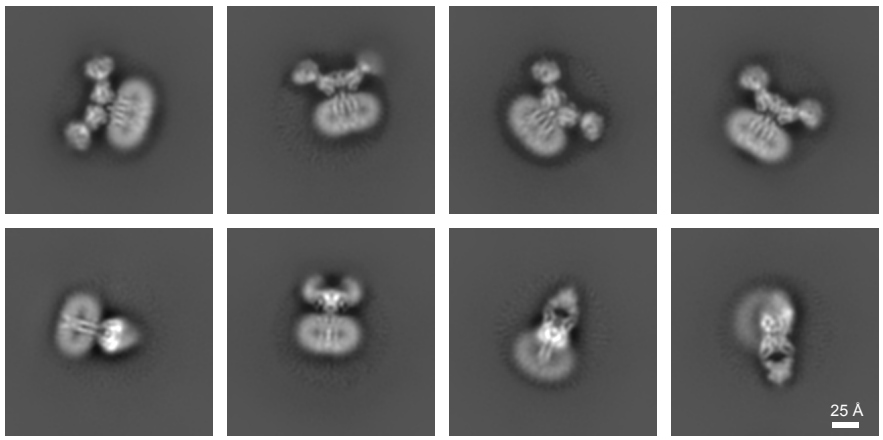
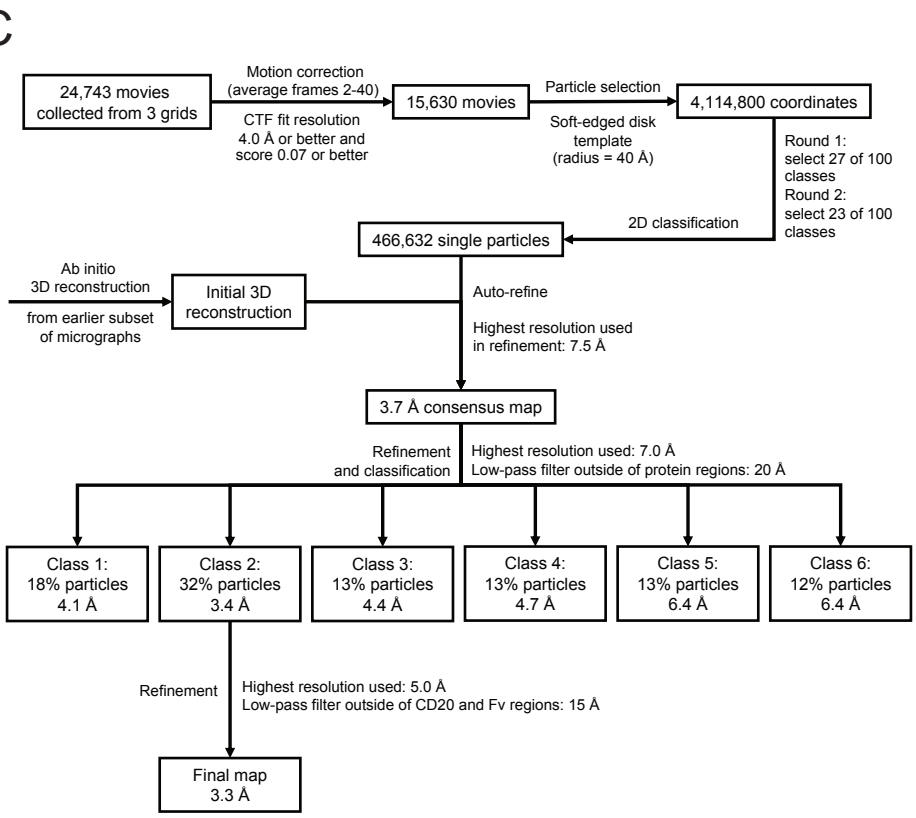
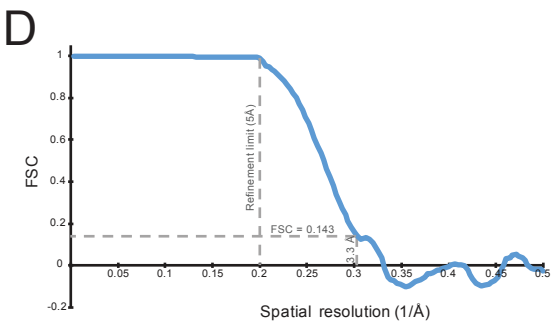

