



OPEN Design of immunogens to present a tumor-specific cryptic epitope

Huafeng Xu^{1,3}✉, Timothy Palpant^{1,3}, Qi Wang¹ & David E. Shaw^{1,2}✉

In many cancers, the epidermal growth factor receptor (EGFR) gene is amplified, mutated, or both. The monoclonal antibody mAb806 binds selectively to cancer cells that overexpress EGFR or express the truncated mutant EGFRvIII, but not to normal cells. This suggests that a promising avenue for developing cancer vaccines may be to design immunogens that elicit mAb806-like antibodies. In this study, we designed immunogens that present the mAb806-binding epitope in the same conformation as in overexpressed or truncated EGFR. We first used molecular dynamics simulations to identify conformations of EGFR in which the residues of the mAb806-binding epitope are accessible. We then designed immunogens by substituting that epitope in place of a structurally similar loop in a different protein and generating mutants that could potentially stabilize the mAb806-binding conformation in this new context. Two mutants in which the epitope remained stable in subsequent simulations were chosen for evaluation *in vitro*. Binding kinetics experiments with these designed immunogens provided strong evidence that the epitope was successfully stabilized in the mAb806-binding conformation, suggesting that they could potentially form the basis of vaccines that elicit cancer-selective antibodies.

Antigens—substances specifically recognized by immune receptors such as antibodies—derive either from external sources, as occurs with infection or vaccination, or from endogenous sources, as in autoimmune disorders. Cancer cells are an endogenous source of antigens, and antibodies that target antigens on the surfaces of cancer cells have been used successfully as cancer therapies in a clinical setting.^{1,2} Current anti-cancer antibodies, however, are often cross-reactive against healthy tissues, which can lead to serious side effects.³ One avenue for improving specificity is to exploit a tendency for cancer cells to express surface proteins in abnormal conformations.^{4,5} These conformations may expose epitopes—the specific regions of antigens that are recognized by antibodies—that are sterically inaccessible in native protein conformations.⁶ (We will refer to such epitopes as cryptic epitopes.) Antibodies that can differentiate between native and abnormal conformations of the same protein have the potential to become a new class of therapeutics with a superior efficacy and safety profile.⁷ Antigens that can elicit such antibodies may both facilitate the development of therapeutic antibodies, and themselves be candidates for cancer vaccines that directly elicit immunological anti-cancer responses in patients.

Given a cryptic epitope and its three-dimensional structure in a non-native protein conformation, epitope-focused design^{8–14} could potentially be used to develop immunogens (antigens that induce antibody production) that present the epitope in the abnormal conformation and thereby elicit antibodies specifically recognizing that conformation. Non-native conformations that arise transiently due to the decreased stability of native conformations—whether caused by pathogenic mutations or compromised protein quality control—are often difficult to determine using experimental structural methods such as X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. Molecular dynamics (MD) simulation¹⁵ could be a viable alternative for generating models of non-native conformations or other transient conformations that expose otherwise inaccessible epitopes.^{16–18}

The mouse monoclonal antibody mAb806 binds selectively to cancer cells that either overexpress the epidermal growth factor receptor (EGFR) or express an exon 2–7 deletion mutant known as EGFRvIII.^{19,20} The 16-residue epitope recognized by mAb806 is accessible only in EGFR proteins that are misfolded as a consequence of overexpression, the exon 2–7 deletion,²¹ or some oncogenic mutations that alter the conformational dynamics of the EGFR ectodomain.²² A crystal structure of the complex between the mAb806 antigen-binding fragment (Fab) and its epitope peptide is available,²³ offering one starting point for structure-based immunogen design. The EGFR conformation to which mAb806 binds, however, remains unknown, and the epitope in mAb806-bound EGFR might have differences in conformation from that in the crystal structure of the peptide-mAb806 complex. Using MD simulations to model the conformation of the epitope in the context of full-length misfolded EGFR could thus increase the potential for designing a successful immunogen.

¹D. E. Shaw Research, New York, NY 10036, USA. ²Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA. ³Present address: Atommap Corporation, New York, NY 10065, USA. ✉email: huafeng@gmail.com; David.Shaw@DEShawResearch.com

In this study, we designed and tested immunogens that present the EGFR cryptic epitope in its mAb806-binding conformation. We first used MD simulations to find conformations of EGFR in which the cryptic epitope was accessible for mAb806 binding. Starting from these conformations, we built structural models of the mAb806-EGFR complex. We then searched the Protein Data Bank for proteins with loop regions that are structurally similar to the epitope in the modeled mAb806-bound EGFR conformations, replaced the loop region of a candidate protein with the epitope, and designed a series of mutants with the aim of stabilizing the epitope in the mAb806-binding conformation. For many of the designed immunogens, the epitope primarily occupied the antibody-binding conformation in subsequent MD simulations. From these immunogens, we selected two for experimental characterization.

Binding kinetics experiments confirmed that these two designed immunogens bound mAb806 with substantially higher association rate constants and binding affinities than the epitope peptide alone, providing strong evidence that the immunogens successfully stabilized the epitope in the mAb806-binding conformation. In preliminary immunization experiments, our designed immunogens were only weakly immunogenic, possibly because their relatively small size (~11 kDa) and four internal disulfide bonds cause them to resemble cyclic peptides, which tend not to be immunogenic. Their ability to present the epitope in the mAb806-binding conformation nevertheless suggests they may be promising starting points for the development of immunogens that elicit a robust immune response of mAb806-like, cancer-selective antibodies. More generally, our immunogen design strategy, in which MD simulations generate models of target epitope conformations that are subsequently assessed for their stability in the context of designed immunogens, may help facilitate the future development of highly selective, and thus safe, therapeutic antibodies, or even cancer vaccines.

Results

To generate a structural model of mAb806 in complex with EGFR, we used long-timescale MD simulations, aided by an adaptive simulated tempering protocol (see Methods for details), to simulate the extracellular portion (ectodomain) of EGFR. The EGFR ectodomain consists of two ligand-binding domains (domains I and II) and two cysteine-rich domains (domains III and IV). In order to reduce the size of the molecular system and thus accelerate our simulations, we omitted domain IV, which is spatially remote from the epitope, and thus, we reasoned, unlikely to affect its conformation or exposure. We performed two such simulations (referred to below as simulations 1 and 2), each starting from a different initial conformation drawn randomly from an equilibration simulation. We then attempted to dock the Fab of mAb806 onto the epitope in EGFR conformations taken from frames of our simulations by aligning the Ca atoms of the epitope in the crystal structure of the mAb806-epitope complex (PDB ID: 3G5V) with those of the epitope in the simulated EGFR conformations, and assessed for each conformation whether steric clashes occurred in the resulting mAb806-EGFR complex.

In its normal, regulated function, EGFR can transition from an auto-inhibited, tethered conformation to an active, untethered conformation in which domains I and II are rotated. Based on their X-ray crystal structures,^{24,25} steric hindrance would prevent mAb806 from binding to either of these conformations (Figure S1). It is possible that the epitope is exposed in an intermediate conformation during this transition,^{6,23} but this appears unlikely given that the specificity of the antibody indicates that it does not bind normally functioning wild-type (WT) EGFR. In simulation 1, EGFR transitioned between the tethered and untethered conformations while hiding the epitope from mAb806 (Fig. 1, Movie S1), further supporting the hypothesis that mAb806 only binds to EGFR in aberrant conformations that are not necessary for its normal function.²¹

In a small number of frames from simulation 2, the epitope, largely adopting the same conformation as in the crystal structure of the mAb806-epitope complex, became accessible to mAb806. When the mAb806 Fab was docked onto the epitope in these frames, there were no steric clashes between the Fab and EGFR (Fig. 2, Movie S1). (A previous study simulating mAb806 binding to EGFR²⁶ used steered MD simulation to promote exposure of the epitope to mAb806 binding.) This allowed us to build a structural model of mAb806 in complex with EGFR. We simulated the mAb806-EGFR complex starting from this structural model—once with simulated tempering (simulation 3) and once without (simulation 4)—and found that the complex remained stable in these simulations.

The mAb806-binding conformation from our MD simulations differs from native EGFR conformations by a rotation in the Ca–C bond of the Arg285 residue; that is, a change in the ψ dihedral angle (Fig. 3). This flips the epitope such that it no longer contacts the adjacent β -sheet (residues Try275 to Arg285) and exposes it for mAb806 binding. Mutationally disrupting this β -sheet destabilizes its contacts with the epitope, making the epitope more accessible to mAb806.²³ In agreement with this finding, we found that the mutant EGFR^{C271A/C283A}, which eliminates a disulfide bond that restrains the β -sheet's conformation, was more prone to exposing the epitope to mAb806 binding in our simulations (simulations 5 and 6; Figure S2). These results offer a plausible explanation for mAb806's selectivity for cells overexpressing EGFR or expressing EGFRvIII: In cells overexpressing EGFR, inadequate chaperone assistance might cause incorrect disulfide bonding patterns in domain II (also known as the CR1 domain), and in cells expressing EGFRvIII, the deletion—of residues 6 to 273—leads to the loss of the disulfide bond between Cys271 and Cys283, potentially disrupting the Tyr275–Arg285 β -sheet and thus exposing the epitope to mAb806 binding.

