

STRUCTURAL BIOLOGY

Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab

Lionel Rougé¹, Nancy Chiang², Micah Steffek³, Christine Kugel⁴, Tristan I. Croll⁵, Christine Tam⁴, Alberto Estevez¹, Christopher P. Arthur¹, Christopher M. Koth¹, Claudio Ciferri¹, Edward Kraft⁴, Jian Payandeh^{1,2*}, Gerald Nakamura^{2*}, James T. Koerber^{2*}, Alexis Rohou^{1*}

Cluster of differentiation 20 (CD20) is a B cell membrane protein that is targeted by monoclonal antibodies for the treatment of malignancies and autoimmune 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 cross-links 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.

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 autoimmune conditions (1). It is ubiquitously expressed on circulating B cells (2) and is predicted to have four transmembrane helices (TMs) with two extracellular loops, ECL1 and ECL2, the second of which is much longer and contains a disulfide bond (fig. 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 (4–6), there exists no high-resolution structural data on any MS4A family member beyond predictions of secondary structure and membrane topology. Although CD20 is the best-studied member of the family, even its oligomeric state is poorly understood: Available evidence suggests that it associates into homo-oligomers and complexes with other proteins (7–9). 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 (7, 10). However, more recent work

showed that CD20⁺ B cells lacking the B cell receptor are unable to initiate calcium signaling, suggesting that CD20 indirectly regulates calcium release downstream from the B cell receptor (11).

CD20-targeted therapies revolutionized the treatment of B cell malignancies and autoimmune 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, 12). Another anti-CD20 mAb, ocrelizumab (OCR; Ocrevus), is now used in the treatment of multiple sclerosis (13). Although all anti-CD20 mAbs act by depleting B cells, they use at least four distinct mechanisms (14): direct cell death, Fc receptor effector functions through antibody-dependent cellular cytotoxicity and phagocytosis, 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, 14). Rituximab is the prototypical type I mAb, characterized by high CDC activity and the ability to cluster CD20 into lipid rafts (12, 15). Other type I mAbs include OCR and ofatumumab [OFA; (16)]. Type II mAbs such as obinutuzumab (OBZ; Gazyva) and tositumomab (Bexxar) exhibit low CDC activity and lack the ability to localize CD20 into lipid rafts but induce higher levels of direct cell death (17).

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 (18), 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 ¹⁷⁰Ala-Asn-Pro-Ser-Glu¹⁷⁴ (¹⁷⁰ANPSE¹⁷⁴) motif of ECL2, whereas at least one type I mAb (OFA) has a completely separate, nonoverlapping epitope involving ECL1 and another part of ECL2 (14, 19). 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 ¹⁷⁰ANPSE¹⁷⁴ bind with different stoichiometries and trigger notably 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° away from the angle at which OBZ approaches the same peptide (4–6). However, in the absence of a structure of full-length CD20 or of a MS4A homolog, it is difficult to speculate how such differences in binding geometry might affect 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 (fig. S1A). After solubilization and purification in the mild detergent glyco-diosgenin (GDN), CD20 was found to be further stabilized by cholesterol hemisuccinate (CHS; fig. 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 (fig. S1, C to E, and table S1). Fab binding to purified CD20 was subsequently evaluated using biolayer interferometry (BLI; Fig. 1A) and surface plasmon resonance (SPR; table S2), revealing sensorgrams consistent with a two-state 2:2 binding of RTX [dissociation constant (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 (Fig. 1B). Cryogenic electron microscopy (cryo-EM) imaging of the RTX:CD20 complex allowed us to determine its structure to a resolution of 3.3 Å (fig. S3), resulting in a near-complete atomic model of the complex (Fig. 1C).

¹Department of Structural Biology, Genentech Inc., South San Francisco, CA 94080, USA. ²Department of Antibody Engineering, Genentech Inc., South San Francisco, CA 94080, USA. ³Department of Biochemical and Cellular Pharmacology, Genentech Inc., South San Francisco, CA 94080, USA. ⁴Department of Biomolecular Resources, Genentech Inc., South San Francisco, CA 94080, USA. ⁵Cambridge Institute for Medical Research, University of Cambridge, Keith Peters Building, Cambridge CB2 0XY, UK. *Corresponding author. Email: rohou.alexis@gene.com (A.R.); koerber.james@gene.com (J.T.K.); nakamura.gerald@gene.com (G.N.); payandeh.jian@gene.com (J.P.)

