

Design of a Water-Soluble CD20 Antigen with Computational Epitope Scaffolding

Zhiyuan Yao ¹, Brian Kuhlman ^{2,3}

¹Department of Pharmacology, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.

²Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.

³Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.

Correspondence

Brian Kuhlman Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA. Email: bkuhlman@email.unc.edu.

Abstract

The poor solubility of integral membrane proteins in water frequently hinders studies with these proteins, presenting challenges for structure determination and binding screens. For instance, the transmembrane protein CD20, which is an important target for treating B-cell malignancies, is not soluble in water and cannot be easily screened against potential protein binders with techniques like phage display or yeast display. Here, we use *de novo* protein design to create a water-soluble mimic of the CD20 dimer (“soluble CD20”). Soluble CD20 replaces the central transmembrane helix of CD20 with a water-soluble helix that dimerizes to form a coiled coil that structurally matches the dimer interface of native CD20 and presents the central extracellular loop of CD20 in a binding competent conformation. Unlike peptides derived from CD20, soluble CD20 binds tightly to monoclonal antibodies that recognize quaternary epitopes on the extracellular face of CD20. We demonstrate that soluble CD20 is easy to produce, remains folded above 60°C, and is compatible with binder screening via yeast display. Our results highlight the ability of computational protein design to scaffold conformational epitopes from membrane proteins for use in binding and protein engineering studies.

Keywords

Epitope scaffolding, transmembrane protein, protein engineering, CD20, computational protein design

Importance

This study demonstrates that computational protein design can be used to create water soluble variants of transmembrane proteins. The water-soluble CD20 mimic created here is compatible with high throughput binding screens and should be useful for developing next generation anti-CD20 therapeutics.

Abbreviations

CMPs: Complex membrane proteins

ECDs: Extracellular Domains

CDC: Complement-Dependent Cytotoxicity

PCD: Programmed Cell Death

mAbs: Monoclonal Antibodies

Fab: fragment antigen-binding

scFv: single chain variable fragment

CD: Circular Dichroism

T_m: Melting Temperature

TM: transmembrane helix

SEC-MALS: Size Exclusion Chromatography-Multi Angle Light Scattering

BLI: Bio-Layer Interferometry

SPR: Surface Plasmon Resonance

RTX: Rituximab

OBZ: Obinutuzumab

RMSD: Root-Mean-Square Deviation

Introduction

Therapeutic antibodies targeting multi-spanning membrane proteins (or complex membrane proteins, CMPs) remain underdeveloped despite their importance in human disease. While antibody therapeutics provide exceptional specificity and affinity, making them a transformative class of therapeutics, their development has largely focused on soluble ligands or transmembrane receptors with large extracellular domains (ECDs) (Stephens and Wilkinson 2024; Crescioli et al. 2024). Despite CMPs comprising approximately 60% of small-molecule drug targets, only 20 among 221 approved therapeutic antibody therapeutics are targeting CMPs, including 13 targeting CD20 (Lim et al. 2010; Casan et al. 2018), 4 targeting CGRP (Aditya and Rattan 2023; Cohen et al. 2022), 1 targeting GPCR5d (Rodriguez-Otero et al. 2024), 1 targeting CCR4 (Zengarini et al. 2024), and 1 targeting Claudin 18.2 (M. A. Shah et al. 2023).

Developing antibodies or other protein binders against CMPs is hindered by challenges associated with producing well-folded CMPs (To'a Salazar et al. 2021; R. Dodd et al. 2020; Hashimoto et al. 2018). CMPs are membrane proteins that span cell membranes with multiple hydrophobic transmembrane domains (Qing et al. 2022). These proteins are critical drug targets due to their role in cell signaling, immune responses, and various human diseases (To'a Salazar et al. 2021; Stephens and Wilkinson 2024; R. B. Dodd, Wilkinson, and Schofield 2018). Access to soluble, well-behaved, conformation correct antigens is important for efficient binder development by animal immunization or high-throughput screening technologies like phage or yeast display. Due to their hydrophobic surfaces, CMPs have low solubility and are often unstable when isolated from the lipid bilayer environment (Qing et al. 2022). Several approaches have been developed to solubilize CMPs (Krishnarjuna and Ramamoorthy 2022), including the use of detergent micelles (Cecchetti et al. 2021; Kotov et al. 2019), proteoliposomes (Yao, Fan, and Yan 2020; Takeda et al. 2015), nanodiscs (Frauenfeld et al. 2016; Ma et al. 2020), or virus-like particles (Tucker et al. 2018). These approaches typically require extensive optimization to balance solubility and protein stability, making it both time-consuming and expensive. An emerging alternative is the creation of water-soluble analogs of MPs. By redesigning hydrophobic transmembrane domains to increase hydrophilicity while

preserving selected epitopes from the MP, these analogs enable binder development with high-throughput screening platforms (Qing et al. 2022). Several previous studies systematically replaced hydrophobic amino acids with hydrophilic residues of similar size and shape (Zhang et al. 2018; Smorodina et al. 2022), and successfully generated solubilized GPCRs (Skuhersky et al. 2021), ion channels (Smorodina et al. 2024), and chemokine receptors (Qing et al. 2023). Recently, AI-based methods for protein design have been used to generate soluble proteins that present selected structural motifs from CMPs (Goverde et al. 2024).

Here, we use computational protein design to create a soluble version of the multi-spanning membrane protein CD20 for use as “bait” in binder discovery campaigns. CD20 is a CMP expressed on B cells and is an effective therapeutic target in the treatment of B-cell malignancies (McLaughlin et al. 1998; Luo et al. 2021). Over the last 20 years, 13 different anti-CD20 antibodies have been developed and approved (Cragg and Glennie 2004; Alduaij and Illidge 2011; Klein et al. 2013). Because native CD20 is insoluble in water, these antibodies were generated through animal immunization with CD20-overexpressing cells as opposed to immunization with soluble recombinant protein. This approach has worked well for generating antibody binders, but it is incompatible with high-throughput technologies such as yeast (Cherf and Cochran 2015) and phage display (Jaroszewicz et al. 2022). These advanced methods not only facilitate the optimization of antibodies derived from animal immunization, but also enable the discovery of non-antibody binders, such as miniproteins (Cao et al. 2022) and DARPins (Plückthun 2015) that can be exceptionally stable (Kortemme 2024) and can be linked to create multi-specific fusion proteins (Huang et al. 2024). The enhanced ability to discover novel CD20 binders would meet the emerging demand for developing next-generation CD20 therapeutics (Dabkowska, Domka, and Firczuk 2024), which focus on multispecific targeting to combat resistance (Beers et al. 2010; Shah et al. 2020; Fousek et al. 2021) and refining well-folded antigen recognition domains essential for chimeric antigen receptor (CAR) therapies (Chen et al. 2023; Zhen et al. 2024).

To engineer a water-soluble version of CD20, we used computational de novo protein design to create a protein scaffold that presents the primary extracellular loop of CD20 in a native-like conformation. Unlike the solubilized native CD20 in detergent micelles, our engineered soluble CD20-mimic is compatible with screening via yeast display. We demonstrate that the CD20 mimic expresses well, is thermostable, and can bind both type I and type II anti-CD20 antibodies with native like affinities and stoichiometries. Our results demonstrate how computational protein design can be used to solubilize transmembrane proteins and the designed CD20 mimic should be a powerful reagent for discovering next-generation anti-CD20 therapeutics.

Results

Design of a soluble CD20 mimic

The CD20 monomer consists of four transmembrane α -helices (TM). The N- and C-termini are located on the intracellular side of the membrane and extracellular loops connect TM1 with TM2 and TM3 with TM4 (Figure 1)(Kumar et al. 2020; Rougé et al. 2020). The first extracellular loop (ECL1) is short and largely buried by the second extracellular loop (ECL2). In the membrane, CD20 forms a homodimer with the majority of the interchain contacts coming from ECL2, TM4 and TM1. On the extracellular face of the homodimer, a cleft is formed between the two copies of ECL2. This cleft and the surrounding residues from ECL2 is a binding site for some anti-CD20 antibodies(Kumar et al. 2020; Rougé et al. 2020). We hypothesized that a de novo designed homodimer that presents residues from ECL2 in the correct orientation relative to each other and forms a native-like cleft would be an effective mimic of the extracellular binding surface of CD20.

The first several residues of TM4 sit above the membrane, serve as an anchor for the C-terminal portion ECL2, and make contacts with TM4 from the partner chain. To stabilize the binding cleft formed by the ECL2/ECL2 interface in a soluble protein, we sought to scaffold ECL2 and the N-terminal portion of TM4 with a de novo designed coiled coil. First,

we generated a large family of idealized dimeric coiled-coils by varying the standard Crick parameters for a coiled-coil, which describe the superhelical twist and radius of a coiled-coil (Grigoryan and DeGrado 2011a). Ten coiled coils were selected, and ECL2 of CD20 cryo-EM structure were grafted onto the N-terminus of the matching coiled coils to construct full-length models of the designed proteins, each containing sixty-eight residues.

