Computational and Structural Biotechnology Journal 20 (2022) 4952–4968

COMPUTATIONAL ANDSTRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Antibody interfaces revealed through structural mining

Yizhou Yin [a,1], Matthew G. Romei [b,1], Kannan Sankar [b,1], Lipika R. Pal [a], Kam Hon Hoi [b], Yanli Yang [b], Brandon Leonard [b], Gladys De Leon Boenig [c], Nikit Kumar [c], Marissa Matsumoto [c], Jian Payandeh [c], Seth F. Harris [c], John Moult [a], Greg A. Lazar [b,*]

[a] Institute for Bioscience and Biotechnology Research and Department of Cell Biology and Molecular Genetics, University of Maryland College Park, College Park, MD, USA
[b] Department of Antibody Engineering, Genentech Inc., South San Francisco, CA, USA
[c] Department of Structural Biology, Genentech Inc., South San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 6 July 2022
Received in revised form 19 August 2022
Accepted 20 August 2022
Available online 31 August 2022

Keywords:
Antibody
Cluster
Crystallography
Homotypic
Interface
Oligomer
Protein data bank
Structure

ABSTRACT

Antibodies are fundamental effectors of humoral immunity, and have become a highly successful class of therapeutics. There is increasing evidence that antibodies utilize transient homotypic interactions to enhance function, and elucidation of such interactions can provide insights into their biology and new opportunities for their optimization as drugs. Yet the transitory nature of weak interactions makes them difficult to investigate. Capitalizing on their rich structural data and high conservation, we have characterized all the ways that antibody fragment antigen-binding (Fab) regions interact crystallographically. This approach led to the discovery of previously unrealized interfaces between antibodies. While diverse interactions exist, β-sheet dimers and variable-constant elbow dimers are recurrent motifs. Disulfide engineering enabled interactions to be trapped and investigated structurally and functionally, providing experimental validation of the interfaces and illustrating their potential for optimization. This work provides first insight into previously undiscovered oligomeric interactions between antibodies, and enables new opportunities for their biotherapeutic optimization.

1. Introduction

Avidity-driven amplification of weak transient protein-protein interactions is a common theme in immunological processes. In some instances, weak interactions are clustered at cell-to-cell synapses, e.g. between T cells and antigen presenting cells or target cells. In other cases, protein-level immune complexation can promote naturally weak monovalent affinities to stronger avidity-driven binding events, such as occurs during B cell receptor (BCR) selection and antibody responses [1]. While the antibody variable region (Fv) commonly binds target antigen with high affinity, monomeric interaction of the fragment crystallizable (Fc) region with effector receptors typically occurs in the μM range where 1:1 binding events are generally inconsequential. Yet within immune complexes, avidity amplifies these interactions into triggers for positive or negative cellular response. A recent illustration of biological selection for avidity-driven triggers is the discovery of a transient homomeric interface in the immunoglobulin G (IgG) Fc region that mediates antibody hexamerization [2]. Nature's ostensible purpose for this interface is the amplification of antibody-mediated complement pathways, which are initiated by interaction of the IgG Fc with pentameric complement protein C1q.

Apart from understanding the contribution of weak antibody interfaces to immunology, their further value is their utility for optimization. Monoclonal antibodies are the most successful class of biotherapeutics, delivering enormous impact for the treatment of cancer, autoimmunity, and other diseases [3]. As a consequence, antibodies have become one of the most highly engineered protein families. Beyond Fv affinity maturation for target binding, native IgG interfaces between immunoglobulin (Ig) domains and between antibody and cognate Fc receptors have been successfully optimized for enhanced activities. Examples include engineering of the IgG CH3 domain for heterodimerization [4] that enables bispecific antibody platforms [5], enhancement of Fc gamma receptor (FcγR)- and complement-mediated effector functions [6–17], and half-life extension through binding optimization to the neonatal Fc receptor FcRn [18–20]. These types of subtle modifications to natural and sometimes weak antibody interfaces have met high success in drug development, with many of these platforms in clinically approved drugs [21–32]. A corollary is that the discovery of

* Corresponding author.
E-mail address: lazarg@gene.com (G.A. Lazar).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.csbj.2022.08.048

new antibody interfaces creates new opportunities for biotherapeutic optimization.

An intriguing aspect of the Fc hexamer discovery is that initial insights were derived from crystal contacts within an IgG structure that had been deposited in the protein data bank (PDB) many years prior [2,33]. Distinguishing true biological assemblies from artifactual crystal contacts is a long-studied problem in structural biology [34,35]. Additional functionally relevant interfaces have been identified from crystal packing, including a Fab homomeric interface [36] and oligomer interfaces in plastocyanin [37]. These examples illustrate that crystal contacts can provide insights into natural biological interfaces, albeit in an ad-hoc manner.

Antibodies offer a unique opportunity to investigate interfaces in crystallographic data in an exhaustive manner. The high degree of structural homology of Ig domains together with the large number of antibody structures in the protein data bank (PDB) provide a rich dataset of interactions. In this study we have comprehensively mined the PDB to search for antibody interactions that may not have been previously realized. This work advances a novel structural informatic approach for the investigation of weak protein–protein interactions, and offers new insights into potential transient interactions in antibodies that may be relevant to immune biology and enable new capabilities for biotherapeutic optimization.

2. Results

2.1. Computational pipeline identifies common interactions from Fab-Fab crystal packing

To search for new antibody interfaces, we computationally analyzed the crystal contacts for all Fab structures in the PDB (Fig. 1a). We leveraged the structural antibody database SAbDab [38], with inclusion criteria of Fab completeness and structure solution by X-ray diffraction. At this stage no additional filter was applied, e.g. for sequence redundancy or resolution, with the rationale of inputting all Fab crystallographic data into the computational method. The structures included apo Fabs as well as those in complex with target antigen, and both were included in the calculations.

Structural analysis of Fabs is facilitated by their high degree of structural and sequence homology; a position in any given Fab is generally structurally equivalent to that same position in virtually all other Fab structures. To enable comparisons, variable heavy (VH) and light (VL) regions were renumbered according to the AHo numbering convention [39], which accounts for sequence length differences to accurately represent structural equivalence. Because the commonly used EU convention [40] for constant regions does not account for gaps across diverse Ig sequence space, we developed a custom constant region numbering convention (Materials and methods). While our computational methods used these more robust conventions, for literature consistency this manuscript labels variable and constant region positions according to Kabat and EU numbering, respectively.

A computational pipeline was built to analyze the crystal contacts for all Fab structures (Fig. 1a). Crystallographic symmetry operators were applied to replicate the asymmetric units and obtain the complete set of non-redundant interfaces within the crystallographic lattice environment. 11,982 total interfaces were identified across 1,456 Fab PDB entries from X-ray diffraction (thus one structure can contribute more than one interface). All inter-Fab residue pairs constituting the interfaces were identified using a 4 or 6 Å distance cutoff. An inter-Fab contact map (Supplementary Fig. S1) revealed patterns both on and off the diagonal, representing homodimeric and heterodimeric interactions, respectively.

2.2. Antibody Fabs oligomerize through common interfaces

Cluster analysis was performed on the 11,982 interfaces to assess recurrence (Fig. 1a). Similarity between all interfaces was calculated in a pairwise manner as the Jaccard index between the sets of interface positions within each interface, followed by hierarchical clustering of similarity values (Materials and methods). From the analysis it was evident that the majority of interface clusters were singletons, i.e. they occur only once in the analyzed set of PDBs. The focus of this study was on interfaces that occur frequently across multiple PDB structures.

To clarify, we distinguish between our computational method and our analysis of the results from that method. Our computational method assembled all available crystallographic interfaces between Fabs and then clustered them. We then analyzed the resulting clusters for how commonly they occur across distinct PDB structures. The results revealed recurrent Fab packing interfaces. Total PDB cluster size (Supplementary Table S1) represents the number of distinct Fab PDBs from the set that contain a given interface at least once in the lattice expansion. However, the PDB set contains structures with identical Fab sequences, and thus the total size for each cluster includes redundant Fabs. Identical sequences (% identity (%ID) = 100) within each cluster were removed, leaving 981 nonredundant (NR) Fabs from the original set of 1,456. NR size divided by 981 nonredundant sequences is referred to as prevalence (Fig. 1b). The most recurrent interface cluster occurs 133 times in the 981 nonredundant Fab PDB structures for a prevalence of roughly 14 %. The next 5 most recurring interfaces occur with a prevalence of 5.7 %, 3.7 %, 3.4 %, 3.0 %, and 2.8 %, respectively (Fig. 1b, Supplementary Table S1). The Fab PDB comprises apo-Fab structures as well as Fab-antigen complexes. Fab-antigen contacts were not part of the calculations, and there was no apparent cluster bias due to the presence or absence of antigen (Fig. 1c). Variability across clusters in terms of percent antigen bound could be due to variance in the data, especially for less prevalent clusters, or in some cases interference of antigen binding by the oligomeric interaction. That stated, broadly speaking roughly half of antibody PDBs contain antigen and the data in Fig. 1c largely follow this global average. These results suggest not only that the observed interfaces are uninfluenced by antigen, but further that they are robust across structures of Fab complexes with diverse target antigens. They also demonstrate that in the majority of clusters the inter-Fab interactions are generally permissive of antigen binding. Cluster members crystallized in a diversity of space groups (Fig. 1c), suggesting no bias of interfaces by crystallographic lattice, and there was no apparent dependence on resolution (not shown).

The observed Fab dimers generally had structural and energetic features of weak transient interfaces, as analyzed using the 'Protein interfaces, surfaces and assemblies' service PISA [41]. With some exceptions, the generally lower buried surfaces of the discovered interfaces (562 ± 314 Å^2) (Fig. 1d) are more similar to "weak transient complexes" (718 ± 195 Å^2) than "obligate" or "permanent homodimers" (1,950 ± 986 Å^2) [42], caveated by the high standard deviations of these values. In contrast, the numbers of Fab interface H-bonds (8 ± 4) (Fig. 1e) are closer to obligate homodimers (10 ± 8) than transient complexes (4 ± 3) [42]. Again caveated by the high deviations, the more robust H-bonding relative to buried surface may be skewed in part by the prevalence of β-sheet dimers, discussed further below.

Each cluster is a distinct Fab-Fab interface. We established a naming convention based on the structurally central residue of each discovered interface (Materials and methods). For this convention, the name of each cluster adopts the format X-#, in which X represents the chain (VH, CH1, VL, CL) and # represents the position of the residue at the structural center of the interface,