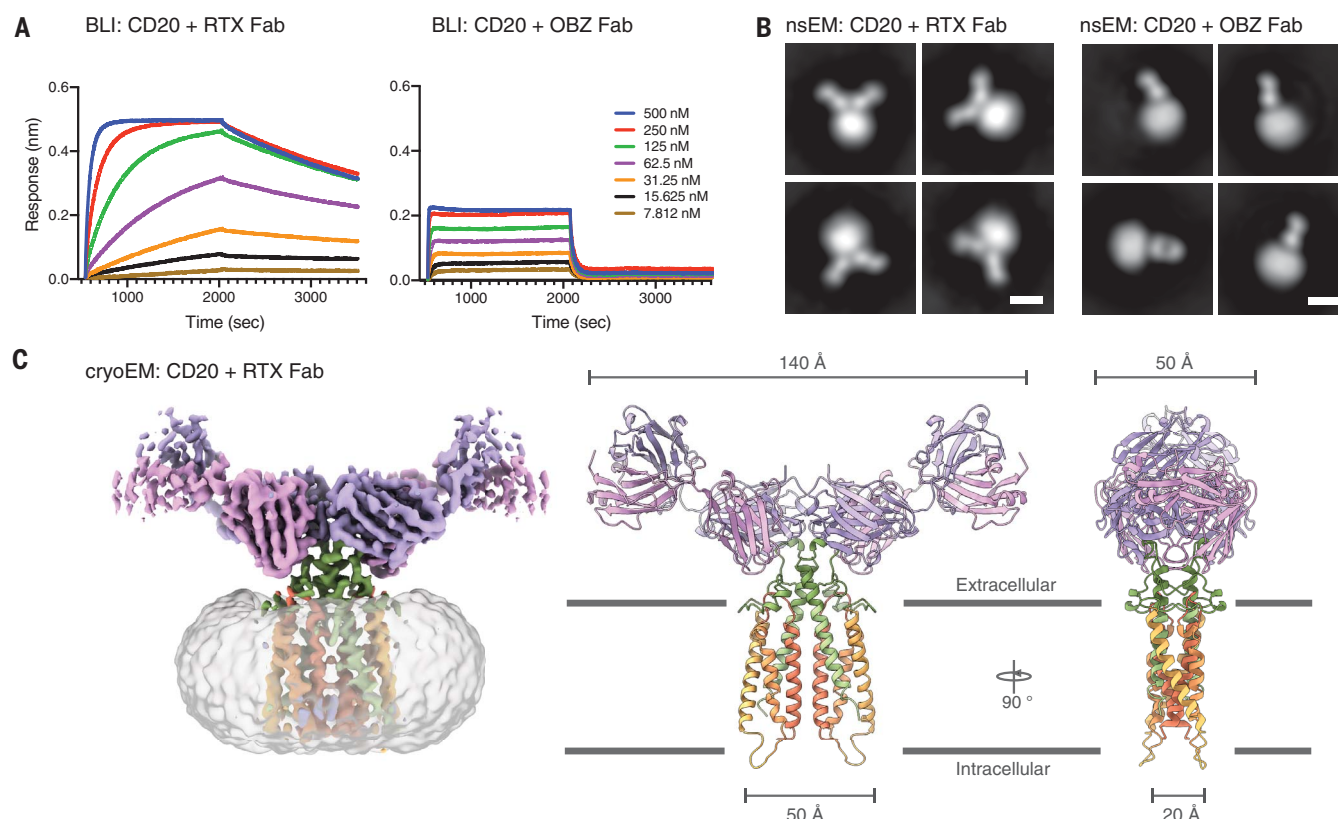


Fig. 1. Characterization of CD20:Fab complexes and cryo-EM structure of CD20:RTX Fab. (A) BLI traces. A serial dilution of RTX (left) or OBZ (right) Fabs was flowed in for the first 1500 s of the experiments, followed by a dissociation step. (B) nsEM of expressed, solubilized, and purified CD20 in complex with RTX (left) or OBZ (right) Fab. Scale bars, 50 Å. (C) Cryo-EM

reconstruction of the CD20:RTX Fab complex at a resolution of 3.3 Å. An isosurface rendering, with the GDN micelle rendered in transparent gray, the RTX Fab heavy chain in purple, and the light chain in pink is shown on the left. Two orthogonal side views (along the plane of the membrane) of a ribbon rendering of the structure are shown on the right.