The free 16-mer epitope peptide (sequence CGADSYEMEEDGVKRC, with a disulfide bond between the cysteines at positions 1 and 16, making it a cyclic peptide) is unstructured, and did not retain its mAb806-binding conformation in our simulations (Figures S3A and 4E), in agreement with a previous NMR experiment.²³ This may limit the ability of the epitope peptide to function as an immunogen to elicit mAb806-like antibodies. To design immunogens that would present the epitope in the mAb806-binding conformation, we used as a starting point the ensemble of conformations generated from our MD simulation of the mAb806-EGFR complex. Using MotifGraft in Rosetta,^{12,27} we searched a subset of PDB structures for scaffolds with an appropriate segment; that

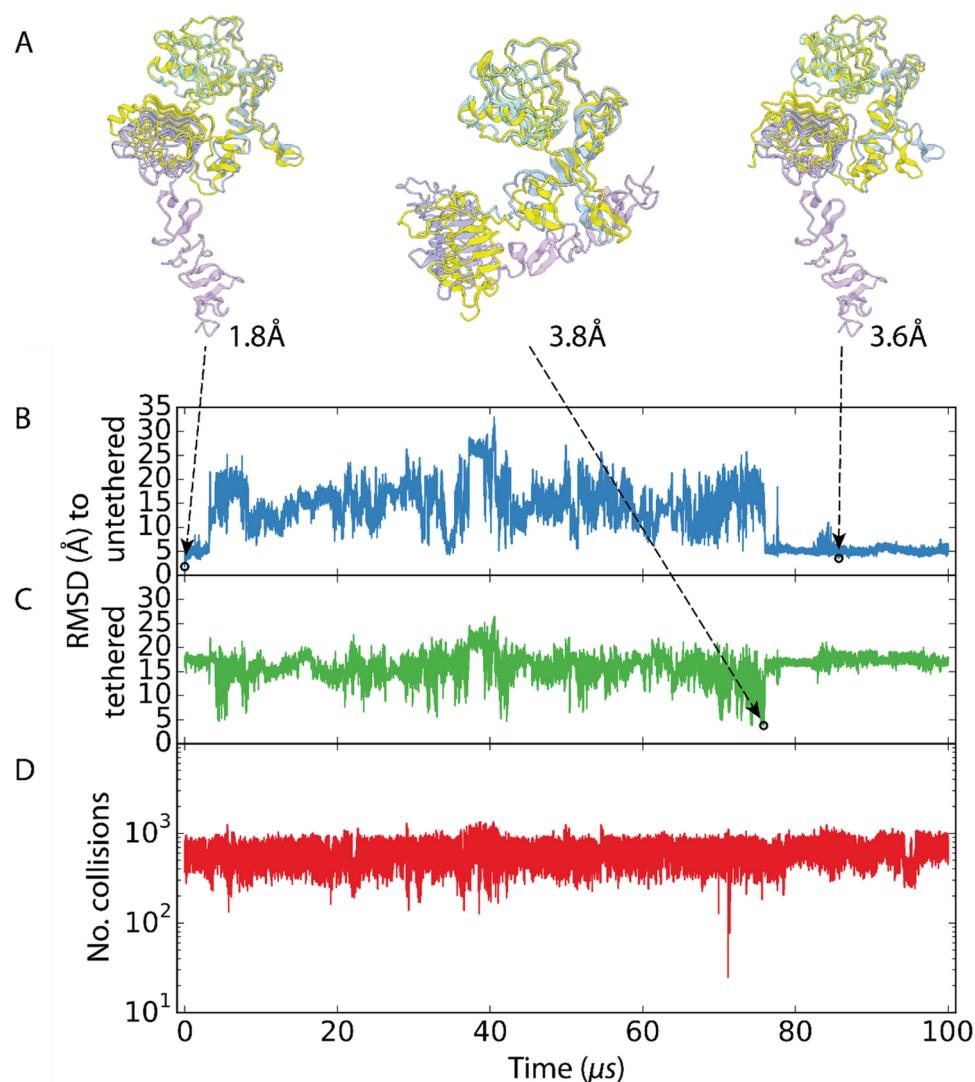


Fig. 1. EGFR can transition between the untethered and tethered conformations without exposing the epitope to mAb806 binding. **(A)** Snapshots of EGFR conformations from simulation 1 (using simulated tempering). Simulated conformations, shown in yellow, are superimposed onto crystal structures of the untethered conformation (left and right; PDB ID: 3NJP) or tethered conformation (center; PDB ID: 1NQL), in the crystal structures, domains I (also known as the L1 domain), II (the CR1 domain), III (the L2 domain), and IV (the CR2 domain) are shown in cyan, light blue, purple, and light purple, respectively. Domain IV was not included in our simulations. The simulation shown starts from an untethered conformation (left), visits a tethered conformation (center), and returns to an untethered conformation (right). The root-mean-square deviation (RMSD) of the Ca atoms of the conformations from the corresponding crystal structures are given beside the conformation. Also shown are time traces of the Ca RMSD from the crystal structures of **(B)** the untethered conformation (blue) and **(C)** the tethered conformation (green). **(D)** The number of EGFR atoms that are in collision (interatomic distance ≤ 1.5 Å) with mAb806 docked onto its epitope remains greater than 20 in the simulation. mAb806 was docked onto each simulated conformation of EGFR by aligning the Ca atoms of the epitope in the crystal structure of the mAb806-epitope complex (PDB ID: 3G5V) onto the epitope in the simulated conformation.

is, a segment that could be replaced with the epitope in a conformation taken from a snapshot of our mAb806-EGFR simulation (see SI for details).

This search identified the cysteine-rich, 93-residue C-terminal domain of the human DKK1 protein (PDB ID: 3S2K, chain C) as a potential scaffold. In four simulated conformations, the 16-residue epitope could be superimposed onto a 16-residue hairpin of this scaffold protein without substantial steric clashes occurring between the Fab and the scaffold protein (Fig. 4). In contrast, performing a search using the crystal structure of the mAb806-epitope complex did not identify DKK1 as a potential scaffold; the epitope in the crystal structure superimposed onto the hairpin segment of DKK1 resulted in unacceptable structural differences (Ca RMSD > 1.5 Å) and steric clashes (No. collisions > 20) (Fig. 4D). This suggests that, for immunogen design in

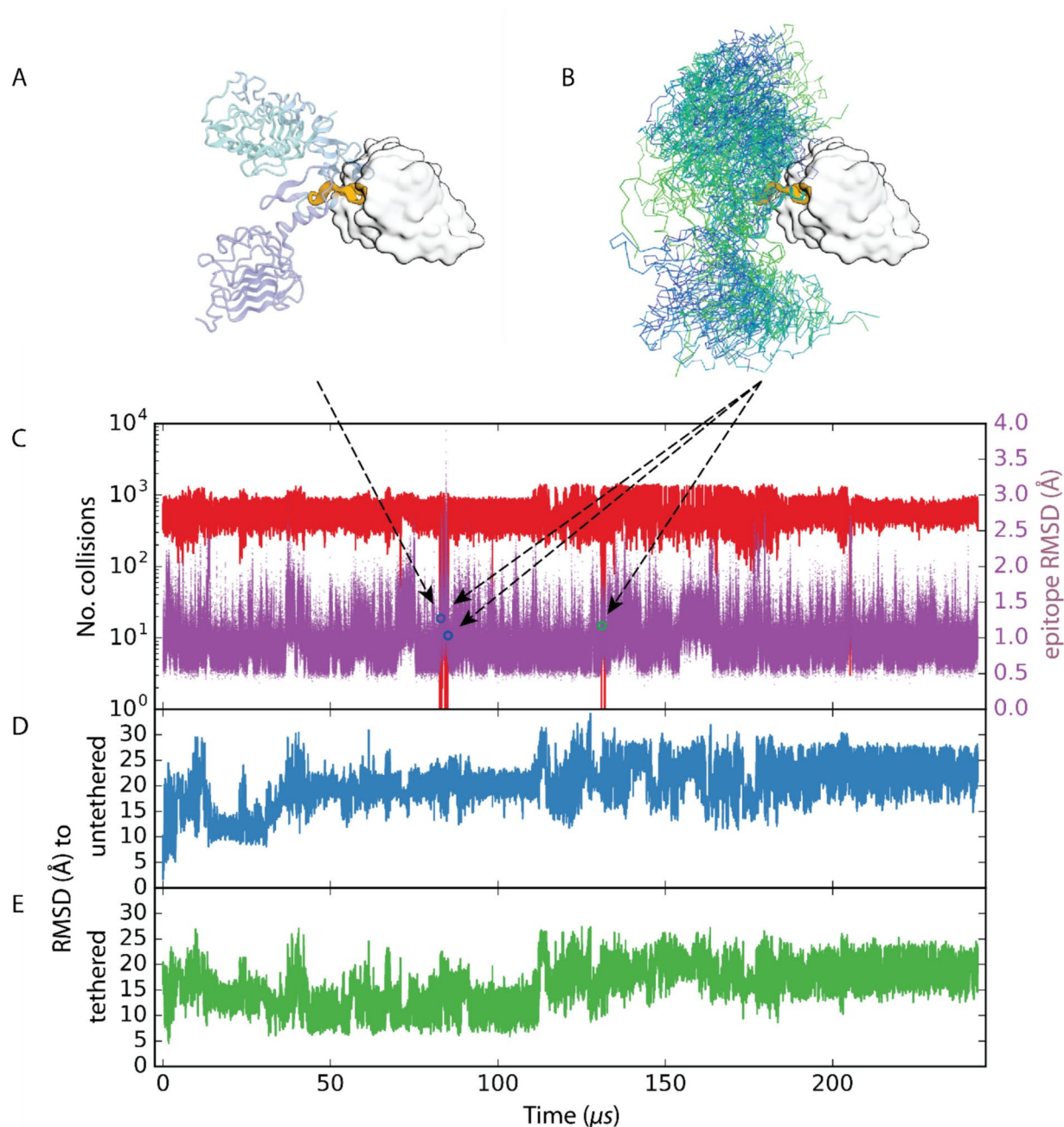


Fig. 2. Local unfolding of EGFR permits mAb806 binding. We considered an EGFR conformation to be binding-competent if it exhibited no steric clashes with a docked mAb806, and the Ca RMSD of the epitope from its conformation in the crystal structures of the mAb806-epitope complex was no greater than 1.5 Å. **(A)** Example of a binding-competent conformation sampled from simulation 2 (using simulated tempering). EGFR is shown as a ribbon diagram, with its I, II, III domains colored as in Fig. 1A. The epitope is highlighted as orange. The docked mAb806 Fab is shown by its molecular surface, in white with black contours. **(B)** Multiple locally unfolded, binding-competent EGFR conformations (shown in backbone trace, colored serially in gradient from blue to green) are sampled in the same MD simulation. **(C)** The number of EGFR atoms in collision with a docked mAb806 (red) and the Ca RMSD of the epitope from its binding conformation (purple) in the simulation. The binding-competent conformations are highlighted by circles colored serially as in panel (B). Also shown are the Ca RMSD from the crystal structures of **(D)** the untethered conformation and **(E)** the tethered conformation in the simulation. The binding-competent conformations are different from both the tethered and untethered conformations.