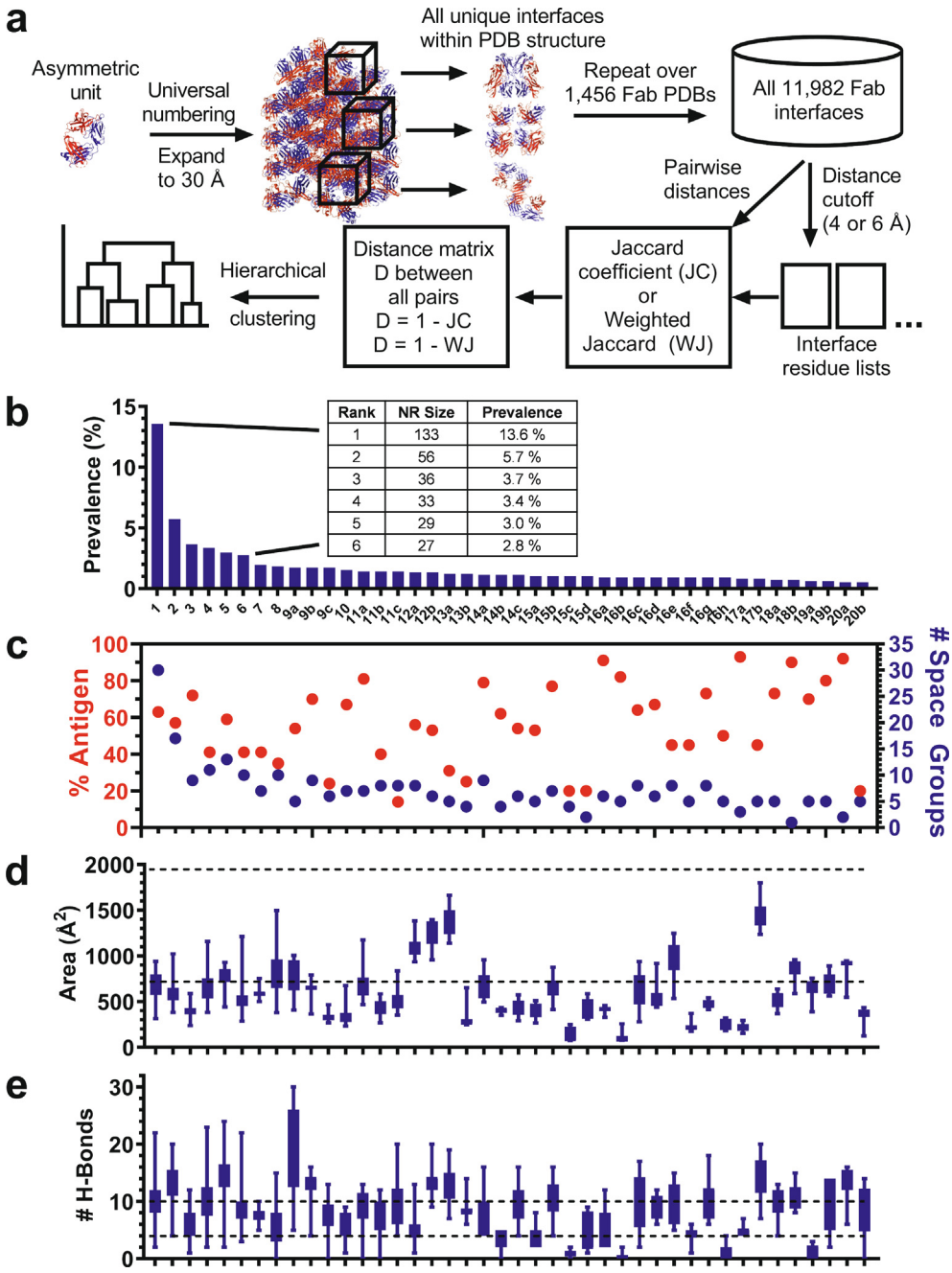


Fig. 1. Antibody Fab contacts and Fab-Fab interfaces revealed from informatic PDB analysis. (a) Computational flow diagram for analyzing inter-Fab interfaces. Fab X-ray structures were universally renumbered and the asymmetric unit was expanded to 30 Å^2. Symmetric oligomers were extracted, and all interacting residue pairs were identified using a distance cutoff of either 4 or 6 Å between any two inter-Fab residues. Similarity between a pair of interfaces was calculated as the Jaccard index or weighted Jaccard index between the sets of interface residue positions within each interface. Hierarchical clustering was performed based on the similarities calculated to result in a dendrogram of Fab interfaces. (b) Recurrent packing Fab-Fab interfaces throughout the collective Fab PDB in order of decreasing prevalence. Prevalence reflects an incidence measure of each interface that is unbiased by the presence of multiple structures of the same Fab sequence in the PDB. The inset provides the nonredundant (NR) size and prevalence values for the 6-most prevalent clusters. (c) Percentage of PDBs within each of the 42 most prevalent clusters that include antigen in the structure (red, left axis), and number of distinct space groups among the PDBs within each cluster (blue, right axis). The plot shows that the lack of cluster bias due the presence of antigen or crystallographic lattice. For example, the six most prevalent interfaces were observed in structures where Fab/antigen complexes make up 63 %, 57 %, 72 %, 41 %, 59 %, and 41 % of the cluster respectively, and that crystallized in 30, 17, 9, 11, 13, and 10 different space groups respectively. (d) Mean buried surface area and (e) number of H-bonds for the 42 most prevalent interfaces, in order of decreasing prevalence. Error bars represent standard deviations. Upper and lower dashed lines correspond to averages for obligate and transient protein complexes respectively [42]. Data in (b-e) are vertically aligned and thus there is correspondence with labeled columns in (b). Numeric values for prevalence and PISA results are provided in Supplementary Table S1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

numbered by Kabat for VH and VL or EU for CH1 and CL. In addition, we designated a structural representative for each cluster, defined as the member with the highest weighted Jaccard similarity index to all other members in the cluster (Materials and Methods). Fig. 2 presents the representative Fab dimer for each cluster with a prevalence of $\geq$ 0.5 %. The interfaces, 42 total, are ranked numerically based on prevalence, with equivalent prevalence designated with arbitrary alphabetic qualifiers. The labels within Fig. 2 include the cluster rank, interface name, prevalence, and buried surface area (described below). Interface positions for all of the

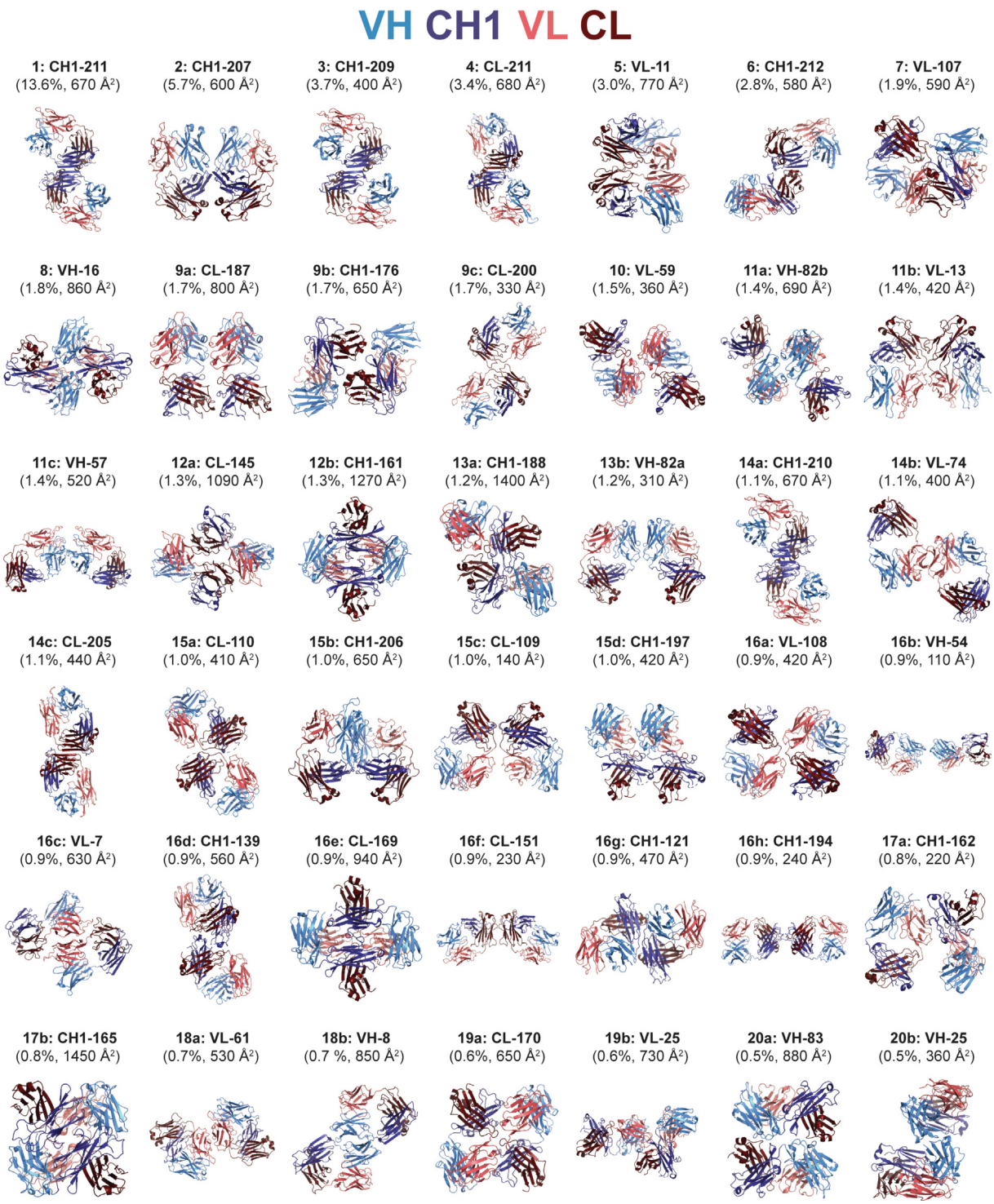


Fig. 2. Representative structures of the 42 most prevalent Fab-Fab interfaces. Domains are colored as follows: VH (light blue), CH1 (dark blue), VL (pink), CL (dark red). The labels for each interface include the cluster rank based on prevalence, interface name based on structurally central residue (Kabat numbering for Fv region, EU number for constant regions), nonredundant prevalence percentage, and mean buried surface area. Interfaces are ranked numerically based on prevalence, with equivalent prevalence being designated with arbitrary alphabetic qualifiers; for example, there are three interfaces that occur in 17 PDB entries for prevalence = 1.7 %, ranked as 9a, 9b, and 9c. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

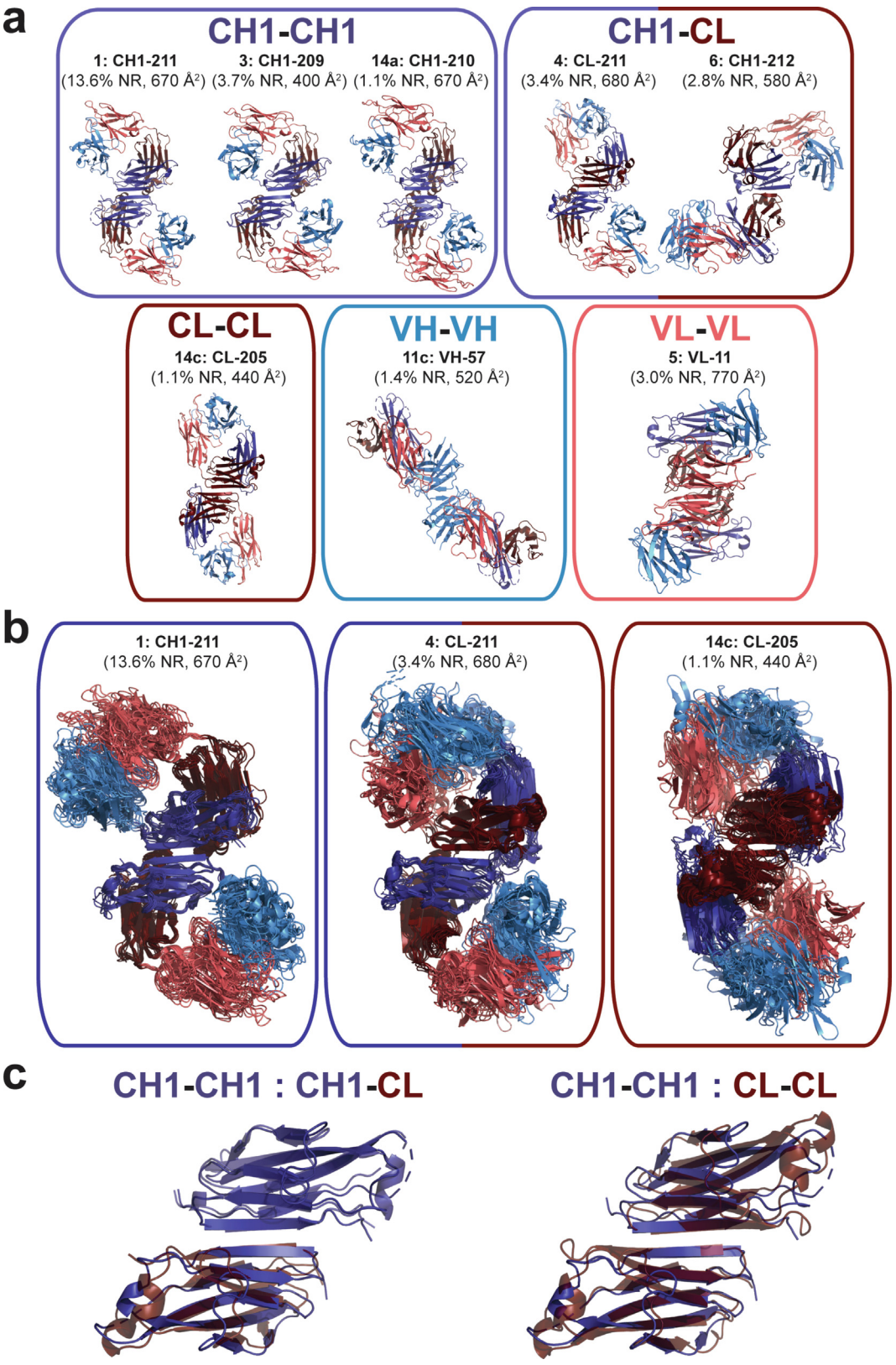
top 42 clusters are provided in Supplementary Table S2. Pymol files for all 42 interfaces are provided in the Supplementary Materials.

2.3. β-sheet dimers and elbow dimers are common Fab interface motifs

The most prevalent interface, labeled CH1-211, is a sheet-extended homodimer mediated through the heavy chain (HC) constant region centered on (EU) position 211 (Fig. 2). This interface is observed in 133 (13.6 %) of the nonredundant Fab PDB structures (Fig. 1b). The extended strand pairing is mediated centrally by the C-terminal strand G (positions 207–214) that effectively doubles the size of the β-sheet. In addition, the interface is complemented by packing of the short helical turn of CL between strands A and B (122–126) against the *N*-terminus of CH1 prior

to strand A (119–122). Altogether this interface buries an average 670 Å^2 of surface calculated across the members of the cluster.