Sequences were designed for the CD20/coiled coil chimeras using the FastDesign protocol in the molecular modeling program Rosetta. FastDesign iterates between rotamer-based sequence optimization and backbone refinement to find low energy sequence/structure pairs (Maguire et al. 2021). The CD20/coiled coil chimeras were designed to form a homodimeric structure, with the antigen epitope loop (ECL2) preserved in its native sequence and conformation, including conservation of the native disulfide between residues 167 and 183 (Figure 1, b). The dimer interface of the designed coiled coil is predominantly stabilized by hydrophobic interactions, forming a tightly packed core that ensures structural stability. In addition to these hydrophobic contacts, a few key hydrogen bonds are observed, contributing to the specificity and further reinforcing the dimerization. When aligned with the native CD20 dimer cryo-EM structure, the designed chimera preserves the ECL2 antigen loop in its native conformation, although the coiled coil regions differ significantly from the native transmembrane helices TM1 and TM4 (Figure 1, c). Eight out of 400 models were selected for experimental validation.

Soluble CD20 forms a stable dimer

The 8 selected design sequences were fused to a N-terminal signal sequence and a C-terminal region that included a FLAG tag, His-tag, and Avitag and were expressed as secreted proteins in Expi293 cells. Following purification with metal affinity chromatography, two of the design sequences produced detectable amounts of protein (Figure S1), with the highest yields observed for design F2 (33.1 mg per liter of cell culture). Design F2 was renamed “soluble CD20”. In the following experiments we make use of three different versions of the soluble CD20 protein. Soluble CD20^{Avi} contains a C-terminal FLAG-tag, His-tag and Avi-tag and was the construct used in the initial screening.

Soluble CD20^{His} includes a N-terminal His-tag followed by a TEV protease site, and soluble CD20^{bacteria} was constructed in a bacterial expression vector (full sequences shown in supplementary table 2). We tested the expression of soluble CD20^{bacteria} in *E. coli* using Shuffle T7 cells. Final yields after purification were 4 mg per liter of cell culture. In contrast to soluble CD20, native CD20 cannot be secreted as a soluble protein from Expi293 cells and yields were low (0.44 mg/L) when expressed in SF9 cells and purified by membrane solubilization (Rougé et al. 2020) (Figure S1,S2).

On a reducing SDS-page gel, soluble CD20^{Avi} runs with an apparent molecular weight of 15 kDa, close to the expected molecular weight (14.9 kDa) of dissociated monomer (Figure 2a). To determine the oligomerization state of soluble CD20 in solution, size-exclusion chromatography with multi-angle light scattering (SEC-MALS) was performed. Soluble CD20^{Avi} eluted as a single peak with a molecular weight of 29.0 ± 0.9 kDa (Figure 2b), consistent with the expected molecular weight for a homodimer 29.8 kDa. Soluble CD20 has a circular dichroism (CD) spectrum consistent with a helical protein and unfolds cooperatively with temperature (Figure 2c). As expected for a homodimer, the thermal stability of soluble CD20 is concentration dependent with a thermal unfolding midpoint of 68.6°C at a protein concentration of 2 μ M (Figure 2d). Overall, biochemical characterization confirmed that the soluble CD20 design was expressed at the expected molecular weight and assembled into a stable dimer.

Soluble CD20 design binds to both type I and type II antibodies with native-like stoichiometry

Previously discovered anti-CD20 mAbs have been categorized into two types: type I, 'rituximab(RTX)-like' mAbs, which induce complement-dependent cytotoxicity (CDC) upon binding by crosslinking CD20 receptor, and type II anti-CD20 mAbs, such as obinutuzumab(OBZ), which do not induce accumulation of CD20 and show relatively little CDC activity (Klein et al. 2013). Cryo-EM structures of the fragment antigen-binding region (Fab) of type I and II mAbs bound to the native CD20 dimer show different stoichiometries and binding geometries (Figure 3a)(Kumar et al. 2020; Rougé et al. 2020). The type I antibody rituximab binds CD20 with two Fab molecules per CD20 dimer. The

complex is stabilized by homotypic Fab-Fab interactions which contribute to a slow off rate of rituximab from the CD20 dimer. In contrast, the type II mAb OBZ binds CD20 with one Fab molecule per CD20 dimer. These varying stoichiometries have been linked to the different functional properties of type I and II anti-CD20 mAbs.

To test if soluble CD20 closely mimics the extracellular epitopes of native CD20, we determined the binding stoichiometries of RTX and OBZ to soluble CD20. Fabs were mixed with soluble CD20 in a molar ratio of 2:2.5 and analyzed with SEC-MALS. With the RTX Fab, a large peak was observed with a measured molecular weight of 115 kDa which is close to the expected molecular weight (119 kDa) for a complex of two Fabs per soluble CD20^{His} dimer (2:2 stoichiometry). With the OBZ Fab, a peak was observed with a measured molecular weight of 77 kDa which is close to the expected molecular weight (71 kDa) for complex of 1 Fab per soluble CD20^{His} dimer, consistent with the 1:2 stoichiometry (Figure 3b). In the RTX experiment, a separate peak was observed for excess amount of soluble CD20 (Figure S3). In the OBZ Fab experiments, a secondary peak was observed for excess Fab. These results indicate that rituximab and OBZ bind soluble CD20 in a manner that closely mimics binding to native CD20.

To further interrogate antibody binding to soluble CD20, gel filtration was used to purify the soluble CD20/Fab complexes, and the samples were examined with mass photometry. The RTX sample displayed a single species in solution with a molecular weight of 113 kDa, consistent with a complex of two Fabs per molecule of soluble CD20^{His} dimer. In contrast, the OBZ sample only displayed a peak at 47 kDa, which aligned with the expected mass of unbound Fab, suggesting that OBZ could be dissociated from soluble CD20^{His} at the experimental conditions used to perform mass photometry. (Fig 3c) Mass photometry requires low protein concentrations for detection and cannot detect species below 30-40 kDa. These experiments were performed at a complex concentration of approximately 50 nM, which is lower than the dissociation constant of OBZ for soluble CD20. Under these conditions, partial dissociation of the soluble CD20/OBZ Fab complex may have occurred, resulting in the observation of a lower molecular weight Fab peak and unbound soluble CD20 unobserved. The mass photometry results demonstrate that soluble CD20 can differentiate between the binding stoichiometries of type I and type II

anti-CD20 antibodies, which consistent with finding by SEC-MALS analysis. This suggests that soluble CD20 accurately replicates the stoichiometric behavior of native CD20 with these two different types of antibodies.

Soluble CD20 binds with native-like affinities to type I and II antibodies

The equilibrium dissociation constants (K_D) between soluble CD20^{Avi} and anti-CD20 mAbs were measured with bio-layer interferometry (BLI) (Figure 4). In these experiments, native CD20, CD20 antigen peptide, or soluble CD20^{Avi} were immobilized on the sensor tips and dipped into solution with varying concentrations of mAbs in an IgG or Fab format. Soluble CD20^{Avi} bound to RTX IgG with an affinity (1.6 nM) that is close to our measured affinity between native CD20 and RTX IgG (1.8 nM) as well as the reported affinity for this interaction (1.7 nM)(Rougé et al. 2020). In contrast, a commonly used antigen peptide derived from the ECL2 loop of CD20 only binds to RTX with an affinity of 2.2 μ M. Soluble CD20^{Avi} displayed affinities for RTX Fab (K_D = 20 nM) and OBZ Fab (68 nM) that are very similar to previously reported K_D s between native CD20 and RTX Fab (21 nM) and OBZ Fab (59 nM). We also observed that RTX Fab dissociates more slowly from soluble CD20^{Avi} than OBZ, consistent with the homotypic interactions formed between RTX Fabs when two Fab molecules bind to the CD20 dimer. Overall, these results indicate that soluble CD20 presents the epitopes for type I and type II antibodies in a manner that closely resembles native CD20.

Soluble CD20 can be used as a high-affinity antigen for yeast surface display

High-throughput display technologies such as phage, yeast, and mammalian display are efficient approaches for screening therapeutic binder libraries. These technologies typically require that the target antigen be stable, pure, and soluble to perform selection, presenting a significant challenge when dealing with membrane proteins. To test if engineered soluble CD20 is an effective target for yeast display experiments, we evaluated native CD20, CD20 primary epitope peptides, and our designed soluble CD20^{Avi} as target antigens in yeast surface display. We quantified the binding efficiency of these target proteins using yeast cells engineered to display a single-chain variable fragment (scFv) derived from rituximab. The expression level of the scFv on the yeast

surface was measured using a C-myc tag at the scFv's C-terminus, and the binding interactions between the displayed scFv and the target proteins were assessed via a fluorescent-labeled streptavidin signal, which binds to the biotinylated target protein (Chao et al. 2006).

The results indicate that a CD20 primary epitope peptide derived from ECL2 of CD20 is a poor target as low binding was detected to the rituximab scFv even at a peptide concentration of 10 μ M. Native CD20 also exhibited low binding efficiency as a target antigen in the yeast display setup. When used as a soluble target protein, native CD20 required solubilization in detergent micelles, which may mask extracellular loops and non-specifically interfere with antibody binding. Despite high binding affinities characterized in biolayer interferometry (BLI) and surface plasmon resonance (SPR), native CD20 solubilized in surfactant micelles was inefficient as an antigen protein in yeast display at a concentration of 0.6 μ M, with only $29.5 \pm 7.8\%$ of yeast cells displaying a positive binding signal. This aligns with the known challenges of using membrane proteins solubilized by detergent micelles in display technologies.

