Energy-guided combinatorial co-optimization of antibody affinity and stability

Ariel Tennenhouse[1], Adva Mechaly[2], Tony Pham[3], Robert J. Oldham[4], Jake Henry[5], Dorry X. Zhao[3], Thomas M. Moon[3], Isabel G. Elliott[4], Omri Porat[1], Jinny Kim[4], Yael Fridmann Sirkis[6], Sveta Gaiduk[1], Jamie Bird[3], Christopher Domalewski[3], Pei-Hsuan Chu[3], Andrew B. Dippel[3], Rebecca Croasdale-Wood[7], Daniel Christ[5], Gilad Kaplan[3], Mark S. Cragg[4], Sarel J. Fleishman[1*]

1. Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel
2. Department of Infectious Diseases, Israel Institute for Biological Research, 7410001 Ness-Ziona, Israel
3. Biologics Engineering, R&D, AstraZeneca, Gaithersburg, MD, USA
4. Antibody & Vaccine Group, Centre for Cancer Immunology, School of Cancer Sciences, Faculty of Medicine, Southampton General Hospital, University of Southampton, Southampton, SO16 6YD, UK
5. Garvan Institute of Medical Research, Sydney, NSW 2010, Australia
6. Protein Analysis Unit, Department of Life Sciences Core Facilities (LSCF), Weizmann Institute of Science, Rehovot, Israel
7. Biologics Engineering, R&D, AstraZeneca, Cambridge, UK
   *correspondence: sarel@weizmann.ac.il

Abstract:
Affinity maturation is an essential process in antibody engineering. Although powerful, it is iterative, time-consuming, and can result in trade-offs, where affinity is gained at the cost of important properties, such as specificity and stability. We present a scalable strategy, called LAffAb, that starts from a crystallographic structure of the antibody-antigen complex and introduces combinations of mutations to optimize its energy. A combinatorial library comprising 7,000 variants with up to nine mutations results in gains of up to 30-fold in affinity while co-optimizing stability. Surprisingly, the library does not converge on a single solution, instead favoring diverse variants with a high mutational load. Small scale screening of 10 designs against a potential drug target results in an order of magnitude affinity improvement with little loss of specificity or stability. We also apply LAffAb to improve the developability of a therapeutic antibody without degrading affinity. We envision that LAffAb can be used to design stable, specific, and high-affinity binders and to improve our understanding of sequence, structure, and function relationships in antibodies.

Main:
Molecular recognition is critical for many biological functions, with antibodies providing a prime example. Due to their tremendous versatility and compatibility with the human immune system, antibodies have been applied as therapeutics against pathological conditions ranging from cancer to neurological and autoimmune diseases and are ubiquitous reagents in all branches of biomedical research[1–3].

Binding affinity is one of several parameters that require optimization for antibodies to be effective. Their binding surfaces are mostly formed by six solvent-exposed loops called the

complementarity-determining regions (CDRs). The animal immune system improves antigen-binding affinity by introducing random mutations, primarily in the CDRs, and selecting the highest-affinity variants for further randomization in an iterative process called affinity maturation[4]. *In vitro* affinity maturation is currently the dominant strategy, often improving affinity by orders of magnitude relative to the parental molecule[5]. But this process is time-consuming and may result in unacceptable trade-offs in which increased affinity is attained only by compromising other properties, such as stability and specificity[6–8]. Beyond these practical concerns, the fact that a critical step in developing antibodies for basic and applied biomedical research still relies on random processes reflects significant gaps in our understanding and control over antibody sequence, structure, and function.

The primary challenge for reliable design in the CDRs is the complexity of their structure. Antibodies often bind antigenic surfaces lacking defined secondary structure and presenting both polar and hydrophobic groups[9]. To form high-affinity and specific interactions, the CDRs must form a complementary surface that typically lacks secondary structure elements. To maintain these conformations even in the absence of the antigen, the CDRs are stabilized by an intricate network of polar and nonpolar interactions within and between the CDRs and with the framework of the antibody variable fragment (Fv)[10,11]. Given this structural complexity, design calculations must carefully weigh the potential gain from forming new interactions with the antigen against the risks of removing crucial interactions for antibody stability and foldability.

Several computational strategies have been developed to address the limitations of antibody affinity maturation[12–14]. Due to the low likelihood of obtaining beneficial multipoint mutants, these approaches have relied on iterative mutational screening of single mutations. In one study, for instance, single-point mutations that improved electrostatic complementarity with the antigen were experimentally screened, and favorable ones were combined to improve affinity[12]. Recently, large language models have been used to improve affinity. Here too, single-point mutations were experimentally screened before generating combinations[13]. To date, language models have been successful when applied to antibodies for which large sets of mutational effects were available[13,14], and simulations suggest that successful training of AI-driven methods for affinity prediction may require two orders of magnitude more affinity measurements than are available in publicly accessible databases[15]. Thus, design methods have been restricted to iterative screening procedures[16]. In the context of developing an antibody therapeutic, additional biophysical properties, including colloidal stability, low self-association and minimal nonspecific binding (collectively termed “developability”) are essential[17]. Antibody engineering is therefore a complex, multidimensional problem that requires balancing multiple biophysical traits that may be mutually exclusive.

Here, we report a non-iterative structure- and energy-based approach, called LAffAb (**L**ibraries of **Aff**ined **A**nti**b**odies), to design multipoint mutants in antibody CDRs that optimize native-state energy. We first tested the reliability of LAffAb by designing and screening a small combinatorial library of CDR mutants. Deep sequencing showed that most designed mutations were allowed and many combinatorial designs were preferred over the parental antibody. Applied at small scale to an antibody of interest for therapeutic develoment, LAffAb improved affinity by 30 fold.

Finally, we targeted the therapeutic antibody Urelumab, and many designs maintained affinity while improving thermostability and reducing undesirable self-association. Unlike conventional computational and experimental methods, LAffAb does not iterate mutagenesis and screening and produces large and diverse sets of gain-of-function mutants without compromising stability.

Results:

Mutation tolerance mapping in the CDRs

We start the design process by mapping the mutational tolerance of the CDRs based on a crystallographic structure of the antibody-antigen complex (Fig. 1, step 1 and 2). Rosetta design calculations are used to mutate each CDR position to Ala, and positions in which mutations increase system energy are flagged as destabilizing (Fig. 1, step 3). Atomistic energy functions tend to penalize buried polar and electrostatic interactions due to desolvation[18], favoring hydrophobic ones in their place; however, buried polar interactions are crucial for foldability, and mutating them may lead to misfolding and dysfunction[19–21]. To mitigate this risk, we also flag positions as immutable if they contribute substantially to hydrogen-bonding energy (Fig. 1, step 4). Following these restrictions, the design workflow follows the rationale of the FuncLib approach[22], which has been successfully used to optimize obligatory and non-obligatory binding surfaces[20,23,24] and enzyme active sites[22,25]. Briefly, we exclude mutations that are rare in a multiple-sequence alignment (MSA) of antibodies in all regions except CDR H3. Because of its diversity, MSAs of H3 are unlikely to be reliable; instead, we use the per-position residue frequencies inferred by the authors of the HUCAL platinum universal antibody library[26]. High-frequency mutations are individually modeled in the context of the antigen-bound structure and scored in Rosetta, and mutations that substantially increase system energy are triaged (Fig. 1, step 5). Our hypothesis is that the pre-filtering steps based on atomistic calculations and observed frequencies remove mutations that destabilize the antibody or deform its backbone, limiting the effects of long-range epistasis between mutations[27]. As an optional step, a few mutations are included in the design if visual inspection determines that they maintain or improve molecular interactions (Supp. Tables 1-3).

To test the reliability of LAffAb, we designed a small library of variants based on the murine anti-hen egg white lysozyme (HEWL) D44.1 antibody[28], starting from a crystallographic structure of the antigen-binding fragment (Fab) bound to its antigen (PDB ID 1MLC). D44.1 has served as a model system for studying affinity maturation for decades[12,23,29], allowing us to benchmark LAffAb against other approaches. Moreover, this is a high-affinity interaction (K_D of approx. 5 nM)[30,31], and D44.1 underwent animal affinity maturation, making it representative of a typical antibody-engineering task.

Mutation tolerance mapping showed that only 18 of 59 CDR positions could be mutated without destabilizing the CDRs or the antigen-antibody complex. The majority of those (15) did not form direct interactions with the antigen (Fig. 1, step 6). D44.1, therefore, presents a case in which most CDR positions are unlikely to tolerate mutations, particularly in the most critical antigen-binding surface. Nevertheless, we hypothesized that mutations in solvent-exposed parts of the CDRs may contribute to affinity indirectly, by stabilizing the conformations of the CDRs, thus promoting the antigen-bound state as observed in the crystallographic structure.