First derivative 350/330nm
0.0002
0.0001
0.0000
-0.0001
-0.0002
-0.0003
-0.0004
T_m (°C)
50
60
70
80
GDN, GDN
GDN, GDN + CHS
GDN + CHS, GDN
GDN + CHS, GDN + CHS

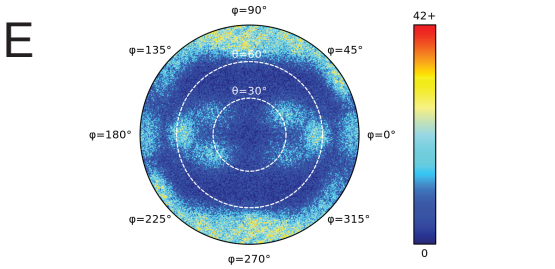
Figure S3

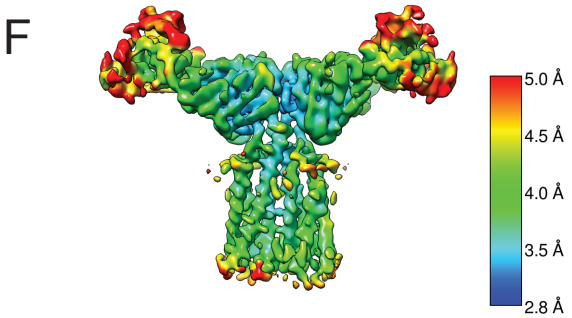

A
250 Å


B
25 Å


C
24,743 movies collected from 3 grids
Motion correction (average frames 2-40)
CTF fit resolution 4.0 Å or better and score 0.07 or better
15,630 movies
Particle selection
Soft-edged disk template (radius = 40 Å)
4,114,800 coordinates
Round 1: select 27 of 100 classes
Round 2: select 23 of 100 classes
2D classification
466,632 single particles
Ab initio 3D reconstruction from earlier subset of micrographs
Initial 3D reconstruction
Auto-refine
Highest resolution used in refinement: 7.5 Å
3.7 Å consensus map
Refinement and classification
Highest resolution used: 7.0 Å
Low-pass filter outside of protein regions: 20 Å
Class 1: 18% particles 4.1 Å
Class 2: 32% particles 3.4 Å
Class 3: 13% particles 4.4 Å
Class 4: 13% particles 4.7 Å
Class 5: 13% particles 6.4 Å
Class 6: 12% particles 6.4 Å
Refinement
Highest resolution used: 5.0 Å
Low-pass filter outside of CD20 and Fv regions: 15 Å
Final map 3.3 Å


D
FSC
Refinement limit (5Å)
FSC = 0.143
3.3 Å
Spatial resolution (1/Å)