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 four-TM arrangement, with its four TMs arranged antiparallel and clockwise when viewed from the extracellular side (Fig. 2, top view). The core of each CD20 monomer presents a compacted rectangular fold measuring ~25 Å by ~20 Å by ~50 Å (Fig. 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 (fig. S4A). Most notable is a constellation of highly conserved, small residues that allow for the close interhelical packing observed in CD20 and are found along TM1 (Gly⁵³, Gly⁶⁰, Gly⁶⁷), TM2 (Gly⁹⁰, Ser⁹⁷, Gly⁹⁸), and TM3 (Ser¹²³, Gly¹³⁰) (fig. S4A). A set of larger residues contributed by TM2 (Tyr⁸⁶, Tyr⁹⁴) and TM4 (Leu¹⁹⁴, Met¹⁹⁷, Ala²⁰¹, Gln²⁰⁴) form the bulk of the tightly packed TM-helical core (Fig. 2B) and are conserved as similarly bulky residues in the MS4A family (fig. 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 four-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 of the MS4A family are extremely diverse in sequence (fig. S4A). In CD20, the first extracellular loop (ECL1) is short and largely shielded by ECL2, leaving only Ile⁷⁶ and Tyr⁷⁷ exposed (Fig. 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 four-helix bundle until it kinks into a single-turn α helix (Asn¹⁵³ to Arg¹⁶⁶), which marks the start of ECL2's extracellular segment (Fig. 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 TM4 (Fig. 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 (¹⁷⁰ANPSE¹⁷⁴) for most anti-CD20 antibodies (14).

A search of the Protein Data Bank 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 (fig. S4B). For example, the transmembrane cavity and cholesterol-coordinating acidic residue present in CD81 (20) are absent in CD20, and although claudins appear to share a similar core four-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 tetramer (7–9). Viewed from the extracellular side, CD20 forms a dimeric double-barrel assembly of approximate dimensions 20 Å by 50 Å (Fig. 2, top view). Two square CD20 subunits abut each other to form a four-helix antiparallel transmembrane coiled coil involving the upper halves of TM1 (Leu⁶¹, Phe⁶², Ala⁶⁵, Leu⁶⁹) and TM4 (Leu¹⁸⁹, Ile¹⁹³, Val¹⁹⁶, Phe²⁰⁰) from each protomer (Fig. 2D).

The residues at this intersubunit 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 interprotomer transmembrane conduction pathway. Additionally, structure-based sequence alignment suggests that a CD20-like dimer interface may be shared across the MS4A family (fig. 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 Ser^{179'}; Gln¹⁸¹ to the backbone amide of Tyr^{161'}) contribute to the interface (Fig. 2C). In total, the CD20 dimer buries 1656 Å² 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 (fig. 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 (¹⁷⁰ANPSE¹⁷⁴) 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 ¹⁷⁰ANPSE¹⁷⁴ (14). In our structure, this core epitope is simultaneously bound by two Fabs, which we denote RTX and RTX' (Fig. 3). The RTX Fab sits atop the CD20 protomer, engaging ECL2 at a shallow angle (~22°) relative to the membrane plane (Fig. 3, side view), likely precluding engagement of CD20 by a single immunoglobulin G (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), where HC indicates heavy chain. The second (RTX') Fab extends its

heavy-chain variable loop 3 (H3) across the dimer interface to present HC.Tyr^{97'}, which interacts with Glu¹⁷⁴ of CD20 (Fig. 3A). At the apex of ECL2, Ser¹⁷³ organizes an extended network of hydrogen bonds spanning from HC.Asn³³, through Glu¹⁷⁴, to HC.Tyr^{97'} (Fig. 3B). This key interaction network is clamped by both Fabs: Ser¹⁷³ is stabilized by HC.Tyr⁵², Glu¹⁷⁴ is sandwiched by HC.Trp^{100b} and HC.Gly^{100'}, and HC.Tyr⁹⁷ is stabilized by HC.Trp^{100b}.