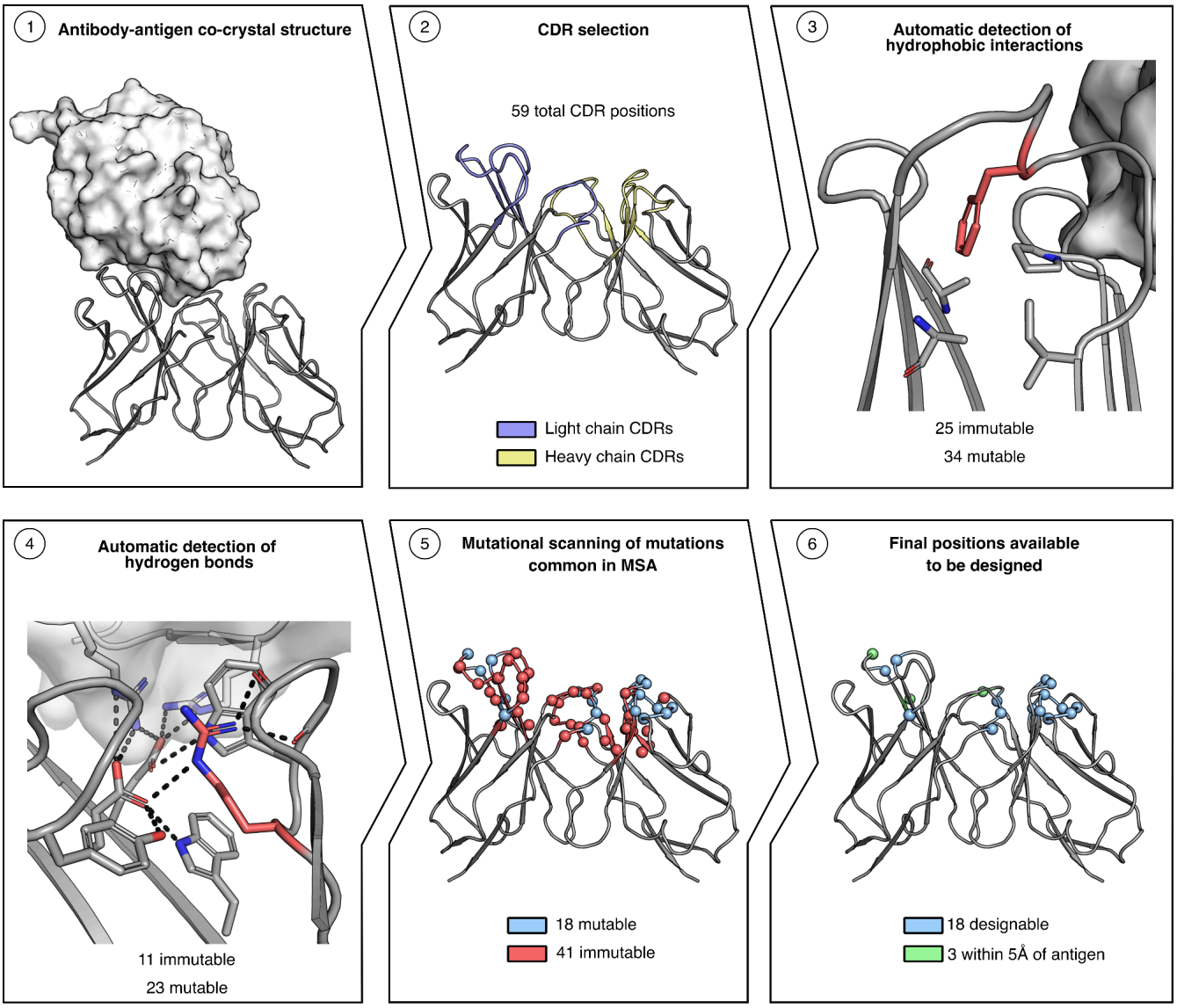

Figure 1: **Computational mutation tolerance mapping in the antibody Fv.** LAffAb begins from a co-crystal structure of an antibody-antigen complex (Step 1). CDR boundaries are automatically determined (Step 2; Supp. Table 4). Atomistic design calculations flag all positions in which mutation to Ala substantially worsens total or hydrogen-bonding energy (Steps 3 and 4, respectively). Then, mutations allowed by phylogenetic screening are modeled at each allowed CDR position, and those that substantially compromise stability are eliminated (Step 5). Mutations were mostly allowed in the periphery of the interface; only three positions in direct contact with the antigen were mutable (Step 6). Models were generated based on the co-crystal structure of D44.1 and HEWL (PDB ID 1MLC ref. 28). In steps 5 and 6, Cα atoms of indicated residues are shown as spheres.

Substantial affinity improvements by one-shot library design

We designed a focused combinatorial library starting from the tolerated sequence space of D44.1. The sequence space of all tolerated mutations at the 18 positions was too large to enumerate (over 1 billion mutation combinations). For computational tractability, we first removed mutations in positions far from the binding surface and ones that introduced noncomplementary charges with the antigen (Supp. Table 1). Only two of the 17 designed point mutations (Asn92$_L$Ala and Ser56$_H$Val) were shared with a previous computational design study

dedicated to improving the affinity of D44.1[12]. A significant challenge for combinatorial-library design is that mutations that are individually tolerated may be destabilizing in combination, in a phenomenon known as negative epistasis[27]. To limit this risk while also minimizing design calculations, we modelled all combinatorial mutants within each CDR separately from the other CDRs (except for L2 and H3, which are proximal and were enumerated in combination), amounting to only 115 atomistic design calculations. Each calculation modeled the combinatorial mutant and iterated sidechain packing and constrained sidechain and backbone minimization. Finally, the low-energy designs from all CDRs were combined, yielding a combinatorial library of 6,912 variants. We verified that >99% of these designs exhibited more favorable Rosetta energy than the parental antibody (Fig. 2A), suggesting that the library would be enriched for stabilized, potentially higher-affinity variants.

We assembled the library into single-chain variable fragments (scFvs) from DNA oligos using GGAssembler[32] and screened them for binding using phage display[33]. After four rounds of panning, 22 individual colonies were sequenced. Surprisingly, we isolated 19 unique designs, none of which was the parental D44.1 (Supp. Table 5). This suggested that despite its small size, the library comprised multiple solutions that were superior to the parent. Furthermore, the selected designs exhibited 4-8 mutations from D44.1 (Supp. Fig. 1A), an unusually large number of mutations in single-shot affinity-maturation campaigns. The designs were assayed for binding to HEWL using monoclonal phage ELISA, and all showed a higher binding signal than D44.1, with some exhibiting dramatic improvements (Supp. Fig. 1B).

We selected four distinct designs with 6-7 mutations from D44.1 for expression as full-length murine IgG2a antibodies (Fig. 2B). After single-step protein A purification, three exhibited purity matching or exceeding that of D44.1 (Supp. Fig. 2A, B). All four exhibited improved expression yields in Chinese hamster ovary (CHO) cells of as much as fourfold (Fig. 2C) and increases in thermal stability relative to D44.1 (Fig. 2D, Supp. Fig. 2C). Finally, surface plasmon resonance (SPR) experiments revealed large gains in affinity, with Des2 exhbiting over 30-fold improvement, reaching 50 pM affinity (Fig. 2E, Supp. Fig. 2D). We concluded that one-shot LAffAb calculations produced diverse variants with many CDR mutations relative to conventional single-shot antibody design and engineering processes, and that the designs exhibited large improvements in affinity together with gains in stability.

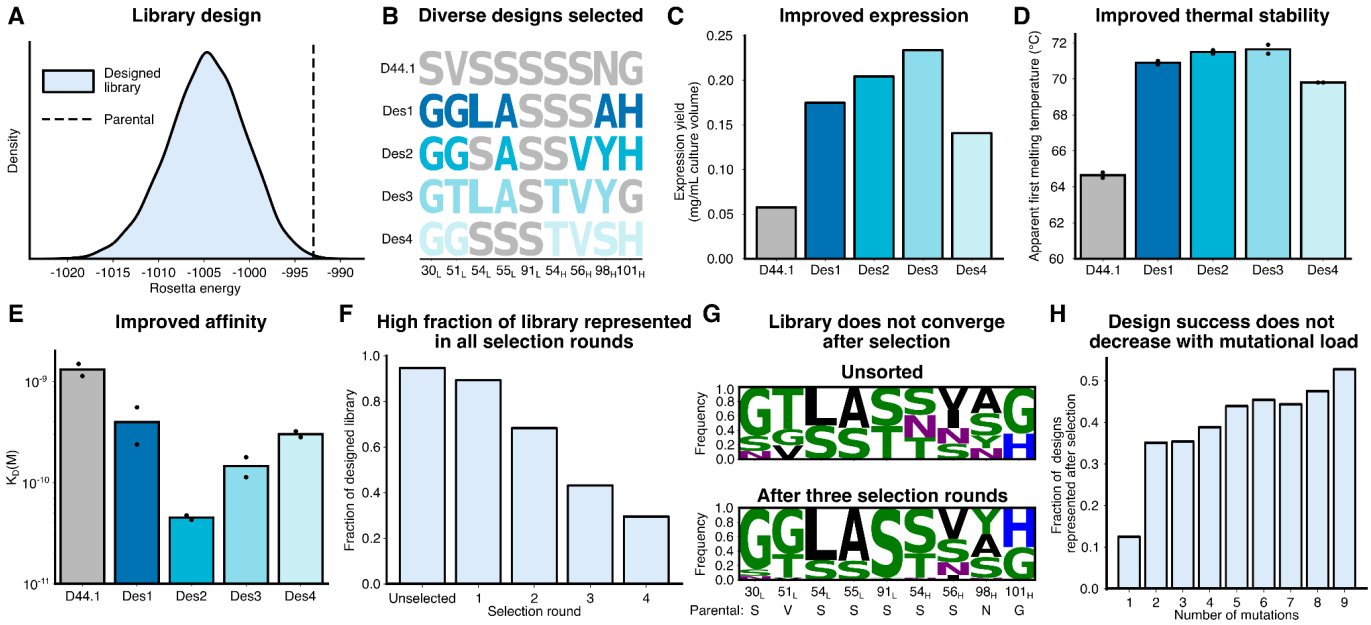

Figure 2: Design and selection of D44.1 variants that co-optimize affinity and stability. (**A**) Rosetta energy distribution of the designed library comprising 6,912 variants. Over 99% of the designed library exhibits improved Rosetta energy compared to D44.1. (**B**) Mutations at each designed position in D44.1. (**C**) Expression yield (corrected by purity) of antibodies formatted as mouse IgG2a. Based on a single measurement. (**D**) Apparent first melting temperatures measured by nano differential scanning fluorimetry (nanoDSF). (**E**) Dissociation constants (K_D) measured by SPR. (**D, E**) Bars represent the mean of two technical replicates (measurements plotted as dots). (**F**) Fraction of the designed library present in at least one read after each phage-selection round. (**G**) Sequence logoplots of the population as a function of the selection round. (**H**) Fraction of the designed library observed in selection round 3 by number of mutations.

<u>Distinct high-affinity solutions revealed by deep sequencing</u>

We applied deep sequencing analysis to understand why so many different multipoint mutants were selected in the designed library. Even after four selection rounds, 30% of the designed sequences exhibited at least one read (Fig. 2F), though the parental D44.1 was purged by the third round. After the third selection round, most designed mutations were allowed, and some were vastly preferred relative to the parental identity (Fig. 2G, Supp. Table 6). Deep sequencing further showed that the fraction of designs with at least one read after three selection rounds showed no decrease as a function of mutational load (Fig. 2H). The results strikingly demonstrate that selecting tolerated point mutations according to energy and structural criteria dramatically reduces the risk of negative epistasis in combinations of individually favorable mutations. They further suggest that the space and diversity of beneficial multipoint mutants in CDRs may be very large.

Given that a large fraction of the library apparently exhibited improved affinity, we sought to understand which individual mutations failed to improve affinity. We calculated the frequency of each mutation in the selected versus unselected populations compared to the parental identity (Supp. Table 6). Ten of the 16 designed single mutations had a normalized enrichment greater than one, suggesting they were beneficial compared to the parental identity. Three others had a normalized enrichment over 0.5, suggesting they were not strongly purged. Only three mutations were strongly counter-selected: Ser91$_L$Thr, Ser54$_H$Asn, and Ser56$_H$Ile (Fig. 2G, Supp. Table 6).