E
φ=90°
φ=135°
φ=45°
θ=60°
θ=30°
φ=180°
φ=0°
φ=225°
φ=315°
φ=270°
42+
0


F
5.0 Å
4.5 Å
4.0 Å
3.5 Å
2.8 Å

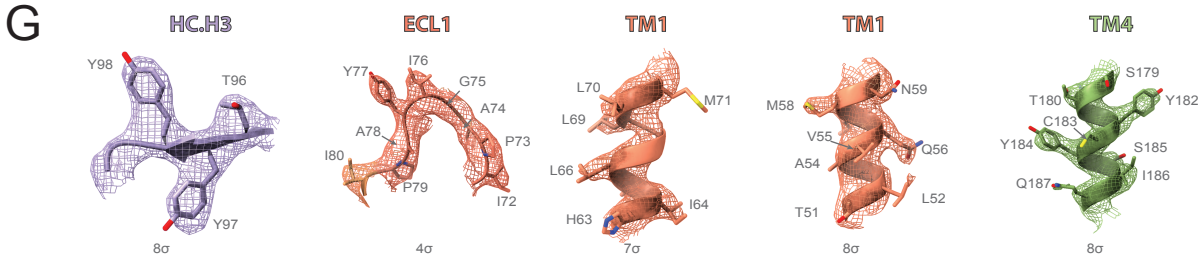
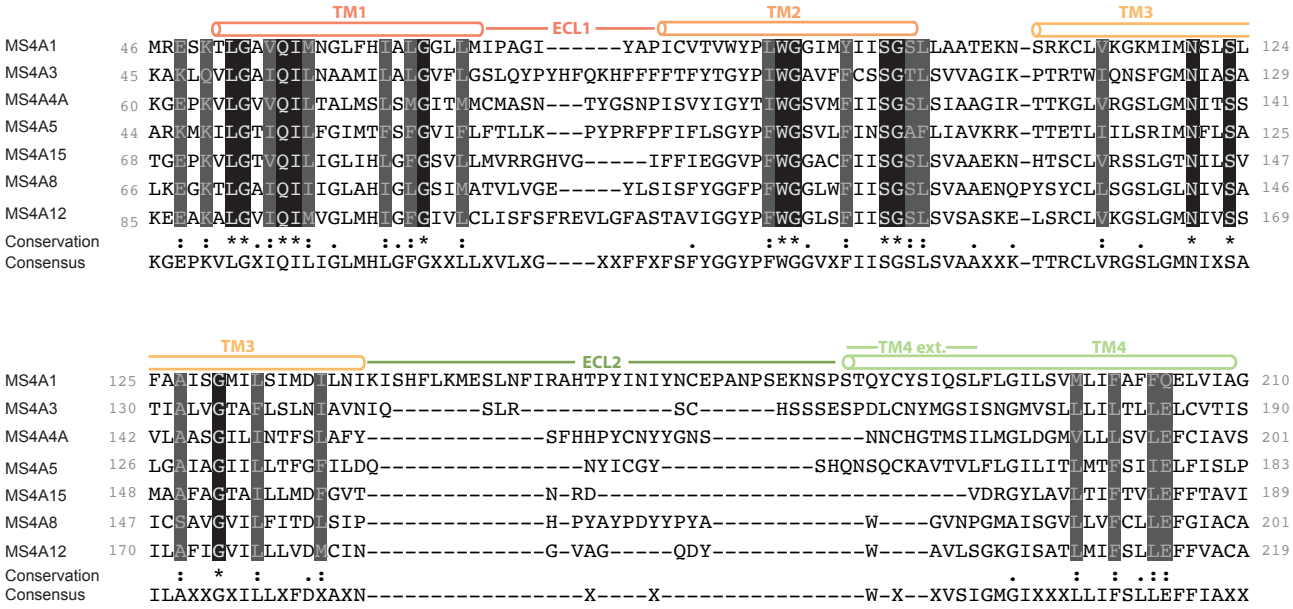

G
HC.H3
Y98
T96
Y97
8σ
ECL1
Y77
I76
G75
A74
A78
P73
I80
P79
I72
4σ
TM1
L70
M71
L69
L66
I64
H63
7σ
TM1
M58
N59
V55
Q56
A54
L52
T51
8σ
TM4
S179
T180
Y182
C183
Y184
S185
Q187
I186
8σ