Strikingly, β-sheet dimers were found to be a common interaction motif, observed in a total of eight clusters (Fig. 3a). Two additional CH1 homodimer interactions are observed as interfaces CH1-209 and CH1-210, which correspond to the 3rd and 14th (14a) most prevalent clusters (3.7 % and 1.1 %, respectively). These dimers are highly similar to CH1-211, containing the same sheet-extended HC homodimer yet with either a 2-residue (CH1-209) or 1-residue (CH1-210) register shift. While interface CH1-210



contains similar light chain (LC) packing arrangement, the greater register shift of CH1-209 separates the dimer units such as to eliminate the LC contacts, resulting in less buried surface (400 Å^2 averaged over the cluster) (Fig. 3a). In addition to CH1 homodimers, sheet-extension interfaces make up two CH1/CL heterodimers (CL-211 and CH1-212), a CL homodimer (CL-205), a VH homodimer (VH-57), and a VL homodimer (VL-11). Altogether, β-sheet dimers make up five of the six most prevalent interfaces, and appear in over 30 % of nonredundant PDB entries. The structural similarity of the interfaces of the members within interface clusters CH1-211, CL-211, and CL-205 are high (Fig. 3b), further illustrating the common structural interaction motif. A high degree of structural homology was observed upon superposition of the structural representatives of the CH1-211 homodimer with the CL-211 heterodimer and CL-205 homodimer (Fig. 3c).

As discussed above and supported by the percentage of cluster members with bound antigen (Fig. 1c), the Fab orientations in most dimers are generally permissive of antigen binding. All of the β-sheet dimers mediated through constant domains position the Fab monomers in an antiparallel orientation, with opposed Fv's such that the two antigen binding regions of dimer face away from each other. In the context of a full-length antibody, using the most *N*-terminal hinge disulfide (EU position Cys226 in human IgG1) as the tether point and a random coil residue distance of 3.5 Å, the C-termini of the two Fab arms are constrained to a maximal distance of ~ 35 Å relative to each other. The calculated distance between the HC C-termini of the two Fabs in all of the β-sheet dimers exceed this value (CH1-211: 41 Å, CH1-209: 47 Å, CH1-210: 37 Å, CL-211: 47 Å, CH1-212: 37 Å, CL-205: 43 Å). This analysis suggests that the orientation of the two Fabs precludes intra-IgG binding, forcing inter-IgG interactions that would ostensibly promote higher-order complexation.

Beyond β-sheet dimers, an additional recurring motif observed are interfaces mediated by the elbow regions between variable and constant domains (Fig. 2, Supplementary Fig. S2). These interactions involve the elbow regions between either two HCs (CH1-207 and CH1-121) or two LCs (CL-200, CL-110, VL-108, CL-109, CL-169, and CL-170). CH1-207 is the second most commonly observed interface, and in total, the eight elbow region interfaces comprise 12.7 % of nonredundant PDB entries. The proximity of both the elbow and the β-turns/loops varies among these eight interfaces. In the case of CH1-207, there are roughly balanced contributions from the variable and constant region loops and turns (Supplementary Table S2). While elbow-elbow interaction was the general theme among these clusters, as a collection there was greater overall structural dissimilarity relative to the group of β-sheet dimers (Supplementary Fig. S2).

2.4. Intra-cluster members show similarities at interface residues

To investigate sequence dependence, interface profiles were generated for the six most prevalent clusters (Fig. 4), as well as the three additional clusters that form β-sheet dimers (VH-57, CH1-210, and CL-205) (Supplementary Fig. S3). The profiles describe the compositional distribution of each cluster based on species (human, mouse, other), light chain (LC) type (kappa or lambda), as well as human and mouse VH and VL subgroups. In addition, the profiles provide intra- versus inter- cluster sequence identity for both the entire Fv as well as only those residues at each corresponding interface. For the Fv, these values reflect the mean pairwise identities for all VH and VL sequences within a cluster (intra) versus across the clusters (inter). A similar comparison is made for the interface, where only those residues that participate in the interface for the designated cluster are compared. Finally, sequence logos provide weighted sequence composition for cluster members at interface residues.

CH1-211 showed no apparent species dependence, being represented by sequences from diverse species, LC type, and VH and VL subgroups (Fig. 4). The diverse subgroup representation is also captured in the similar Fv intra- vs inter- cluster identities (~50 %). Similar overall results were observed for the other top six clusters, with most demonstrating diverse representation of species, LC, and Fv subgroups (Fig. 4). Notable biases are the Fv subgroup trends of CH1-207 and to lesser extent VL-11. These are the two clusters among the top six where the Fv contributes substantially to the interface, specifically the VH for the former and VL for the latter (Fig. 2). These clusters are comprised principally of human subgroup IGHV3 for CH1-207 and human subgroup IGKV1 for VL-11. The consensus interface residues based on their sequence logos (Fig. 4) are well-represented in these germlines.

The interface residues of CH1-211 are roughly as similar among members of the cluster (intra identity 68 %) as they are to those same residues among members of all other clusters (inter identity 64 %) (Fig. 4). In contrast, three of the top clusters, CH1-209, CL-211, and CH1-212, display high intra-cluster interface identity (94 %, 94 %, and 96 %, respectively) relative to lower inter-cluster identities (70 %, 58 %, and 72 %, respectively) (Fig. 4). CL-211 is notable also for its bias towards lambda LCs (Fig. 4). While kappa LCs are overrepresented in the PDB set overall (roughly 85 % are kappa, 15 % are lambda), nearly all of the members of CL-211 are lambda. Deeper investigation revealed that kappa LCs contain a proline at position 204 that terminates the *N*-terminus of the last β-strand. In contrast, lambda LCs contain a Thr at 204 that, together with a shorter *N*-terminal loop, enable an extended β-strand that permits β-sheet heterodimerization with the HC. Consistent with this observation, the CL-CL homodimer interface CL-205 also showed a bias for lambda (Supplementary Fig. S3). Overall, the diverse nature of the sequences at a global level yet trends for some clusters at an interface level, suggest motifs that may provide further insight into the interactions as well as serve as guides for design.

2.5. Disulfide engineering enabled Fab dimers to be trapped

While some of the observed interfaces may be artifacts of crystallization, we experimentally investigated the prospect that some

Fig. 3. β-sheet dimers are commonly observed Fab oligomers. (a) Structural similarity between sheet-extended Fab oligomers that interact as CH1-CH1 homodimers (top left), CH1-CL heterodimers (top right), CL-CL homodimers (bottom left), VH-VH homodimers (bottom middle), and VL-VL homodimers (bottom right). Domains are colored as follows: VH (light blue), CH1 (dark blue), VL (pink), CL (dark red). The labels for each interface include the cluster rank based on prevalence, interface name based on structurally central residue, nonredundant prevalence percentage, and mean buried surface area. (b) Fab β-sheet dimers are structurally similar across cluster members. Interface residues of 10 cluster members were superimposed for Fab oligomers within the CH1-211 homodimer (left), CL-211 heterodimer (middle), and CL-205 homodimer (right). RMSD values were calculated that reflect alignment of Cα atoms within stacked CH1 and/or CL domains for the 10 representative cluster members depicted. For example, for CH1-211 both CH1 domains of one structure were together aligned to both CH1 domains of another structure. Calculations were run in Pymol without an outlier cutoff. The reported RMSD is an average of the matrix produced from aligning 10 structures to each other (CH1-211: CH1-CH1, 1.6 Å; CL-211: CH1-CL, 0.9 Å; CL-205: CL-CL, 2.3 Å). (c) Constant domain superposition of CH1-CH1 homodimer with CH1-CL heterodimer (left) and CL-CL homodimer (right). β-sheet residues at the Fab-Fab dimer interface were superimposed for the structural representatives of clusters CH1-211 and CL-211 (left), or CH1-211 and CL-205 (right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

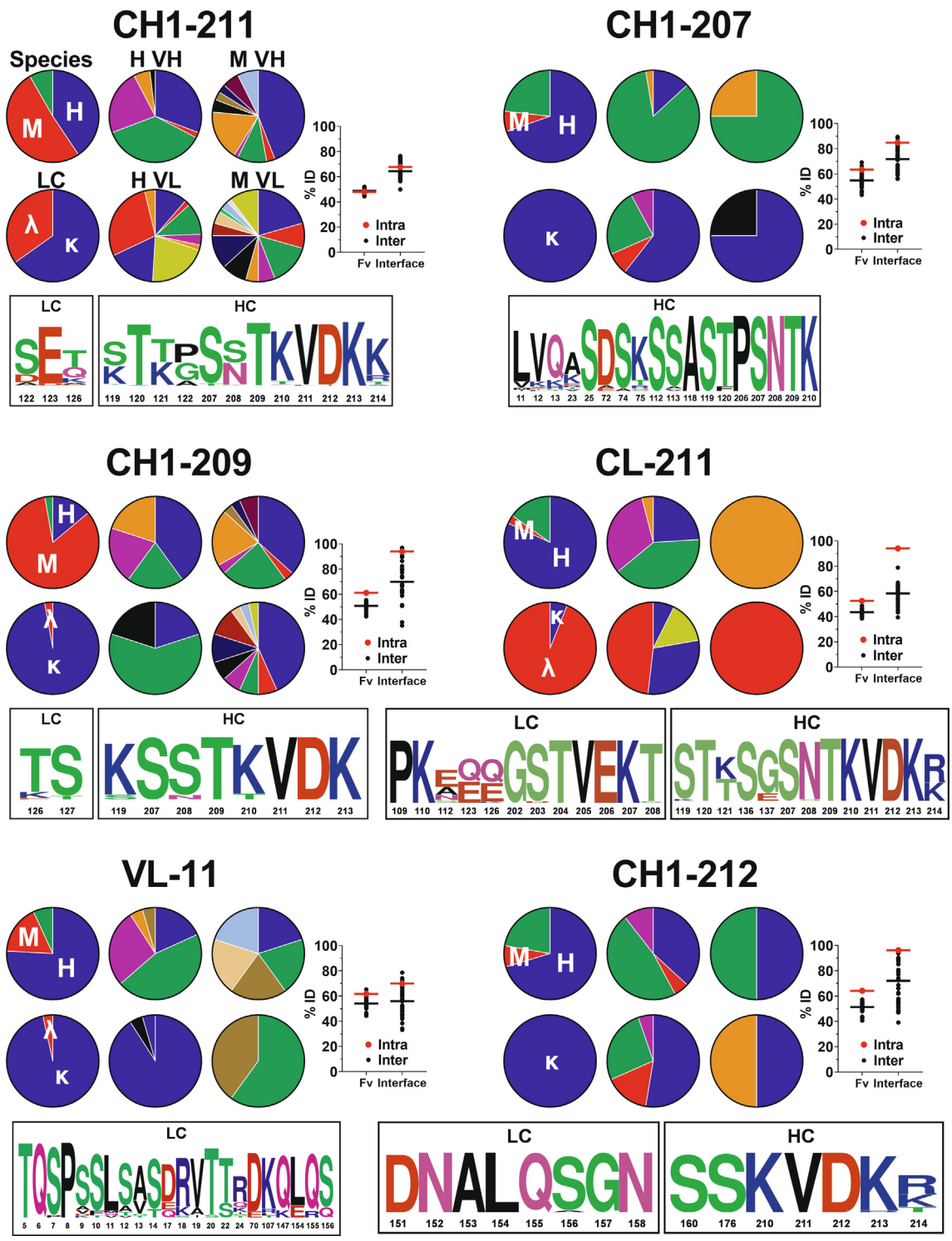


Fig. 4. Interface profiles of the six most prevalent clusters. For all interfaces, the upper left-most pie chart provides the sequence distribution of cluster members based on species: human (blue, H), mouse (red, M), and other (e.g., rat, rabbit, etc., green). The lower left-most pie chart provides the sequence distribution of cluster members based on light chain type: kappa (blue, κ), and lambda (red, λ). The four right-most pie charts provide the sequence distribution of human VH (H VH), human VL (H VL), mouse VH (M VH), and mouse VL (M VL) subgroups. The labels above the pie charts in the upper left profile (CH1-211) are not repeated in the other charts for visual simplicity. The %ID plot on the right provides the intra (red) versus inter (black) cluster sequence identity for both the entire variable region (Fv) as well as only those residues at the interface as shown in the sequence logo and Supplementary Table S2. For the Fv, these values reflect the mean pairwise identities for all VH and VL sequences within the cluster (intra) versus the mean pairwise identities between each member of the cluster and all other members of all other clusters (inter). A similar comparison is made for each interface, where %ID reflects the mean pairwise identity of the residues at the interface of a given cluster aligned with those same residues for each member of the cluster (intra) versus all other members of all other clusters (inter). The sequence logo at the bottom provides weighted sequence composition at interface residues for members of the designated cluster, with numbering according to Kabat and EU conventions for the Fv and constant regions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