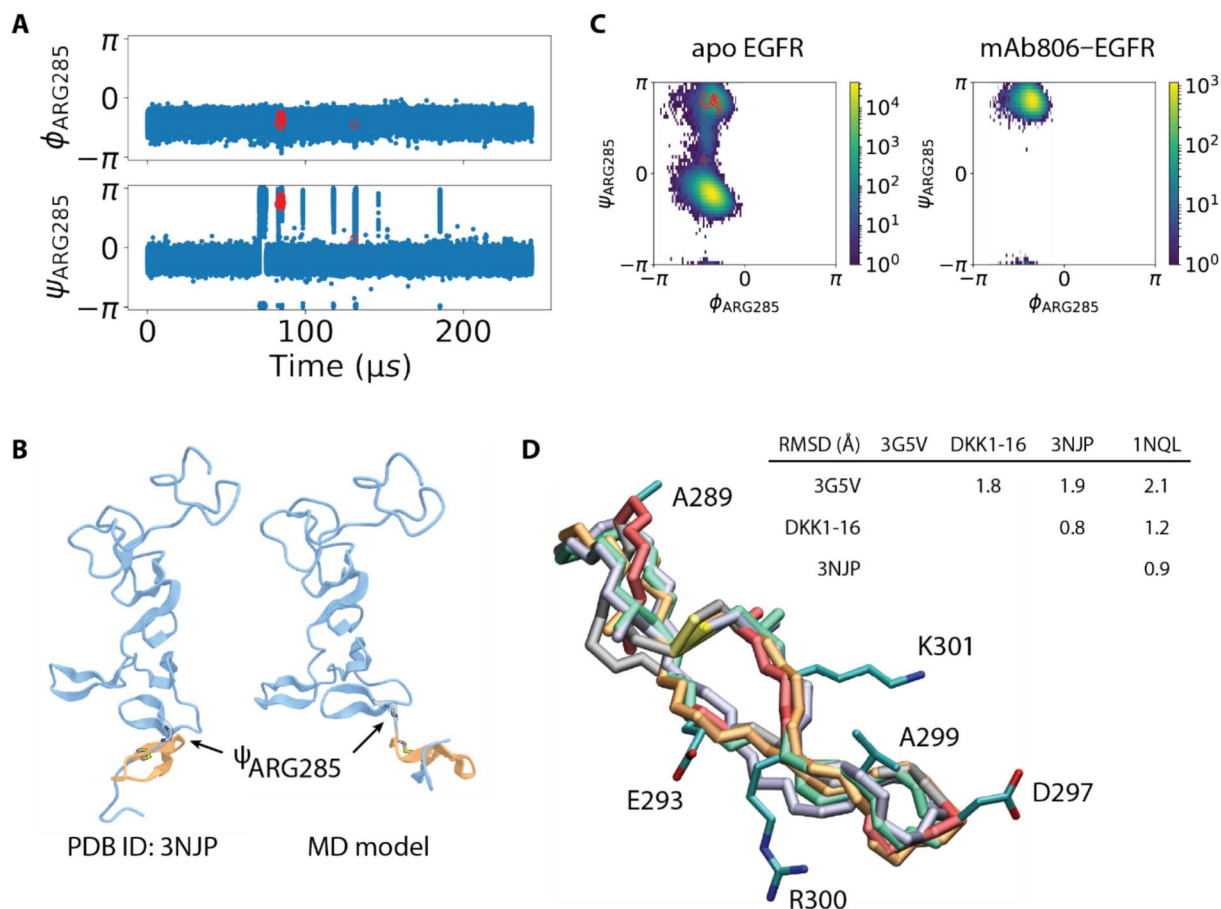
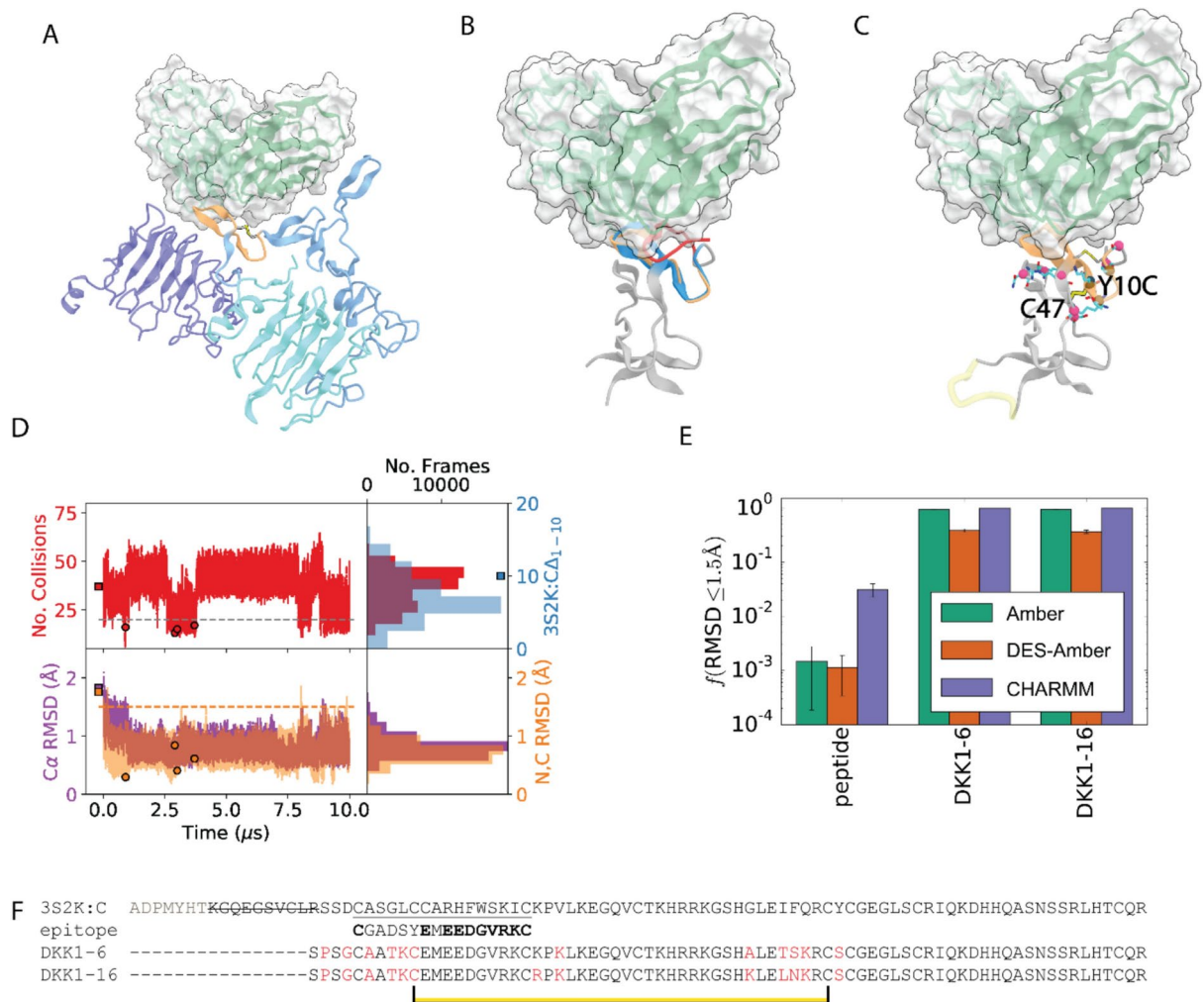


Fig. 3. Computational model of the mAb806-binding conformation of EGFR. **(A)** Time trace of the backbone ϕ and ψ dihedral angles of residue Arg285 in simulation 2. The snapshots in which the mAb806 Fab could be docked onto the epitope without steric clash are indicated by red circles. The epitope was only accessible when the ψ dihedral departed from its value in the native conformations. **(B)** Domain II in the native, untethered conformation, and in the mAb806-accessible conformation generated in simulation 2. In the mAb806-binding conformation, the epitope (orange) flips around Arg285 and breaks its contact with the preceding β -sheet (Tyr275 to Arg285), exposing itself. **(C)** The Ramachandran plots of the backbone ϕ, ψ dihedral angles of Arg285 in the simulation of the apo EGFR ectodomain (simulation 1) and in the simulation of the mAb806 Fab-EGFR complex (simulation 5). The bound Fab prevents Arg285 from flipping back to the native conformation. **(D)** The mAb806-epitope in crystallographic and modeled conformations. The backbone traces of different conformations are colored as follows: the untethered conformation (PDB ID: 3NJP) in light green, the tethered conformation (PDB ID: 1NQL) in ice blue, the epitope peptide in complex with mAb806 (PDB ID: 3G5V) in gray, with the residues in contact with mAb806 highlighted in pink, and the epitope in the designed DKK1-16 in orange. The epitope structure in DKK1-16 is taken from a snapshot in the simulation of the mAb806 Fab-EGFR complex (simulation 4), except for the side chains of the redesigned residues. Sidechains in contact with mAb806 are shown as sticks. The RMSD between Ca atoms in the different conformations are shown in the table. The structure of the 16-mer epitope peptide bound to mAb806 has notable differences from the structures of the epitope in EGFR.

general, including a conformational ensemble of an epitope from MD simulations can increase the chance of successfully identifying a suitable grafting scaffold.