Figure S4

A


TM1
ECL1
TM2
TM3
MS4A1 46 MRESKTLGAVQIMNGLFHIALGGLLMIPAGI------YAPICVTVWYPLWGGIMIISGSLLAATEKN-SRKCLVKGKMIMNSLSL 124
MS4A3 45 KAKLQVLGAIQILNAAMILALGVFLGSLQYPYHFQKHFFFFTFYTGYPIWGAVFECSSGTLSVVAGIK-PTRTWIQNSFGMNIASA 129
MS4A4A 60 KGEPKVLGVVQILTALMSLSMGITMMCMASN---TYGSNPISVYIGYTIWGSVMEIISGSLSIAAGIR-TTKGLVRGSLGMNITSS 141
MS4A5 44 ARKMKILGTIQILFGIMTFSFGVIFLFTLLK---PYPRFPFIFLSGYPFWGSVLEINSGAFLIAVKRK-TTETLIILSRIMNFLSA 125
MS4A15 68 TGEPKVLGTVQILIGLIHLGFGSVLLMVRRGHVG-----IFFIEGGVPFWGGACEIISGSLSVAAEKN-HTSCLVRSSLGTNILSV 147
MS4A8 66 LKEGKTLGAIQIIIGLAHLGLGSIMATVLVGE-----YLSISFYGGFPFWGGLWEIISGSLSVAAENQPYSYCLLSGSLGLNIVSA 146
MS4A12 85 KEEAKALGVIQIMVGLMHLGFGIVLCLISFSFREVLGFASTAVIGGYPFWGGLSEIISGSLSVSASKE-LSRCLVKGSLGMNIVSS 169
Conservation
Consensus KGEPKVLGXIQILIGLMHLGFGXXLLXVLXG----XXFFXFSFYGGYPFWGGVXFIISGSLSVAAXXK-TTRCLVRGSLGMNIXSA
TM3
ECL2
TM4 ext.
TM4
MS4A1 125 FAAISGMILSIMDILNIKISHFLKMESLNFIRAHTPYINIYNCEPANPSEKNSPSTQYCYSIQSLFLGILSVMLIFAFFQELVIAG 210
MS4A3 130 TIALVGTALSLNIAVNIQ-------SLR----------SC------HSSSESPDLCNYMGSISNGMVSLLLIFTLLELCVTIS 190
MS4A4A 142 VLAASGILINTFSLAFY----------SFHHPYCNYYGNS----------NNCHGTMSILMGLDGMVLLLSVLEFCIAVS 201
MS4A5 126 LGAIAGIILLTFGFILDQ----------NYICGY----------SHQNSQCKAVTVLFLGILITLMTFSIIELFISLP 183
MS4A15 148 MAAFAGTALLLMDFGVT----------N-RD----------VDRGYLAVLTIFTVLEFFTAVI 189
MS4A8 147 ICSAVGVILFITDLSIP----------H-PYAYPDYYPYA----------W----GVNPGMAISGVLLVFCLLEFGIACA 201
MS4A12 170 ILAFIGVILLLVDMCIN----------G-VAG-----QDY----------W----AVLSGKGISATLMIFSLLEFFVACA 219
Conservation
Consensus ILAXXGXILLXFDXAXN----------X----X----------W-X--XVSIGMGIXXXLLIFSLLEFFIAXX