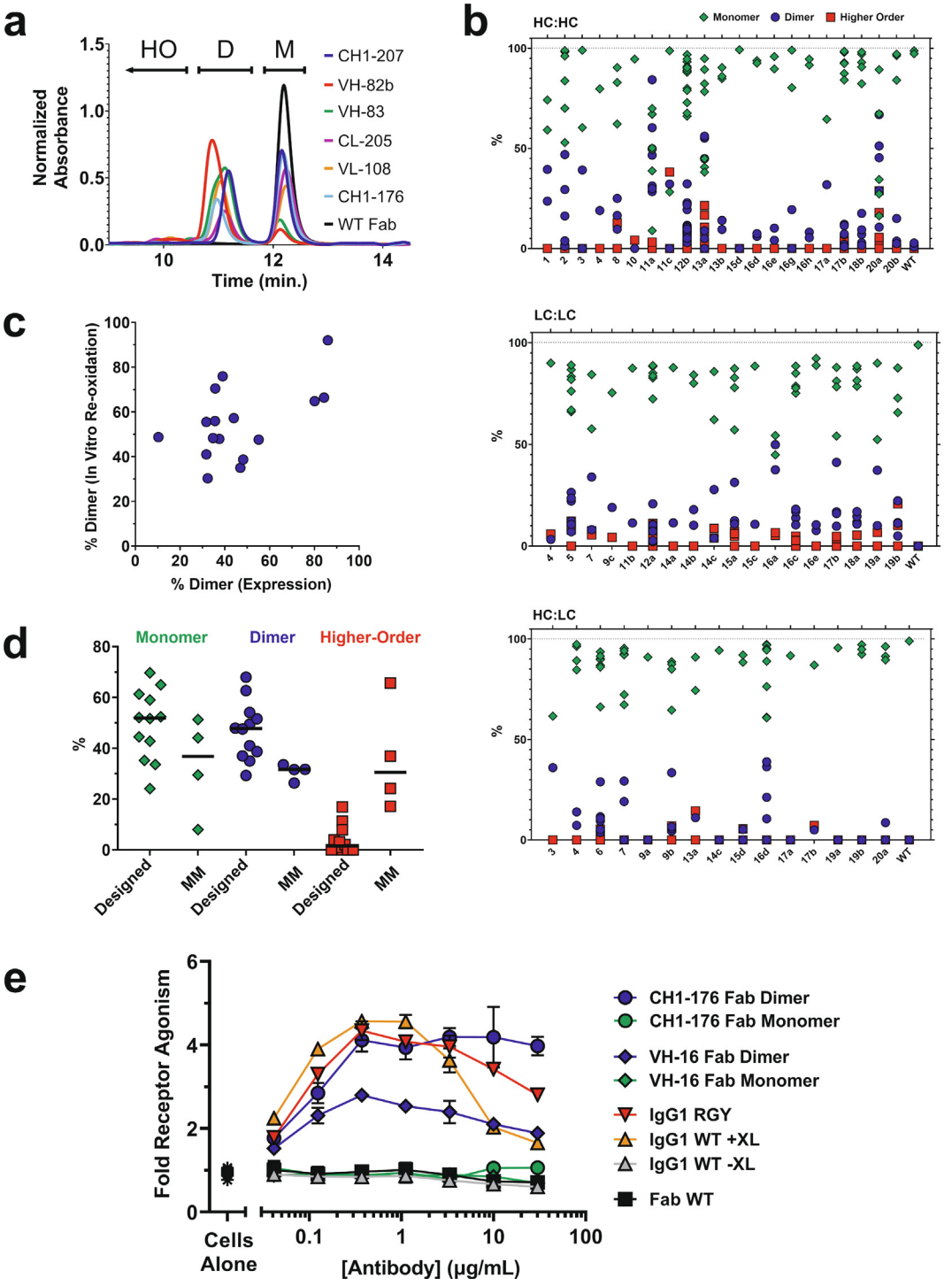


Fig. 5. Fab dimers can be trapped with engineered disulfides. (a) Representative *SEC* chromatograms of select cysteine variant and WT anti-Her2 Fabs after affinity chromatography. Peaks are labeled as monomer (M), dimer (D), and higher-order (HO). (b) Summary of monomer (green diamonds), dimer (blue circles), and higher order (red squares) species for all cysteine variants by *SEC* post-affinity chromatography. Data represent percentage of species based on integration of *SEC* peaks. Multiple symbols for a given cluster represent multiple designed cysteine variants. The identities of all variants and associated numeric *SEC* data are provided in Supplementary Tables S3-S5. (c) Correlation between expression and in vitro coupling data for disulfide-trapped dimers. (d) Summary of data from in vitro coupling of selected cysteine variant Fabs. The data reflect the % monomer (green diamonds), dimer (blue circles), and higher-order (red squares) species after assembly for the designed and mismatched (MM) variants run as negative controls. For MM variants the pairings of HCs and LCs with single cysteine mutations were scrambled in four separate combinations. The identities of all variants and associated numeric data are provided in Supplementary Table S6. (e) OX40 receptor activation by variant and WT Fab and full-length IgG antibodies based on an NFκB luciferase reporter assay. VH-16 represents Fab cysteine variant VH(S113C) / CH1(G178C) and CH1-176 represents Fab cysteine variant VH(P14C) / CL(D151C). Fold receptor activation represents normalized RLU relative to cells alone. For the IgG1 versions, +XL and -XL indicate the presence or absence of secondary crosslinker antibody, respectively. RGY represents a triple Fc variant that promotes IgG hexamer formation. No crosslinker was used for the RGY version or the Fab monomer or dimer versions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

may reflect natural albeit transient Fab oligomeric interactions in solution. Given their weak nature, we explored whether dimers could be trapped via engineered disulfides. Starting from the representative structures of each of the 42 most prevalent clusters (Fig. 2) we performed in silico cysteine scanning analysis to explore all possible residue pairs at each interface for distance and geometry that may be conducive to disulfide bond formation (Materials and methods). A total of 179 residue pairs were identified, spanning 77 HC-HC, 54 LC-LC, and 48 HC-LC pairs. While the range of designed variants per cluster was as high as 14, the median was 2 variants per cluster. Cysteine variants were engineered into the Fab region of the anti-human epidermal growth factor receptor 2 (Her2) antibody trastuzumab [43]. Variant and wildtype (WT) Fab antibodies were expressed in human embryonic kidney

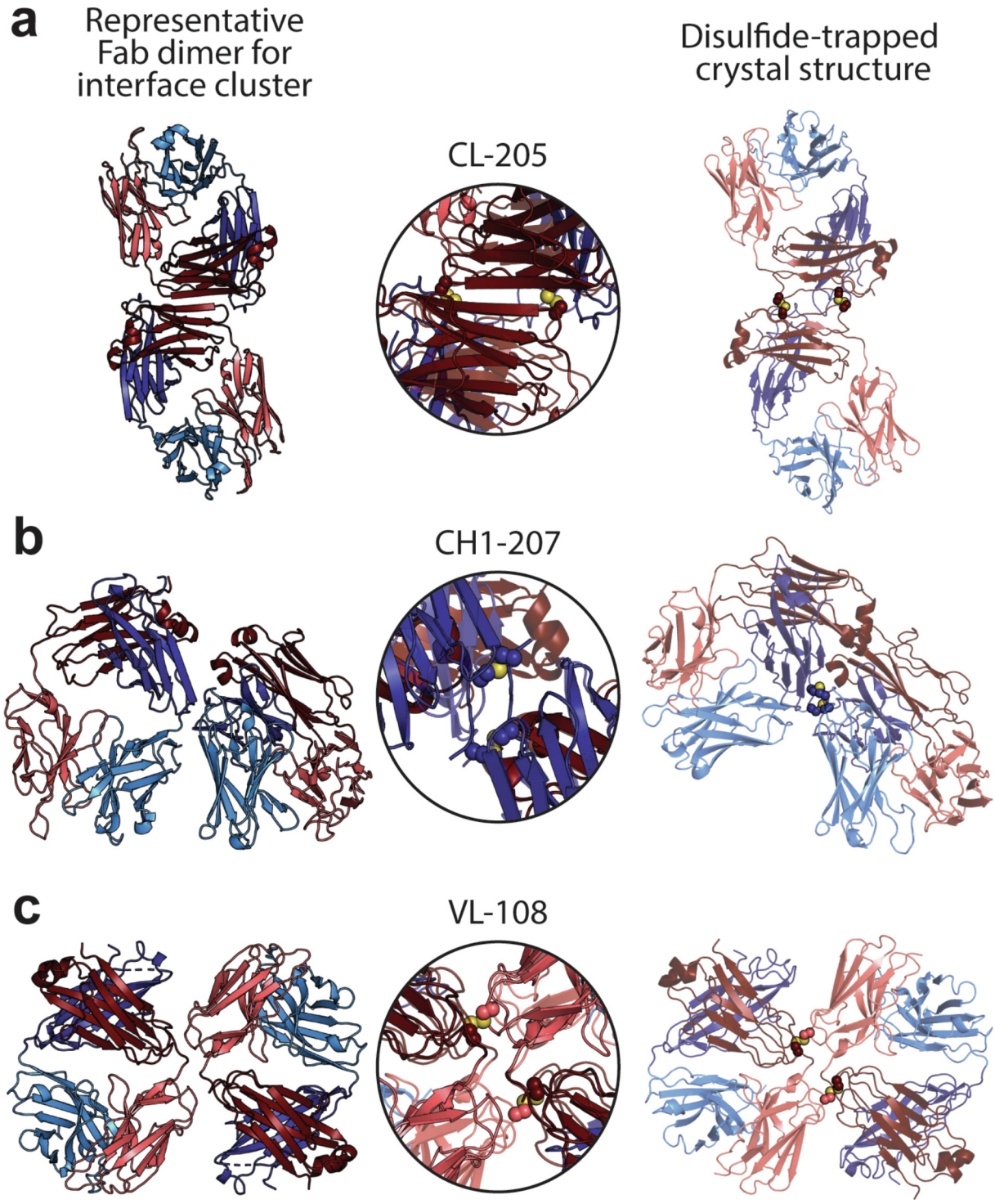


Fig. 6. Disulfide-trapped Fab dimers affirm conformations of PDB mining. All Fabs comprise the variable region of the anti-Her2 antibody trastuzumab. Fab structures represent the β-sheet-extended interface cluster CL-205 [variant CL(S202C) / CL(S208C)] (a) and the elbow region clusters CH1-207 [variant CH1(S119C) / CH1(G122C)] (b) and VL-108 [variant VL(G16C) / CL(D170C)] (c). Fab domain colors match those in Fig. 2, with the informatically-mined structure at 0% transparency (left) and the experimentally-determined structure at 30% transparency (right). Sulfur atoms participating in disulfide bonds are shown in yellow. Overlays of the interface (middle) for (a) and (c) were configured by aligning the α carbons of the Fab dimers from each structure, whereas the overlay for (b) used a single Fab for alignment (see Results). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(HEK) 293 cells, purified by affinity resin, and analyzed by size exclusion chromatography (*SEC*) (Fig. 5a). While peaks corresponding to monomeric Fab species were generally similar in elution, assigned dimer species showed greater variation in retention time, possibly reflecting differences in Fab dimer geometries (Fig. 2) and as a consequence hydrodynamic radii. Analysis of the full set of analytical data revealed that many variants showed discrete dimeric species, in addition to monomeric Fab (Fig. 5b, Supplementary Tables S3-S5).

The presence of discrete dimer species in many of the designed variants supports the computational analysis. Yet many of the cysteine variants were > 95 % monomeric, which is a common expression profile for Fabs without engineered disulfides, as evidenced by WT (Fig. 5a,b). These monomeric constructs represent internal controls, suggesting that transient non-random contact in solution as well as proper geometry are required to form the stable oligomeric species. There were generally low amounts of higher order species across the cysteine variant set (Fig. 5b), which may be a consequence of the cellular quality control that exists for antibodies [44], as well as possible differences in affinity purification. Overall, while the expression and purification results provided a good screen for dimer formation, the multi-factorial nature of the process made it ill-suited to investigate the variants more deeply. In order to study the solution properties of the variants in a more controlled experiment, a subset of variants with the most prominent dimer formation were selected to explore whether dimers could be assembled in vitro. The monomeric Fab species of this variant subset were reduced in solution followed by gradual re-oxidation with dehydroascorbic acid. Upon re-oxidation, discrete dimer species were formed, similar to and correlated with the results from in vivo expression (Fig. 5c, Supplementary Table S6). As a control, the pairings of HCs and LCs with single cysteine mutations were scrambled in four separate combinations, referred to a mismatched (MM) variants (Supplementary Table S6). Upon refolding, all mismatched combinations displayed substantial higher order oligomers that indicate nonspecific disulfide formation (Fig. 5d, Supplementary Table S6). These results suggest that for the designs, ordered contact influences Fabs toward stable dimer species in solution, further supporting the presence of specific albeit transient interactions.

2.6. Fab dimers potentiate antibody functional activity

A representative set of 26 cys-engineered variants were selected to explore functional application. Cysteines for each variant were introduced into the Fab of an antibody referred to as 3C8 that is an agonist of the receptor OX40 (CD134), which we have previously shown to provide a sensitive system for detecting antibody-mediated receptor clustering [45]. Due to low expression yields or variable monomer/dimer ratios, many of the variants were not characterized further. A subset of 5 variants that expressed well and had a favorable monomer/dimer profile were chosen for further purification and separation of species. Final monomer and dimer samples for each variant had high purity and were stable in solution (i.e, they did not interconvert over time based on analytical *SEC*). Affinity measurements by Biacore indicated that a subset of dimeric versions did not retain binding (data not shown), possibly due to steric clash of antigen binding in the context of the coupled Fab dimer. However, two variants, VH (S113C)/CH1(G178C) at the VH-16 interface and VH(P14C)/CL (D151C) at the CH1-176 interface, maintained their affinities for OX40 (K_D's for variants and WT $\sim$ 10 nM, Supplementary Fig. S4).