Visual inspection of the epitope-grafted scaffold led us to introduce an additional disulfide bond by changing the Tyr at residue 10 of the immunogen to Cys (Fig. 4C). Subsequently, 17 mutants were computationally generated using RosettaRemodel²⁸ in order to resolve steric clashes and form new specific interactions that would stabilize the designed protein conformation (Fig. 4C). For each designed mutant, MD simulations totaling 10 to 100 μs of simulation time were carried out using three different force fields (see Methods). For the majority of the mutants, the epitope retained or predominantly adopted the mAb806-binding conformation in our simulations (Figure S3B). After visual inspection of the structures of the designs, we chose to pursue two, DKK1-6 and DKK1-16 (Fig. 4F), for additional experimentation.

Both designed proteins expressed well in Sf21 cells. The purified proteins were characterized by 1D proton NMR and showed signatures indicating folded proteins (Figure S4A). The similarity between the spectra of DKK1-6 and DKK1-16 suggests that the two proteins adopt the same fold (Figure S4A). The proteins were



induced to unfold by the reduction of their disulfide bonds, suggesting that those disulfide bonds are important in maintaining the proteins' correct fold (Figure S4B).

We characterized the binding of the designed immunogens against ch806, a chimeric version of mAb806 with a human fragment crystallizable (Fc) region. As measured by surface plasmon resonance (SPR), both DKK1-6 and DKK1-16 had substantially higher association rate constants for binding to ch806 than the free epitope peptide [Cyclo(1,16)] H₂N-CGADSYEMEEDGVRKC-OH (Figs. 5A and S5). Their association rate constants ($> 10^6 \text{ M}^{-1} \text{ s}^{-1}$) are in the range for diffusion-limited antibody-antigen binding,²⁹ suggesting that the epitope in the designed immunogens is stabilized in the ch806-binding conformation,³⁰ consistent with our MD simulation results (Fig. 4E, Fig. 5A inset). Microscale thermophoresis (MST) was then used to confirm that the designed immunogens have higher affinity than the free epitope peptide for ch806 in bulk solution, suggesting that the improved binding measured by SPR was not an artifact of binding to the surface-attached antibody (Fig. 5B). DKK1-16 was also shown to form a stable complex with the mAb806 Fab (Figs. 5, C and D). Taken together, these results suggest that the designed immunogens likely stabilize the epitope in the mAb806-binding conformation.

We next tested the immunogenicity of the designed immunogens. Given their small size and stabilizing internal disulfide bonds, we expected that the designed immunogens would be conformationally stable and would not denature during the immunization protocol. Rabbits were immunized with DKK1-6, DKK1-16, and the epitope peptide ([Cyclo(1,16)] H₂N-CGADSYEMEEDGVRKCGGS(KAoa)-amide) conjugated to keyhole limpet hemocyanin (KLH), a carrier protein (Fig. 6A), and four bleeds in total were collected. For each rabbit, we assessed the antisera from each bleed, as well as polyclonal antibodies (pAbs) affinity-purified from the third-bleed sera, by enzyme-linked immunosorbent assay (ELISA). Compared to the rabbits immunized with peptide-KLH, the rabbits immunized with DKK1-6 and DKK1-16 developed weaker antisera, and developed them more slowly, against both the eliciting immunogen and the epitope peptide (Figs. 6B). Although all of the purified pAbs had high affinities against the eliciting immunogens (Fig. 6B), the antisera from the rabbits immunized with the designed immunogens generally had lower affinities than the antisera from the rabbits immunized with peptide-KLH conjugate (Fig. 6C), which might be attributable to a lower abundance of elicited antibodies resulting from a weaker immune response. The pAbs from rabbit 2842, which was immunized with peptide-KLH conjugate, were affinity-purified using the epitope peptide and showed high affinity for the epitope peptide. In contrast, the pAbs elicited by, and affinity-purified using, the designed immunogens showed much

Fig. 4. The design of immunogens. (A) Model of the mAb806-EGFR ectodomain complex, generated from an MD simulation, used as the starting point for immunogen design. The 16-residue epitope is highlighted in orange, while domains I, II, and III of EGFR are shown as a ribbon diagram and colored as in Fig. 1A. The Fab of mAb806 is shown as a ribbon diagram overlaid with its molecular surface. (B) The C-terminal domain of human DKK1 (PDB ID: 3S2K, chain C), shown as a ribbon diagram and colored gray, was identified as a potential scaffold, allowing four epitope conformations—taken from an MD simulation of the mAb806-EGFR complex, such as that in (A)—to be superimposed onto the hairpin structure of residues C14–C29 (blue). Visual inspection suggests that the first ten N-terminal residues in the scaffold, shown in red, sterically clash with the antibody and should be deleted from the graft. (C) Visual inspection suggests that the mutation Y10C may introduce a disulfide bond, shown as yellow sticks, with Cys47. RosettaDesign was used to identify additional mutations that might stabilize the conformation of the designed immunogen. The residues mutated in the final design of DKK1-16 are highlighted by purple spheres. The loop that was missing in the crystal structure but modeled by Rosetta is shown in yellow. (D) The design was enabled by the MD simulation, which yielded EGFR-mAb806 Fab conformations in which the epitope could be grafted into the scaffold protein DKK1. The top panel shows the steric clashes between the scaffold protein DKK1 with the epitope-mAb806 complex, when the latter is superimposed onto the former by aligning the Ca atoms of the epitope onto the Ca atoms of the replacement segment C14–C29. The time trace is shown on the left, and corresponding histogram on the right. The light blue bars in the histogram on the right show the number of atoms that would clash if the first ten N-terminal residues in DKK1 were removed (right y-axis). The bottom panel shows the Ca RMSD (purple, left y-axis) and the N, C, and Ca RMSD values (orange, right y-axis) of the N- and C-terminal residues between the epitope and the replacement segment. The conformational snapshots that were successfully grafted onto the DKK1 scaffold protein are highlighted in circles. The values corresponding to the crystal structure of the epitope-mAb806 complex (PDB ID: 3G5V) are shown as squares. (E) In our MD simulations of the designed immunogens DKK1-6 and DKK1-16 using three different force fields (Amber: Amber99SB*-ILDN + TIP3P water model; DES-Amber: DES-Amber SF1.0 + TIP4P-D water model; CHARMM: CHARMM22* + TIP3P-CHARMM water model), the epitope mostly occupies the mAb806-binding conformation (Ca RMSD from the crystal structure ≤ 1.5 Å), whereas the free epitope peptide (cyclized due to the C1–C16 disulfide bond) rarely visits the binding conformation. The error bars were generated by dividing each simulation into five non-overlapping time intervals and computing the standard deviations between those intervals. (F) Amino acid sequences of the two designed immunogens DKK1-6 and DKK1-16 aligned to that of the scaffold DKK1 protein and the epitope C287–C302 in EGFR. The gray letters in the N-terminus of the scaffold protein indicate the unresolved residues in the crystal structure (PDB ID: 3S2K), and the struck-through letters signify the ten residues in the scaffold protein that were deleted in the designed immunogens to avoid steric clashes with mAb806. The bold-faced residues in the epitope form specific interactions with mAb806 (except for the two terminal cysteines, which form the critical disulfide bond), and were not considered for mutation in the remodeling step. The residues mutated in the remodeling step are highlighted in red. The connection between Cys10 and Cys47 signifies the manually designed disulfide bond.

weaker affinity for the epitope peptide (Fig. 6D). This might be partly because only a small fraction of the pAbs were targeting the epitope, while the majority of the pAbs were targeting other parts of the immunogen.

We next tested how strongly the antisera and pAbs could bind to a set of five different tumor cell lines, including two that express normal levels of EGFR (U87MG and A549), two that overexpress EGFR (A431 and U87MG^{wt} EGFR; U87MG^{wt} EGFR is a stable engineered U87MG cell line overexpressing EGFR, Figure S6), and a stable engineered U87MG cell line expressing EGFRvIII (U87MG^{EGFRvIII}, Figure S6). We validated the EGFR profile of these cells by measuring ch806 binding to them: Both fluorescence activated cell sorting (FACS) and On-Cell Western assays showed that ch806 selectively binds to the latter three cell lines, with undetectable binding to the first two (Figs. 7A and S7A). Quantitative analysis of the On-Cell Western titration curves (Figure S7B) for ch806 and cetuximab,³¹ a monoclonal antibody that binds to normal EGFR, suggests that ch806 binds to a subpopulation of EGFR on the cell surface (Figure S7C). These results are in agreement with previous experimental findings.^{19,21}

All the antisera and pAbs from our immunizations bound to the A431, U87MG^{wt} EGFR, and U87MG^{EGFRvIII} cells with much weaker affinity than did ch806 (Fig. 7). The weak binding of the antisera and pAbs from rabbit 2842 to these cells contrasts with their high affinities for the epitope peptide, suggesting that the pAbs elicited by the unstructured epitope peptide do not recognize the mAb806-binding conformation of the epitope in EGFR. Notably, the weak binding of all the pAbs to U87MG^{wt} EGFR was only partially inhibited by ch806 (Fig. 7B insets), which suggests that a subpopulation of the antibodies might have bound to a partial epitope that is exposed in normal EGFR, and that their binding did not compete with that of ch806, which primarily binds to misfolded EGFR. (It is unlikely that pAbs elicited by the designed immunogens bound to U87MG cells because of antibodies reactive against DKK1, as U87MG has been shown to have undetectable expression of the *dkk1* gene³²). Nevertheless, our finding that ch806 partially inhibits the binding of the pAbs indicates that at least a subpopulation of the antibodies is competing with ch806 for the same epitope.