B

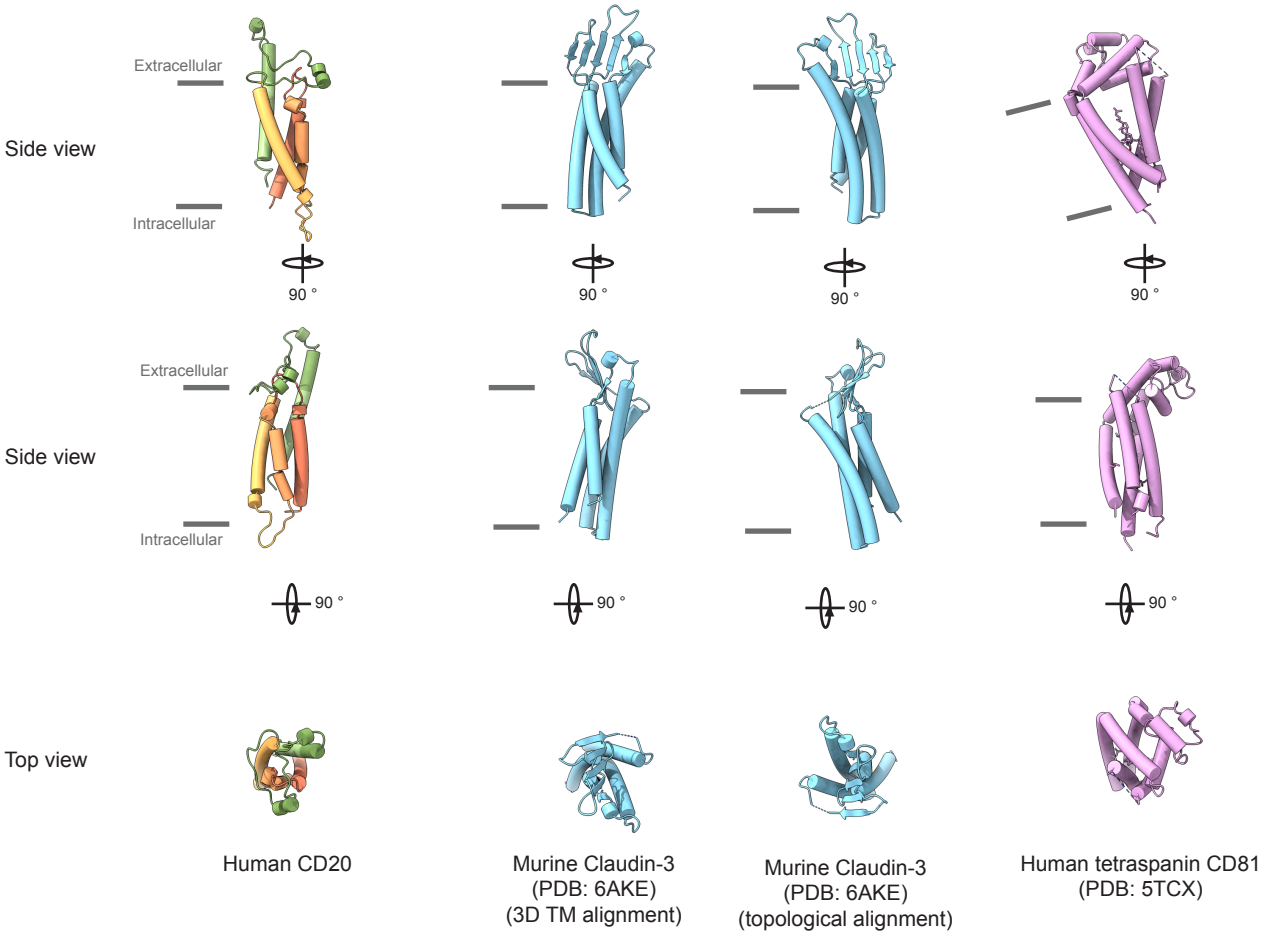

Side view
Extracellular
Intracellular
90 °
Side view
Extracellular
Intracellular
90 °
Top view
Human CD20
Murine Claudin-3
(PDB: 6AKE)
(3D TM alignment)
Murine Claudin-3
(PDB: 6AKE)
(topological alignment)
Human tetraspanin CD81
(PDB: 5TCX)