In a nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) luciferase reporter assay utilizing OX40+ Jurkat cells, the 3C8 antibody has no activity as a bivalent IgG on its own, but is a strong agonist of OX40 signaling when extrinsically crosslinked with a secondary antibody (Fig. 5e). An Fc-engineered triple variant E345R/E430G/S440Y (RGY) of this IgG that promotes hexamerization [2] agonizes receptor without reliance on extrinsic crosslinker. Strikingly, Fab dimer versions of the two cysteine-linked variants were capable of agonizing OX40 in the absence of crosslinking, in contrast to inactive monomer versions of the same variants (Fig. 5e). The ability of the cys-linked Fab dimers to activate receptor signaling despite their bivalency may be due to their ability to engage receptor with a particular geometry and/or orientation. Modeling based on the previously solved structure of 3C8 Fab bound to OX40 [45] did not provide insight into why these two particular architectures enable intrinsic agonist activity. While the mechanism requires further study, these results illustrate how new interfaces may be used to engineer novel geometries and orientations into antibodies in order to enable activities of therapeutic relevance.

2.7. Disulfide-trapped Fab dimers are structurally similar to informatically-mined conformations

To validate proper disulfide trapping of the interfaces and further support the results, we solved the X-ray structures at < 3 Å resolution for a representative subset of disulfide-trapped Fab dimers containing the trastuzumab Fv. Experimental structures included one representative of the β-sheet dimer class CL-205 (Fig. 6a), and two representatives of the elbow dimer class CH-207 (Fig. 6b) and VL-108 (Fig. 6c). Electron density maps were consistent with the presence of engineered disulfide bonds at the expected sequence positions (Supplementary Fig. S5). Experimentally-determined and informatically-mined structures were topologically similar, with good superposition at the interface sites (Fig. 6). The most notable alignment discrepancy across the three can be seen for CH1-207. In this structure, the disulfide bond pulls the HC elbow loops closer together on one end. While this perturbation rotates the left Fab with respect to the right Fab, the interface region remains largely intact, suggesting that this discrepancy may be due to disulfide bond constraints rather than the formation of an altered solution-phase configuration. Overall, these results provide strong evidence that the observed solution-phase dimers homogeneously resemble the discovered dimer conformation rather than a mixture of interface-independent conformations from non-specific disulfide pairing.

3. Discussion

Weak transient protein–protein interactions play fundamental roles in diverse biological processes [46–48]. While best characterized in the context of intracellular signaling cascades, transient interactions are also relevant extracellularly, including during immunological recognition where they can convert the avidity of a cell synapse or antibody complex into a cooperative trigger for cellular activation or inhibition. While obligate interactions are readily investigated with direct biochemical and structural methods such as crystallography and electron microscopy, transient interactions are low in abundance and harder to identify, requiring sensitive and sometimes indirect techniques such as the yeast two-hybrid system, fluorescence resonance energy transfer, nuclear magnetic resonance spectroscopy, and split protein complementation [46,49]. While in rare instances weak interactions have been gleaned from crystallographic structures, it can be difficult to distinguish between true biological interfaces and crystal packing artifacts [34,35].

We have explored a novel structural informatics approach to search for previously uncharacterized interfaces in antibodies. We selected antibodies for three reasons. First, the avid nature of

immune complexation provides a biological rationale for the existence of as yet undescribed interfaces that could tune immune response. Second, the virtual one-to-one residue equivalence of antibodies across a large structural data set enables mining of interaction patterns for commonality. Effectively, prevalence in the current work serves as a signal-to-noise parameter that may suggest biological relevance. Finally, we are interested in discovering new antibody interfaces for their potential in biotherapeutic engineering. Monoclonal antibodies are one of the most successful classes of drugs across a myriad of medical needs, with the 100th antibody drug recently approved [50]. While the most effective antibody drugs have historically been native IgGs, recent years have witnessed an acceleration in development and approval of engineered versions optimized for activity [3]. Rather than de novo sites, the most clinically successful enhancements are modest mutational modifications of natural and often weak interactions, either intra-IgG or between antibodies and cognate receptors. The logic flows that innovation of new capabilities in antibody therapeutics is best served by the discovery of new natural antibody interfaces.

Our characterization of the crystal packing arrangements of antibody Fab regions resulted in a diversity of interfaces, with 42 represented in 5 or more PDB entries of nonredundant sequence. The low frequency of most of these together with the generally weak structural and energetic features suggest that many may be crystal artifacts. Rigorously speaking, the current work does not conclusively establish biological relevance for any of the described interactions. With that caveat aside, support for biological relevance is increased by both high prevalence and the existence of shared motifs across multiple results, namely the recurrence of β-sheet dimers and interaction at Fab elbow regions. β-sheet dimers are by far the most commonly observed motif, making up five of the six most prevalent interfaces and present in 30 % of nonredundant PDB entries, yet with a diversity of oligomeric and regional architectures. In all cases, the anti-parallel orientations of the paired Fabs, together with their C-terminal distances that exceed hinge flexibility, would preclude intra-Ig Fab interaction, forcing inter-Ig interactions that would promote immune complexation. This notion is supported by the existence of naturally occurring examples where intermolecular β-sheets associate to form protein heterodimers, homodimers, and larger oligomers [51]. Homodimeric β-sheets are observed naturally, for example in ParB [52], transthyretin [53,54], and Ras-binding domain of c-Raf1 [55]. β-sheet homodimerization has also been successfully used as a template for de novo protein design [56], and is the commonly observed arrangement in the structures of macrocyclic β-sheet peptides [57]. The intrinsic preference of exposed β-strands to pair is illustrated most dramatically in the aggregation of amyloid fibrils [58], and it has been proposed that naturally occurring proteins use negative design to avoid edge-to-edge association [59]. Indeed a "generic hypothesis" advanced by Dobson and Karplus suggests that extended β-sheets in amyloid structure are an inherent characteristic of polypeptide chains rather than unique to a specific sequence or structure [60]. This model is consistent with our observation of intermolecular β-sheet dimers across varying domains and orientations of the antibody Fab. It is tempting to speculate that the Ig fold, in addition to providing loops for diversity and stable scaffolding to support that diversity, also offers antibodies the additional and previously unrealized benefit of avid transient self-association through inter-strand interactions.

On the contrary, while homotypic antibody interactions may enhance immune response, unintended self-associations can lead to steric crowding and could negatively affect antigen targeting and reduce activity. Another concern relates to therapeutic developability [61,62]. Self-interaction could contribute to solubility problems and/or promote higher viscosity. Previous studies have investigated aggregation prone regions in antibodies [63–65]. While beyond the scope of the current work, the potential impact of the described interactions, and their corresponding sequence-dependencies, on antibody solution properties and developability requires further study.

Our disulfide engineering experiments provided further support for the discovered interactions and enabled direct structural investigation. Antibody expression and purification are influenced by diverse factors, including for example cellular quality control and possible differences in purification. Consequently, we utilized the analytical *SEC* results from production more as a screen rather than a method to investigate the variants in detail. Variants that showed significant dimer by expression were advanced to an in vitro oxidation experiment, that together with gel band quantitation, was more controlled. Overall, the experimental results suggest that disulfide-trapped species are a result of predisposed contact in solution that, together with favorable geometry, promote coupling of specific homotypic dimers. To explore the functional potential of these interfaces we leveraged a therapeutically relevant antibody-receptor system that is sensitive to oligomeric interaction. The anti-OX40 results represent a first attempt at exploiting this work for optimization, illustrating that the discovered interfaces can serve as novel engineering sites to enhance biotherapeutic properties.

While further biological validation of the discovered interfaces is needed, our results suggest a possible previously unknown structural feature of antibodies, and one that would be well-suited for avidity-driven immune response. Transient β-sheet dimers could boost antigen affinity at the BCR level during clonal selection. Transient interfaces could also be relevant at the IgG level for the enhancement of antigen neutralization and Fc-mediated effector functions. The often repetitive and multivalent nature of pathogenic targets, as well as the multiclonal nature of antibody response, have provided evolutionary pressure for the multivalency of isotypes such as immunoglobulins M and A (IgM and IgA) [66,67] and IgG hexamerization [2]. Such selective pressure has also resulted in more nuanced valency tricks such as chain swap and Fab-dimerization observed in anti-human immunodeficiency virus (HIV) and -severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) antibodies [68–70], homotypic interactions involved in antibodies against plasmodium [71], as well as in the context of therapeutic antibodies [36,72,73]. While inter-IgG β-sheet interactions are weak, they could become relevant energetic drivers on the cell surface or in the context of a solution immune complex where the effective concentration of IgG may be high. In this manner, the environment of a cell surface or solution immune complex may be akin to a biomolecular condensate [74]. While condensates have typically been characterized in the context of intracellular biology, extracellular examples are known, for example contributing to protein assembly in the extracellular matrix and cell–cell adhesion [75,76]. From this perspective, the Fab crystal lattice may be a proxy for how antibodies behave in their condensed native biological environments. In this light the results here, derived from holistic analysis of antibody structural packing data, provide fresh mechanistic insight into how condensation of immune complexes, either at the cell surface or in solution, may enable amplification of immune interactions into immunological response.

4. Materials and methods

4.1. Dataset from SAbDab

All Fab structure coordinates and the corresponding *meta*-information were downloaded from the SAbDab database [38].

Species and germline information was cross-checked with information extracted from IMGT/3Dstructure-DB (https://www.imgt.org/3Dstructure-DB/) [77,78]. Fab PDB structures were then filtered to ensure first that each Fab structure was complete (e.g. both heavy and LCs were longer than 180 residues), and second that the structure was solved by X-ray diffraction with valid symmetry information (e.g "SMTRY" lines in "REMARK 290" section).

4.2. Computational pipeline for identification of antibody interfaces

An in-house computational pipeline was developed to identify symmetric oligomers and their corresponding interfaces. A universal numbering scheme was used for both variable and constant domains to effectively compare the sequence and structures of symmetric oligomers and interfaces across different Fab structures. Unlike a "bottom-up" strategy (e.g. PISA) where all molecular contacts in the crystal lattice are first extracted and later assembled to complexes, we used a "top-down" strategy that first identifies all existing symmetric oligomers and then tracks their corresponding interfaces. Main components of the computational pipeline are detailed below. Access to the computational methods is available by contacting the authors.

4.2.1. Universal numbering scheme of Fabs

To analyze sequences and structures across diverse Fabs, a universal numbering scheme was adopted to renumber all positions in each Fab structure. Such universal numbering assigns structurally equivalent positions in different Fabs with the same number. For the variable domain (VH and VL), this was achieved by using the ANARCI program [79] with the AHo convention [39], a numbering scheme designed to account for gaps in the variable domain to most accurately reflect structural equivalence. For the Fab constant domain (CH1 and CL), a customized numbering scheme referred to as Universal Constant Numbering (UCN) was created as follows. (1) Germline genes (IGHC, IGKC, and IGLC) representing species that appear in the SAbDab dataset were collected from the IMGT database (https://www.imgt.org/vquest/refseqh.html). They were paired with Fab structures containing the highest sequence similarity. Immunoglobulin E (IgE) sequences were excluded due to lack of available structures in SAbDab. (2) The seven most conserved β-strands corresponding to the protein core were manually identified based on the multiple sequence alignment and structure alignment (if available) of the germline sequences and their paired structures. Interestingly, structures of these β-strands superimpose well even between CH1 and CL. (3) By using these seven intermittent β-strands as fixed regions and accounting for gaps in-between, the representative germline sequences were manually aligned and numbered. The sequence alignment was stored in .stockholm format. (4) Hidden Markov Model (HMM) profile libraries were compiled for each species and germline gene in the manual sequence alignments using the *hmmbuild* and *hmmpress* commands of the HMMER program. (5) All Fab constant domain sequences in the structure dataset were then searched against these pre-computed HMM profile libraries using the *hmmscan* command of the HMMER program, and were mapped to the corresponding positions in the best hit.

4.2.2. Identification of interfaces

The asymmetric unit in each Fab PDB structure was expanded into a crystal lattice block by running a PyMOL script that calls the *symexp* command, which expands around the original asymmetric unit up to 30 Å. For computing efficiency, the expanded crystal lattice block was then trimmed down to at most 20 closest Fab monomer structures around the original asymmetric unit, which was large enough to capture almost all potential symmetric interfaces.

To exhaustively identify all interfaces between Fab monomers in each crystal lattice block, all pairs of Fab monomers were examined for all interfaces above 100 Å^2 (two-sides) after stripping away water and other small molecules. An undirected graph was then built using Fab monomers as nodes and interfaces as edges. An oligomer here is presented as a connected sub-graph, in which every node has at least 1 edge to other nodes in the same sub-graph. By doing this, the problem of searching for all existing oligomers is transformed to searching for all connected sub-graphs. An in-house merging algorithm was developed to efficiently solve this problem.