In contrast, the designed soluble CD20, which retained most antigenic features without the need for surfactant micelles, exhibited high binding efficiency as a target protein in yeast display. Among yeast cells expressing positive scFv, $98.1 \pm 1.1\%$ showed a positive binding signal at a soluble CD20 concentration of 22 nM. Further experiments were conducted with soluble CD20 concentrations ranging from 600 nM to 0.8 nM in 1:3 titrations to assess the binding efficiency on rituximab scFv-displayed yeast cells. Remarkably, $43.1 \pm 8.2\%$ of cells demonstrated binding at the lowest tested concentration of 0.8 nM. These findings underscore the efficacy of soluble CD20 as a high-performance target protein in yeast display settings.

Discussion

There is long standing interest in methods that can be used to redesign complex membrane proteins (CMPs) to be folded and soluble in water (Rawlings 2016). One

approach is to redesign hydrophobic amino acids on the surface of the transmembrane segments of the protein to be polar amino acids (Roosild and Choe 2005; Smorodina et al. 2024). Recent AI-based methods for sequence design have made this strategy more robust and have been used to solubilize GPCRs, the claudin-like fold, and a 6 transmembrane protein with the rhomboid protease fold (Goverde et al. 2024). One attractive aspect of this approach is that the protein remains full-length and if the design process is successful the overall fold of the protein should match its structure when embedded in the membrane. However, there are cases where it may only be important to maintain key structural features from the extracellular (or intracellular) components of the protein. In our study, our aim was to present the primary extracellular loop of CD20 as a homodimer with interchain positioning that mimics the wild-type protein. We were able to achieve this goal by scaffolding the loop with a *de novo* designed coiled-coil. Useful aspects of this approach include the high stability and solubility of *de novo* designed proteins and the ability to cut away regions of the membrane protein what may be problematic for downstream studies (Goverde et al. 2024). Also, this is a design approach that will benefit significantly from recent advances in motif-scaffolding with generative AI methods such as RFDiffusion (Watson et al. 2023).

Since full length CD20 is challenging to produce and is not easily adopted to screening or immunization studies, it has been common to use peptides and peptide fusion derived from the extracellular loops of CD20 as surrogates for the full-length protein. It has been hypothesized that peptides of this type could be used to elicit an anti-CD20 antibody response and provide active immunotherapy against autoimmune diseases (Perosa et al. 2005; Li et al. 2006; Liu et al. 2016). Our results highlight that these peptides are unlikely to be good mimics of CD20. In BLI experiments with a cyclic peptide derived from ECL2 the measured K_d for rituximab was only 2 μ M, while soluble CD20 and native CD20 bind rituximab with a K_d of 2 nM. The similar affinity of soluble CD20 and native CD20 for rituximab suggests that the cleft formed between two copies of ECL2 is the primary structural feature contributing to binding affinity.

The antibody binding studies with rituximab and obinutuzumab indicate that soluble CD20 presents ECL2 in a native-like manner. We also tested binding with the anti-CD20

antibody ofatumumab. The binding epitope for ofatumumab includes residues 71-80 on ECL1 and residues 146-160 on ECL2. Since soluble CD20 does not include ECL1, it is not surprising that only modest affinity was observed for ofatumumab via BLI, and soluble CD20 and ofatumumab failed to form a stable complex when probed with SEC-MAS (Figure S4). This suggests that current engineered soluble CD20 can only be used to discover antibodies or antibody-like molecules that interact with ECL2 only.

In summary, we have shown that computational protein design can be used to present a critical dimer epitope from CD20 in a native-like conformation in a protein that is stable and soluble in water. Furthermore, we demonstrated that soluble CD20 is amenable to yeast display experiments and binds effectively to anti-CD20 binders presented on the surface of yeast. We anticipate that soluble CD20 will be useful for discovering non-antibody binders against CD20 and will be a useful reagent for characterizing newly discovered anti-CD20 antibodies. Our study highlights the use of computational design for epitope scaffolding to create water-soluble analogs of challenging transmembrane proteins, potentially broadening the scope of drug development.

Materials and Methods

Rosetta computational design methods for Soluble CD20 Dimer

The cryo-EM structure of full-length CD20 dimer (PDB code: 6VJA) served as the starting point for the design process. We designed chimeras that scaffolded the ECL2 loop and the replaced transmembrane domain with a de novo coiled coil. First, a large family of idealized dimeric coiled-coils with 42-residues were generated. In PyRosetta, we described the superhelical twist and radius of a coiled-coil (Grigoryan and DeGrado 2011a). We then probed which of these coiled coils most closely align structurally ($C\alpha$ RMSD) with residues from the n-terminal portions of TM4 in the CD20 dimer. More specifically we calculated the root mean square deviation (RMSD) between the $C\alpha$ atoms on residues 1 and 2 from both chains of the idealized coiled coil with the $C\alpha$ atoms on residues 184 and 185 (residue numbering from PDB 6VJA) from both chains of CD20. Ten coiled coils, which superimpose with low RMSD (≤ 0.25 Å) were selected to build full length models of the designed proteins. These models were constructed by grafting residues 158-185 from CD20 onto the N terminus of the matching coiled coils, resulting in protein chains with sixty-eight residues.

Sequences were designed for the CD20/coiled coil chimeras using the FastDesign protocol in the molecular modeling program Rosetta. FastDesign iterates between rotamer-based sequence optimization and backbone refinement to find low energy sequence/structure pairs (Maguire et al. 2021). During these simulations the sequence of the ECL2 was fixed as the native sequence and the residues in the coiled coil were allowed to vary to the following amino acids ADEIKLNQRSTV. The native disulfide between residues 167 and 183 in the ECL2 was maintained during the simulations. To preserve sequence symmetry between the two chains of the homodimer, symmetry restraints were applied to designable residues. The total energy of the design model as calculated with the full atom energy function (energy function beta_nov15) binding energy normalized by interface area (“dG_separated/dSASA”). From 400 design trajectories (40 for each coiled-coil template), 8 sequences were selected for experimental characterization. The scripts used for the design runs are provided in the supplementary materials (supplementary table 1).

Cloning, expression, and purification of soluble CD20

DNA encoding the design sequences were synthesized (Twist biology) and cloned into a modified pαH mammalian expression vector. Two different soluble CD20 constructs, soluble CD20^{His} and soluble CD20^{Avi} (supplementary table 2) were produced. Soluble CD20 was expressed in EXPI293F cells (Thermo Fisher Scientific, catalog A14527) with a human serum albumin signal peptide for secretion to the culture medium. Constructs were transiently transfected according to the manufacturer's protocol with a modified cell density of 3.0×10^6 cells/ml. EXPI293F cells were maintained at 37°C with 8% CO₂ at 250 rpm without antibiotics. Enhancers were added according to the manufacturer's protocol 20 hours after transfection. The cell cultures were harvested 96 hours after transfection and the supernatants were separated from the cells by spinning at 3000 rfc (4°C) and passing through 0.4µm filters. The supernatants were then mixed with penta-Ni resin (Marvelgent) under gentle shaking at 4 °C for 30 minutes to allow protein binding. After incubation, resin was washed with 2.5 column volumes (CV) of 10mM imidazole in high salt Tris buffer (1M NaCl, 50mM Tris base in PH 8), 1 column volumes(CV) of 25mM imidazole in high salt tris buffer and 1 CV of 50mM Imidazole in PBS, and elution was performed by 4 steps of 1 CV PBS with 500mM Imidazole.

Escherichia coli Shuffle T7 cells (Thermo Fisher Scientific, USA) were used for protein expression due to their high efficiency and compatibility with T7 promoter-driven expression systems. The gene of interest was cloned into the expression vector pET-28a(+), containing an N-terminal His-tag as described in supplementary table 2 Soluble CD20^{bacteria}. Plasmid DNA was transformed into competent Shuffle T7 cells using the heat-shock method. A single colony was picked and grown in 5 mL of LB medium with the Ampicillin overnight at 37°C, with shaking at 250 rpm. The overnight culture was used to inoculate 50ml LB and then 1L of fresh TB culture. Cells were induced with IPTG(final concentration of 0.5 mM) at OD600 of 0.6-0.8. The temperature was then dropped to 16°C and shake at 250rpm overnight. Cells were harvested by centrifugation at 4,000 x g for 30 minutes at 4°C. The cell pellet was frozen and stored at -80 °C. Cells were lysed in lysis buffer (500mM NaCl, 50mM Tris HCl pH 7.5, 10mM Imidazole, PMSF 1mM,

Pepstatin 10 μ M, Bestatin 1 μ M, Leupeptin 10 μ M) by sonication on ice (5 seconds on, 5 seconds off, at 70% amplification for 5 minutes).

The lysate was clarified by centrifugation at 12,000 x g for 30 minutes at 4°C. The supernatant was batch bind with penta-Ni resin (Marvelgent) under gentle shaking at 4 °C for 30 minutes to allow protein binding. The column was washed with 20 column volumes of high salt wash buffer (1M NaCl, 50mM Tris HCl pH 7.5, 10mM Imidazole) , 2.5 column volume of wash buffer (500mM NaCl, 50mM Tris HCl pH 7.5, 25mM Imidazole) to remove non-specifically bound proteins. The target protein was eluted with elution buffer (500mM NaCl, 50mM Tris HCl pH 7.5, 500mM Imidazole).