Most encouraging were the results for rabbit 2759, which was immunized with the designed immunogen DKK1-16. Its antisera and purified pAbs, in addition to binding to the epitope peptide (Fig. 6, B and D), showed moderate preference in binding to cell lines overexpressing EGFR or expressing the mutant EGFRvIII over cell lines with normal EGFR expression (Fig. 7, A and B), and its pAbs compete against ch806 in binding to cells overexpressing EGFR (Fig. 7B, inset). Taken together, these results suggest that the immunogen elicited the

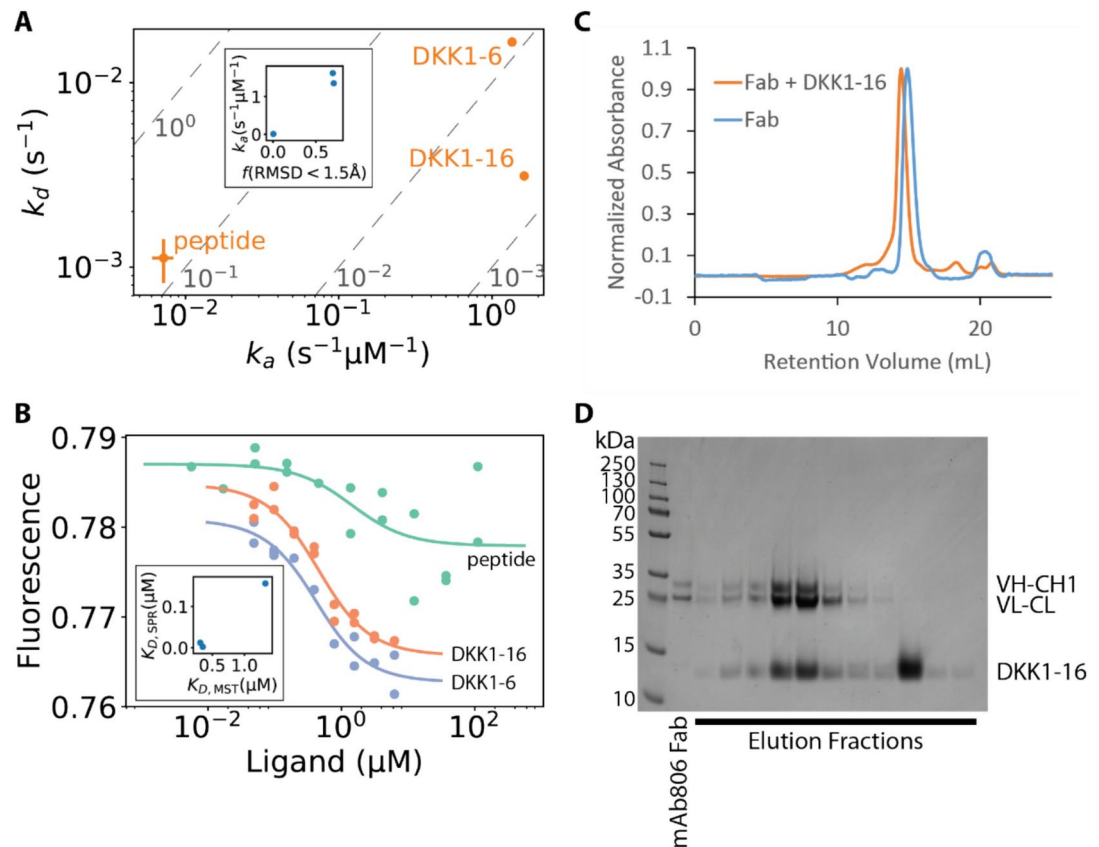


Fig. 5. Experimental characterization of the binding of designed immunogens to the ch806 antibody. **(A)** Surface plasmon resonance measurements show that the designed immunogens DKK1-6 and DKK1-16 bind to ch806 orders-of-magnitude faster than the free epitope peptide, suggesting that the epitopes in the designed immunogens have been stabilized in the binding-competent conformation. (k_a , k_d and K_D values are presented in Table S1.) The binding affinities are constant along the dashed diagonal lines, with the corresponding equilibrium dissociation constants, K_D , indicated in gray. The high conformational stability of the epitope in the designed immunogens, as estimated from MD simulations (Fig. 4E), is consistent with the fast association rate constants of mAb806 binding determined by SPR (inset). **(B)** Microscale thermophoresis (MST) measurements of the ch806-binding affinities for the epitope peptide (green, $K_D = 1.33\text{ }\mu M$), DKK1-6 (blue, $K_D = 0.31\text{ }\mu M$), and DKK1-16 (orange, $K_D = 0.35\text{ }\mu M$). The inset compares the K_D values determined by SPR and MST. **(C)** Co-elution of mAb806 Fab and DKK1-16 in analytical size-exclusion chromatography. The association between the Fab and the designed immunogen led to a shift in the retention volume that corresponds to a difference in molecular weight of $\sim 10\text{ kDa}$, in agreement with the molecular weight of DKK1-16. **(D)** The Fab and designed immunogen in the co-elution peak were confirmed by SDS-PAGE. VH-CH1 and VL-CL are the heavy and the light chains, respectively, of the Fab. These results demonstrate that the designed immunogen indeed forms a stable complex with mAb806 in solution.

desired, albeit weak, response: the development of mAb806-like antibodies that recognize the target epitope in misfolded EGFR.

Discussion

MD simulations have been used previously to identify cryptic binding pockets that arise from conformational plasticity of the protein^{33,34}. In the present study, we propose that they might also be useful in identifying cryptic epitopes associated with aberrant protein conformations. Our MD simulations of EGFR in complex with the antibody mAb806 have suggested a plausible, atomistically detailed model of how mAb806 selectively binds to misfolded EGFR. Using epitope conformations sampled by MD simulations, we identified a suitable scaffold—one that could not have been found using the crystal structure alone—on which we have designed immunogens that present the mAb806-recognizing epitope in its antibody-binding conformation. Structural stabilization of such immunogens might be necessary in order to elicit mAb806-like antibodies, as the epitope peptide on its own is unstructured. This is supported by the results of our preliminary immunization experiment, in which antibodies elicited by the epitope peptide had high affinity for the peptide, but only weakly bound to overexpressed EGFR or EGFRvIII on tumor cell lines.

Despite being stabilized, our designed immunogens, DKK1-6 and DKK1-16, appeared to suffer from weak immunogenicity and an inability to focus the antibody response against the target epitope, potentially in part