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 (Fig. 3D). Completely distinct from the classic ECL2 turret epitope ¹⁷⁰ANPSE¹⁷⁴, this secondary epitope is recognized primarily by light-chain (LC) 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 Pro^{178'}, 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 Å² (4), but these L1-ECL1/2 interactions increase the contact surface area by almost 50%, to ~640 Å² (Fig. 4B), suggesting that this secondary epitope likely contributes substantially 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 (Fig. 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 by means of the D gene segment (fig. S6A), seems essential to the Fab-Fab interaction: Its C β and C γ 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 Glu^{174'} of the contralateral CD20' protomer (Fig. 3, A and B). Two key additional Fab:Fab interactions are mediated by HC.Ser^{31'}, whose backbone and side chain directly engage HC.Gly⁹⁹ while its side chain contacts LC.Tyr⁴⁹ and HC.Tyr⁹⁸ (Fig. 3C). Overall, this Fab:Fab homotypic interface buries 375 Å² of solvent-exposed area (Fig. 4B).

Our structure has thus unveiled a composite CD20-Fab' epitope with three components:

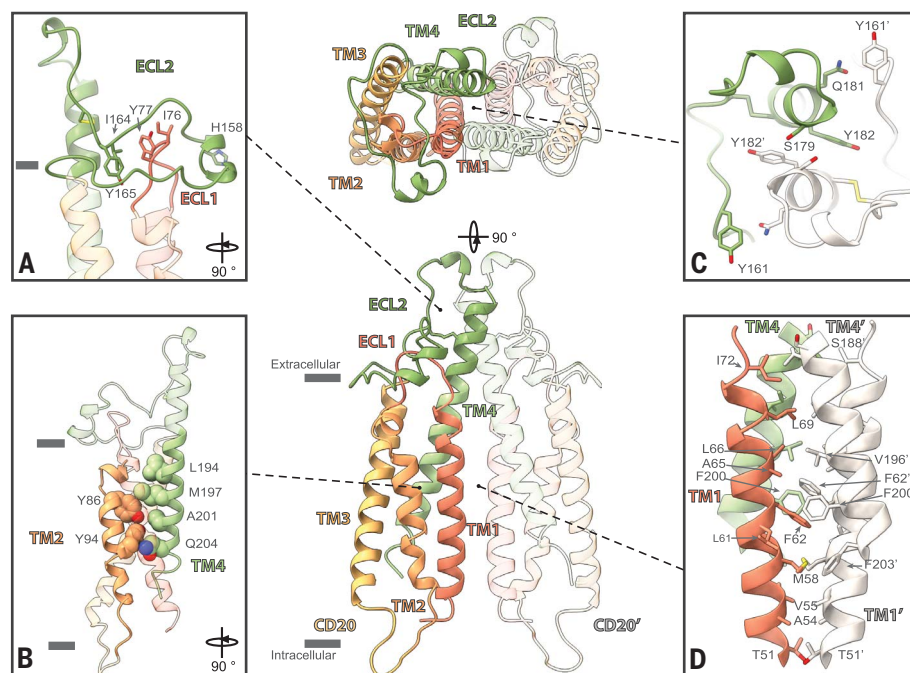


Fig. 2. The CD20 dimer is a compact double-square-barrel structure. Ribbon diagrams of the CD20 structure, with RTX omitted and one of the CD20 protomers transparent for clarity. (A) The short loop ECL1 (red), between TM1 and TM2, is almost entirely surrounded by the first half of ECL2 (green). (B) The 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 and D) The extensive dimeric interface of CD20 involves the extracellular domain (C) as well as TM1 and TM4 (D). In (A), (B), and the center diagram, the gray bars indicate the boundaries of the membrane region. Single-letter abbreviations for amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. Oxygen, nitrogen, and sulfur atoms are colored in red, blue, and yellow, respectively.

the primary epitope on ECL2 (¹⁷⁰ANPSE¹⁷⁴), a secondary ECL1-ECL2 epitope, and a direct homotypic Fab:Fab interface. The total composite buried surface area of ~1000 Å² is comparable to traditional 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 (Fig. 4C). We used charge reversals and bulky side chains expected to disrupt these molecular interfaces and measured their effects on CDC (Fig. 4E) and on IgG binding to purified CD20 (SPR; table S2), ECL2 epitope peptide (SPR; table S3), and CD20⁺ cells (flow cytometry; Fig. 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 (Fig. 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 [median inhibitory concentration (IC₅₀) > 10 times higher] relative to wild-type RTX. LC.Ser³¹Asp had a notably stronger phenotype, which resulted in a ~100-fold reduction in Fab:CD20 binding and ~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; Fig. 3, A and B). The HC.Tyr⁹⁷Ser mutation, which we predict may destabilize H3, effectively abolishes target engagement and CDC activity, whereas 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 > 10 times 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 epi-

tope and Fab:Fab interface contribute to the distinctive binding properties and high CDC activity of RTX.