4.2.3. Rotation angles and axis

A transformation matrix was computed between every possible pair of Fab monomers (regardless of contact or not) by calling *align* and *get_object_matrix* PyMOL commands. The rotation angle and axis were mathematically determined from the rotation matrix R using the following equations:

$$1)\ R=\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad V=\begin{bmatrix} Vx \\ Vy \\ Vz \end{bmatrix}$$
$$2)\ Tr = 1 + 2\cos\theta$$
$$3)\ Vx = h - f \qquad Vy = c - g \qquad Vz = d - b$$

Where V is the rotation axis vector, θ is the rotation angle, and Tr is the trace of R. θ is between 0, 180°, which is a sufficient range to effectively test the rotational symmetry. Specially, when θ is 180°, to determine the rotation axis, there are the following different possibilities:

$$V=\begin{bmatrix} Ux \\ sUy \\ tUz \end{bmatrix},$$

$$\begin{cases} Ux = \sqrt{(a+1)/2} \\ Uy = \sqrt{(e+1)/2}, \\ Uz = \sqrt{(i+1)/2} \end{cases}$$
$$s=\begin{cases} 1 \ if\ b+d \geq 0 \\ -1 \ if\ b+d < 0 \end{cases}$$
$$t=\begin{cases} 1 \ if\ c+g > 0 \\ -1 \ if\ c+g < 0 \\ 1 \ if\ c+g = 0,\ h+f \geq 0 \\ -1 \ if\ c+g = 0,\ h+f < 0 \end{cases}$$

Due to the limited computation precision, imperfect PDB structures, and pseudo symmetry, very similar rotations were treated as identical rotations. To identify them, a pair of rotations were first compared by angles with a cutoff of 5°, and then compared by the angle between their rotation axes with a cutoff of 5°. The angle between rotation axes was calculated using the more numerically stable method:

$$\theta = \tan^{-1}\frac{\|U \times V\|}{U \cdot V}$$

Where θ is the angle between axes, and U and V are the vectors of the two axes. In two special cases, 0° rotations were all dropped, and 180° inter rotation axes (opposite directions) were considered as the same axes.

4.2.4. Determination of symmetric oligomers

The last step in the pipeline used the computed rotation information to determine whether a given Fab oligomer is symmetric and in which types. The approach explored every observed rotation in one oligomer. A symmetric rotation allows an oligomer to superimpose to itself. Because a full-atom RSMD calculation on

all symmetric rotations for all Fab oligomers was prohibitively expensive, subunits were first represented as centroids that reflected their position and orientation with minimal coordinates. This allowed a fast calculation with a permissive low-resolution root-mean-square deviation (RMSD) threshold to filter out the vast majority of obvious non-symmetric rotations. A full-atom RMSD calculation was applied to the remaining oligomers. Both centroid and full-atom RMSD were required to be < 5 Å for a given rotation to be considered symmetric. Finally, identical symmetric oligomers were identified and collapsed using the same centroid to full-atom RMSD strategy.

4.3. Clustering analysis

An interface I was described using the set of interacting residues between them (identified by using a distance cutoff R). Analysis was performed using a distance cutoff of 4 Å and 6 Å.

$$I = \{(i,j) \; \forall \; i,j \text{ such that } d(F_i, F_j) \leq R\},$$

The similarity between two interfaces I^m and I^n is defined as the Jaccard index between the sets of interfaces as follows:

$$J^{m,n} = \frac{|I^m \cap I^n|}{|I^m \cup I^n|}$$

A variant of the analysis utilized a weighted Jaccard similarity, in which each of the two contact pair sets is a real number vector V, with its elements derived from the distance of the corresponding contact pair (and value 0 for the non-contact residue pairs) as follows.

$$V_i^m = 1 - \frac{d_i}{R} \; \forall \; i \in I^m$$

Then, the weighted Jaccard similarity was calculated as.

$$W^{m,n} = \frac{\sum \min_i (V_i^m, V_i^n) \forall \; i = 1, 2, \cdots |I^m|}{\sum \max_i (V_i^m, V_i^n) \forall \; i = 1, 2, \cdots |I^m|}$$

The above similarity values were then used to generate distance matrices (using 1 – J or 1 – W as the values) for distance cutoffs of 4 Å and 6 Å. Hierarchical clustering was then performed using single linkage in R using the 'hclust' function. The resulting dendrograms were cut at different heights and the similarities of various cluster members were compared to decide on the optimal cutoff height. The clusters obtained by cutting the Weighted Jaccard index based dendrogram (and interface distance cutoff of 4 Å) at a height of 0.3 was selected for further analysis. Based on the hierarchical clustering, each interface was assigned to an interface cluster.

4.4. Cluster representatives and naming

For each member of a cluster, the average of all weighted Jaccard similarity indices to all other members was calculated. The representative member of each cluster was defined as the interface that had the highest average weighted Jaccard index to all other members.

The name of a given cluster represents the residue at the structural center of the Fab-Fab interface. The name adopts the format X-#, in which X represents commonly used abbreviations for Fab chains (VH, VL, CH1, CL) and # represents the residue number according to either the Kabat or EU numbering convention for variable or constant regions, respectively. The central residue for each cluster was determined as follows. Relevant interface residues for each cluster's representative member were determined by applying an interface distance cutoff of 5 Å. The xyz coordinates of the α-carbon for each interface residue were averaged to obtain the structural average of the interface. For each residue in the interface, the distance between the given residue's α-carbon and the structural average xyz coordinates was determined. The residue with the shortest distance to the central xyz coordinates was used for the naming convention. Duplication was avoided in a small number of instances by choosing the second closest residue to the structural center.

4.5. Sequence analysis

For the antibody Fabs in the dataset, all-against-all pairwise alignments were determined for the different Fab regions listed here: VH-only, VL-only, CH1-only, CL-only, and VH/VL-Fv-only. Dynamic programming, with matching residue score at 1, mismatching score at −1, gap-opening score at −6, and gap-extension score at −3, was used to determine the respective alignments. Referencing the pairwise alignment results, Levenshtein distance was calculated to determine the difference between the pairwise-compared regions. Percent-identity was then calculated as the subtraction from 100 % by the difference percentage. In the case of Fv percent-identity, the respective VH and VL percent-identities were determined separately; then, the average of the VH- and VL- percent-identities was reported as the identity for the compared Fv regions. In the case of interface percent identity, sequence alignments were performed only at the residue positions of a given interface for each interface (Supplementary Table S2).

4.6. PISA analysis

Structural features of interfaces were analyzed using the 'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute. (https://www.ebi.ac.uk/pdbe/prot_int/pistart.html) [41]. PISA interfaces were downloaded from the PDBePISA site (https://www.ebi.ac.uk/pdbe/pisa/) and mapped to in-house generated oligomer interfaces. Each such oligomer interface consists of single or multiple disconnected PISA interfaces between different chains, and each PISA interface can be included partly or fully, depending on the overlapping interface residues. Buried surface area, number of H-bonds, and number of salt bridges for each interface were calculated as the summation of its component PISA interfaces. In order to accurately analyze the in-house interfaces, we applied two filtering processes for each oligomer interface: (1) removal of PISA interface patches involving antigens, keeping only antibody-antibody interfaces and (2) removal of PISA interfaces for which there was poor overlap (<0.5) between PISA interface and in-house interface residues.

4.7. Disulfide design

The representative structure of each cluster was used to perform an in silico cysteine scanning simulation using the cysteine scanning module [80] in the BioLuminate suite (Schrodinger Inc.). Default parameters were used, with residues within 5 Å allowed to be flexible (flex_dist = 5.0) while performing stability calculations. Potential disulfide pairs were identified as those residue pairs with a Cβ-Cβ distance within 5.0 Å, irrespective of whether the calculations yielded favorable energies or not. The list of all potential disulfide residue pairs identified through this analysis is summarized in Supplementary Tables S3-S5.

4.8. DNA construction and protein production

The anti-Her2 and anti-OX40 antibodies used in this work have been described previously [43,45]. Molecular biology to generate

the Fab disulfide variants was carried out using gene synthesis (Genewiz). DNAs encoding LCs and Fab HCs in the pRK mammalian expression vector were cotransfected into Expi293 cells for expression. Fabs were purified using CaptureSelect CH1-XL affinity resin (Thermo, 194346201L) followed by *SEC* using a HiLoad 16/600 Superdex 200 column. Fab monomer and dimer fractions were pooled separately during *SEC* purification to isolate the desired oligomeric species. Protein quality was assessed by *SEC* using a Waters xBridge BEH200A *SEC* 3.5um (7.8 × 300 mm) column (Waters,176003596) on a Thermo Fisher Scientific Ultimate 3000 HPLC system, running Chromeleon Chromatography Data System (CDS) software for system operation and data analysis. For characterization of oligomeric species following expression, small aliquots of affinity purified Fabs were loaded onto the Waters column, and relative percentages of oligomers were calculated using the area under the chromatogram for each peak. Peak integration is based on methods that are internal to the CDS software, and due to variances across the full chromatogram the values for selected monomer, dimer, and higher order peaks do not necessarily add up to 100 %. Molecular weight of all Fabs was confirmed by LC/MS. Fabs were stored in a buffer consisting of 20 mM histidine acetate and 150 mM NaCl at pH 5.5.

4.9. Affinity measurements

Solution binding was assessed using a Biacore T200 instrument (GE). Fabs were captured using either a Series S Protein L chip (Cytiva, 29205138) or a human Fab capture reagent immobilized on a Series S CM5 chip (Cytiva, 29,104,988 and 29234601) according to the manufacturer's specifications. A 4-fold serial dilution of OX40 starting at 100 nM (G&P Bioscience, FCL0103) was prepared in HBS-P+ buffer (10 mM HEPES, 150 mM NaCl, 0.05 % v/v surfactant P20, Cytiva, BR100671) and injected for 5 min, followed by a 6 min dissociation period. Affinity constants were obtained through kinetic fitting using the Biacore Evaluation Software (GE).

4.10. In vitro oxidation

Anti-Her2 Fab disulfide variants were exchanged into a phosphate-buffer saline (PBS) solution at pH 7.4 and concentrated to 0.5 mg/ml using Amicon Ultra centrifugal filters with a 10 kDa molecular weight cutoff (Millipore, UFC901096). Dithiothreitol (Thermo, R0861) was added to 5 mM from a 40 mM stock solution in water. Samples were incubated at 37 °C for 2 h to fully reduce all cysteines. Dithiothreitol was removed through buffer exchange into PBS using an Amicon Ultra centrifugal filters with a 10 kDa molecular weight cutoff. The temperature of the samples was kept at 4 °C during the buffer exchange process. A tenfold molar excess of the oxidizing agent (L)-dehydroascorbic acid (Sigma-Aldrich, 261556) was then added to each sample from a 5 mM stock solution in PBS, and samples were incubated for one week at room temperature. The oxidation reactions were stopped and any remaining free sulfurs were capped upon addition of *N*-ethylmaleimide (Pierce, 23030) to 5 mM from a 100 mM stock solution in water. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) of the re-oxidized samples was performed on a 4–15 % Criterion TGX precast midi protein gel (Bio-Rad, 5671085) under non-reducing conditions. Bands corresponding to oligomeric species were quantified using ImageJ software and normalized so that monomer + dimer + higher order equaled 100 %. No precipitation was observed during the experiment, and no bands were observed at the top of the gels, suggesting that all Fabs assembled solubly.

4.11. Activity assay

The OX40 agonist assay was performed as previously described [45]. Briefly, OX40 overexpressing Jurkat cells engineered with an NFκB luciferase reporter were seeded at 80,000 cells/well in 20 μl RPMI containing 1 % L-glutamine and 10 % HI FBS in a 384-well tissue culture plate (Corning Inc., 3985BC). Anti-OX40 antibody formats were serially diluted threefold in media starting at 30 μg/ml, and 10 μl of the concentrated antibodies were added to each well. For the conditions with crosslinking, 10 μl of AffiniPure goat anti-human IgG Fcγ fragment specific antibody (for IgG samples, Jackson ImmunoResearch Laboratories Inc, 109–005-098) in media was added to yield a 1:1 molar ratio with each antibody dilution. For conditions without crosslinker, 10 μl of media was added to each antibody dilution. The plates were then incubated for 16–18 hrs under 5 % CO_2 at 37 °C. 40 μl of Bright Glo (Promega, E2610) was then added to each well and incubated with shaking at room temperature for 5 min. Luminescence was detected using a Perkin Elmer Envision plate reader. For larger panels of antibodies, automation of this assay was developed using a Tecan Fluent. All activity data from this assay is reported as fold change over control well without antibody.

4.12. Crystallographic structure solution

The purified dimers of the anti-Her2 Fab disulfide variants were concentrated to 10 mg/ml in PBS. Crystallization trials were performed using the sitting-drop vapor diffusion method with commercially available sparse-matrix screens in a 96-well format. The final crystallization condition for the anti-Her2 Fab disulfide variant CH1(S119C) / CH1(G122C) (interface CH1-207) contained 0.8 M sodium succinate. The condition for variant VL(G16C) / CL (D170C) (interface VL-108) contained 1 % PEG 2000 and 100 mM HEPES pH 8.0. Finally, the condition for variant CL(S202C) / CL (S208C) (interface CL-205) contained 15 % w/v PEG 3350 and 100 mM sodium succinate. Crystals of the dimers were preserved for data collection by brief soaking in a cryo-protectant buffer (25 % glycerol added to the reservoir solution), followed by rapid immersion into liquid nitrogen. Diffraction data were collected at the Advanced Light Source (ALS) beam line 5.0.2 for CH1-207 and VL-108, while data for CL-205 were collected at the Northeastern Collaborative Access Team (NE-CAT) beamline 24IDC.