Figure S5

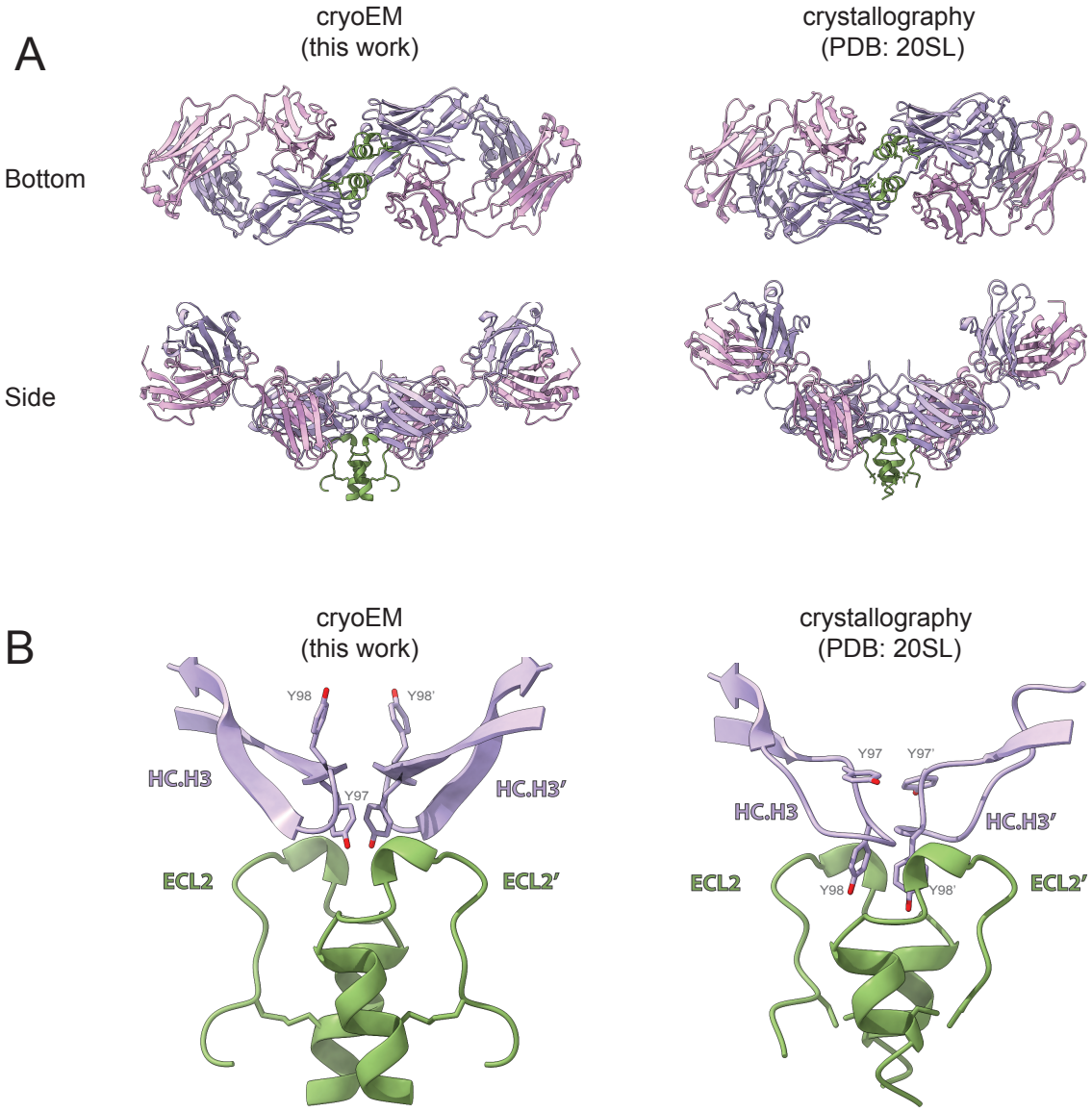

A
cryoEM
(this work)
crystallography
(PDB: 20SL)
Bottom
Side
B
cryoEM
(this work)
crystallography
(PDB: 20SL)
Y98
Y98'
Y97
HC.H3
HC.H3′
ECL2
ECL2′
Y97
Y97'
HC.H3
HC.H3′
ECL2
ECL2′
Y98
Y98'