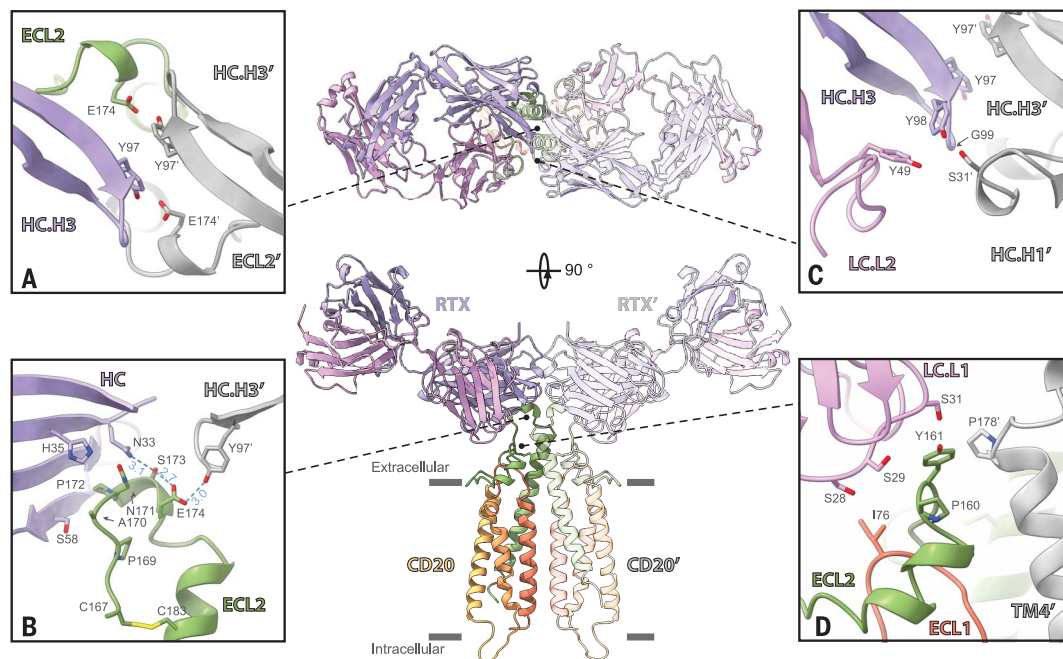
Full-length RTX cross-links CD20 dimers into higher-order assemblies

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 (Fig. 1C). This suggests that 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 (fig. 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 (Fig. 5A). These closed-ring assemblies feature CD20 dimers and Fc domains splayed outward, linked by pairs of Fab arms that position the Fab-Fc hinges of RTX on a circle of approximate diameter 150 Å (Fig. 5, A and B). Because of the marked similarity with our three-dimensional structure of the CD20:RTX Fab complex, we were able to generate a model of these rings

Fig. 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. In the center diagram, the gray bars indicate the boundaries of the membrane region. (A) Top view of the center of the complex, where HC.Tyr⁹⁷ mediates Fab:Fab and Fab:CD20' contacts. (B) The canonical RTX epitope ¹⁷⁰ANPSE¹⁷⁴, 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, gray). (C) Additional Fab:Fab contacts between heavy-chain loops H3 and H1 and light-chain loop L2. (D) A secondary epitope consisting of ECL1 and ECL2 is contacted by RTX's LC loop 1.