Data were reduced with Global Phasing's autoProc [81] using XDS [82]. Datasets for CH1-207 and VL-108 were defined with elliptical anisotropic resolution as implemented in the STARANISO procedure [83]. Structure determination was done by molecular replacement with Phaser [84] using a search model from a prior internal antibody Fab structure separated into VH, VL, CH1, and CL subdomains. Electron density maps were consistent with the presence of the engineered disulfide bonds at the expected sequence positions (Supplementary Fig. S5) and these were introduced during manual model building in Coot [85]. Of note, the initial solution for VL-108 was determined in a 6-fold symmetry group, but presented only 1 molecule per asymmetric unit in that context with the neighboring linked protomer positioned by apparent crystal symmetry. However, the inability to adequately model the disulfide bond of interest across the special symmetry position and the potential twinning led us to reduce the space group symmetry to P3121 with two Fab molecules per asymmetric unit and to model twinning in the data that had presented as the higher symmetry. Similarly, the CL-205 dataset initially appeared to be in space group P21 with two Fabs per asymmetric unit but had strong indicators of translational pseudo-symmetry (Phenix xtriage [86] analysis) and the Rfree value did not decrease much below 30 % in refinement efforts. The data were re-processed in P1 and the molecular replacement search repeated to identify 4

Fab molecules per asymmetric unit, and refinement in this setting with amplitude-based twin estimates in REFMAC [87] provided approximately 4 % reduction in R factors. The CH1-207 data were mercifully more straightforward. Models were refined with cycles of REFMAC [87], Phenix [86], or Buster [88] to reasonable statistics (Supplementary Table S7). CL-205 has 4 Fab molecules per asymmetric unit in space group P1, the engineered interface between the two Fabs, which are assigned chains H (heavy) and L (light) and the second Fab with chains A (heavy) and B (light) and analogously for the neighboring C/D and E/F chain Fabs. The CH1-207 structure also displays 4 Fab molecules/asu, with engineered interfaces between Fabs HL and AB and Fabs CD and EF. Finally, VL-108 contains 2 Fab molecule/asu in a P3121 space group setting, twinned, assigned chains HL and AB. Coordinates and structure factors are deposited with the PDB under accession codes 7T97 (CH1-207), 7T98 (VL-108), and 7T99 (CL-205).

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Greg Lazar reports financial support was provided by Genentech USA Inc South San Francisco. Matthew G. Romei, Kannan Sankar, Kam Hon Hoi, Yanli Yang, Brandon Leonard, Gladys De Leon Boenig, Nikit Kumar, Marissa Matsumoto, Jian Payandeh, Seth F. Harris reports financial support was provided by Genentech USA Inc South San Francisco. Greg Lazar reports a relationship with Genentech USA Inc South San Francisco that includes: employment and equity or stocks. Matthew G. Romei, Kannan Sankar, Kam Hon Hoi, Yanli Yang, Brandon Leonard, Gladys De Leon Boenig, Nikit Kumar, Marissa Matsumoto, Jian Payandeh, Seth F. Harris reports a relationship with Genentech USA Inc South San Francisco that includes: employment and equity or stocks.

Acknowledgements

The authors thank Ingrid Kim, Farzam Farahi, Peter Luan, and Nina Kim for technical contributions, and Kiran Mukhyala for thoughtful input. The synchrotron radiation sources for structure determination are supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-05CH11231 (Advanced Light Source) and contract no. DE-AC02-06CH11357 (Advanced Photon Source).

Author contributions

G.A.L. conceived the study, and M.M., J.P., S.F.H., J.M., and G.A.L. supervised the project and experiments. Y.Y, M.G.R., K.S., L.R.P, and K.H.H. performed computational development and analysis. M.G. R., K.S., Y.Y., B.L., G.D.L.B., and N.K. executed laboratory experiments. G.D.L.B. and S.F.H. executed structural studies. Y.Y, M.G.R., K.S., K.H.H., S.F.H., and G.A.L. wrote the manuscript, with input from all authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.08.048.

References

[1] Lingwood D, McTamney PM, Yassine HM, Whittle JRR, Guo X, Boyington JC, et al. Structural and genetic basis for development of broadly neutralizing influenza antibodies. Nature 2012;489:566–70. https://doi.org/10.1038/nature11371.

[2] Diebolder CA, Beurskens FJ, de Jong RN, Koning RI, Strumane K, Lindorfer MA, et al. Complement Is Activated by IgG Hexamers Assembled at the Cell Surface. Science 2014;343:1260–3. https://doi.org/10.1126/science.1248943.

[3] Carter PJ, Lazar GA. Next generation antibody drugs: pursuit of the "high-hanging fruit". Nat Rev Drug Discov 2018;17:197–223. https://doi.org/10.1038/nrd.2017.227.

[4] Ridgway JBB, Presta LG, Carter P. 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. Protein Eng Des Sel 1996;9:617–21. https://doi.org/10.1093/protein/9.7.617.

[5] Spiess C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. Mol Immunol 2015;67:95–106. https://doi.org/10.1016/j.molimm.2015.01.003.

[6] Shields RL, Namenuk AK, Hong K, Meng YG, Rae J, Briggs J, et al. High Resolution Mapping of the Binding Site on Human IgG1 for FcγRI, FcγRII, FcγRIII, and FcRn and Design of IgG1 Variants with Improved Binding to the FcγR*. J Biol Chem 2001;276:6591–604. https://doi.org/10.1074/jbc.m009483200.

[7] Shields RL, Lai J, Keck R, O'Connell LY, Hong K, Meng YG, et al. Lack of Fucose on Human IgG1 N-Linked Oligosaccharide Improves Binding to Human FcγRIII and Antibody-dependent Cellular Toxicity*. J Biol Chem 2002;277:26733–40. https://doi.org/10.1074/jbc.m202069200.

[8] Lazar GA, Dang W, Karki S, Vafa O, Peng JS, Hyun L, et al. Engineered antibody Fc variants with enhanced effector function. P Natl Acad Sci Usa 2006;103:4005–10. https://doi.org/10.1073/pnas.0508123103.

[9] Richards JO, Karki S, Lazar GA, Chen H, Dang W, Desjarlais JR. Optimization of antibody binding to FcγRIIa enhances macrophage phagocytosis of tumor cells. Mol Cancer Ther 2008;7:2517–27. https://doi.org/10.1158/1535-7163.mct-08-0201.

[10] Chu SY, Vostiar I, Karki S, Moore GL, Lazar GA, Pong E, et al. Inhibition of B cell receptor-mediated activation of primary human B cells by coengagement of CD19 and FcγRIIb with Fc-engineered antibodies. Mol Immunol 2008;45:3926–33. https://doi.org/10.1016/j.molimm.2008.06.027.

[11] Stavenhagen JB, Gorlatov S, Tuaillon N, Rankin CT, Li H, Burke S, et al. Fc Optimization of Therapeutic Antibodies Enhances Their Ability to Kill Tumor Cells In vitro and Controls Tumor Expansion In vivo via Low-Affinity Activating Fcγ Receptors. Cancer Res 2007;67:8882–90. https://doi.org/10.1158/0008-5472.can-07-0696.

[12] Umaña P, Jean-Mairet J, Moudry R, Amstutz H, Bailey JE. Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. Nat Biotechnol 1999;17:176–80. https://doi.org/10.1038/6179.

[13] Mimoto F, Katada H, Kadono S, Igawa T, Kuramochi T, Muraoka M, et al. Engineered antibody Fc variant with selectively enhanced FcγRIIb binding over both FcγRIIaR131 and FcγRIIaH131. Protein Eng Des Sel 2013;26:589–98. https://doi.org/10.1093/protein/gzt022.

[14] Mimoto F, Igawa T, Kuramochi T, Katada H, Kadono S, Kamikawa T, et al. Novel asymmetrically engineered antibody Fc variant with superior FcγR binding affinity and specificity compared with afucosylated Fc variant. Mabs 2013;5:229–36. https://doi.org/10.4161/mabs.23452.

[15] Niwa R, Shoji-Hosaka E, Sakurada M, Shinkawa T, Uchida K, Nakamura K, et al. Defucosylated Chimeric Anti-CC Chemokine Receptor 4 IgG1 with Enhanced Antibody-Dependent Cellular Cytotoxicity Shows Potent Therapeutic Activity to T-Cell Leukemia and Lymphoma. Cancer Res 2004;64:2127–33. https://doi.org/10.1158/0008-5472.can-03-2068.

[16] Idusogie EE, Wong PY, Presta LG, Gazzano-Santoro H, Totpal K, Ultsch M, et al. Engineered Antibodies with Increased Activity to Recruit Complement. J Immunol 2001;166:2571–5. https://doi.org/10.4049/jimmunol.166.4.2571.

[17] Moore GL, Chen H, Karki S, Lazar GA. Engineered Fc variant antibodies with enhanced ability to recruit complement and mediate effector functions. Mabs 2010;2:181–9. https://doi.org/10.4161/mabs.2.2.11158.

[18] Lee C-H, Kang TH, Godon O, Watanabe M, Delidakis G, Gillis CM, et al. An engineered human Fc domain that behaves like a pH-toggle switch for ultra-long circulation persistence. Nat Commun 2019;10:5031. https://doi.org/10.1038/s41467-019-13108-2.

[19] Zalevsky J, Chamberlain AK, Horton HM, Karki S, Leung IWL, Sproule TJ, et al. Enhanced antibody half-life improves in vivo activity. Nat Biotechnol 2010;28:157–9. https://doi.org/10.1038/nbt.1601.

[20] Dall'Acqua WF, Kiener PA, Wu H. Properties of Human IgG1s Engineered for Enhanced Binding to the Neonatal Fc Receptor (FcRn)*. J Biol Chem 2006;281:23514–24. 10.1074/jbc.m604292200.

[21] Rugo HS, Im S-A, Cardoso F, Cortés J, Curigliano G, Musolino A, et al. Efficacy of Margetuximab vs Trastuzumab in Patients With Pretreated ERBB2-Positive Advanced Breast Cancer. Jama Oncol 2021;7:573–84. https://doi.org/10.1001/jamaoncol.2020.7932.

[22] Nordstrom JL, Gorlatov S, Zhang W, Yang Y, Huang L, Burke S, et al. Anti-tumor activity and toxicokinetics analysis of MGAH22, an anti-HER2 monoclonal antibody with enhanced Fcγ receptor binding properties. Breast Cancer Res 2011;13:R123. https://doi.org/10.1186/bcr3069.

[23] Ollila TA, Sahin I, Olszewski AJ. Mogamulizumab: a new tool for management of cutaneous T-cell lymphoma. Oncotargets Ther 2019;12:1085–94. https://doi.org/10.2147/ott.s165615.

[24] Tobinai K, Klein C, Oya N, Fingerle-Rowson G. A Review of Obinutuzumab (GA101), a Novel Type II Anti-CD20 Monoclonal Antibody, for the Treatment of