Generation of CD20 peptide and native CD20

The CD20 epitope peptide NIYN(CEPANPSEKNPSTQYC)YSIQ-K(Biot)-NH₂ was synthesized and purified by the UNC High-throughput Peptide Synthesis and Array Facility. Native CD20 was recombinantly expressed and purified by cell membrane solubilization. Native CD20 constructs were synthesized (Twist biology) and cloned into a pFastBac vector. Recombinant baculovirus was generated using the Bac-to-Bac® Baculovirus Expression System (invitrogen). Sf9 cells were infected for protein production and harvested 72 hrs post-infection. The sequence of native CD20 was taken from the cryo-EM structure study (Rougé et al. 2020). Native CD20s were expressed in Sf9 cells that was conducted using a baculovirus encoding for recombinant Multidrug resistance-associated protein 4(MRP4) generated from a pFastBac-his6-TEV-CD20-FLAG-AviTag.

Native CD20 was purified by solubilizing the membranes from a large scale Sf9 cell culture. 2L Sf9 cells were spin down 3200 g for 15–30 min at 4 C and resuspended in 100ml 25 mM Tris pH 7.5, 150 mM NaCl with complete EDTA-free protease-inhibitor cocktail tablets (Roche). The cell suspension was homogenized by 30-40 strokes in a 40 mL Dounce homogenizer on ice, then spun down at 8,000 rpm for 20 min. Supernatant that contains the cell membrane component was separated by spinning at 40,000 rpm at 4 °C for 1 hour. Isolated membrane components were resuspended in buffer with 1% (w/v) GDN (Anatrace) and 0.2% (w/v) cholesterol hemisuccinate in 25 mM Tris pH 7.5, 150 mM NaCl with surfactant as suggested (Rougé et al. 2020). Solubilization was achieved using

gentle agitation for at least 4 hours at 4 °C. Ultracentrifugation at 40,000 rpm at 4 °C for 30 min was used to remove insoluble debris. Clarified supernatant was used for 1st Nickel-His purification with penta-Ni resin (Marvelgent). Supernatant were incubated with pre-equilibrated penta-Ni resin under gentle shaking at 4 °C overnight and then washed with 10 column volumes of detergent buffer (25 mM Tris pH 7.5, 150 mM NaCl, 0.02% GDN) plus 10 mM imidazole and eluted with detergent buffer plus 250mM imidazole. The eluted samples were loaded onto a Superdex S200 10/300 GL column (GE Healthcare) pre-equilibrated with detergent buffer, and major fractions were collected for the 2^{ed} TEV cutting purification. These fractions were inculcated with His-tagged TEV protease to cleave the tags, followed by passage through a fresh Ni-NTA column equilibrated with detergent buffer. Untagged CD20 was recovered in the flow-through and wash fractions after washing and elution. The eluted samples were loaded onto a Superdex S200 10/300 GL column (GE Healthcare) pre-equilibrated with detergent buffer, and the major fractions were collected. These fractions were used in 2^{ed} TEV cutting purification. Fractions were incubated overnight with His-tagged TEV protease prepared in-house to cleave the tags and then passed through a new penta-Ni column that equilibrated with detergent buffer. After washing and elution steps, as previously described, the flow-through and wash fractions containing the untagged CD20 were collected.

Assessment of De Novo CD20 protein stability with Circular dichroism (CD)

Each purified protein was centrifuged at 14,000 × g for 10 min at 4°C and the supernatant diluted to 0.5uM, 2uM, 20uM and 40Um in PBS. Circular dichroism spectra and temperature melts were acquired with a Jasco CD J-815 C. Thermal unfolding was monitored by ramping at 1°C intervals from 20°C to 90°C and measuring the CD signal at 222 nm during the temperature ramp. Molar ellipticity was calculated with the following expression: $[\theta]=\theta/(10 \times c \times l)$ where l is the path length in cm. The midpoint for thermal

unfolding was determined by fitting the data to the temperature dependent Gibbs-Helmholtz Equation using non-linear regression tools in Microsoft Excel.

Bio-Layer Interferometry (BLI) measurements for binding affinity

Bio-Layer Interferometry (BLI) assays were performed to determine the binding affinity between selected antibodies and native or de novo designed CD20 receptors. Biotinylated native or soluble CD20 was immobilized on streptavidin SA biosensors (FortéBio / Molecular Devices) at a concentration of 30ug/ml. The biosensors were moved to wells with 5-fold serial dilutions of antibodies to assess association and dissociation. One additional reference well with unloaded biosensors was mixed with the highest antibody concentration well and were used to normalize data. The association and dissociation of 5 concentrations of antibodies were used to fit kinetic data to a 2:1 binding mode to determine the affinity constant (KD) using the association (k_{on}) and dissociation (k_{off}) rates using software (Octet® Analysis Studio). For de novo CD20, kinetic buffer (kinetics buffer 10X (No. 18-1105) was used for all washing, dilution and sensor activation. For native CD20, PBS buffer containing 0.02% GDN supplemented was use for all washing, dilution and sensor activation.

Size-Exclusion Chromatography-Multiangle Light Scattering (SEC-MALS)

Purified soluble CD20 and CD20 Fab antibodies were mixed in a molar ratio 1.5:1 and incubated for 30 min at 4°C with rotation to form complexes. Proteins were then injected onto a superdex column (Superdex 200 Increase 10/300 GL Cytiva 28-9909-44) equilibrated in PBS at 0.5ml/min and probed with an inline light scattering detector (Wyatt DAWN HELEOS II static light scattering system).

Mass photometry

Purified soluble CD20 and CD20 Fab antibodies were mixed in a ratio 1.5:1 and incubated for 30 min at 4°C with rotation to form complexes, and then further purified with gel filtration (Superdex 200 Increase 10/300 GL Cytiva 28-9909-44). The purified complexes were prepared at a concentration of 17 μ M in PBS buffer and then diluted to a concentration of 170 nM immediately before mass photometry. 2 μ L of sample were loaded into a 10 μ L buffer-containing well on clean coverslips. The data were acquired using a Refeyn One mass photometer, Refeyn AcquireMP for 60s and analyzed with the DiscoverMP (v2.3) software.

Yeast surface display

The single-chain variable fragment (scFv) of rituximab was cloned into the pCTCON2 yeast display vector, which includes a myc-tag for expression verification. *Saccharomyces cerevisiae* strain EBY100 competent cell was prepared and transformed with the pCTCON2-scFv plasmid using electroporation following a standard transformation protocol (Chao et al. 2006). Transformed yeast cells were recovered on YPD media and resuspended in SDCAA media. A single colony was picked and cultured twice for 48 hours in 5 mL SDCAA medium at 30°C. 5×10^7 cells were used to inoculate 5 ml SGCAA culture and grown for 20 hours at 20°C, 220 rpm for induction.

Binding assays were performed using three forms of biotinylated CD20 antigens: CD20 peptide (10 μ M), native CD20 (0.6 μ M), and a serial dilution of soluble CD20 ranging from 0.8 nM to 0.6 μ M. The binding signal of CD20 antigens were detected through Alexa Fluor™ 647-conjugated streptavidin (Invitrogen, S21374) that binds to the biotinylated CD20 antigens. Surface expression of the scFv was confirmed by flow cytometry using FITC-conjugated anti-Myc tag antibody (Abcam, ab1394). After induction, 5×10^6 yeast cells displaying rituximab scFv were collected and were washed twice with PBSF buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na₂HPO₄, and 2 mM KH₂PO₄, supplemented with 0.1% BSA). Cells were incubated with FITC-conjugated anti-Myc tag antibody for 1 hour at 4°C with shaking, followed by three washes with PBSF buffer. CD20 antigens were preloaded with Alexa Fluor™ 647-conjugated streptavidin at a molar ratio of 4:1 for 30

minutes at room temperature before being incubated with yeast cells for 1 hour at 4°C with shaking. After incubation, cells were washed three times with PBSF buffer.

Flow cytometry was performed on a Thermo Fisher Attune NxT instrument. The expression signal (myc-positive) and the antigen binding signal was detected in the FITC channel and Alexa Fluor™ 647 channel respectively. Control samples, that labeled with fluorescent Streptavidin conjugate but without antigen, were used to establish background fluorescence levels. Flow cytometry data were analyzed using FlowJo software. The binding positive rate of each CD20 antigens were quantified as the percentage of double-positive cells (scFv-expressing and antigen-binding cells) among total scFv-expressing cells.

Supplementary material description

Supplementary table 1. Pyrosetta and Rosetta protocols used to design soluble CD20.

Supplementary table 2. List of designed soluble and native CD20 protein sequences and constructs.

Supplementary table 3. List of anti-CD20 antibodies used in soluble expression and yeast display.

Supplementary Figure 1. Expression of all designed soluble CD20 sequences and native CD20

Supplementary Figure 2. Expression and purification of Native CD20 with membrane solubilization

Supplementary Figure 3. Components of distinct peaks from size exclusion chromatography with soluble CD20 and Fab antibody mixtures.

Supplementary Figure 4. Binding affinity and binding complex characterization of ofatumumab Fab and soluble CD20.

Acknowledgements

The authors thank Kiara Thompson for help in expressing soluble CD20 in *e.coli*, Nicole Hajicek and Runfan Yang in John Sondek's lab for advice regarding membrane protein expression in SF9 cells, and Matthew Cummins for advice and discussion of the project. We are grateful to Julia Koehler Leman for her help in manuscript writing. We thank the UNC Macromolecular Interactions Facility and UNC Flow Cytometry Core Facility for advice and instrumentation. This work was supported by the NIH grant R35GM131923 (BK).