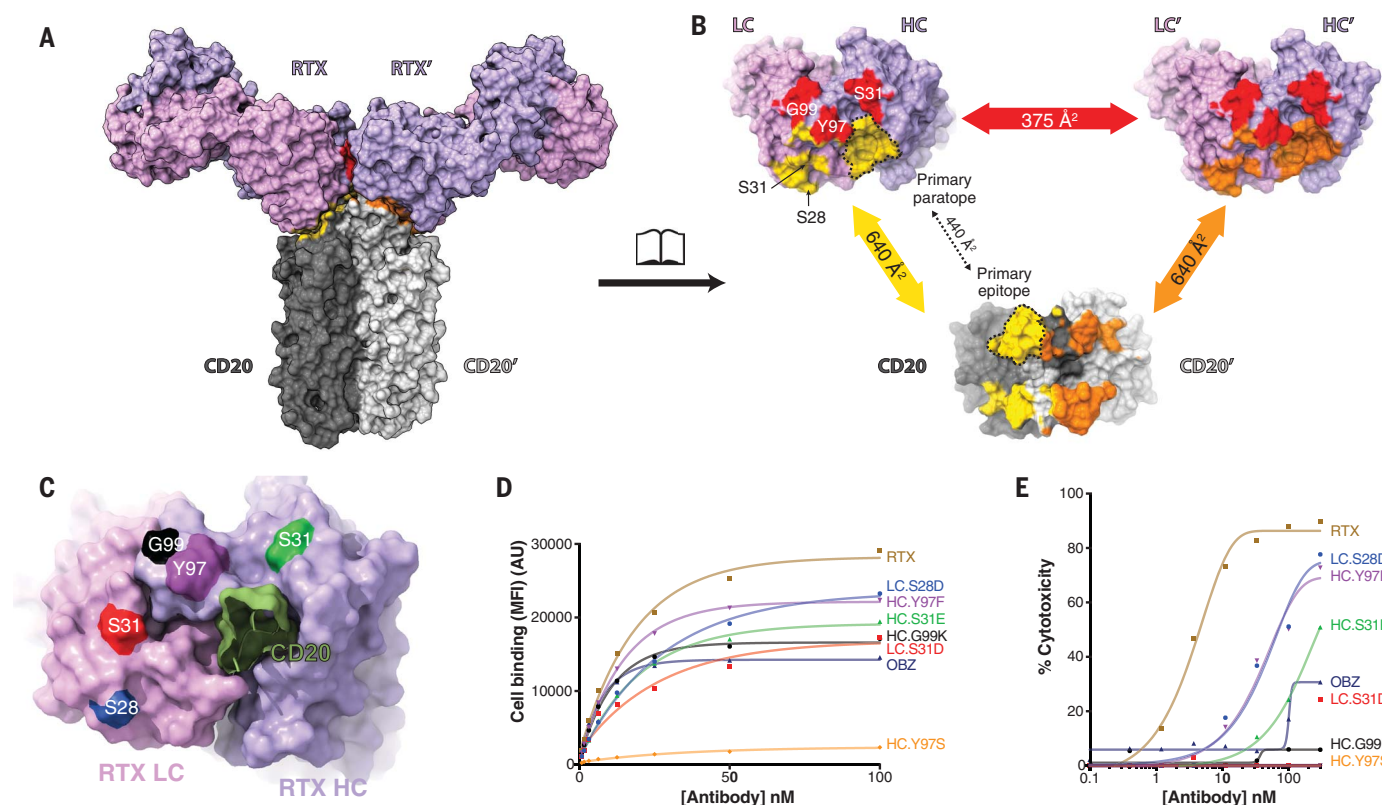


Fig. 4. Multiple CD20:RTX and RTX:RTX interactions enable cell binding and CDC. (A) Surface representation of the CD20:RTX Fab complex. The surfaces buried by complex formation are colored in yellow (CD20:RTX), orange (CD20':RTX'), or red (RTX:RTX'). (B) Open-book representation of the same surfaces, with surface area measurements for each buried surface indicated. Residues mutated as part of this study are labeled at the top left and are not involved in the primary paratope. The primary

epitope (dashed line) only accounts for less than half (440 out of 1015 Å²) 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 and E) Cell binding (D) and CDC (E) of RTX mutants are plotted as a function of antibody concentration and compared with wild-type RTX and OBZ. MFI, mean fluorescence intensity; AU, arbitrary units.

as present on the nsEM grids (Fig. 5, B and C). To understand how these assemblies might relate to RTX function, we endeavored to build a model of a CD20:RTX IgG assembly as it might occur on a cell. This was achieved by rotating each CD20:Fab complex 90° while keeping the ends of the Fab domains in close proximity to each other (Fig. 5D, left). This modeled assembly exhibits precisely the dimensions that would be required for the three Fc domains to “fold in” (Fig. 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 (25). The resulting model of a CD20:RTX:C1 complex (Fig. 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 autoimmune dis-

eases, but its structure and function have remained unknown. In contrast to the prevailing view that CD20 is a tetramer (14), 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 (18). 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 owing 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 (Fig. 5, A and 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 toward the formation of the hexameric Fc platforms necessary for complement recruitment (25). In