Patients with B-Cell Malignancies. Adv Ther 2017;34:324–56. https://doi.org/10.1007/s12325-016-0451-1.
[25] Ghazi A, Trikha A, Calhoun WJ. Benralizumab – a humanized mAb to IL-5Rα with enhanced antibody-dependent cell-mediated cytotoxicity – a novel approach for the treatment of asthma. Expert Opin Biol Th 2011;12:113–8. https://doi.org/10.1517/14712598.2012.642359.
[26] Kolbeck R, Kozhich A, Koike M, Peng L, Andersson CK, Damschroder MM, et al. MEDI-563, a humanized anti–IL-5 receptor α mAb with enhanced antibody-dependent cell-mediated cytotoxicity function. J Allergy Clin Immun 2010;125:1344–1353.e2. https://doi.org/10.1016/j.jaci.2010.04.004.
[27] Vijayaraghavan S, Lipfert L, Chevalier K, Bushey BS, Henley B, Lenhart R, et al. Amivantamab (JNJ-61186372), an Fc Enhanced EGFR/cMet Bispecific Antibody, Induces Receptor Downmodulation and Antitumor Activity by Monocyte/Macrophage Trogocytosis. Mol Cancer Ther 2020;19:2044–56. https://doi.org/10.1158/1535-7163.mct-20-0071.
[28] Robbie GJ, Criste R, Dall'Acqua WF, Jensen K, Patel NK, Losonsky GA, et al. A Novel Investigational Fc-Modified Humanized Monoclonal Antibody, Motavizumab-YTE, Has an Extended Half-Life in Healthy Adults. Antimicrob Agents Ch 2013;57:6147–53. https://doi.org/10.1128/aac.01285-13.
[29] Gaudinski MR, Coates EE, Houser KV, Chen GL, Yamshchikov G, Saunders JG, et al. Safety and pharmacokinetics of the Fc-modified HIV-1 human monoclonal antibody VRC01LS: A Phase 1 open-label clinical trial in healthy adults. Plos Med 2018;15:e1002493.
[30] Salles G, Długosz-Danecka M, Ghesquières H, Jurczak W. Tafasitamab for the treatment of relapsed or refractory diffuse large B-cell lymphoma. Expert Opin Biol Th 2021;21:1–9. https://doi.org/10.1080/14712598.2021.1884677.
[31] de Latour RP, Brodsky RA, Ortiz S, Risitano AM, Jang JH, Hillmen P, et al. Pharmacokinetic and pharmacodynamic effects of ravulizumab and eculizumab on complement component 5 in adults with paroxysmal nocturnal haemoglobinuria: results of two phase 3 randomised, multicentre studies. Brit J Haematol 2020;191:476–85. https://doi.org/10.1111/bjh.16711.
[32] Sheridan D, Yu Z-X, Zhang Y, Patel R, Sun F, Lasaro MA, et al. Design and preclinical characterization of ALXN1210: A novel anti-C5 antibody with extended duration of action. PLoS ONE 2018;13:e0195909.
[33] Saphire EO, Parren PWHI, Pantophlet R, Zwick MB, Morris GM, Rudd PM, et al. Crystal Structure of a Neutralizing Human IgG Against HIV-1: A Template for Vaccine Design. Science 2001;293:1155–9. https://doi.org/10.1126/science.1061692.
[34] Capitani G, Duarte JM, Baskaran K, Bliven S, Somody JC. Understanding the fabric of protein crystals: computational classification of biological interfaces and crystal contacts. Bioinformatics 2016;32:481–9. https://doi.org/10.1093/bioinformatics/btv622.
[35] Elez K, Bonvin AMJJ, Vangone A. Distinguishing crystallographic from biological interfaces in protein complexes: role of intermolecular contacts and energetics for classification. BMC Bioinf 2018;19:438. https://doi.org/10.1186/s12859-018-2414-9.
[36] Tamada T, Shinmi D, Ikeda M, Yonezawa Y, Kataoka S, Kuroki R, et al. TRAIL-R2 Superoligomerization Induced by Human Monoclonal Agonistic Antibody KMTR2. Sci Rep-Uk 2015;5:17936. https://doi.org/10.1038/srep17936.
[37] Crowley PB, Matias PM, Mi H, Firbank SJ, Banfield MJ, Dennison C. Regulation of Protein Function: Crystal Packing Interfaces and Conformational Dimerization. Biochemistry-Us 2008;47:6583–9. https://doi.org/10.1021/bi800125h.
[38] Dunbar J, Krawczyk K, Leem J, Baker T, Fuchs A, Georges G, et al. SAbDab: the structural antibody database. Nucleic Acids Res 2014;42:D1140–6. https://doi.org/10.1093/nar/gkt1043.
[39] Honegger A, Plückthun A. Yet Another Numbering Scheme for Immunoglobulin Variable Domains: An Automatic Modeling and Analysis Tool. J Mol Biol 2001;309:657–70. https://doi.org/10.1006/jmbi.2001.4662.
[40] Kabat EA, Wu TT, Perry HM, Foeller C, Gottesman KS. Sequences of proteins of immunological interest. NIH Pulication: Fifth Edition; 1991.
[41] Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. J Mol Biol 2007;372:774–97. https://doi.org/10.1016/j.jmb.2007.05.022.
[42] Luo J, Liu Z, Guo Y, Li M. A structural dissection of large protein-protein crystal packing contacts. Sci Rep-Uk 2015;5:14214. https://doi.org/10.1038/srep14214.
[43] Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL, et al. Humanization of an anti-p185HER2 antibody for human cancer therapy. Proc National Acad Sci 1992;89:4285–9. https://doi.org/10.1073/pnas.89.10.4285.
[44] Feige MJ, Buchner J. Principles and engineering of antibody folding and assembly. Biochimica Et Biophysica Acta Bba - Proteins Proteom 2014;1844:2024–31. https://doi.org/10.1016/j.bbapap.2014.06.004.
[45] Yang Y, Yeh SH, Madireddi S, Matochko WL, Gu C, Sanchez PP, et al. Tetravalent biepitopic targeting enables intrinsic antibody agonism of tumor necrosis factor receptor superfamily members. Mabs 2019;11:996–1011. https://doi.org/10.1080/19420862.2019.1625662.
[46] Ozbabacan SEA, Engin HB, Gursoy A, Keskin O. Transient protein–protein interactions. Protein Eng Des Sel 2011;24:635–48. https://doi.org/10.1093/protein/gzr025.
[47] Humphris EL, Kortemme T. Design of Multi-Specificity in Protein Interfaces. Plos Comput Biol 2007;3:e164.
[48] Garcia-Seisdedos H, Empereur-Mot C, Elad N, Levy ED. Proteins evolve on the edge of supramolecular self-assembly. Nature 2017;548:244–7. https://doi.org/10.1038/nature23320.
[49] Romei MG, Boxer SG. Split Green Fluorescent Proteins: Scope, Limitations, and Outlook. Annu Rev Biophys 2013;48:1–26. https://doi.org/10.1146/annurev-biophys-051013-022846.
[50] Mullard A. FDA approves 100th monoclonal antibody product. Nat Rev Drug Discov 2021;20:491–5. https://doi.org/10.1038/d41573-021-00079-7.
[51] Guharoy M, Chakrabarti P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. Bioinformatics 2007;23:1909–18. https://doi.org/10.1093/bioinformatics/btm274.
[52] Schumacher MA, Funnell BE. Structures of ParB bound to DNA reveal mechanism of partition complex formation. Nature 2005;438:516–9. https://doi.org/10.1038/nature04149.
[53] Monaco H, Rizzi M, Coda A. Structure of a complex of two plasma proteins: transthyretin and retinol-binding protein. Science 1995;268:1039–41. https://doi.org/10.1126/science.7754382.
[54] Prapunpoj P, Leelawatwattana L. Evolutionary changes to transthyretin: structure–function relationships. Febs J 2009;276:5330–41. https://doi.org/10.1111/j.1742-4658.2009.07243.x.
[55] Nassar N, Horn G, Herrmann CA, Scherer A, McCormick F, Wittinghofer A. The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with RaplA and a GTP analogue. Nature 1995;375:554–60. https://doi.org/10.1038/375554a0.
[56] Stranges PB, Machius M, Miley MJ, Tripathy A, Kuhlman B. Computational design of a symmetric homodimer using β-strand assembly. Proc National Acad Sci 2011;108:20562–7. https://doi.org/10.1073/pnas.1115124108.
[57] Cheng P-N, Pham JD, Nowick JS. The Supramolecular Chemistry of β–Sheets. J Am Chem Soc 2013;135:5477–92. https://doi.org/10.1021/ja3088407.
[58] Greenwald J, Riek R. Biology of amyloid: structure, function, and regulation. Structure 2010;18:1244–60. https://doi.org/10.1016/j.str.2010.08.009.
[59] Richardson JS, Richardson DC. Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc National Acad Sci 2002;99:2754–9. https://doi.org/10.1073/pnas.052706099.
[60] Dobson CM, Karplus M. The fundamentals of protein folding: bringing together theory and experiment. Curr Opin Struc Biol 1999;9:92–101. https://doi.org/10.1016/s0959-440x(99)80012-8.
[61] Xu Y, Wang D, Mason B, Rossomando T, Li N, Liu D, et al. Structure, heterogeneity and developability assessment of therapeutic antibodies. Mabs 2018;11:1–26. https://doi.org/10.1080/19420862.2018.1553476.
[62] Jarasch A, Koll H, Regula JT, Bader M, Papadimitriou A, Kettenberger H. Developability Assessment During the Selection of Novel Therapeutic Antibodies. J Pharm Sci 2015;104:1885–98. https://doi.org/10.1002/jps.24430.
[63] van der Kant R, Karow-Zwick AR, Durme JV, Blech M, Gallardo R, Seeliger D, et al. Prediction and Reduction of the Aggregation of Monoclonal Antibodies. J Mol Biol 2017;429:1244–61. https://doi.org/10.1016/j.jmb.2017.03.014.
[64] Li W, Prabakaran P, Chen W, Zhu Z, Feng Y, Dimitrov DS. Antibody Aggregation: Insights from Sequence and Structure. Antibodies 2016;5:19. https://doi.org/10.3390/antib5030019.
[65] Wang X, Das TK, Singh SK, Kumar S. Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. Mabs 2009;1:254–67. https://doi.org/10.4161/mabs.1.3.8035.
[66] Kumar N, Arthur CP, Ciferri C, Matsumoto ML. Structure of the secretory immunoglobulin A core. Science 2020;367:1008–14. https://doi.org/10.1126/science.aaz5807.
[67] Li Y, Wang G, Li N, Wang Y, Zhu Q, Chu H, et al. Structural insights into immunoglobulin M. Science 2020;367:1014–7. https://doi.org/10.1126/science.aaz5425.
[68] Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, et al. Antibody Domain Exchange Is an Immunological Solution to Carbohydrate Cluster Recognition. Science 2003;300:2065–71. https://doi.org/10.1126/science.1083182.
[69] Wu Y, West AP, Kim HJ, Thornton ME, Ward AB, Bjorkman PJ. Structural Basis for Enhanced HIV-1 Neutralization by a Dimeric Immunoglobulin G Form of the Glycan-Recognizing Antibody 2G12. Cell Reports 2013;5:1443–55. https://doi.org/10.1016/j.celrep.2013.11.015.
[70] Williams WB, Meyerhoff RR, Edwards RJ, Li H, Manne K, Nicely NI, et al. Fab-dimerized glycan-reactive antibodies are a structural category of natural antibodies. Cell 2021;184:2955–2972.e25. https://doi.org/10.1016/j.cell.2021.04.042.
[71] Kucharska I, Thai E, Srivastava A, Rubinstein JL, Pomès R, Julien J-P. Structural ordering of the Plasmodium berghei circumsporozoite protein repeats by inhibitory antibody 3D11. Elife 2020;9:e59018.
[72] Rougé L, Chiang N, Steffek M, Kugel C, Croll TI, Tam C, et al. Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab. Science 2020;367:1224–30. https://doi.org/10.1126/science.aaz9356.
[73] Leonard B, Sankar K, Romei MG, Tse MJ, Do N, Yang Y, et al. Antibody homotypic interactions are encoded by germline light chain complementarity determining region 2. P Natl Acad Sci Usa 2022;119:e2201562119. 10.1073/pnas.2201562119.
[74] Feng Z, Chen X, Wu X, Zhang M. Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications. J Biol Chem 2019;294:14823–35. https://doi.org/10.1074/jbc.rev119.007895.
[75] Reichheld SE, Muiznieks LD, Keeley FW, Sharpe S. Direct observation of structure and dynamics during phase separation of an elastomeric protein. Proc National Acad Sci 2017;114:E4408–15. https://doi.org/10.1073/pnas.1701877114.

[76] Chiu Y-P, Sun Y-C, Qiu D-C, Lin Y-H, Chen Y-Q, Kuo J-C, et al. Liquid-liquid phase separation and extracellular multivalent interactions in the tale of galectin-3. Nat Commun 2020;11:1229. https://doi.org/10.1038/s41467-020-15007-3.
[77] Ehrenmann F, Kaas Q, Lefranc M-P. IMGT/3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC. IgSF and MhcSF Nucleic Acids Res 2009;38:D301–7. https://doi.org/10.1093/nar/gkp946.
[78] Kaas Q, Ruiz M, Lefranc M. IMGT/3Dstructure-DB and IMGT/StructuralQuery, a database and a tool for immunoglobulin, T cell receptor and MHC structural data. Nucleic Acids Res 2004;32:D208–10. https://doi.org/10.1093/nar/gkh042.
[79] Dunbar J, Deane CM. ANARCI: antigen receptor numbering and receptor classification. Bioinformatics 2016;32:298–300. https://doi.org/10.1093/bioinformatics/btv552.
[80] Salam NK, Adzhigirey M, Sherman W, Pearlman DA. Structure-based approach to the prediction of disulfide bonds in proteins. Protein Eng Des Sel 2014;27:365–74. https://doi.org/10.1093/protein/gzu017.
[81] Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr Sect D Biological Crystallogr 2011;67:293–302. https://doi.org/10.1107/s0907444911007773.
[82] Kabsch WXDS. Acta Crystallogr Sect D Biological Crystallogr 2010;66:125–32. https://doi.org/10.1107/s0907444909047337.
[83] Tickle IJ, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, et al. STARANISO. Cambridge United Kingdom: Global Phasing Ltd.; 2018.
[84] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr 2007;40:658–74. https://doi.org/10.1107/s0021889807021206.
[85] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr Sect D Biological Crystallogr 2004;60:2126–32. https://doi.org/10.1107/s0907444904019158.
[86] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Sect D Biological Crystallogr 2010;66:213–21. https://doi.org/10.1107/s0907444909052925.
[87] Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Crystallogr Sect D Biological Crystallogr 1997;53:240–55. https://doi.org/10.1107/s0907444996012255.
[88] Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, et al. BUSTER version 2.11.4. Global Phasing Ltd, Cambridge United Kingdom 2011.