Figure S6

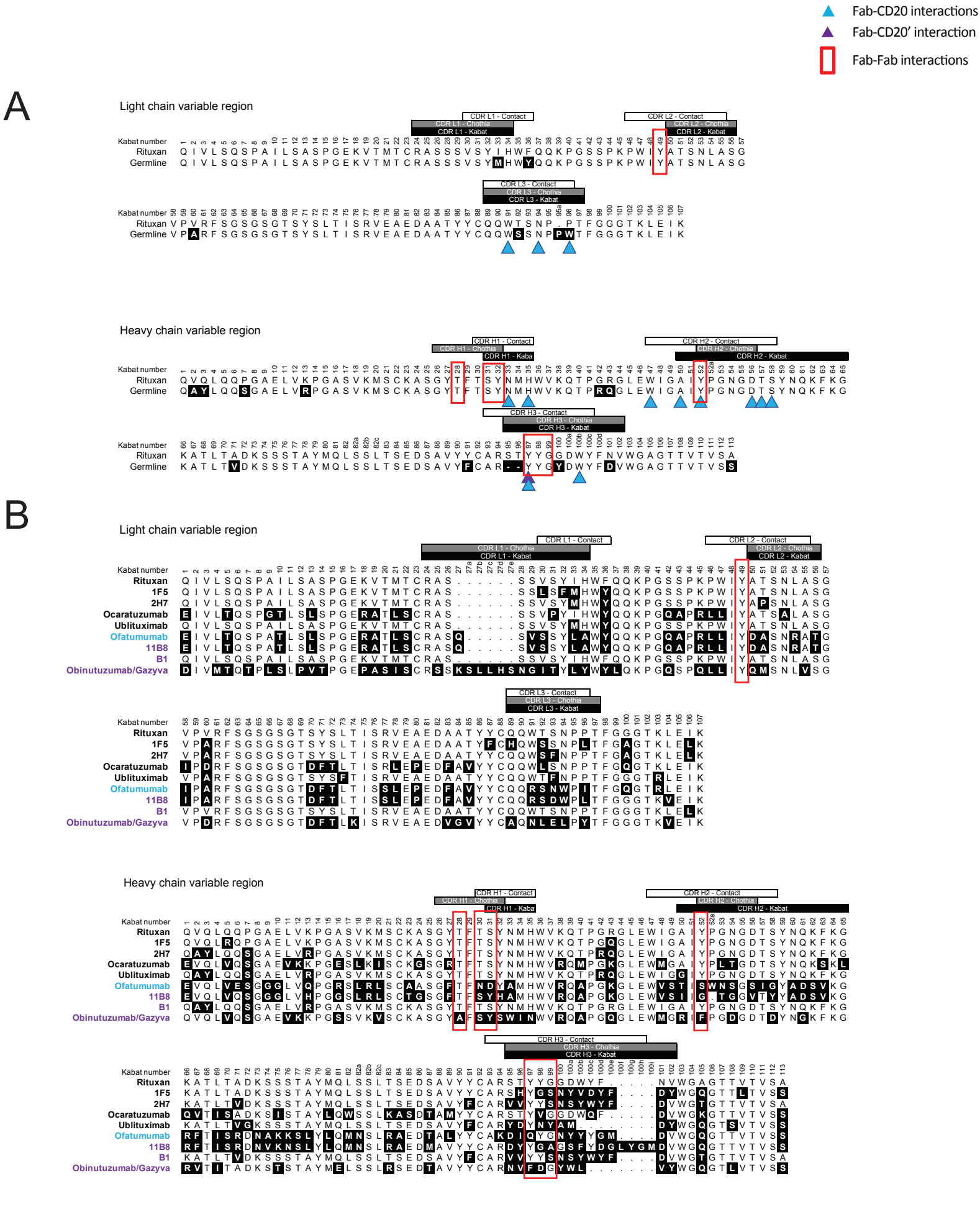
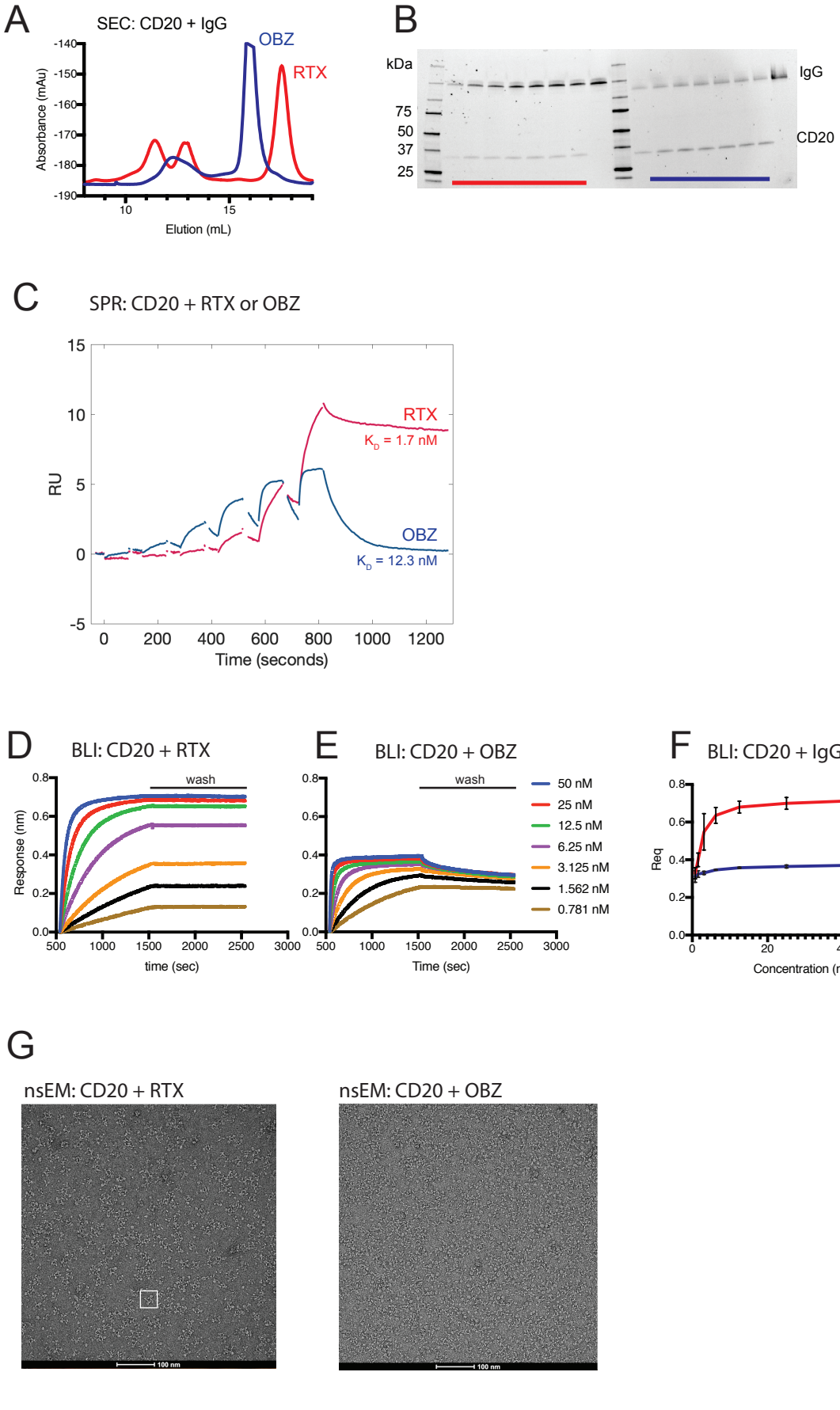