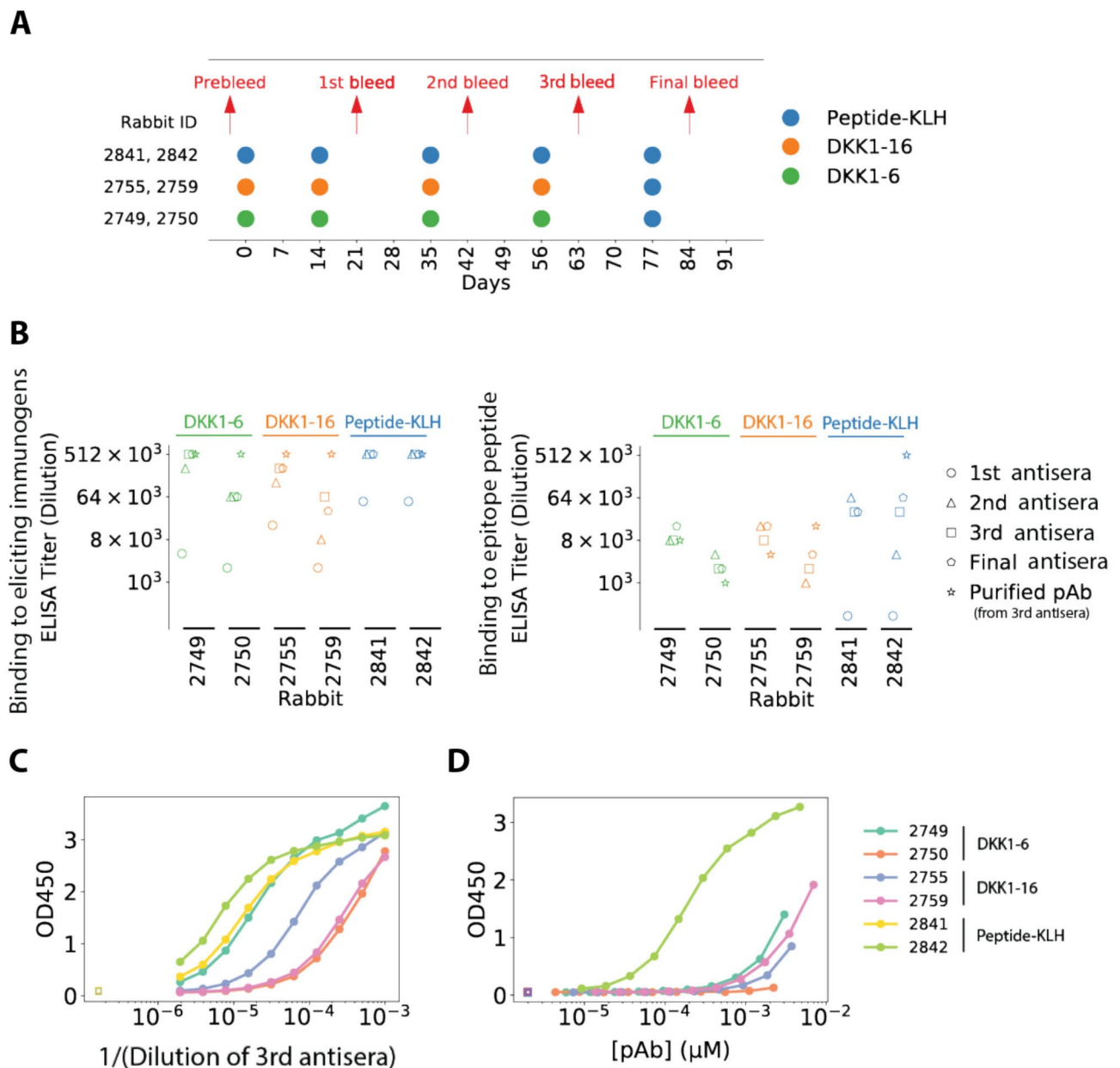


Fig. 6. The designed immunogens are capable of eliciting antibodies that bind to the epitope peptide. **(A)** Rabbit immunization protocol. Two rabbits were immunized with each immunogen on the indicated days (rabbits 2749 and 2750 received DKK1-6, rabbits 2755 and 2759 received DKK1-16, and rabbits 2841 and 2842 received peptide-KLH). Pre-immunization sera (prebleed) were obtained three days prior to the first immunization, and antisera were obtained on the indicated days. Polyclonal antibodies were affinity-purified from the third bleed. All rabbits received peptide-KLH as the final boost to stimulate antibody response against the mAb806 epitope. **(B)** ELISA analysis of the antisera and purified pAbs. The markers indicate the titers of binding to the eliciting immunogen (left) or to the epitope peptide [Cyclo(5,20)] H_2N -SGGGCGADSYEMEEDGVRKC-OH (right). The titer is the highest dilution at which the readouts are at least twice that of the blank (antibody-free) solution. Eliciting immunogens are indicated above the plots. The final antisera and the purified pAbs from the rabbits immunized with the designed immunogens all showed clear but weak binding to the epitope peptide, compared to much stronger binding to the eliciting immunogen. **(C)** ELISA readout of binding to the eliciting immunogens at different dilutions of the third-bleed antisera (from which the pAbs were purified). The humoral immune response to the peptide-KLH immunogen was robust in both rabbits, whereas the response to the designed immunogens tended to be weak (with the exception of rabbit 2749). **(D)** ELISA readout of binding to the epitope peptide at different concentrations of the purified polyclonal antibodies. The pAbs elicited by the epitope-peptide KLH-conjugate and affinity purified by the epitope peptide bound to the epitope peptide with high affinity, whereas those elicited and affinity purified by the designed immunogens had much lower affinity for the epitope peptide.

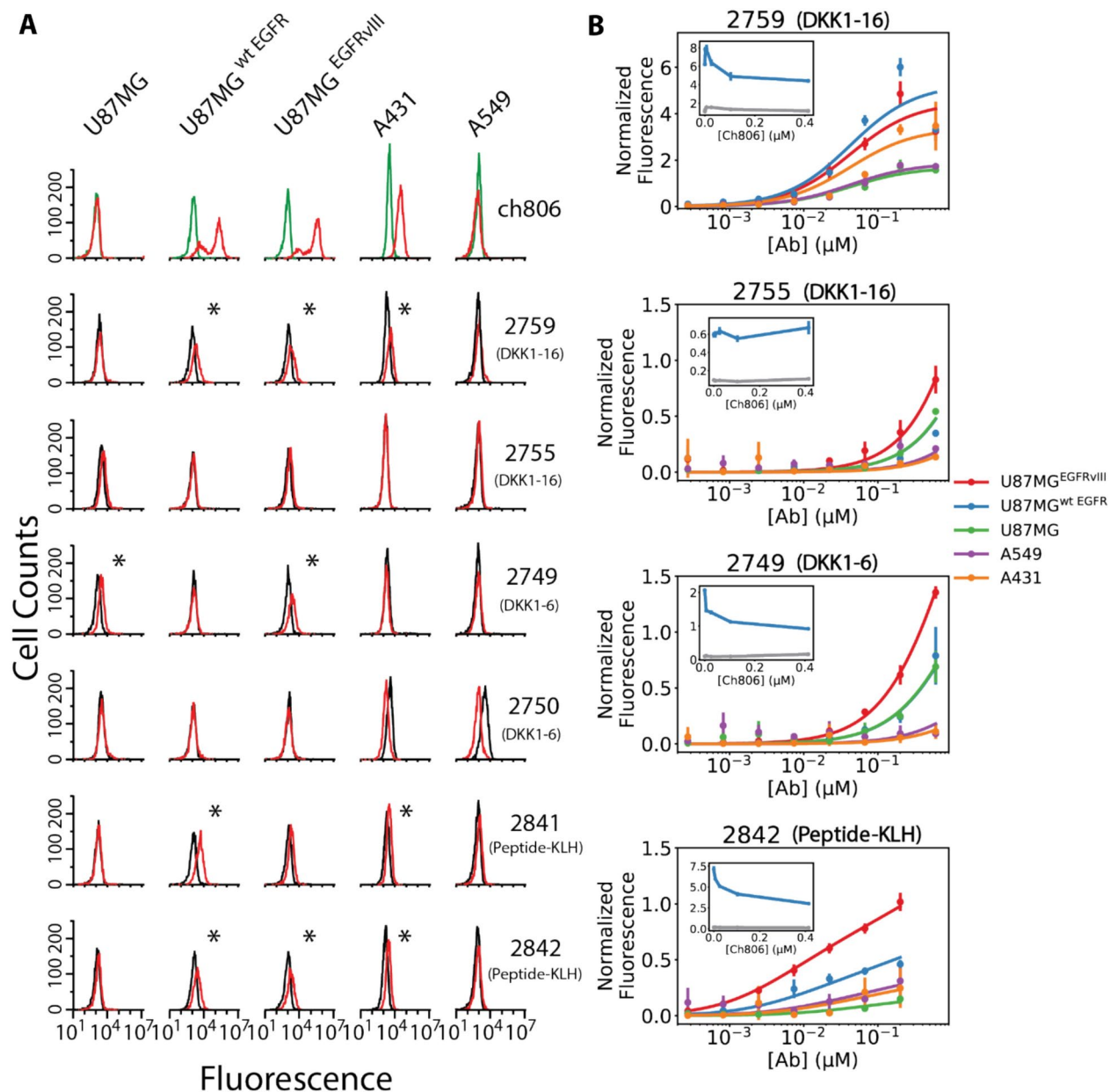


Fig. 7. Rabbit polyclonal antibodies elicited by the epitope-peptide keyhole limpet hemocyanin (KLH) conjugate and the designed immunogens bound weakly, and in some cases not at all, to tumor cells. **(A)** FACS analysis of the final antisera from the indicated rabbits (with eliciting immunogens in parentheses) binding to tumor cell lines. None of the antisera exhibited as high a binding affinity to any of the cell lines as the monoclonal ch806 binding to the three cell lines U87MG^{wt} EGFR, U87MG^{EGFRvIII}, and A431. Nevertheless, small shifts in the fluorescence distributions—indicative of some binding—were observed for some pairs of cell lines and antisera, as indicated by the asterisks. In particular, the antisera from rabbit 2759 and 2842 exhibited binding to the same three cell lines as ch806. The red lines are the binding data for the monoclonal antibody ch806 or the antisera, the green lines are those for human IgG1 control (Biolegend, Catalog No. 403102), and the black lines are for the pre-immunization bleed of the corresponding animals. **(B)** For the indicated rabbits (eliciting immunogens in parentheses), characterization of the binding of purified pAbs to the tumor cell lines by an On-Cell Western assay. All pAbs demonstrated some binding to the cell line expressing EGFRvIII mutant, suggesting that they targeted an epitope that was more exposed in this mutant. All pAbs elicited by the designed immunogens, however, showed binding only at high concentrations, consistent with their low ELISA titers against the epitope peptide. Insets (in which the y-axis is also normalized fluorescence): Ch806 was able to compete against the binding of the polyclonal antibodies from rabbits 2759, 2749, and 2842 to U87MG^{wt} EGFR cells, suggesting that these polyclonal antibodies contain antibody molecules that bind to the same EGFR epitope on the tumor cells as the ch806 antibody. The rabbit pAb concentrations in these competition experiments were 67 nM (blue) and 0 nM (gray). The inhibition of binding by the ch806 antibody was only partial even at high ch806 concentrations, suggesting that the pAbs may have recognized a partial segment of the ch806-epitope that is inaccessible in normal EGFR conformations.

because some of the antibodies recognized the scaffold portion of the immunogen. It is also possible that the germline antibodies in rabbits are less likely than those in mice to develop through affinity maturation to recognize the epitope of mouse-derived mAb806 (or mAb175, which has a similar sequence and is known to bind the same epitope²³); different animal species do not always produce antibodies that recognize the same epitope. In addition, it remains unclear whether polyclonal antibodies that recognize the target epitope will also bind to it in misfolded EGFR (in the case of pAbs from rabbit 2759, this appears true), as antibodies may approach the epitope from different angles, some of which may be sterically excluded in this context. Nonetheless, our designed immunogens successfully presented the mAb806-epitope in its antibody-binding conformation, and may thus offer a useful starting point for further development. Improvements in the immunogen designs—for example designing larger protein immunogens that incorporate multiple units of the cryptic epitope—along with modifications to the immunogen composition^{35–37} and immunization protocols^{38,39} could enhance immunogenicity and focus the antibody response against the epitope,¹¹ potentially leading to immunogens that can overcome the obstacles outlined above and robustly elicit mAb806-like, cancer-selective antibodies.