We evaluated the Rosetta system energy of all single-point mutations, but they did not correlate with normalized enrichment (Supp. Fig. 3A), excluding this computed metric as a predictor. We therefore analyzed finer Rosetta energy-based filters for clues to explain the negative impact of the counter-selected mutations. Ser91$_L$Thr was in a position flagged by the computational alanine scan, because it weakens a hydrogen bond to light chain His34 (Fig. 3A, right panel). We added this deleterious mutation to test how sensitive the selection process is to conservative mutations. Structural analysis suggested that the other two deleterious mutations each introduced strain to the local structure that is relieved in modeling through a subtle backbone change (0.4 and 0.6 Å between Cα atoms of the parental structure and design model of the single-point mutant) or a change in the conformation of the antigen (Fig. 3A, left panels). By contrast, the 13 favorable or neutral mutations did not abrogate hydrogen bonds, and all mutations with enrichment better than the parental identity (Supp. Table 6) did not alter the backbone by more than 0.4 Å (Supp. Table 7, Supp. Fig. 3B). Surprisingly, whereas Ser56$_H$Ile was strongly counterselected, Ser56$_H$Val was tolerated. The dramatic influence of a single methyl emphasizes the need for accurate atomistic design methods (compare Fig. 3A bottom left panel and Fig. 3B left panel). Additionally, two neighboring small-to-large mutations, Asn98$_H$Tyr, and Gly101$_H$His, may stabilize the backbone of CDR H3 by forming additional interactions with nearby positions and by reducing the desolvation penalty of the partly buried Asn in the parental D44.1 (Fig. 3B bottom right panel). Interestingly, Gly101$_H$His was not strongly favored as an individual mutation according to the deep sequencing results (Fig. 2G, Supp. Table 6), highlighting the potential of combining several weakly enhancing mutations to generate large gains in affinity.

We concluded that, in addition to selecting mutations that exhibit low energy, CDR design calculations should avoid mutations predicted to distort the local backbone structure, cause a change in the conformation of the antigen, or remove internally stabilizing hydrogen bonds. These structure- and energy-based rules can be applied to any antibody-antigen co-crystal complex to dramatically enrich for beneficial combinations of mutations. Indeed, with the benefit of hindsight, had we excluded the three deleterious mutations, the size of the library would have been reduced to 1,728 variants, of which 1,381 (80%) were observed in deep sequencing after three selection rounds.

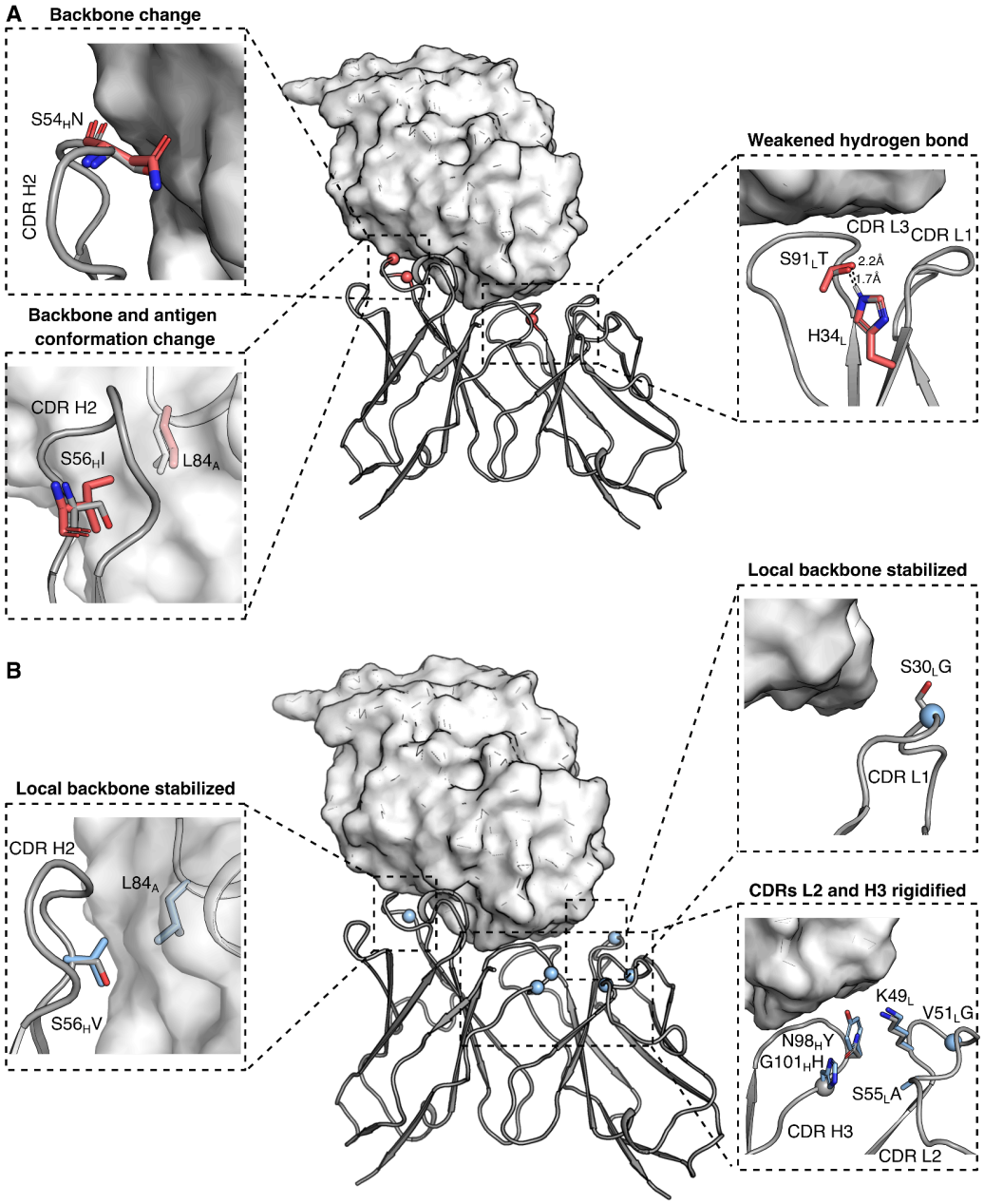

Figure 3: Structural underpinnings of favorable and deleterious mutations. (**A**) Ser91$_L$Thr (right panel) weakens a hydrogen bond with His34$_L$. Ser54$_H$Asn (top left panel) shifts the position of the main chain backbone atoms, potentially introducing strain into CDR H2. Similarly, Ser56$_H$Ile (bottom-left panel) changes the position of the main chain backbone atoms. Furthermore, the mutation forces a change in the conformation of Leu84 in HEWL to accommodate the larger residue. (**B**) Ser56$_H$Val likely stabilizes CDR H2 because of its β-branched side chain (left panel). A cluster of small-to-large mutations likely stabilizes CDR H3, particularly by improving interactions with Lys49 of CDR L2. (**A and B**) The middle structure shows the Rosetta-relaxed model structure of the D44.1 Fv in complex with HEWL shown as a surface map. The Cα atoms of the mutations are shown in red (deleterious mutations) or blue (favorable) spheres in the middle panel. Inset panels show the Rosetta-relaxed model in gray and mutations in colors. Spheres in the inset panels represent glycines. Subscript L denotes light chain, subscript H denotes heavy chain, and subscript A denotes antigen (HEWL).

Large gains in affinity through low-throughput screening

We next tested whether small-scale application of LAffAb to an unrelated antibody-antigen complex would also improve affinity. We applied the design process described above to the 6G08 antibody, which targets the sole inhibitory crystallizable fragment (Fc) receptor, Fc gamma Receptor IIb (FcγRIIb)[34], an attractive drug target due to its role in regulating the immune response[34]. However, for some applications, antibodies targeting FcγRIIb must also avoid binding to FcγRIIa, a key activating FcγR with 93% sequence identity to FcγRIIb in the extracellular domain[35] (Fig. 4A). Antibodies fulfilling these criteria were discovered using phage display, and in co-administration experiments, these increased the effectiveness of target-cell-deleting antibodies in mouse models[34]. Likely due to the harsh requirement to avoid binding to the homologous FcγRIIa during selection, the resulting affinities were low, and one specific antibody, 6G08, had particularly low affinity. 6G08, therefore, presented a stringent test of the ability of LAffAb to improve affinity without compromising specificity.

We applied computational mutation tolerance mapping to the co-crystal structure of 6G08 bound to FcγRIIb[36] (PDB ID 5OCC). Unlike the D44.1-HEWL complex, our calculations suggested many stabilizing mutations in the antibody-antigen interface. We speculate that the reason for the difference between the two antibodies is that 6G08 had not undergone affinity maturation[34]. We enumerated all combinations of point mutations in the CDRs and selected ten designs with six mutations relative to 6G08 and 3-10 mutations relative to one another (Supp. Table 8). Some positions were flagged by the computational alanine scan, but were included because they improved the system and/or hydrogen-bonding energy (Supp. Table 2).

Eight of the ten designs bound FcγRIIb and two exhibited much improved affinity of up to 30-fold (Fig. 4B, Supp. Fig. 4A). The two designs had four mutations relative to each other, suggesting that LAffAb was again able to produce distinct high-affinity solutions (Fig. 4C). Structural modeling showed that mutations in the two successful designs were predicted to form new interactions with the antigen, fill cavities in the antibody-antigen interface, and stabilize the CDR backbones (Fig. 4D).

To verify that the designs function in a physiological setting, they were assayed for binding to FcγRIIb-expressing CD19+ B cells within isolated peripheral blood mononuclear cells (PBMCs) by flow cytometry (Fig. 4E), and both showed greatly improved binding relative to 6G08. Additionally, the designs were tested for binding to CHO-K1 cells overexpressing the homologous off-target FcγRIIa (Fig. 4F), and both showed only weak binding, verifying that specificity was mostly maintained. Furthermore, both designs saturated binding to the CD19+ B cells at 1 μg/mL, and at this concentration, both showed very weak binding to the CHO-K1 cells, suggesting there is a significant window in which the designs can be given without detectable binding to FcγRIIa, even when FcγRIIa is overexpressed. The two designs with improved affinity maintained thermal stability (Supp. Fig. 4B and C) and over 80% of the expression titer when expressed in CHO cells (Supp. Fig. 4D). Thus, the affinity of 6G08 for its intended target was enhanced without compromising stability and while mostly maintaining specificity with respect to a very close, evolutionarily conserved homolog of its target.