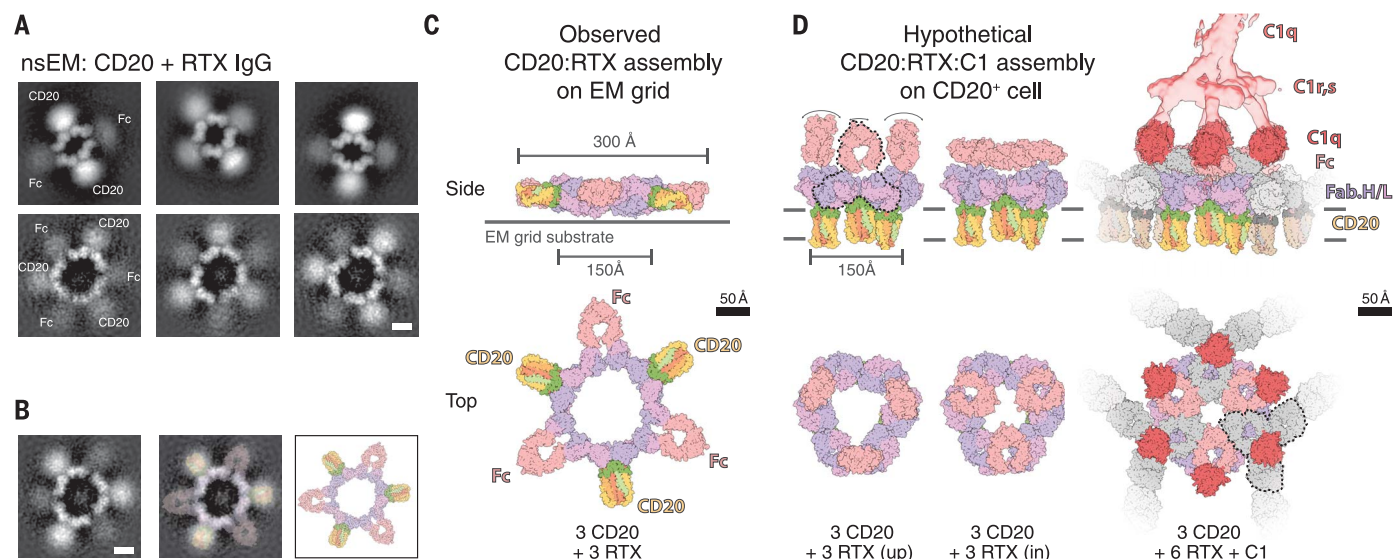


Fig. 5. RTX cross-links CD20 into circular superassemblies. (A) Average nsEM images of CD20 incubated with full-length RTX show cyclical higher-order structures of involving 2-to-2 (top row; diameter of 250 Å) or 3-to-3 (bottom row; diameter of 300 Å) CD20-to-RTX complexes. The RTX Fc domains appear disordered, presumably because of IgG hinge flexibility. Scale bar, 50 Å. (B and C) Interpretation of an nsEM class average of a 3-to-3 assembly. Scale bar in (B), 50 Å. (D) Proposed model for CD20:RTX supercomplex formation and complement recruitment. During nsEM experiments, the IgGs and solubilized CD20s are coplanar (C). Modeling these higher-order assemblies as they might

occur at the surface of CD20⁺ cells requires rotating the CD20:Fab complexes 90° [(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 (gray) would complete the Fc hexamer formation and enable recruitment of C1q [(D), right]. Dashed lines outline IgG molecules. The following models were used: structure from present work (RTX Fab:CD20 complex), EMDB-4232 (EM map of C1:Fc complex), and Protein Data Bank (PDB) 6FCZ (Fc domains and C1q head domains) (25).

a simplistic model, we speculate that dimeric RTX:CD20 building blocks (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 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 that 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 (fig. S6A). This observation suggests that the homotypic Fab:Fab interaction was inherent to the progenitor RTX B cell before somatic hypermutation. These residues are also highly conserved among the RTX-like type I mAbs (fig. S6B), most of which are also mouse-derived. Similar Fab:Fab contacts mediate crystal packing of the isolated RTX Fab-ECL2 peptide complex (4) (fig. 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 preformed dimers before 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 (26, 27) and cross-linking-independent activation of TRAIL-R2 (28). 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