Conflict of interest statement

The authors have no conflict of interests related to this research.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon request.

Reference

- Alduaij, Waleed, and Tim M. Illidge. 2011. "The Future of Anti-CD20 Monoclonal Antibodies: Are We Making Progress?" *Blood* 117 (11): 2993–3001. <https://doi.org/10.1182/blood-2010-07-298356>.
- Beers, Stephen A., Ruth R. French, H. T. Claude Chan, Sean H. Lim, Timothy C. Jarrett, Regina Mora Vidal, Sahan S. Wijayaweera, et al. 2010. "Antigenic Modulation Limits the Efficacy of Anti-CD20 Antibodies: Implications for Antibody Selection." *Blood* 115 (25): 5191–5201. <https://doi.org/10.1182/blood-2010-01-263533>.
- Cao, Longxing, Brian Coventry, Inna Goreshnik, Buwei Huang, William Sheffler, Joon Sung Park, Kevin M. Jude, et al. 2022. "Design of Protein-Binding Proteins from the Target Structure Alone." *Nature* 605 (7910): 551–60. <https://doi.org/10.1038/s41586-022-04654-9>.
- Cecchetti, Cristina, Jannik Strauss, Claudia Stohrer, Claire Naylor, Edward Pryor, Jeanette Hobbs, Simon Tanley, Adrian Goldman, and Bernadette Byrne. 2021. "A Novel High-Throughput Screen for Identifying Lipids That Stabilise Membrane Proteins in Detergent Based Solution." *PLOS ONE* 16 (7): e0254118. <https://doi.org/10.1371/journal.pone.0254118>.
- Chao, Ginger, Wai L. Lau, Benjamin J. Hackel, Stephen L. Sazinsky, Shaun M. Lippow, and K. Dane Wittrup. 2006. "Isolating and Engineering Human Antibodies Using Yeast Surface Display." *Nature Protocols* 1 (2): 755–68. <https://doi.org/10.1038/nprot.2006.94>.
- Chen, Ximin, Laurence C. Chen, Mobina Khericha, Xiangzhi Meng, Emma Salvestrini, Amanda Shafer, Neha Iyer, et al. 2023. "Rational Protein Design Yields a CD20 CAR with Superior Antitumor Efficacy Compared with CD19 CAR." *Cancer Immunology Research* 11 (2): 150–63. <https://doi.org/10.1158/2326-6066.CIR-22-0504>.
- Cherf, Gerald M., and Jennifer R. Cochran. 2015. "Applications of Yeast Surface Display for Protein Engineering." *Methods in Molecular Biology (Clifton, N.J.)* 1319:155–75. https://doi.org/10.1007/978-1-4939-2748-7_8.

- Cragg, Mark S., and Martin J. Glennie. 2004. "Antibody Specificity Controls in Vivo Effector Mechanisms of Anti-CD20 Reagents." *Blood* 103 (7): 2738–43. <https://doi.org/10.1182/blood-2003-06-2031>.
- Dabkowska, Agnieszka, Krzysztof Domka, and Malgorzata Firczuk. 2024. "Advancements in Cancer Immunotherapies Targeting CD20: From Pioneering Monoclonal Antibodies to Chimeric Antigen Receptor-Modified T Cells." *Frontiers in Immunology* 15 (April). <https://doi.org/10.3389/fimmu.2024.1363102>.
- Dodd, Roger B., Trevor Wilkinson, and Darren J. Schofield. 2018. "Therapeutic Monoclonal Antibodies to Complex Membrane Protein Targets: Antigen Generation and Antibody Discovery Strategies." *BioDrugs* 32 (4): 339–55. <https://doi.org/10.1007/s40259-018-0289-y>.
- Dodd, Roger, Darren J. Schofield, Trevor Wilkinson, and Zachary T. Britton. 2020. "Generating Therapeutic Monoclonal Antibodies to Complex Multi-Spanning Membrane Targets: Overcoming the Antigen Challenge and Enabling Discovery Strategies." *Methods, Membrane Protein Tools for Drug Discovery*, 180 (August):111–26. <https://doi.org/10.1016/j.ymeth.2020.05.006>.
- Fousek, Kristen, Junji Watanabe, Sujith K. Joseph, Ann George, Xingyue An, Tiara T. Byrd, Jessica S. Morris, et al. 2021. "CAR T-Cells That Target Acute B-Lineage Leukemia Irrespective of CD19 Expression." *Leukemia* 35 (1): 75–89. <https://doi.org/10.1038/s41375-020-0792-2>.
- Frauenfeld, Jens, Robin Löving, Jean-Paul Armache, Andreas F.-P. Sonnen, Fatma Guettou, Per Moberg, Lin Zhu, et al. 2016. "A Saposin-Lipoprotein Nanoparticle System for Membrane Proteins." *Nature Methods* 13 (4): 345–51. <https://doi.org/10.1038/nmeth.3801>.
- Goverde, Casper A., Martin Pacesa, Nicolas Goldbach, Lars J. Dornfeld, Petra E. M. Balbi, Sandrine Georgeon, Stéphane Rosset, et al. 2024. "Computational Design of Soluble and Functional Membrane Protein Analogues." *Nature* 631 (8020): 449–58. <https://doi.org/10.1038/s41586-024-07601-y>.
- Grigoryan, Gevorg, and William F. DeGrado. 2011a. "Probing Designability via a Generalized Model of Helical Bundle Geometry." *Journal of Molecular Biology* 405 (4): 1079–1100. <https://doi.org/10.1016/j.jmb.2010.08.058>.

- Hashimoto, Yosuke, Wei Zhou, Kohtaroh Hamauchi, Keisuke Shirakura, Takefumi Doi, Kiyohito Yagi, Tatsuya Sawasaki, Yoshiaki Okada, Masuo Kondoh, and Hiroyuki Takeda. 2018. "Engineered Membrane Protein Antigens Successfully Induce Antibodies against Extracellular Regions of Claudin-5." *Scientific Reports* 8 (1): 8383. <https://doi.org/10.1038/s41598-018-26560-9>.
- Huang, Buwei, Mohamad Abedi, Green Ahn, Brian Coventry, Isaac Sappington, Cong Tang, Rong Wang, et al. 2024. "Designed Endocytosis-Inducing Proteins Degrade Targets and Amplify Signals." *Nature*, September, 1–9. <https://doi.org/10.1038/s41586-024-07948-2>.
- Jaroszewicz, Weronika, Joanna Morcinek-Orłowska, Karolina Pierzynowska, Lidia Gaffke, and Grzegorz Węgrzyn. 2022. "Phage Display and Other Peptide Display Technologies." *FEMS Microbiology Reviews* 46 (2): fuab052. <https://doi.org/10.1093/femsre/fuab052>.
- Klein, Christian, Alfred Lammens, Wolfgang Schäfer, Guy Georges, Manfred Schwaiger, Ekkehard Mössner, Karl-Peter Hopfner, Pablo Umaña, and Gerhard Niederfellner. 2013. "Epitope Interactions of Monoclonal Antibodies Targeting CD20 and Their Relationship to Functional Properties." *mAbs* 5 (1): 22–33. <https://doi.org/10.4161/mabs.22771>.
- Kortemme, Tanja. 2024. "De Novo Protein Design-From New Structures to Programmable Functions." *Cell* 187 (3): 526–44. <https://doi.org/10.1016/j.cell.2023.12.028>.
- Kotov, Vadim, Kim Bartels, Katharina Veith, Inokentij Josts, Udaya K. Tiruttani Subhramanyam, Christian Günther, Jörg Labahn, et al. 2019. "High-Throughput Stability Screening for Detergent-Solubilized Membrane Proteins." *Scientific Reports* 9 (1): 10379. <https://doi.org/10.1038/s41598-019-46686-8>.
- Krishnarjuna, Bankala, and Ayyalusamy Ramamoorthy. 2022. "Detergent-Free Isolation of Membrane Proteins and Strategies to Study Them in a Near-Native Membrane Environment." *Biomolecules* 12 (8): 1076. <https://doi.org/10.3390/biom12081076>.
- Kumar, Anand, Cyril Planchais, Rémi Fronzes, Hugo Mouquet, and Nicolas Reyes. 2020. "Binding Mechanisms of Therapeutic Antibodies to Human CD20." *Science* 369 (6505): 793–99. <https://doi.org/10.1126/science.abb8008>.