Methods

Molecular dynamics simulations

For the starting structures of our EGFR simulations (simulations 1, 2, 5, and 6), four frames were selected randomly from an equilibration simulation of an untethered EGFR ectodomain crystal structure (PDB ID: 3NJP); the double mutation C271A/C283A was introduced for simulations 5 and 6. The CR2 domain (also known as domain IV; residues Val481 through Thr614) was omitted from the ectodomain in these simulations to reduce the size of the molecular system and accelerate conformational dynamics. mAb806 was docked onto EGFR in frames selected at a regular time interval from simulations 1, 2, 5, and 6, by aligning the Ca atoms of the epitope in the crystal structure of the mAb806-epitope complex (PDB ID: 3G5V) onto the epitope in the simulated EGFR conformation. We then used selected frames from simulation 2 to generate a structural model of the mAb806-EGFR complex, from which we initiated two additional simulations (simulations 3 and 4). In all six of these simulations, the DES-Amber SF1.0 force field⁴⁰ was used for the protein, together with the TIP4P-D water model.⁴¹ Simulations 1, 2, 3, 5, and 6 used Times Square sampling⁴² (an adaptive variant of simulated tempering^{43,44}) so that the systems traversed a temperature range between 310 and 350 K, thereby accelerating sampling of the protein conformations; simulation 4 was run at a constant temperature of 310 K. All systems were maintained at a pressure of 1 atm.

Isobaric-isothermal simulations for the free epitope peptide and each of the designed immunogens were performed at a temperature of 310 K and a pressure of 1 atm, using three different force fields: 1) amber99SB*-ILDN⁴⁵ (which builds on other modifications^{46,47} to Amber99⁴⁸) with the TIP3P water model,⁴⁹ 2) DES-Amber SF1.0 with the TIP4P-D water model, and 3) Charmm22*⁵⁰ with the TIP3P-CHARMM water model.⁵¹

In all simulations, bond lengths to hydrogen atoms were constrained using an implementation⁵² of M-SHAKE,⁵³ electrostatic interactions were computed using the *u*-series method,⁵⁴ and Lennard-Jones interactions were calculated using a cutoff of 12 Å. The pressure and temperature were controlled using the Martyna-Tobias-Klein (MTK) algorithm⁵⁵ and the Nosé-Hoover thermostat,⁵⁶ respectively, applied using the multigrator framework.⁵⁷ The RESPA multiple time scale integrator⁵⁸ was used to propagate the dynamics, updating the bonded, Lennard-Jones, and short-range electrostatic forces every 2.5 fs, and the long range electrostatic forces every 5 fs. All simulations were performed on the Anton 2 specialized supercomputer.⁵⁹

Expression and purification of DKK1-6 and DKK1-16

DKK1-6 and DKK1-16 constructs, including an N-terminal His6-SUMO tag, were expressed in Sf21 insect cells grown in Insect-XPRESS medium (Lonza) at 27 °C for 66 h. To facilitate disulfide bond formation, the proteins were secreted in the medium. After clarification of the medium by centrifugation, the proteins were isolated using HisPur Ni-NTA resin (ThermoFisher) equilibrated in buffer A (20 mM Tris/HCl, 500 mM NaCl, 20 mM imidazole, pH 7.1). The resin was washed extensively with buffer A, followed by elution with buffer B (20 mM Tris/HCl, 500 mM NaCl, 500 mM imidazole, pH 7.1). Subsequently, the proteins were dialyzed into buffer C (20 mM Tris/HCl, 50 mM NaCl, pH 7.1) in the presence of Ulp1 protease, in order to cleave the His6-SUMO tag. The cleaved proteins were further purified by cation exchange chromatography using HiTrap SP HP column (GE Healthcare) equilibrated in buffer C. After washing with buffer C, the proteins were eluted by applying a linear gradient from 0 to 100% of buffer D (20 mM Tris/HCl, 1 M NaCl, pH 7.1) in 30 column volumes. Elution fractions containing pure DKK1-6 and DKK1-16, as determined by SDS-PAGE, were deglycosylated by incubation with PNGase F (New England BioLabs) at 37 °C for one hour. Deglycosylated proteins were purified to homogeneity by size exclusion chromatography using a HiLoad Superdex 75 PG column (GE Healthcare) equilibrated in Dulbecco's Phosphate buffered saline (DPBS; Gibco). Subsequently, endotoxins were removed using Pierce High-Capacity Endotoxin Removal Resin (Thermo Scientific). The proteins were flash frozen in liquid nitrogen and stored at a concentration of approximately 10 mg/mL at –80 °C.

Stable cell lines

Stable cell lines of U87MG^{wt EGFR} and U87MG^{EGFR^{III}} were produced by Genscript under a research contract. Briefly, the full-length cDNA of human EGFR and mutant EGFR^{III} were synthesized and subcloned into the pLenti-puro vector, and the recombinant lentivirus was packaged in 293T cells. U87MG cells were transduced with the recombinant lentivirus, centrifuged at 800 g for 120 min at 20 °C, incubated for 24 h, and then cultured in fresh complete growth medium for an additional 24 h. The transduced cells were selected for four days in medium containing 0.75 µg/mL (for U87MG^{wt EGFR}) or 1 µg/mL (for U87MG^{EGFR^{III}}) puromycin. Elevated expression levels of EGFR and EGFR^{III} were confirmed by quantitative polymerase chain reaction (qPCR) and by western blot analysis (Figure S6). U87MG, A431, and A549 cell lines were obtained from ATCC.

Surface plasmon resonance

SPR binding data were collected on a ProteOn XPR36 biosensor using GLM sensor chips (Bio-Rad). Antibody ch806, which was produced by LakePharma (now part of Curia), was captured onto a Protein A-coated sensor at six different surface densities, starting with a 1:50 dilution of the stock concentration of 1.57 mg/ml, and then in a threefold dilution series. All studies were run in 10 mM HEPES, 150 mM NaCl, 0.005% tween-20, 3.4 mM EDTA, and 0.2 mg/mL BSA at pH 7.4 and 25 °C. The epitope peptide, DKK1-6, and DKK1-16 were tested at a flow rate of 100 µL/min in a three-fold concentration series up to the highest concentrations of 4 µM, 1 µM, and 333 nM, respectively; these are the highest concentrations for each at which no significant non-specific binding was observed. Between testing of the different analytes (i.e., epitope peptide, DKK1-6, and DKK1-16), the sensor surface was regenerated and ch806 recaptured. Interspot referenced response data from six different surface densities of antibodies were globally fit to a 1:1 binding stoichiometry to extract estimates of the binding rate constants.

Microscale thermophoresis

The fluorescent label AlexFluor-647 (ThermoFisher) was covalently attached by N-hydroxysuccinimide (NHS) coupling to the N-terminal primary amine of the ch806 antibody. The experiments were performed on the Monolith NT.115 instrument using the RED detector, with the ch806 concentration at 70 nM. The epitope peptide was tested in a 12-point, three-fold serial dilution beginning at a 1000 µM peptide concentration, whereas DKK1-6 and DKK1-16 were tested in a 12-point, two-fold serial dilution beginning at 72.5 µM and 54.5 µM, respectively. The buffer used in the assays contained 20 mM HEPES and 150 mM NaCl, with pH 7.5. After a 30-min incubation, the samples were loaded into MST NT.115 standard glass capillaries and a microscale thermophoresis (MST) analysis was performed using 60% LED power and 40% MST power. At high concentrations, analytes can stick to the capillary walls, leading to artificial changes in the thermophoresis. We thus omitted the data points at high concentrations (> 111 µM for the epitope peptide, > 9.4 µM for DKK1-6, and > 6.7 µM for DKK1-16) from our estimations of the binding affinities.

One-dimensional NMR spectroscopy

Proteins were measured in DPBS buffer with the addition of 10% D₂O. Proton 1D spectra were acquired at 25 °C. Water suppression sequence (Watergate) was used to eliminate water resonance at 4.7 ppm. 128 scans were recorded for each spectrum. For TCEP-reduced samples, the proteins were incubated with 2 mM TCEP at room temperature for 20 min prior to measurements.

Immunization of rabbits

Two New Zealand rabbits were immunized with each immunogen. 1 mL of serum from each rabbit was collected via ear vein three days before the first immunization (prebleed). In the first immunization, 200 µg of immunogen, mixed with Complete Freund's Adjuvant (CFA), was injected subcutaneously at four spots on the rabbit's back. In subsequent immunizations, 200 µg of immunogen, mixed with Incomplete Freund's Adjuvant (IFA), was similarly injected. 100 µL of antisera were drawn from each rabbit via ear vein one week after the second and third immunizations for tests of antigen-binding by ELISA. Polyclonal antibodies were affinity-purified from the 10–15 mL of antisera drawn one week after the fourth immunization; the antisera from rabbits immunized by the designed antigens were purified using the corresponding protein immunogens coupled to a Sepharose 4B column, whereas the antisera from rabbits immunized by epitope-peptide KLH-conjugate were purified using the epitope peptide [Cyclo(5,20)] H₂N-SGGGCGADSYEMEEDGVRKC-OH coupled to a Sepharose 4B column. The antibody concentrations were determined by absorbance at 280 nm with a NanoDrop 2000 spectrophotometer.

Enzyme-linked immunosorbent assay (ELISA)

ELISA plates were coated by incubating with 100 µL per well of antigen solutions (pH 7.4) at a concentration of 4 µg/mL for two hours at 37 °C or overnight at 4 °C. The plates were then washed three times with 200 µL of washing buffer (PBS buffer containing 0.05% Tween 20), and incubated with 200 µL blocking buffer (PBS buffer containing 0.05% Tween 20 and 0.01 g/mL bovine serum albumin) for two hours at 37 °C to block non-specific binding in the coated wells. The antisera or purified polyclonal antibodies were diluted with the above blocking buffer, 100 µL of the solution was added to each well, and the plate was incubated for one hour at 37 °C or overnight at 4 °C. The plate was washed three times with 200 µL of washing buffer. Horseradish peroxidase-conjugated anti-rabbit IgG antibody was diluted in the blocking buffer, 100 µL of the solution was added to each well, and the plate was incubated for 30 min at 37 °C. The plate was washed five times with 200 µL of washing buffer. 100 µL of TMB reagent was added to each well, and after sufficient color development (15–30 min), 100 µL of stopping buffer (1 M HCl) was added to each well. The absorbance was measured at 450 nm with a plate reader.

Data availability

The molecular dynamics (MD) trajectories described in this work are available for non-commercial use through contacting trajectories@deshawresearch.com.