- M. J. E. Marshall, R. J. Stopforth, M. S. Cragg, *Front. Immunol.* **8**, 1245 (2017).
- P. Stashenko, L. M. Nadler, R. Hardy, S. F. Schlossman, *J. Immunol.* **125**, 1678–1685 (1980).
- L. Eon Kuek, M. Leffler, G. A. Mackay, M. D. Hulett, *Immunol. Cell Biol.* **94**, 11–23 (2016).
- J. Du et al., *J. Biol. Chem.* **282**, 15073–15080 (2007).
- G. Niederfellner et al., *Blood* **118**, 358–367 (2011).
- J. Du et al., *Mol. Immunol.* **45**, 2861–2868 (2008).
- J. K. Bubien, L. J. Zhou, P. D. Bell, R. A. Frizzell, T. F. Tedder, *J. Cell Biol.* **121**, 1121–1132 (1993).
- M. J. Polyak, J. P. Deans, *Blood* **99**, 3256–3262 (2002).
- M. J. Polyak, H. Li, N. Shariat, J. P. Deans, *J. Biol. Chem.* **283**, 18545–18552 (2008).
- T. W. Kuijpers et al., *J. Clin. Invest.* **120**, 214–222 (2010).
- C. A. Walshe et al., *J. Biol. Chem.* **283**, 16971–16984 (2008).
- D. G. Maloney et al., *Blood* **90**, 2188–2195 (1997).
- X. Montalban et al., *N. Engl. J. Med.* **376**, 209–220 (2017).
- C. Klein et al., *MAbs* **5**, 22–33 (2013).
- J. P. Deans, S. M. Robbins, M. J. Polyak, J. A. Savage, *J. Biol. Chem.* **273**, 344–348 (1998).
- W. G. Wierda et al., *J. Clin. Oncol.* **28**, 1749–1755 (2010).
- E. Mössner et al., *Blood* **115**, 4393–4402 (2010).
- H. T. C. Chan et al., *Cancer Res.* **63**, 5480–5489 (2003).
- J. Du, H. Yang, Y. Guo, J. Ding, *Mol. Immunol.* **46**, 2419–2423 (2009).
- B. Zimmerman et al., *Cell* **167**, 1041–1051.e11 (2016).
- S. Nakamura et al., *Nat. Commun.* **10**, 816 (2019).
- R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, R. MacKinnon, *Nature* **415**, 287–294 (2002).
- F. Li, J. Liu, Y. Zheng, R. M. Garavito, S. Ferguson-Miller, *Science* **347**, 555–558 (2015).

24. T. Ramaraj, T. Angel, E. A. Dratz, A. J. Jesaitis, B. Murney, *Biochim. Biophys. Acta* **1824**, 520–532 (2012).
25. D. Ugurlar *et al.*, *Science* **359**, 794–797 (2018).
26. K. Imkeller *et al.*, *Science* **360**, 1358–1362 (2018).
27. D. Oyen *et al.*, *Sci. Adv.* **4**, eaau8529 (2018).
28. T. Tamada *et al.*, *Sci. Rep.* **5**, 17936 (2015).

ACKNOWLEDGMENTS

Molecular graphics and analyses were 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 department at Genentech for support throughout the project. Thanks to the Peptide Synthesis group in the department of Early Discovery Biochemistry. We thank the Research Material Group for antibody expression as well as the Wu lab in Protein Chemistry

and the Antibody Production and Automation Technologies group for antibody purification. **Funding:** T.I.C. was supported by a Principal Research Fellowship from the Wellcome Trust to R. 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.M.K., A.R.; CD20 construct design: C.M.K., L.R.; CD20 protein production: C.K., C.T., E.K.; Electron microscopy: A.E., C.P.A., C.C.; EM image processing: L.R., A.R.; Model building and refinement: A.R., T.I.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, differential scanning fluorimetry, and 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.I.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 Protein Data Bank (PDB) 6VJA and EMD-21212, respectively.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/367/6483/1224/suppl/DC1
Materials and Methods

Figs. S1 to S7

Tables S1 to S4

References (29–40)

[View/request a protocol for this paper from Bio-protocol.](#)

29 October 2019; accepted 6 February 2020

Published online 20 February 2020

10.1126/science.aaz9356