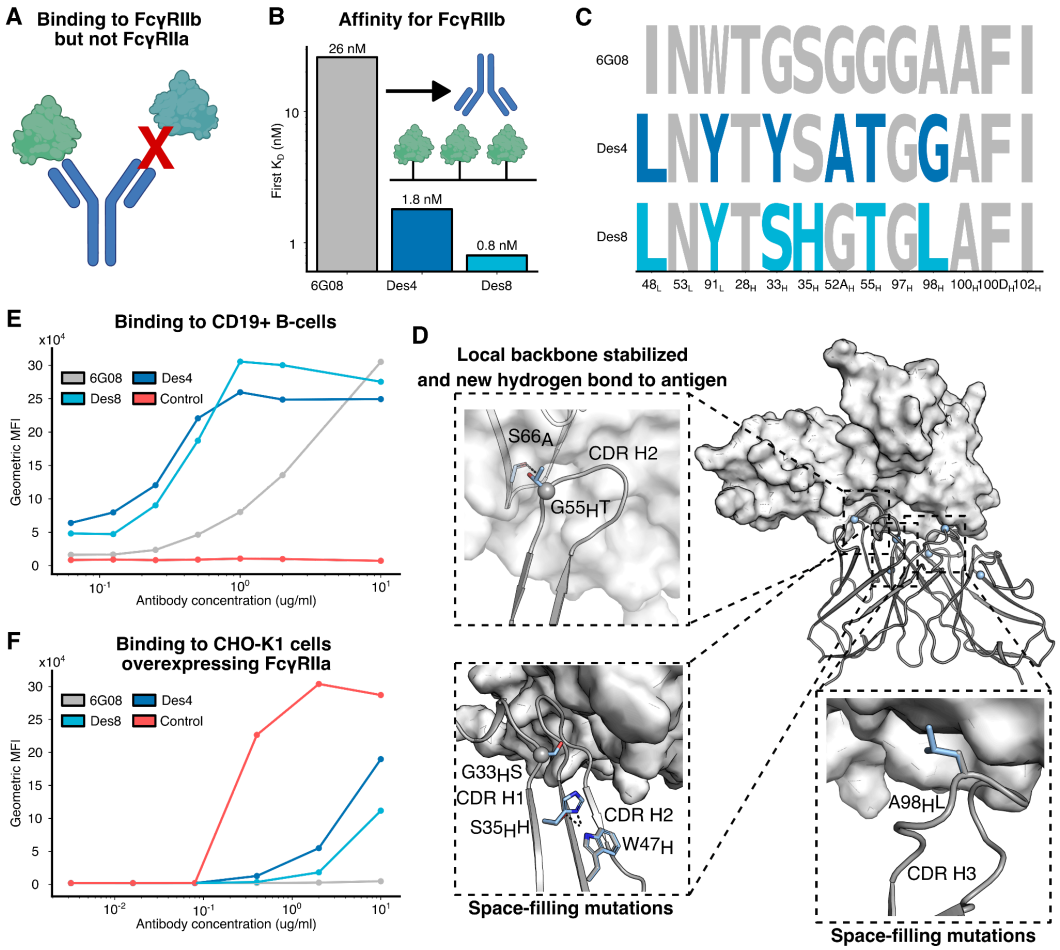

Figure 4: Improving affinity through small-scale experimental screening of antibody 6G08. (**A**) The desired outcome of the design process is high-affinity binding to FcγRIIb and not its close homolog FcγRIIa. Scheme made with BioRender. (**B**) Dissociation constants measured using SPR with FcγRIIb immobilized and antibodies flown over the chip. Inset scheme made using BioRender. Based on a single measurement. (**C**) Mutations at each designed position in 6G08 and the two successful designs. (**D**) Structural reasoning for affinity-enhancing mutations in Des8. Gly55$_H$Thr likely stabilizes CDR H2 because of its β-branched side chain and makes a hydrogen bond with the backbone of Ser66 in FcγRIIb (top left panel). Two small-to-large mutations likely fill a hole in the antibody-antigen interface. Ser35$_H$His likely strengthens a hydrogen bond to Trp47$_H$ (bottom left panel). Ala98$_H$Leu fills a cavity in the antibody-antigen interface (right panel). The middle structure shows the Rosetta-relaxed model structure of the 6G08 Fv in complex with FcγRIIb shown as a surface map. The Cα atoms of the mutations in Des8 are shown in blue spheres in the middle panel. Inset panels show the Rosetta-relaxed model in gray and mutations in colors. Spheres in the inset panels represent glycines. Subscript L denotes light chain, subscript H denotes heavy chain, and subscript A denotes antigen (FcγRIIb). (**E**) Binding to B-cells isolated from PBMCs. B-cells naturally express FcγRIIb. (**F**) Binding to FcγRIIa using CHO-K1 cells overexpressing FcγRIIa. (**E and F**) Two independent experiments were performed using either directly conjugated reagents or detection with a fluorescently labelled secondary antibody. The same trend was observed in both experiments. The results using secondary detection are presented.

Developability improvement of a therapeutic antibody

The large space of high-affinity and diverse variants LAffAb generates presents an opportunity to select designs predicted to maintain affinity while improving other desirable properties. We applied LAffAb to Urelumab, a high-affinity therapeutic antibody that targets the immune receptor 4-1BB (K_D 8 nM) and is known to exhibit high reversible self-association[17]. Self-association and nonspecific binding are correlated with low and high Fv pI, respectively, and the pI of Urelumab is 5.9, outside the desirable range (6.5-8.5)[37]. We selected designs with pI in the desirable range (Supp. Table 10), and as a control, we also selected high-energy designs comprising mutations that passed the initial filtering.

We tested 38 low- and 14 high-energy designs for binding to 4-1BB and in several developability assays measuring aggregation, non-specific binding, and reversible self-association (Supp. Table 11). The low-energy designs had significantly higher thermal stability than the high-energy ones, but most other criteria were indistinguishable among these two groups (Supp. Fig. 5). 24 low-energy designs maintained affinity within twofold of Urelumab (Fig. 5A, Supp. Fig. 6), and nearly all (36) showed improvements in thermal stability (Fig. 5B). We classified each design as passing or failing developability based on resistance to aggregation, reversible self-association, and non-specific binding (Supp. Fig. 7). 17 passed all assays and six of these exhibited comparable affinity to Urelumab (Fig. 5A). The six high-affinity designs that passed all developability tests were diverse, comprising 4-6 mutations from Urelumab and at least two mutations relative to one another (Fig. 5D).

We analyzed the individual mutations that comprised the designs to identify mutations leading to improvements or losses in each assay. Asp100_BAsn in the heavy chain significantly improved stability and developability compared to designs containing the parental Asp but compromised affinity (Fig. 5A-C). Asp100_BTyr improved stability and developability without compromising affinity (Fig. 5A-C). In both cases, the improvements in developability were driven by significant reduction in reversible self-association and nonspecific binding (Supp. Figs. 7B and C).

Ser30Lys in the heavy chain significantly improved aggregation resistance while maintaining affinity (Fig. 5A and Supp. Fig 7A), but increased reversible self-association (Supp. Fig. 7B) and nonspecific binding (Supp. Figs. 7C, D). Despite these liabilities in designs containing Lys30, four of six high-affinity and developable designs contained Lys30 (Fig. 5D). This suggests that other mutations were able to compensate for the non-specific binding and self-association seen with Lys30 while maintaining the benefits in aggregation resistance.

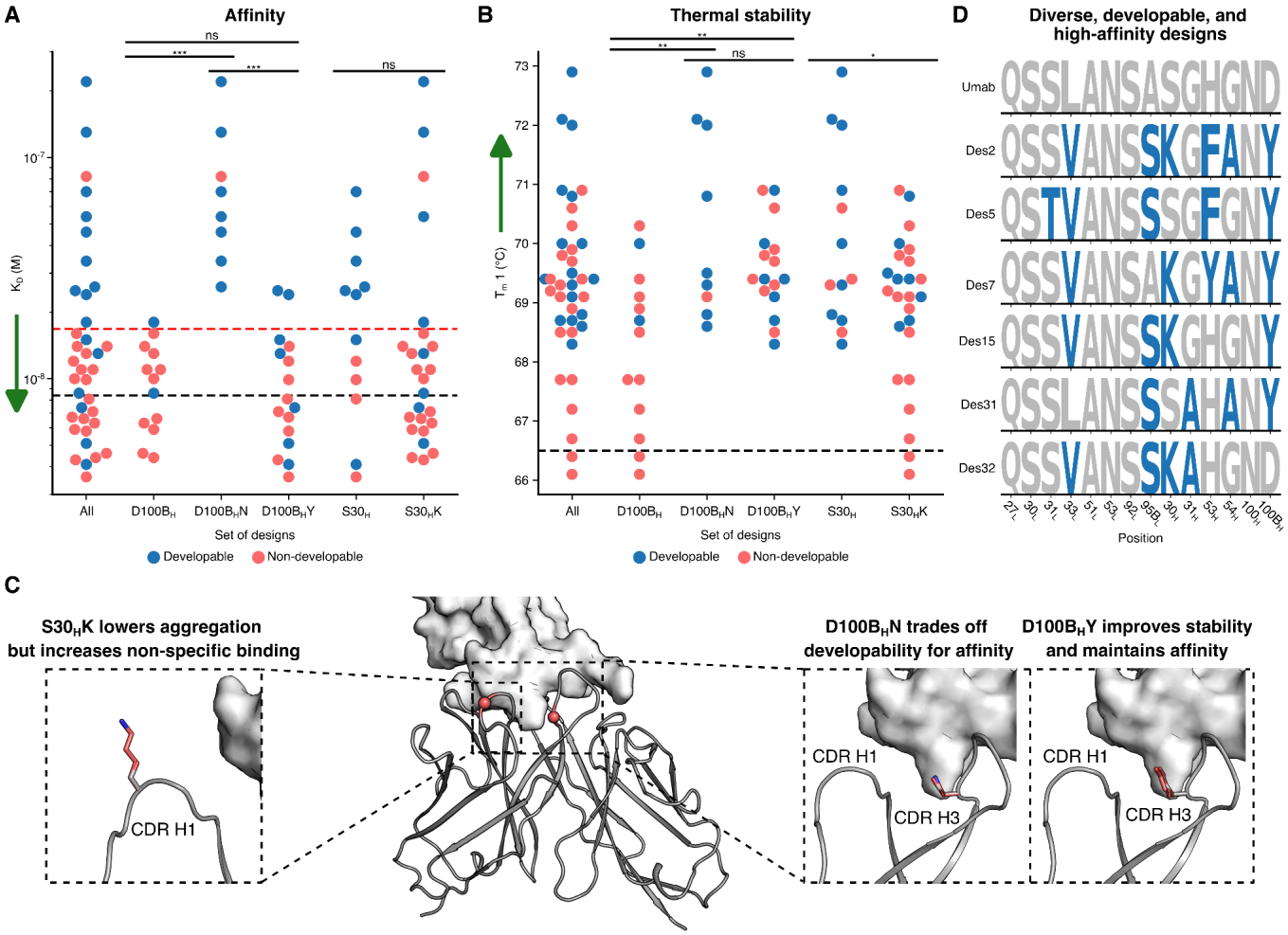

Figure 5: Improving developability through medium-scale experimental screening of Urelumab. (**A**) Affinities of different design subsets measured by SPR. Red dashed line represents a twofold decrease in the affinity of Urelumab. Sensorgrams shown in Supp. Fig. 6. (**B**) Thermal stability of different design subsets measured by nano differential scanning fluorimetry (nanoDSF). (**A,B**) Black dashed line represents the value of Urelumab. ***, **, *: *p*-value <= 0.001, 0.01, 0.05, respectively, calculated using an independent *t*-test. Dots represent the means of three technical repeats. Green arrows represent the preferred direction in each assay. (**C**) Structural reasoning for developability effects of selected mutations. The middle structure shows the Rosetta-relaxed model structure of the Urelumab Fv in complex with 4-1BB shown as a surface map. The Cα atoms of the selected mutations are shown in red spheres in the middle panel. Inset panels show the Rosetta-relaxed model in gray and mutations in red. (**D**) Mutations at each designed position in Urelumab (Umab) and the six high-affinity, developable designs. (**A-D**) Subscript H denotes heavy chain.