- Li, Meng, Zhen Yan, Wei Han, and Yingqi Zhang. 2006. “Mimotope Vaccination for Epitope-Specific Induction of Anti-CD20 Antibodies.” *Cellular Immunology* 239 (2): 136–43. <https://doi.org/10.1016/j.cellimm.2006.05.005>.
- Liu, Jinny L., Dan Zabetakis, Ellen R. Goldman, and George P. Anderson. 2016. “Selection and Characterization of Single Domain Antibodies against Human CD20.” *Molecular Immunology* 78 (October):146–54. <https://doi.org/10.1016/j.molimm.2016.09.013>.
- Luo, Chengxin, Guixian Wu, Xiangtao Huang, Yanni Ma, Yali Zhang, Qiuyue Song, Mingling Xie, et al. 2021. “Efficacy and Safety of New Anti-CD20 Monoclonal Antibodies versus Rituximab for Induction Therapy of CD20+ B-Cell Non-Hodgkin Lymphomas: A Systematic Review and Meta-Analysis.” *Scientific Reports* 11 (1): 3255. <https://doi.org/10.1038/s41598-021-82841-w>.
- Ma, Yanbin, Yao Ding, Xianqiang Song, Xiaochuan Ma, Xun Li, Ning Zhang, Yunpeng Song, et al. 2020. “Structure-Guided Discovery of a Single-Domain Antibody Agonist against Human Apelin Receptor.” *Science Advances* 6 (3): eaax7379. <https://doi.org/10.1126/sciadv.aax7379>.
- Maguire, Jack B., Hugh K. Haddox, Devin Strickland, Samer F. Halabiya, Brian Coventry, Jermel R. Griffin, Surya V. S. R. K Pulavarti, et al. 2021. “Perturbing the Energy Landscape for Improved Packing during Computational Protein Design.” *Proteins: Structure, Function, and Bioinformatics* 89 (4): 436–49. <https://doi.org/10.1002/prot.26030>.
- McLaughlin, P., A. J. Grillo-López, B. K. Link, R. Levy, M. S. Czuczman, M. E. Williams, M. R. Heyman, et al. 1998. “Rituximab Chimeric Anti-CD20 Monoclonal Antibody Therapy for Relapsed Indolent Lymphoma: Half of Patients Respond to a Four-Dose Treatment Program.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 16 (8): 2825–33. <https://doi.org/10.1200/JCO.1998.16.8.2825>.
- Perosa, Federico, Elvira Favoino, Maria Antonietta Caragnano, and Franco Dammacco. 2005. “CD20 Mimicry by a mAb Rituximab-Specific Linear Peptide: A Potential Tool for Active Immunotherapy of Autoimmune Diseases.” *Annals of the New*

- York Academy of Sciences* 1051 (1): 672–83.
<https://doi.org/10.1196/annals.1361.112>.
- Plückthun, Andreas. 2015. “Designed Ankyrin Repeat Proteins (DARPs): Binding Proteins for Research, Diagnostics, and Therapy.” *Annual Review of Pharmacology and Toxicology* 55:489–511. <https://doi.org/10.1146/annurev-pharmtox-010611-134654>.
- Qing, Rui, Shilei Hao, Eva Smorodina, David Jin, Arthur Zalevsky, and Shuguang Zhang. 2022. “Protein Design: From the Aspect of Water Solubility and Stability.” *Chemical Reviews* 122 (18): 14085–179.
<https://doi.org/10.1021/acs.chemrev.1c00757>.
- Qing, Rui, Mantian Xue, Jiayuan Zhao, Lidong Wu, Andreas Breitwieser, Eva Smorodina, Thomas Schubert, et al. 2023. “Scalable Biomimetic Sensing System with Membrane Receptor Dual-Monolayer Probe and Graphene Transistor Arrays.” *Science Advances* 9 (29): eadf1402.
<https://doi.org/10.1126/sciadv.adf1402>.
- Rawlings, Andrea E. 2016. “Membrane Proteins: Always an Insoluble Problem?” *Biochemical Society Transactions* 44 (3): 790–95.
<https://doi.org/10.1042/BST20160025>.
- Roosild, Tarmo P., and Senyon Choe. 2005. “Redesigning an Integral Membrane K⁺ Channel into a Soluble Protein.” *Protein Engineering, Design and Selection* 18 (2): 79–84. <https://doi.org/10.1093/protein/gzi010>.
- Rougé, Lionel, Nancy Chiang, Micah Steffek, Christine Kugel, Tristan I. Croll, Christine Tam, Alberto Estevez, et al. 2020. “Structure of CD20 in Complex with the Therapeutic Monoclonal Antibody Rituximab.” *Science* 367 (6483): 1224–30.
<https://doi.org/10.1126/science.aaz9356>.
- Shah, Nirav N., Bryon D. Johnson, Dina Schneider, Fenlu Zhu, Aniko Szabo, Carolyn A. Keever-Taylor, Winfried Krueger, et al. 2020. “Bispecific Anti-CD20, Anti-CD19 CAR T Cells for Relapsed B Cell Malignancies: A Phase 1 Dose Escalation and Expansion Trial.” *Nature Medicine* 26 (10): 1569–75.
<https://doi.org/10.1038/s41591-020-1081-3>.

- Skuhersky, Michael A., Fei Tao, Rui Qing, Eva Smorodina, David Jin, and Shuguang Zhang. 2021. "Comparing Native Crystal Structures and AlphaFold2 Predicted Water-Soluble G Protein-Coupled Receptor QTY Variants." *Life* 11 (12): 1285. <https://doi.org/10.3390/life11121285>.
- Smorodina, Eva, Igor Diankin, Fei Tao, Rui Qing, Steve Yang, and Shuguang Zhang. 2022. "Structural Informatic Study of Determined and AlphaFold2 Predicted Molecular Structures of 13 Human Solute Carrier Transporters and Their Water-Soluble QTY Variants." *Scientific Reports* 12 (1): 20103. <https://doi.org/10.1038/s41598-022-23764-y>.
- Smorodina, Eva, Fei Tao, Rui Qing, Steve Yang, and Shuguang Zhang. 2024. "Computational Engineering of Water-Soluble Human Potassium Ion Channels through QTY Transformation." *Scientific Reports* 14 (1): 28159. <https://doi.org/10.1038/s41598-024-76603-7>.
- Stephens, Amberley D., and Trevor Wilkinson. 2024. "Discovery of Therapeutic Antibodies Targeting Complex Multi-Spanning Membrane Proteins." *BioDrugs* 38 (6): 769–94. <https://doi.org/10.1007/s40259-024-00682-1>.
- Takeda, Hiroyuki, Tomio Ogasawara, Tatsuhiko Ozawa, Atsushi Muraguchi, Pei-Ju Jih, Ryo Morishita, Motokazu Uchigashima, et al. 2015. "Production of Monoclonal Antibodies against GPCR Using Cell-Free Synthesized GPCR Antigen and Biotinylated Liposome-Based Interaction Assay." *Scientific Reports* 5 (1): 11333. <https://doi.org/10.1038/srep11333>.
- To'a Salazar, Georgina, Ziyi Huang, Ningyan Zhang, Xue-Guang Zhang, and Zhiqiang An. 2021. "Antibody Therapies Targeting Complex Membrane Proteins." *Engineering* 7 (11): 1541–51. <https://doi.org/10.1016/j.eng.2020.11.013>.
- Tucker, David F., Jonathan T. Sullivan, Kimberly-Anne Mattia, Christine R. Fisher, Trevor Barnes, Manu N. Mabila, Rona Wilf, et al. 2018. "Isolation of State-Dependent Monoclonal Antibodies against the 12-Transmembrane Domain Glucose Transporter 4 Using Virus-like Particles." *Proceedings of the National Academy of Sciences* 115 (22): E4990–99. <https://doi.org/10.1073/pnas.1716788115>.

- Watson, Joseph L., David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, Woody Ahern, et al. 2023. “De Novo Design of Protein Structure and Function with RFdiffusion.” *Nature* 620 (7976): 1089–1100. <https://doi.org/10.1038/s41586-023-06415-8>.
- Willis, Sharon, Candice Davidoff, Justin Schilling, Antony Wanless, Benjamin J. Doranz, and Joseph Rucker. 2008. “Virus-like Particles as Quantitative Probes of Membrane Protein Interactions.” *Biochemistry* 47 (27): 6988–90. <https://doi.org/10.1021/bi800540b>.
- Yao, Xia, Xiao Fan, and Nieng Yan. 2020. “Cryo-EM Analysis of a Membrane Protein Embedded in the Liposome.” *Proceedings of the National Academy of Sciences* 117 (31): 18497–503. <https://doi.org/10.1073/pnas.2009385117>.
- Zhang, Shuguang, Fei Tao, Rui Qing, Hongzhi Tang, Michael Skuhersky, Karolina Corin, Lotta Tegler, et al. 2018. “QTY Code Enables Design of Detergent-Free Chemokine Receptors That Retain Ligand-Binding Activities.” *Proceedings of the National Academy of Sciences* 115 (37): E8652–59. <https://doi.org/10.1073/pnas.1811031115>.
- Zhen, Xia, Qihan Jin, Zhilin Long, Yexuan He, Fuyi Liu, Chengfang Sun, Jingyang Liao, et al. 2024. “Computational Design of Protein Binders That Boost the Antitumour Efficacy of CAR T Cells.” *Nature Biomedical Engineering*, October, 1–2. <https://doi.org/10.1038/s41551-024-01263-x>.