Fab-CD20 interactions
Fab-CD20' interaction
Fab-Fab interactions
A
Light chain variable region
CDR L1 - Contact
CDR L1 - Chothia
CDR L1 - Kabat
CDR L2 - Contact
CDR L2 - Chothia
CDR L2 - Kabat
CDR L3 - Contact
CDR L3 - Chothia
CDR L3 - Kabat
Kabat number
Rituxan
Germline
Heavy chain variable region
CDR H1 - Contact
CDR H1 - Chothia
CDR H1 - Kaba
CDR H2 - Contact
CDR H2 - Chothia
CDR H2 - Kabat
CDR H3 - Contact
CDR H3 - Chothia
CDR H3 - Kabat
B
Light chain variable region
Rituxan
1F5
2H7
Ocaratuzumab
Ublituximab
Ofatumumab
11B8
B1
Obinutuzumab/Gazyva
Heavy chain variable region

Figure S7


A
SEC: CD20 + IgG
OBZ
RTX
Absorbance (mAu)
-140
-150
-160
-170
-180
-190
10
15
Elution (mL)
B
kDa
75
50
37
25
IgG
CD20
C
SPR: CD20 + RTX or OBZ
RU
15
10
5
0
-5
RTX
K_D = 1.7 nM
OBZ
K_D = 12.3 nM
0
200
400
600
800
1000
1200
Time (seconds)
D
BLI: CD20 + RTX
wash
Response (nm)
0.8
0.6
0.4
0.2
0.0
500
1000
1500
2000
2500
3000
time (sec)
E
BLI: CD20 + OBZ
wash
50 nM
25 nM
12.5 nM
6.25 nM
3.125 nM
1.562 nM
0.781 nM
Time (sec)
F
BLI: CD20 + IgG
Req
+ RTX
K_D = 1.7 nM
+ OBZ
K_D = 3.3 nM
0
20
40
60
Concentration (nM)
G
nsEM: CD20 + RTX
nsEM: CD20 + OBZ
100 nm
100 nm