Discussion:

Mutation tolerance mapping is key to the success of LAffAb, ensuring that the mutations are individually allowed and retain the CDR backbone conformation. Maintaining the backbone conformation is crucial to enable combining multiple mutations that do not antagonize one another but is complicated by the irregularity and sensitivity of the CDRs to mutations. The careful mutation tolerance mapping in LAffAb identifies a “smooth” local sequence space where mutations are mostly structurally and energetically compatible and complementary with one another, limiting negative epistasis[27]. Thus, LAffAb is distinguished from previous natural, experimental, and computational maturation methods that prioritize single mutations without explicitly assessing their impact on the structure. Furthermore, by optimizing the energy of the

antigen-bound structure subject to constraints that favor foldability, our approach, in effect, co-optimizes affinity and stability.

Our results demonstrate that CDR positions may contribute to affinity, even if they do not interact directly with the antigen. This is important because interfacial positions are typically less tolerant to mutation. The three case studies presented very different scenarios: in D44.1, almost no antigen-binding position was mutable; in 6G08, both high-affinity designs contained interfacial mutations; and in Urelumab many designs contained interfacial mutations. Furthermore, we applied LAffAb to design 12 variants of the cetuximab antibody[38], which binds epidermal growth factor receptor (EGFR), and all designs exhibited significant decreases in binding affinity. We speculate that this failure resulted from abrogating an elaborate network of ordered water molecules surrounding CDR H3 (Supp. Fig. 8). The crystal structure originally used for design (PDB entry: 6ARU) did not show these water molecules whereas a higher resolution structure (PDB entry: 1YY9[38]) did. Our study thus highlights the need for accurate bound structures to predict mutational tolerance. It further highlights the remarkable structural diversity in antibody-antigen complexes, and how different structures impose dramatically different mutational constraints.

For decades, antibody affinity optimization has been a critical step in antibody engineering that relied on iterative maturation processes[39]. LAffAb is the first approach to our knowledge that enables directly optimizing affinity and other important biophysical properties without iterations. Notably, our study highlights the effectiveness of using structure and energy-based criteria to generate designs and to explain their success and failure. We envision that mutational tolerance mapping can be further improved by leveraging experimental data accumulated in academia and pharma companies[40], thereby developing general ML/AI methods based on energy criteria for improving antibody design methodology. The high pace of advances in antibody modeling and design methods in recent years are encouraging that reliable and automated methods for one-shot optimization of antibody affinity and developability may be within reach.

Methods:

Computational methods:

All Rosetta design simulations used git v.d9d4d5dd3fd516db1ad41b302d147ca0ccd78abd of the Rosetta biomolecular modeling software, which is freely available to academics at http://www.rosettacommons.org. All RosettaScripts[41] XMLs can be found at https://github.com/Fleishman-Lab/LAffAb_public.

Selecting a set of allowed CDR mutations:

We defined the CDRs once for all antibodies in this study based on previous definitions[30,42,43] and visual inspection. CDR definitions can be found in Supp. Table 4. For each antibody, a position-specific scoring matrix (PSSM) was computed as described previously[44] for the light and heavy chain variable regions separately. Mutations were allowed at each CDR position outside of CDR H3 if they had a PSSM score >= -2. Within CDR H3, mutations were allowed

from position 95-100 if they appear in human CDR H3s at a given position >=4% of the time based on per-position residue frequency data[26]. All mutations were allowed in positions 101 and 102.

As a first step, the antibody Fv structure in complex with the antigen was relaxed through cycles of sidechain and harmonically restrained backbone minimization and combinatorial sidechain packing in the entire Fv (see xmls/Relax.xml). Each CDR position was then mutated to alanine and scored using the REF15[18] score function (see xmls/alascan.xml) and the REF15 score function with the weights of all terms not related to hydrogen bonding set to zero (see xmls/alascan_Hbond.xml). Positions were considered designable if the mutation to alanine had $\Delta\Delta G$ <=3.5 REU above the parental using the REF15 score function and <= 0.6 REU using the hydrogen bond only score function. Each mutation allowed by the PSSM or CDR H3 residue frequency data was then individually modeled and scored in Rosetta (see xmls/filterscan.xml). Mutations were also modeled at positions flagged as making critical interactions, and individual mutations were allowed in these positions only if they maintained or improved these interactions (by visual inspection). For D44.1, only Ser91Thr in the light chain was allowed in a flagged position (Supp. Table 1). For 6G08, Trp91Tyr in the light chain, Ser35His in the heavy chain, $Phe100_{D}Leu$ in the heavy chain, and Ile102Val in the heavy chain were allowed in flagged positions (Supp. Table 2). For Urelumab, no mutations were allowed in flagged positions (Supp. Table 3).

<u>Selecting combinatorial CDR mutants:</u>
Before combinatorial enumeration, completely solvent-exposed mutations, mutations far from the interface, and mutations that introduced unmatched charges were removed. (Supp. Tables 1-3). For D44.1, all combinations of allowed mutations within each CDR were enumerated. CDR L2 and CDR H3 were enumerated together because of their structural proximity. Each mutant was modeled in Rosetta, including combinatorial sidechain packing, and the backbone and sidechains of all residues were minimized, subject to harmonic restraints on the Cα coordinates of the entire protein (xmls/combinatorial_enumeration.xml). Each design was labeled as positive if the all-atom energy was more favorable than that of the parental antibody and negative otherwise. Mutations were removed if the variant for that mutation had $\Delta\Delta G$ > 1 REU.

For D44.1, an EpiNNet[45,46] neural network model was then trained to predict which designs exhibited lower-energy than the parental antibody-antigen complex for each CDR (or L2 and H3 together). In short, a multi-layer perceptron classifier with a single hidden layer the size of the number of selected positions was trained on all the data. Each classifier was trained for 20,000 iterations on a one-hot encoded representation of the sequence data to classify whether a design exhibited lower energy than the parental antibody. Next, the single-point mutations were ranked according to their probability of being lower-energy than the parent. Mutations were selected for the library by iteratively adding the top-ranked mutations until the fraction of designs exhibiting lower energy than the parental antibody-antigen complex significantly dropped. All combinations of the selected mutations for each CDR were then enumerated. Each mutant was modeled in Rosetta following the same procedure.

For 6G08, all combinations of allowed mutations within the entire Fv were enumerated. Each mutant was modeled in Rosetta, including combinatorial sidechain packing, and the backbone and sidechains of all residues were minimized, subject to harmonic restraints on the Cα coordinates of the entire protein (see xmls/combinatorial_enumeration.xml). The designs were ranked by all-atom energy. Only designs with fewer than seven mutations were considered. For 6G08, positions in the antibody-antigen interface were defined using the InterfaceByVector Residue Selector in Rosetta on the antibody-antigen crystal structure (see xmls/interface.xml). Asn53 in the light chain and Gly50 in the heavy chain were manually removed from the description of the interface despite being included in the positions selected by InterfaceByVector. Designs were clustered so that each design had at least three interface mutations from all other designs, and the top ten clustered designs ranked by total score were selected. For Urelumab, the pI of the Fv of each design was calculated using the ProtParam[47] module of BioPython[48], and only designs with pI greater than 6.5 and less than 8.5 were considered. Designs that introduce N-glycosylation sites (NX[ST] where X is not Pro) were purged. 44 positive designs and 15 negative designs were selected. For each, the designs were ranked according to Rosetta energy (with low-energy designs at the top of the list for positive designs and the bottom for negative designs). Four (one) designs with six mutations, ten (three) with five mutations, and thirty (11) with four mutations were chosen for the positive and negative pools. Within each group, designs were clustered such that there were at least three, two, and two mutations between designs. After design selection, if any positions had less than one third of designs containing the parental identity, 0.1 REU was subtracted from all designs containing the parental identity in these positions and designs were selected again. This process was iterated until all parental identities were present in at least one third of the designs.

Analysis of Nanopore sequencing:

The pod5 files output from nanopore were basecalled using the Dorado duplex basecaller with the super-accurate basecalling model. The resulting basecalled reads were then filtered to remove simplex reads that were already accounted for as duplex reads using SAMtools view[49]. The filtered reads were then demultiplexed using Dorado demux, and the adaptors were trimmed using Dorado trim.

Each read was aligned to a FASTA sequence of each of the constant segments between designed variable regions (six constant segments total) using MiniMap2[50] with a minimizer kmer length of 21 (-k 21) and a minimizer window size of 11 (-w 11). The alignments were filtered to exclude unmapped reads and non-primary alignments using SAMtools view[49] (-F 3844). PySam[49] was used to iterate through the alignments for each database and filter the reads based on several properties. The query for each alignment needed to be over 850 base pairs and less than 900 base pairs to be considered. Additionally, the query for each alignment needed to have an average PhRed[51] score >= 20. Each alignment needed to be less than five base pairs shorter than the length of each constant segment and less than five base pairs longer than the length of each constant segment. The designed variable segments of each read were extracted based on the positions of the corresponding constant segments from the alignments. Variable segments with a high probability (>= 0.9) of having no errors based on the

per-base pair PhRed scores were searched against a database of the designed variable segments, and reads with a perfect match to each designed variable segment were considered.

Analysis of individual mutations based on Nanopore sequencing data:
The frequencies of each identity at each position at each selection round were calculated as follows:

$$frequency\ (mutation_i,\ position_j,\ round_z) = \frac{number\ of\ reads\ containing\ mutation_i\ at\ position_j\ at\ round_z}{number\ of\ total\ reads\ at\ round\ z}$$

Enrichment was calculated as follows:

$$enrichment\ (mutation_i,\ position_j,\ round_z) = \frac{frequency\ (mutation_i, position_j, round_z)}{frequency\ (mutation_i, position_j, round_0)}$$

Normalized enrichment was calculated as follows:

$$normalized\ enrichment\ (mutation_i,\ position_j,\ round_z) = \frac{enrichment\ (mutation_i, position_j, round_z)}{frequency\ (parental\ identity,\ position_j, round_0)}$$

These frequencies were used to make the logoplot in Fig. 2G and the frequencies and enrichments are summarized in Supp. Table 6.

Mutational analysis of Urelumab designs:
An EpiNNet[45,46] neural network model was trained to predict the value in each experiment for each design. In short, a multi-layer perceptron regressor with a single hidden layer the size of the number of selected positions was trained for 20,000 iterations on a one-hot encoded representation of the sequence data. Next, each pair of mutations was ranked by the difference in their predicted values for each experiment. Positions with large differences were flagged for visual inspection.