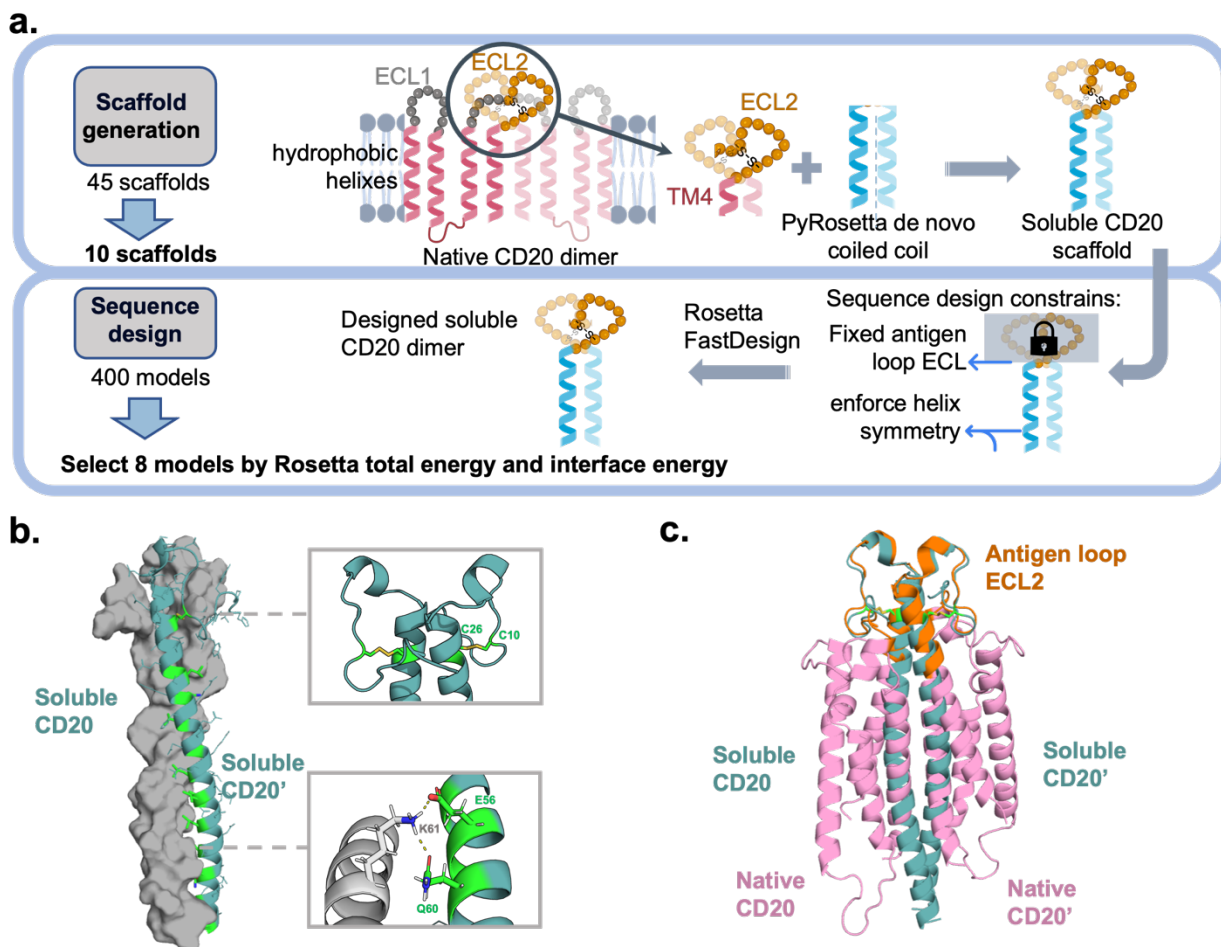


Figure 1. Design of a water soluble CD20 dimer mimic using Rosetta Design. a) The ECL2 loop of CD20 was grafted onto dimeric coiled coils and regions outside the antigen loop were sequence optimized with the Rosetta FastDesign protocol (Maguire et al. 2021). b) Final design model of soluble CD20 with one chain shown as a gray surface and one chain represented in ribbon mode. The native disulfide from the ECL2 loop is conserved in the design (top right inset) and the coiled coil dimer is stabilized by hydrophobic interactions and hydrogen bonding (lower right inset). c) A superposition of soluble CD20 (blue) and native CD20 (pink).

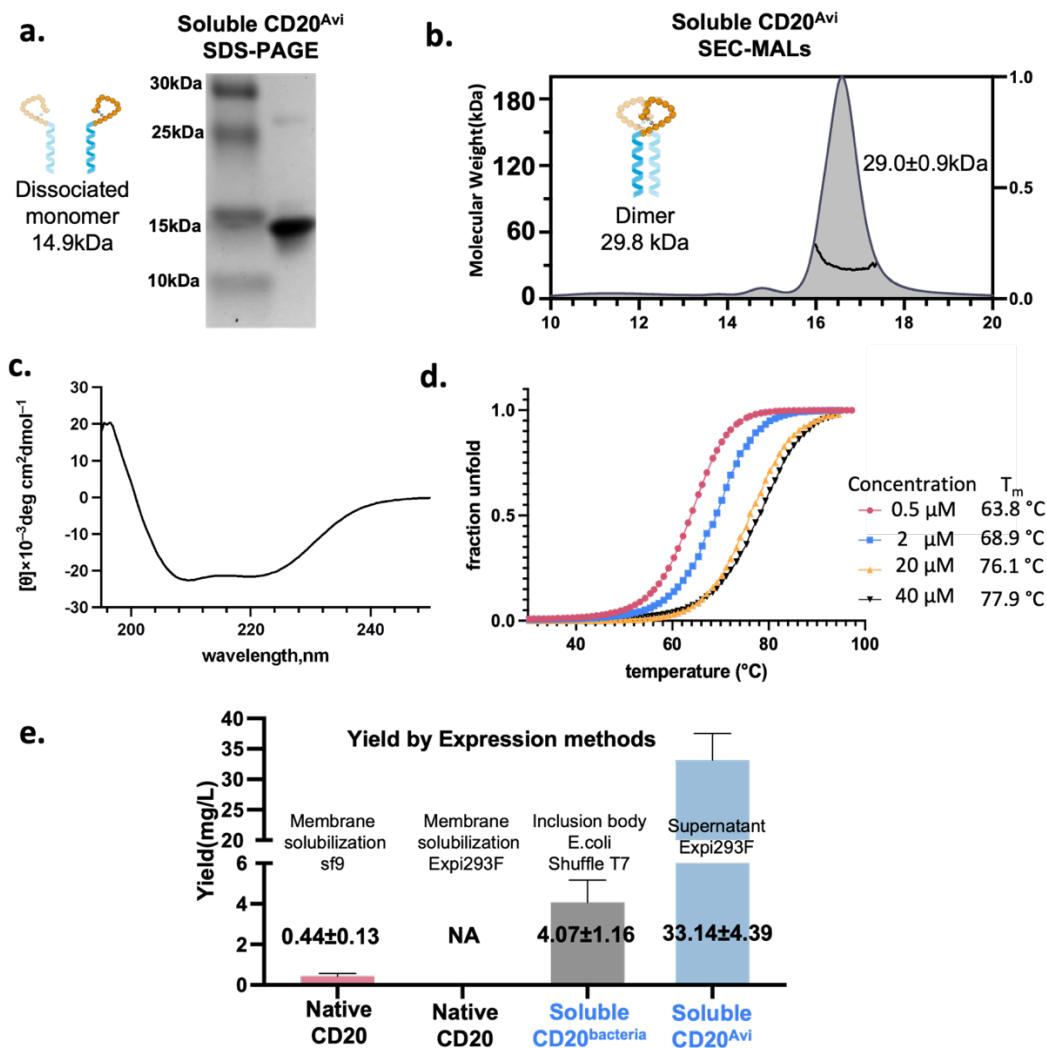


Figure 2. Expression and biophysical characterization of soluble CD20 as a stable dimer. a) Purified CD20^{Avi} visualized with SDS-PAGE. b) SEC-MALS profile of soluble CD20 shows a single peak around 29 kDa, which corresponds to its expected molecular weight of 29.8 kDa as a dimer. c) Circular dichroism spectra and d) thermal meltsTM of CD20^{Avi} at the indicated protein concentrations. CD data suggest that De novo CD20 adopts an α -helical secondary structure T_m increases in higher concentration suggests De novo CD20 assembles as oligomer and dissociates when temperatures increase. e) Expression yields of selected native and soluble CD20 constructs from sf9, Expi293F, or E. coli cells. Native CD20 is unattainable from Expi293 supernatant. While de novo CD20 achieves expression yield at 33mg/L.