Antibody numbering:
Antibody numbering was performed using the AbNum[52] web server (http://www.bioinf.org.uk/abs/abnum/). Kabat numbering[53] is used throughout the work.

Distance calculations:
Distance calculations were performed using PyMOL v2.3.5.

Data visualization:
Visualization of protein structures was done using PyMOL v3.1.0. Protein surface maps were generated using the APBS[54] plug-in in PyMOL. Data visualization was performed using matplotlib v3.1.0 and seaborn v0.9.0. Data was organized using pandas v0.24.2. Logo plots were created using Logomaker[55].

Significance tests:
Significance tests were performed using the ttest_ind function in SciPy.

Experimental methods:

Assembly of DNA encoding for the D44.1 library:

GGAssembler[32] was used to optimize a set of Golden-Gate[56] overhangs to minimize the total cost of ordered oligos required to cover all mutations in the library without introducing unwanted mutations. This resulted in several variable and constant segments, with and without mutations. A gene encoding the D44.1 sequence formatted as a single-chain variable fragment was ordered from Twist BioScience. Constant segments were PCR-amplified from the gene with primers adding *BsaI* recognition sites. These and all other DNA fragments were purified using the Monarch® Spin PCR & DNA Cleanup Kit (NEB T1130). Variable segments were ordered as single-stranded degenerate oligos from IDT. The single-stranded oligos were double-stranded by a short PCR (KAPA HiFi HotStart ReadyMix; Roche 07958935001) with a single primer and purified using the Monarch kit. The concentration of each segment was measured using NanoDrop One (Thermo Scientific), and TapeStation was used to assess the purity and size of each segment. A Golden-Gate assembly was conducted using the manufacturer's specifications for the light and heavy chains separately. Briefly, all segments are added at an equal amount, without the vector, and assembled using T4-ligase and BsaI-HF-v2 using cycles of 16°C and 37°C (New-England Biolabs, M202 and R3733, respectively). The resulting assemblies were PCR amplified using the KAPA HiFi HotStart ReadyMix (Roche Diagnostics 07958927001) and then combined using overlap-extension PCR, which also added the linker connecting the heavy and light chains. The final product was used to construct a phage-display library as described previously[57]. Briefly, the scFv fragments and library plasmid[58] were digested with *NcoI-HF (*NEB R3193) and *NotI-HF* (NEB R3189) restriction enzymes and ligated with T4 DNA ligase (NEB M0202). Transformation of the desalted ligation into *E. coli* TG1 electro-competent cells (Lucigen 60502-1), led to $4x10^7$ transformants, greater than 1,000-fold coverage of the 7,000 variant library.

Phage display selections of the D44.1 library against HEWL:

Panning of the phage display library against plate-coated HEWL (3 µg/ml, Sigma 62970) was carried out as described[57]. Four panning cycles were performed. The buffer was alternated between each cycle between 2% skim milk (Sigma 70166) and 2% Bovine serum albumin (Sigma A9418) in 1X PBS (Sartorius 02-023-5A) with 0.05% Tween 20 (Sigma P1379). The number of input phage at each cycle decreased ($1x10^{11}$ for the first cycle, $5x10^{10}$ for the second cycle, and $1x10^{10}$ for the third and fourth cycles), and the number of washes increased with each cycle (from three in the first cycle to six for the fourth cycle). In an attempt to improve the k_{off} of the resultant antibodies, a parallel panning regime was carried out from the second panning cycle onward, where the last wash of the bound phages in each cycle was elongated (30, 60 and 120 minutes for the second, third and fourth cycles, respectively) and soluble HEWL (1, 1.5 and 2µM) was added.

Monoclonal phage ELISA:

Direct ELISA was performed as described previously[59]. In short, HEWL (Sigma 62970) was used to coat (3 µg/ml) maxisorb nunc-immuno plates (Thermo 44204). Equilibrated, titrated phage were added to the blocked plate, washed, and bound phage were detected with

HRP-conjugated anti-M13 antibody (Sino Biological 11973) followed by TMB substrate (Millipore ES001).

Antibody production:

D44.1 and its designs:
The anti-HEWL antibodies were produced by GenScript. The amino acid sequences for the light and heavy chains of D44.1 and the four designs were back-translated into DNA and codon optimized for expression in CHO cells using GeneSmart. The DNA was synthesized and cloned into an optimized murine IgG2a expression vector for both the light and heavy chains. The plasmids were maxiprepped to yield transfection-grade plasmids using GenScript Proprietary in-house columns and reagents. The plasmids were then transiently co-transfected at a 1:1 ratio into suspension CHO-Express cell cultures. Briefly, suspension CHO cells were grown in serum-free expression medium containing GenScript proprietary in-house buffers and supplements in Erlenmeyer Flasks at 37°C with 5% CO_2 on an orbital shaker. One day before transfection, the cells were seeded at 3-6x10^6 cells/mL in an Erlenmeyer flask. On the day of transfection, DNA and transfection reagent were mixed at an optimal ratio and then added to the flask with cells ready for transfection. The cell culture supernatants were collected by centrifugation at 3,000 rpm for 30 min when the cell viability was less than 50%. The supernatants were filtered and loaded onto AmMag Protein A Magnetic Beads (GenScript L00695) and incubated for the appropriate time. The beads were washed and eluted using GenScript Proprietary in-house columns and reagents. The eluted fractions were pooled and buffer exchanged to 1X PBS pH 7.2. The purified protein was analyzed by SDS-PAGE analysis using a 4%~20% gradient SDS-PAGE gel (GenScript M42012) to determine the molecular weight and purity. Each sample was run in reducing sample buffer (60 mM Tris-HCl, 2% SDS, 6% Glycerol, 0.1% bromophenol blue, 50 mM DTT, pH 6.8) and non-reducing sample buffer (50 mM Tris-HCl, 1% SDS, 6% Glycerol, 0.1% bromophenol blue, 50 mM NEM, pH 6.8). The endotoxin level of protein was also detected using a Kinetic Turbidimetric Assay (Bioendo). The concentration was determined by measuring A280 using a NanoDrop Ultra (ThermoScientific). The final yield was calculated by multiplying the concentration by the total volume of the eluate. The final yield was then corrected by the purity measured using SDS-PAGE.

6G08 and its designs:
The variable heavy and light chain sequences of the antibodies were cloned into pEE6.4 expression vectors (Lonza) containing human IgG1 constant domains. Plasmid DNA was prepared using a DNA miniprep kit (Qiagen), and constructs were verified by Sanger sequencing (Source BioScience). ExpiCHO cells (Thermo Fisher) were cultured at 37°C with 8% CO2 and shaking at 125 rpm. Transient transfections were performed at a density of 6.0 X 10^6 cells/ml following the manufacturer's protocol, and cultures were harvested 7 days post-transfection. Antibodies were purified using a Protein A MabSelect SuRe column (Cytiva) on an AKTA Start chromatography system. Purity and integrity were confirmed by gel electrophoresis and HPLC, and endotoxin levels were measured for downstream applications.

Urelumab and its designs:
The heavy- and light-chain variable domains of Urelumab and its designs were codon-optimized, synthesized as double-stranded DNA fragments, and assembled into human IgG1 expression vectors by Golden Gate cloning. Transient expression was performed in ExpiCHO-S cells at a viable cell density of 6×10^6 cells/mL using the ExpiCHO Expression System High-Titer workflow at 3 mL scale with minor modifications (Thermo Fisher, A29133). At 18–22 h post-transfection, cultures were shifted to 32°C, 5% CO2, and supplemented with the ExpiCHO Feed/Enhancer mixture according to the manufacturer's instructions. Supernatants were harvested on day 10 post-transfection, clarified by centrifugation and filtration through a 1.2 µm AcroPrep membrane (Cytiva, 97035). Antibodies were purified from clarified supernatants by tip-based Protein A affinity purification using MabSelect PrismA resin (IMCS, 04T-H6R80A-1-50-96). Samples were loaded by repeated aspiration–dispense cycles, washed with PBS (pH 7.2), and eluted with 0.1 M glycine-HCl (pH 2.8; Pierce, 21009), with immediate neutralization by Tris-HCl to a final pH of 7.2. Supernatant clarification and Protein A purification were executed on a Hamilton Vantage platform under fully automated control. Six low-energy and one high-energy design either did not express or had low monomer content, and were excluded from further analysis.

<u>NanoDSF:</u>

D44.1 and its designs:
The first apparent melting temperature of the antibodies was determined using the Prometheus Panta instrument (NanoTemper). Full-length antibodies formatted as IgG2a were diluted to 0.2 mg/mL in 1X PBS. The temperature was ramped from 20 °C to 100 °C at 1.0 °C per minute. The first apparent melting temperatures were identified using the Prometheus Panta software (v. 1.4.4) with default parameters.

6G08 and its designs:
Thermal stability of antibodies was assessed by nano differential scanning fluorimetry (nanoDSF) using the Prometheus NT.48 (NanoTemper Technologies, Munich, Germany). Samples were loaded into Prometheus standard capillaries at approximately 1 mg/mL and exposed to thermal stress from 15°C to 95°C, with a ramp of 1°C/min. Intrinsic protein fluorescence was measured at 330 nm and 350 nm. Melting temperature (T_m) was calculated by PR.ThermControl software (NanoTemper Technologies) using first derivative analysis of the 350/330 nm fluorescence ratio plotted against temperature.

Urelumab and its designs:
The thermal stability of the antibodies was determined using the Prometheus Panta instrument (NanoTemper). Antibodies were diluted to 0.25 mg/mL in PBS pH 7.2, and the temperature was ramped from 25 °C to 95 °C at 1 °C/min. The T_m1 values were identified using the Prometheus Panta software (v. 1.4.4) with default parameters.

SPR:

D44.1 and its designs:
SPR experiments on the anti-HEWL antibodies were carried out on a Biacore 8K instrument (Cytiva) at 25 °C with PBST (1X PBS + 0.05% Tween-20). For binding analysis, each antibody was immobilized to a flow cell of a protein A sensor chip (Cytiva 29127555) at a concentration of 5 µg/mL (corrected by purity from SDS-PAGE) in 1X PBS for 60 sec at a flow rate of 10 µL/min. A reference flow cell for each channel was not loaded with antibody. Approximately 100-200 response units of each antibody were immobilized. Samples of different HEWL (Sigma 62970) concentrations (10, 5, 2.5, 1.25, 0.63, 0.31, and 0.16, and 0.08 nM) were injected over the surface at a flow rate of 30 µL/min for 240 s, and the chip was washed with PBST for 600 sec at a flow rate of 30 µL/min. After each injection, surface regeneration was performed with a 30-second injection of 10 mM glycine pH 1.5 at a flow rate of 10 µL/min, and reached the same baseline. The sensogram for each flow cell without antibody was subtracted from the sensogram for the flow cell with antibody from the same channel. The acquired sensograms were analysed using the device's software (v.4.0.8.19879). The values for k_{off} were first fit independently with a 1:1 dissociation model with k_{off} fit globally and offset and drift set to a constant value of 0. The fit values for k_{off} were then used to perform full kinetic fits to a 1:1 binding model with global fitting of k_{on}, R_{max}, and t_c with standard initial values. RI and drift were set to a constant 0, and k_{off} was set as a constant to the values fit previously. The experiment was repeated using different samples of HEWL and different dilutions of the antibodies. The 2.5, 1.25, 0.63, 0.31, and 0.16 nM samples were used for fitting.