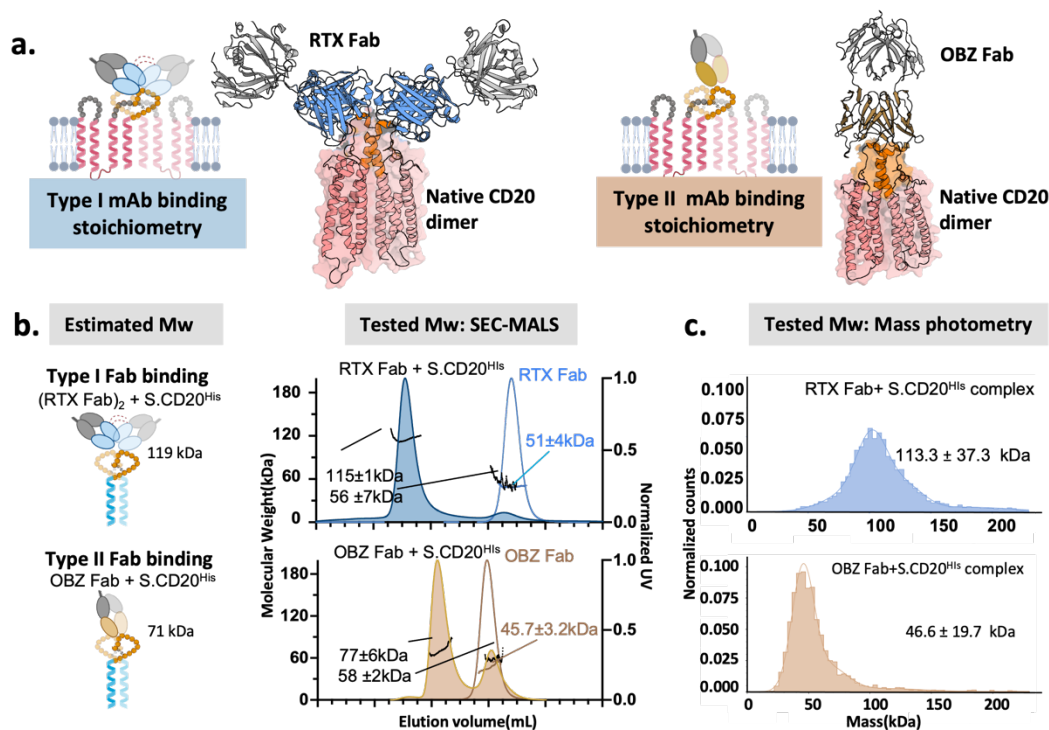


Figure 3. Soluble CD20 mimicked the binding stoichiometry of type I and type II antibodies a) Rituximab (RTX) is a type I anti-CD20 antibody that binds native CD20 with 2:2 stoichiometry. Obinutuzumab (OBZ) is a type II anti-CD20 antibody that binds native CD20 with 1:2 (Fab:CD20) stoichiometry. b) RTX and OBZ form complexes with soluble CD20^{His} and the measured molecular weights via SEC-MALS match the expected binding stoichiometries. c) Mass photometry experiments with soluble CD20^{His} mixed with RTX or OBZ. Protein concentrations were 50 nM.

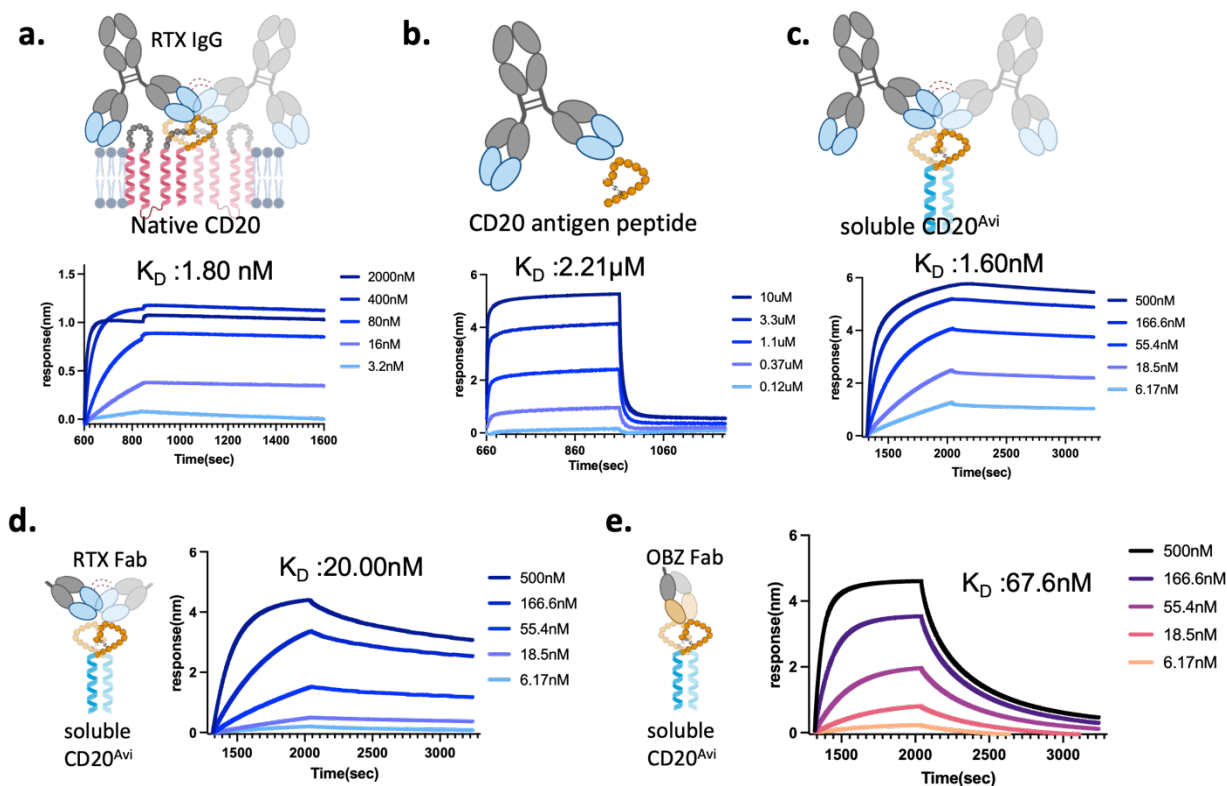


Fig. 4. Soluble CD20 mimics the binding affinity of types I and II antibodies. BLI experiments with a) native CD20, b) CD20 peptide and c) soluble CD20^{Avi} immobilized on the BLI sensor tips and incubated with the indicated concentrations of RTX IgG. BLI experiments with soluble CD20^{Avi} binding to Fabs from d) RTX or e) OBZ. The measured binding affinity of soluble CD20 to RTX and OBZ is 20.00nM and 67.6, which is in proximity to native CD20 binding affinity 21.4nM and 58.8nM reported in literature. (Rougé et al. 2020; Kumar et al. 2020)

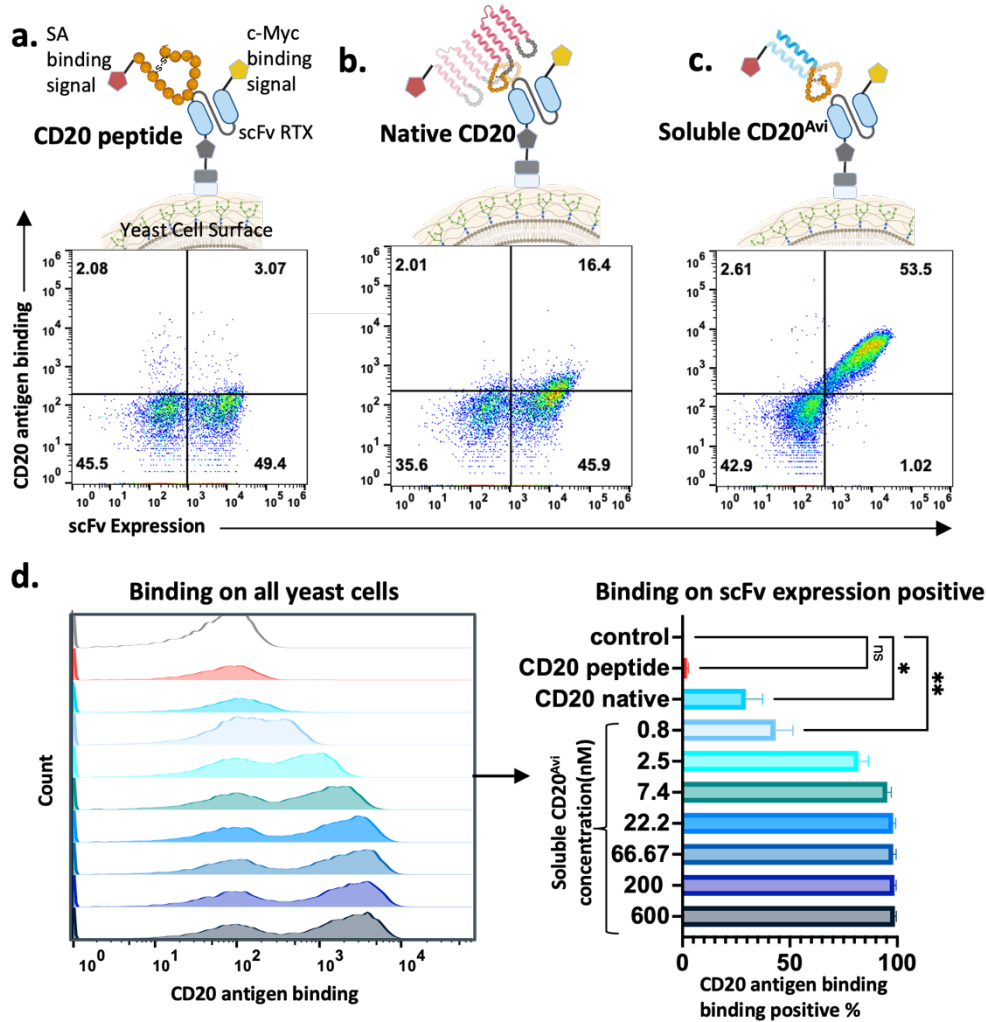


Figure 5. Soluble CD20 as a high-affinity antigen in yeast surface display. a,b,c) Yeast cells expressing a single-chain variable fragment (scFv) antibody derived from rituximab were tagged with a myc-tag to indicate expression levels. CD20 antigen forms — biotinylated CD20 peptide, native CD20, and soluble CD20 — were loaded onto yeast at concentrations of 10 μ M, 0.6 μ M, and 22.2nM, respectively. Binding was detected using fluorescently tagged streptavidin (SA). d) Concentration-dependent binding of soluble CD20 on yeast surface display. Binding-positive ratios represent the fraction of yeast cells with a positive SA binding signal among those positive for scFv expression. Control yeast cells, incubated with fluorescent SA alone, served to establish background fluorescence levels.