6G08 and its designs:
SPR was performed using a Biacore T100 upgraded to T200 sensitivity (Cytiva) with a flow cell temperature set to 25°C. An anti-HIS antibody was immobilised on a Biacore CM5 chip using the His capture kit (Cytiva 28995056). Recombinant human FcγRIIb (R&D Systems) at 0.5µg/ml was captured with a target of 100 RU. Antibodies at indicated concentrations were then flowed at 30 µL/ml for 4 minutes, followed by a 5-minute dissociation with HBS-EP+ (Cytiva) used as running buffer. Data was analysed using Biacore Evaluation software and fit to a bivalent model.

Urelumab and its designs:
SPR measurements were performed on an LSA^{XT} (Carterra) at 37°C using a capture-based non-regenerative multi-cycle kinetic experiment format. To prepare the sensor surface, the SPR system was exchanged into a running buffer of 10 mM HEPES pH 7.6, 150 mM NaCl, 3 mM EDTA, 0.05% P-20. Next, a new CMDP sensor chip (Carterra) was primed with six pulses of 200 mM Sodium Borate, 1 M NaCl, pH 9.0 (180s exposure time). Next, the surface was activated with a 1:1 mixture of 400 mM EDC and 100 mM NHS (Thermo) for 7 minutes. Next, an anti-human Fc polyclonal (Southern Biotech) was prepared at 100 ug/mL in 10 mM Sodium Acetate pH 4.5 and injected over the surface for 7 minutes to prepare a saturated "lawn" of capture reagent. Finally, all remaining NHS-esters were deactivated by the addition of 0.1 M ethylene diamine diluted in 200 mM sodium borate pH 9.0. For measurement of binding kinetics for the antibody cohort, the system was exchanged into an experimental running buffer consisting of 10 mM HEPES pH 7.6, 150 mM NaCl, 3 mM EDTA, 0.05% P-20, 0.01% w/v BSA.

Antibodies were prepared as 1 ug/mL dilution in the experimental running buffer and captured on the first quadrant using the 96-print head with an exposure time of 240s. Following this, the single flow cell was docked and a series of 10 blank injections followed by a set of 7 increasing concentrations of 4-1BB (Acro) ranging from 0.2 nM - 1000 nM (3-fold dilution) diluted into running buffer. The association phase was set to 180s followed by a dissociation phase of 600s for each blank and antigen concentration injected. Following injection of the last antigen concentration, The surface was regenerated by 3 pulses of 10 mM Glycine pH 1.7 (30s exposure time). Data was processed using a double-referencing format in Kinetics software (Carterra) and data were fit to a global 1-1 binding model. Data that significantly deviated from a 1-1 model fit or showed no discernable binding were excluded from analysis.

Flow cytometry staining of B-cells:

Leukocyte cones were obtained from healthy donors from the Southampton NHS Blood and Transplant Service. Use was approved by the University of Southampton Faculty of Medicine Ethics Committee and East of Scotland Research Ethics Service, Research ethical committee reference number 16/ES/0048. PBMC were isolated using lymphoprep (Stemcell Technologies). $2.5x10^5$ PBMC were stained with 1µl anti-CD19 PE-Cy7 (Biolegend) and 6G08 antibodies at the indicated concentrations for 30 minutes at 4°C. Cells were washed, and 0.5 µl PE goat anti-human IgG (Jackson Immunoresearch) was added for 30 minutes at 4°C. After washing, samples were run on a Cytoflex flow cytometer (Beckman Coulter). Data was analysed using FlowJo v10 (BD Biosciences) by gating on single cells, identifying lymphocytes based on FSC vs. SSC, and gating on CD19+ B cells.

Flow cytometry staining of CHO-K1 cells overexpressing FcγRIIa:

CHO-K1 cells stably expressing hFcγRIIa (H131) were generated as described previously[60] and were maintained in RPMI + 10% FCS (Sigma-Aldrich), glutamine (2mM), sodium pyruvate (1mM), penicillin (100 U/ml), streptomycin (100 ug/ml) and 10 ug/ml puromycin (Invivogen), all Life Technologies unless otherwise stated. Cells were stained with 6G08 antibodies as above, washed, and run on a Cytoflex flow cytometer (Beckman Coulter).

CSI-BLI of Urelumab and its designs:

CSI-BLI measurements were performed on an Octet RH16 (Sartorius) with Octet Anti-Human IgG Quantitation (AHQ) Biosensors (Sartorius, Cat# 18-5001). Adalimumab and bococizumab formatted as hIgG1 were used as negative and positive controls, respectively. Test articles and controls were diluted in 1x Octet Kinetics buffer (Sartorius, Cat# 18-1105) to 3 µg/mL for loading and 150 µg/mL for association steps. Human Fc fragment (Jackson Immuno Research, 009-000-008) was diluted to 42.4 µg/mL in 1x Octet Kinetics buffer for blocking. Sensors were regenerated and neutralized using glycine pH 1.5 (Cytiva, Cat# BR100354) and 1x Octet Kinetics buffer, respectively. Samples were prepared in 384-well tilted bottom Octet plates (Sartorius, Cat# 18-5166) with 50 µL per well. Octet AHQ Biosensors were soaked in 1x Octet Kinetics buffer for at least 10 min before measurements. The experiment included baseline (60 sec in 1x Octet Kinetics buffer), loading (3 µg/mL antibodies for 100 sec), blocking (human Fc for 180 sec), baseline (60 sec in 1x Octet Kinetics buffer), association (150 µg/mL antibodies for 300 sec), dissociation (30 sec in 1x Octet Kinetics buffer), and regeneration. Data was

processed in Octet Analysis Studio software with alignment to baseline and Savitzky-Golay filtering applied. CSI-BLI fold change was determined by normalizing the association response to that of the negative control, adalimumab.

Accelerated stability of Urelumab and its designs:
Test antibodies were normalized to 1 mg/mL in PBS pH 7.2 with 10 mM EDTA and incubated for 14 days at either 4°C or 45°C in 96-well plates (Armadillo PCR Plate, Thermo Scientific AB2396). Samples were filtered with 0.2 um filters (AcroPrep Advance 96-well plates, 350 uL, wwPTFE membrane), then 5 uL was injected onto a UPLC Agilent 1290 Infinity II LC system equipped with a Waters ACQUITY Premier Protein SEC column (250Å, 1.7 µm, 150 mm). The mobile phase used was 100 mM sodium phosphate, 100 mM sodium sulfate, 0.05% w/w sodium azide, pH 6.8, with a flow rate of 0.5 mL/min. The monomer, aggregate, and fragment percentages for each antibody were determined using ChemStation software (Agilent). The change in monomer, aggregate, and fragment content was calculated from the difference between each antibody incubated at 4°C versus 45°C.

BSA ELISA of Urelumab and its designs:
Black, 384-well high binding plates (Greiner Bio-One, Cat# 781077) were coated with blocking buffer (PBS pH 7.2 + 0.5% BSA) at 95 µL/well and incubated overnight at 4°C. After coating, plates were washed with PBS pH 7.2. Antibody samples and controls were diluted to 15 µg/mL in blocking buffer and added at 12.5 µL/well in quadruplicates. Adalimumab and bococizumab formatted as hIgG1 were used as negative and positive controls, respectively. Plates were incubated for 1 h at room temperature (RT), washed three times with PBS pH 7.2, and anti-human IgG (Fc specific)-peroxidase (Sigma-Aldrich, Cat# A0170) was added (0.3 µg/mL dilution in blocking buffer) at 12.5 µL/well. After 1 h incubation at RT, plates were washed three times with PBS pH 7.2. QuantaBlu working solution (ThermoFisher, Cat# 15169) was prepared by mixing 1 part QuantaBlu Stable Peroxide Solution with 9 parts Substrate Solution, equilibrated to RT, and added at 25 µL/well. After 20 min incubation at RT, the reaction was stopped with QuantaBlu Stop Solution at 25 µL/well. Fluorescence readings were taken on an Envision plate reader with excitation and emission at 325 nm and 420 nm, respectively. BSA scores were calculated by normalizing the average fluorescence to the average fluorescence of the positive control, bococizumab.

Ovalbumin binding of Urelumab and its designs:
Ovalbumin binding was measured using the PAIA Biotech developability kit (PA-DEV-OVA; PAIA Biotech). Adalimumab and bococizumab formatted as hIgG1 were used as negative and positive controls, respectively. Samples were diluted to 10 ug/mL in ddH_2O, then 40 uL of PAIA PSP mix and 20 uL of antibody were combined in the 384-well PA-DEV-OVA plate in duplicate. Samples were shaken for 30 mins at RT at 2200 rpm, then centrifuged at 500xg for 1 minute. Fluorescence readings were taken on an Envision plate reader using bottom read with excitation and emission at 620 and 685, respectively. Relative ovalbumin binding was calculated by normalizing the average signal between buffer only signal (no resin binding) and zero signal (full resin binding).

Contributions:

A.T. and S.J.F. conceived the idea. A.T. and S.J.F. wrote the paper, with input from all authors. A.T. developed the affinity optimization algorithm and designed all amino acid sequences. A.M. assembled the D44.1 library and performed the phage display screening. A.T. and Y.F.S. performed biophysical characterization of the anti-HEWL antibodies and A.T., Y.F.S, S.J.F, and T.M.M analyed the data. S.G. prepared the anti-HEWL sequences for deep sequencing. A.T. performed the deep sequencing experiment, and A.T. and S.J.F. analyzed the data. R.J.O., I.G.E., and J.K. performed experimental work related to the 6G08 designs, and A.T., S.J.F., R.O., I.G.E., and M.S.C. analyzed the data. T.P., C.D., J.B., P-H.C., and D.X.Z. performed experimental work related to the Urelumab designs, and A.T., S.J.F., T.P., T.M.M., A.B.D., and G.K. analyzed the results. O.P. and A.T. performed the mutational analysis of the Urelumab designs, and O.P., A.T., and S.J.F. analyzed the results. J.H. performed the experimental work related to the Cetuximab designs, and J.H., D.C., A.T., and S.J.F. analyzed the data. S.J.F., M.S.C., and D.C. were responsible for funding acquisition.

Acknowledgments:

We thank the Membrane Protein Laboratory at the Diamond Light Source for access to the Prometheus NT.48 instrument. The Membrane Protein Laboratory is funded by a grant from the Wellcome Trust (223727/Z/21/Z) with additional support provided by Diamond Light Source and the Research Complex at Harwell, both Instruct-ERIC centers. We thank Dina Listov, Olga Khersonsky, Rebecca Wilen, Inbar Cahana, Mimi Schwartz, Lucas Krauss (Weizmann Institute of Science), Ahuva Nissim (Queen Mary University of London and Weizmann Institute of Science), Shahar Harel (Technion), and Hayden Fisher (ESRF) for critical discussions on the work. A.T. was supported by Teva Pharmaceutical Industries Ltd as part of the Israeli National Forum for BioInnovators (NFBI) and by the Azrieli Foundation under the Azrieli Fellows Program. I.G.E. and M.S.C. were supported by a grant from Cancer Research UK, DRCPGM\100039. The study was supported by a European Research Council Advanced Grant (101140394) to S.J.F.

Competing interests:

M.S.C. acts as a retained consultant for BioInvent International, consults for several other biotech companies, and receives institutional payments and royalties from antibody licenses. He has received research funding from BioInvent, GSK, iTeos, UCB, Surrozen, and Roche. S.J.F. and A.T. are named inventors on patents related to antibody design. S.J.F. is a paid advisor on protein design. T.P., C.D., J.B., P-H.C., D.X.Z., A.B.D., T.M.M., R.C-W. and G.K. are all AstraZeneca employees and may or may not hold AstraZeneca stock.

Data availability:

Amino acid sequences of the Fvs of the four D44.1, ten 6G08, and 59 Urelumab designs characterized as IgG are available in Supp. Tables 12-14. The D44.1 deep sequencing data will be made available upon publication.

References:

1. Strohl, W. R. Structure and function of therapeutic antibodies approved by the US FDA in 2024. *Antib. Ther.* (2025) doi:10.1093/abt/tbaf014.
2. Bradbury, A. & Plückthun, A. Reproducibility: Standardize antibodies used in research. *Nature* **518**, 27–29 (2015).
3. Chan, A. C., Martyn, G. D. & Carter, P. J. Fifty years of monoclonals: the past, present and future of antibody therapeutics. *Nat. Rev. Immunol.* (2025) doi:10.1038/s41577-025-01207-9.
4. Victora, G. D. & Nussenzweig, M. C. Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).
5. Hu, D. *et al.* Effective optimization of antibody affinity by phage display integrated with high-throughput DNA synthesis and sequencing technologies. *PLoS One* **10**, e0129125 (2015).
6. Julian, M. C., Li, L., Garde, S., Wilen, R. & Tessier, P. M. Efficient affinity maturation of antibody variable domains requires co-selection of compensatory mutations to maintain thermodynamic stability. *Sci. Rep.* **7**, (2017).
7. Shehata, L. *et al.* Affinity maturation enhances antibody specificity but compromises conformational stability. *Cell Rep.* **28**, 3300–3308.e4 (2019).
8. Rabia, L. A., Desai, A. A., Jhajj, H. S. & Tessier, P. M. Understanding and overcoming trade-offs between antibody affinity, specificity, stability and solubility. *Biochem. Eng. J.* **137**, 365–374 (2018).
9. Zhou, J., Le, C. Q., Zhang, Y. & Wells, J. A. A general approach for selection of epitope-directed binders to proteins. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2317307121 (2024).
10. Foote, J. & Winter, G. Antibody framework residues affecting the conformation of the hypervariable loops. *J. Mol. Biol.* **224**, 487–499 (1992).
11. Makabe, K. *et al.* Thermodynamic consequences of mutations in vernier zone residues of a humanized anti-human epidermal growth factor receptor murine antibody, 528. *J. Biol. Chem.* **283**, 1156–1166 (2008).
12. Lippow, S. M., Wittrup, K. D. & Tidor, B. Computational design of antibody-affinity improvement beyond in vivo maturation. *Nat. Biotechnol.* **25**, 1171–1176 (2007).
13. Hie, B. L. *et al.* Efficient evolution of human antibodies from general protein language models. *Nat. Biotechnol.* **42**, 275–283 (2024).
14. Shanker, V. R., Bruun, T. U. J., Hie, B. L. & Kim, P. S. Unsupervised evolution of protein and antibody complexes with a structure-informed language model. *Science* **385**, 46–53 (2024).
15. Hummer, A. M., Schneider, C., Chinery, L. & Deane, C. M. Investigating the volume and diversity of data needed for generalizable antibody-antigen ΔΔG prediction. *Nat. Comput. Sci.* **5**, 635–647 (2025).
16. Frey, N. C. *et al.* Lab-in-the-loop therapeutic antibody design with deep learning. *bioRxiv* (2025) doi:10.1101/2025.02.19.639050.
17. Jain, T. *et al.* Biophysical properties of the clinical-stage antibody landscape. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 944–949 (2017).
18. Alford, R. F. *et al.* The Rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theory Comput.* **13**, 3031–3048 (2017).
19. Baran, D. *et al.* Principles for computational design of binding antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 10900–10905 (2017).
20. Netzer, R. *et al.* Ultrahigh specificity in a network of computationally designed protein-interaction pairs. *Nat. Commun.* **9**, 5286 (2018).
21. Khersonsky, O. & Fleishman, S. J. What Have We Learned from Design of Function in Large Proteins? *BioDesign Research* **2022**, 1–11 (2022).

22. Khersonsky, O. *et al.* Automated design of efficient and functionally diverse enzyme repertoires. *Mol. Cell* **72**, 178–186.e5 (2018).
23. Warszawski, S. *et al.* Optimizing antibody affinity and stability by the automated design of the variable light-heavy chain interfaces. *PLoS Comput. Biol.* **15**, e1007207 (2019).
24. VanDrisse, C. M., Lipsh-Sokolik, R., Khersonsky, O., Fleishman, S. J. & Newman, D. K. Computationally designed pyocyanin demethylase acts synergistically with tobramycin to kill recalcitrant Pseudomonas aeruginosa biofilms. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
25. Listov, D. *et al.* Complete computational design of high-efficiency Kemp elimination enzymes. *Nature* **643**, 1421–1427 (2025).
26. Prassler, J. *et al.* HuCAL PLATINUM, a synthetic Fab library optimized for sequence diversity and superior performance in mammalian expression systems. *J. Mol. Biol.* **413**, 261–278 (2011).
27. Lipsh-Sokolik, R. & Fleishman, S. J. Addressing epistasis in the design of protein function. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2314999121 (2024).
28. Braden, B. C. *et al.* Three-dimensional structures of the free and the antigen-complexed Fab from monoclonal anti-lysozyme antibody D44.1. *J. Mol. Biol.* **243**, 767–781 (1994).
29. Cauerhff, A., Goldbaum, F. A. & Braden, B. C. Structural mechanism for affinity maturation of an anti-lysozyme antibody. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3539–3544 (2004).
30. Tennenhouse, A. *et al.* Computational optimization of antibody humanness and stability by systematic energy-based ranking. *Nat. Biomed. Eng.* **8**, 30–44 (2024).
31. Hutchinson, M. *et al.* Toward enhancement of antibody thermostability and affinity by computational design in the absence of antigen. *MAbs* **16**, 2362775 (2024).
32. Hoch, S. Y. *et al.* GGAssembler: Precise and economical design and synthesis of combinatorial mutation libraries. *Protein Sci.* **33**, e5169 (2024).
33. Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. Making antibodies by phage display technology. *Annu. Rev. Immunol.* **12**, 433–455 (1994).
34. Roghanian, A. *et al.* Antagonistic human FcγRIIB (CD32B) antibodies have anti-tumor activity and overcome resistance to antibody therapy in vivo. *Cancer Cell* **27**, 473–488 (2015).
35. Frampton, S. *et al.* Fc gamma receptors: Their evolution, genomic architecture, genetic variation, and impact on human disease. *Immunol. Rev.* **328**, 65–97 (2024).
36. Sutton, E. J. *et al.* Evaluating anti-CD32b F(ab) conformation using molecular dynamics and small-angle X-ray scattering. *Biophys. J.* **115**, 289–299 (2018).
37. Makowski, E. K. *et al.* Optimization of therapeutic antibodies for reduced self-association and non-specific binding via interpretable machine learning. *Nat. Biomed. Eng.* **8**, 45–56 (2024).
38. Li, S. *et al.* Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7**, 301–311 (2005).
39. Low, N. M., Holliger, P. H. & Winter, G. Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *J. Mol. Biol.* **260**, 359–368 (1996).
40. Li, B. *et al.* PROPERMAB: an integrative framework for in silico prediction of antibody developability using machine learning. *MAbs* **17**, 2474521 (2025).
41. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6**, e20161 (2011).
42. MacCallum, R. M., Martin, A. C. & Thornton, J. M. Antibody-antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.* **262**, 732–745 (1996).
43. Dondelinger, M. *et al.* Understanding the significance and implications of antibody numbering and antigen-binding surface/residue definition. *Front. Immunol.* **9**, 2278 (2018).
44. Goldenzweig, A. *et al.* Automated structure- and sequence-based design of proteins for

high bacterial expression and stability. *Mol. Cell* **70**, 380 (2018).
45. Weinstein, J. Y. *et al.* Designed active-site library reveals thousands of functional GFP variants. *Nat. Commun.* **14**, 2890 (2023).
46. Lipsh-Sokolik, R. *et al.* Combinatorial assembly and design of enzymes. *Science* **379**, 195–201 (2023).
47. Gasteiger, E. *et al.* Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* 571–607 (Humana Press, Totowa, NJ, 2005).
48. Cock, P. J. A. *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
49. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
50. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
51. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185 (1998).
52. Abhinandan, K. R. & Martin, A. C. R. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* **45**, 3832–3839 (2008).
53. Kabat, E. A. & Wu, T. T. Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann. N. Y. Acad. Sci.* **190**, 382–393 (1971).
54. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041 (2001).
55. Tareen, A. & Kinney, J. B. Logomaker: beautiful sequence logos in Python. *Bioinformatics* **36**, 2272–2274 (2020).
56. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
57. Noy-Porat, T. *et al.* A panel of human neutralizing mAbs targeting SARS-CoV-2 spike at multiple epitopes. *Nat. Commun.* **11**, 4303 (2020).
58. Azriel-Rosenfeld, R., Valensi, M. & Benhar, I. A human synthetic combinatorial library of arrayable single-chain antibodies based on shuffling in vivo formed CDRs into general framework regions. *J. Mol. Biol.* **335**, 177–192 (2004).
59. Mechaly, A. *et al.* Novel phage display-derived anti-abrin antibodies confer post-exposure protection against abrin intoxication. *Toxins (Basel)* **10**, 80 (2018).
60. Tutt, A. L. *et al.* Development and characterization of monoclonal antibodies specific for mouse and human fcγ receptors. *J. Immunol.* **195**, 5503–5516 (2015).