Received: 30 May 2024; Accepted: 12 March 2025

Published online: 02 April 2025

References

- Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nat. Rev. Cancer* **12**(4), 278–287 (2012).
- Weiner, L. M., Surana, R. & Wang, S. Antibodies and cancer therapy: versatile platforms for cancer immunotherapy. *Nat. Rev. Immunol.* **10**(5), 317–327 (2010).
- Hansel, T. T., Kropshofer, H., Singer, T., Mitchell, J. A. & George, A. J. T. The safety and side effects of monoclonal antibodies. *Nat. Rev. Drug Discov.* **9**(4), 325–338 (2010).
- Gan, H. K., Burgess, A. W., Clayton, A. H. A. & Scott, A. M. Targeting of a conformationally exposed, tumor-specific epitope of EGFR as a strategy for cancer therapy. *Cancer Res.* **72**(12), 2924–2930 (2012).
- Kamal, A. et al. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* **425**(6956), 407–410 (2003).
- Johns, T. G. et al. Identification of the epitope for the epidermal growth factor receptor-specific monoclonal antibody 806 reveals that it preferentially recognizes an untethered form of the receptor. *J. Biol. Chem.* **279**(29), 30375–30384 (2004).
- Glabe, C. G. Conformation-dependent antibodies target diseases of protein misfolding. *Trends Biochem. Sci.* **29**(10), 542–547 (2004).
- Castro, K. M., Scheck, A., Xiao, S. & Correia, B. E. Computational design of vaccine immunogens. *Curr Opin. Biotechnol.* **78**, 102821 (2022).
- Sesterhenn, F., Yang, C., Bonet, J., Cramer, J. T., Wen, X., Wang, Y., Chiang, C. I., Abriata, L. A., Kucharska, I., Castoro, G., Vollers, S. S., Galloux, M., Dheilly, E., Rosset, S., Corthésy, P., Georgeon, S., Villard, M., Richard, C. A., Descamps, D., Delgado, T., Oricchio, E., Rameix-Welti, M. A., Más, V., Ervin, S., Eléouët, J. F., Riffault, S., Bates, J. T., Julien, J. P., Li, Y., Jardetzky, T., Krey, T., & Correia, B. E. De novo protein design enables the precise induction of RSV-neutralizing antibodies. *Science* (New York, N.Y.) **368**(6492), eaay5051 (2020).
- Correia, B. E., Ban, Y.-E.A., Holmes, M., Xu, H., Ellingson, K., Kraft, Z., Carrico, C., Boni, E., Sather, D. N., Zenobia, C., Burke, K. Y., Bradley-Hewitt, T., Bruhn-Johannsen, J. F., Kalyuzhnyi, O., Baker, D., Strong, R. K., Stamatos, L., & Schief, W. R. Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure* **19**(9), 1116–1126 (2010).
- Correia, B. E. et al. Proof of principle for epitope-focused vaccine design. *Nature* **507**(7491), 201–206 (2014).
- Ofek, G. et al. Elicitation of structure-specific antibodies by epitope scaffolds. *Proc. Natl. Acad. Sci.* **107**(42), 17880–17887 (2010).
- Kulp, D. W. & Schief, W. R. Advances in structure-based vaccine design. *Curr. Opin. Virol.* **3**(3), 322–331 (2013).
- Domitzer, P. R., Ulmer, J. B. & Rappuoli, R. Structure-based antigen design: a strategy for next generation vaccines. *Trends Biotech.* **26**(12), 659–667 (2008).
- Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. & Shaw, D. E. Biomolecular simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.* **41**, 429–452 (2012).
- Casalino, L. et al. Breathing and Tilting: Mesoscale Simulations Illuminate Influenza Glycoprotein Vulnerabilities. *ACS Central Science* **8**(12), 1646–1663 (2022).
- Sikora, M. et al. Computational epitope map of SARS-CoV-2 spike protein. *PLoS Comput. Biol.* **17**(4), e1008790 (2021).
- Guarra, F. & Colombo, G. Computational Methods in Immunology and Vaccinology: Design and Development of Antibodies and Immunogens. *J. Chem. Theory Comput.* **19**(16), 5315–5333 (2023).
- Jungbluth, A. A. et al. monoclonal antibody recognizing human cancers with amplification/overexpression of the human epidermal growth factor receptor. *Proc. Natl. Acad. Sci.* **100**(2), 639–644 (2003).
- Johns, T. G. et al. Novel monoclonal antibody specific for the de2-7 epidermal growth factor receptor (EGFR) that also recognizes the EGFR expressed in cells containing amplification of the EGFR gene. *Int. J. Cancer* **98**(3), 398–408 (2002).
- Johns, T. G. et al. The antitumor monoclonal antibody 806 recognizes a high-mannose form of the EGF receptor that reaches the cell surface when cells over-express the receptor. *FASEB J.* **19**(7), 780–782 (2005).
- Orellana, L., Thorne, A. H., Lema, R., Gustavsson, J., Parisian, A. D., Hospital, A., Cordeiro, T. N., Bernadó, P., Scott, A. M., Brun-Heath, I., Lindahl, E., Cavenee, W. K., Furnari, F. B. & Orozco, M. (2019). Oncogenic mutations at the EGFR ectodomain structurally converge to remove a steric hindrance on a kinase-coupled cryptic epitope. *Proc. Natl. Acad. Sci.* **116**(20), 10009–10018 (2019).
- Garrett, T. P. J. et al. Antibodies specifically targeting a locally misfolded region of tumor associated EGFR. *Proc. Natl. Acad. Sci.* **106**(13), 5082–5087 (2009).
- Ferguson, K. M. et al. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Molecular Cell* **11**(2), 507–517 (2003).
- Lu, C. et al. Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. *Molecular and Cellular Biology* **30**(22), 5432–5443 (2010).
- Ng, Y. Z., Kannan, S., Lane, D. P., Fuentes, G. & Verma, C. S. mAb806 binding to epidermal growth factor receptor: a computational study. *Proteins* **83**(1), 153–168 (2015).
- Azoitei, M. L. et al. Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science* **334**(6054), 373–376 (2011).
- Huang, P.-S. et al. RosettaRemodel: a generalized framework for flexible backbone protein design. *PLoS ONE* **6**(8), e24109 (2011).
- Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* **109**(3), 839–860 (2009).
- Schmidt, A. G. et al. Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc. Natl. Acad. Sci.* **110**(1), 264–269 (2013).
- Goldstein, N. I., Prewett, M., Zuklys, K., Rockwell, P. & Mendelsohn, J. Biological efficacy of a chimeric antibody to the epidermal growth factor receptor in a human tumor xenograft model. *Clin. Cancer Res.* **1**(11), 1311–1318 (1995).
- Shou, J. et al. Human Dkk-1, a gene encoding a Wnt antagonist, responds to DNA damage and its overexpression sensitizes brain tumor cells to apoptosis following alkylation damage of DNA. *Oncogene* **21**(6), 878–889 (2002).
- Kuzmanic, A., Bowman, G. R., Juarez-Jimenez, J., Michel, J. & Gervasio, F. L. Investigating cryptic binding sites by molecular dynamics simulations. *Accounts of Chemical Research* **53**(3), 654–661 (2020).
- Shan, Y. et al. How does a small molecule bind at a cryptic binding site? *PLoS Computational Biology* **18**(3), e1009817 (2022).
- Roldao, A., Mellado, M. C. M., Castilho, L. R., Carrondo, M. J. T. & Alves, P. M. Virus-like particles in vaccine development. *Expert Rev. Vaccines* **9**(10), 1149–1176 (2010).
- Reed, S. G., Orr, M. T. & Fox, C. B. Key roles of adjuvants in modern vaccines. *Nat. Med.* **19**(12), 1597–1608 (2013).
- Kanekiyo, M. et al. Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies. *Nature* **499**(7456), 102–108 (2013).
- Vaine, M. et al. Improved induction of antibodies against key neutralizing epitopes by Human Immunodeficiency Virus type 1 gp120 DNA prime-protein boost vaccination compared to gp120 protein-only vaccination. *J. Virol.* **82**(15), 7369–7378 (2008).
- Haynes, B. F., Kelsoe, G., Harrison, S. C. & Kepler, T. B. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. *Nat. Biotechnol.* **30**(5), 423–433 (2012).
- Piana, S., Robustelli, P., Tan, D., Chen, S. & Shaw, D. E. Development of a Force Field for the Simulation of Single-Chain Proteins and Protein-Protein Complexes. *J. Chem. Theory Comput.* **16**(4), 2494–2507 (2020).
- Piana, S., Donchev, A. G., Robustelli, P. & Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **119**(16), 5113–5123 (2015).

42. Predescu, C. et al. Times Square sampling: an adaptive algorithm for free energy estimation. *J. Comput. Graphical Stat.* **33**(3), 1073–1083 (2024).
43. Marinari, E. & Parisi, G. Simulated tempering: A new Monte Carlo scheme. *Europhys Lett.* **19**, 451–458 (1992).
44. Lyubartsev, A. P., Martsinovski, A. A., Shevkunov, S. V. & Vorontsov-Velyaminov, P. N. New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles. *J. Chem. Phys.* **96**, 1776–1783 (1992).
45. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**(8), 1950–1958 (2010).
46. Hornak, V. et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**(3), 712–725 (2006).
47. Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **113**(26), 9004–9015 (2009).
48. Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?. *J. Comput. Chem.* **21**(12), 1049–1074 (2000).
49. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
50. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization?. *Biophysical J.* **100**(9), L47–L49 (2011).
51. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**(18), 3586–3616 (1998).
52. Lippert, R. A. et al. A common, avoidable source of error in molecular dynamics integrators. *J. Chem. Phys.* **126**(4), 046101 (2007).
53. Kräutler, V., van Gunsteren, W. F. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **22**, 501–508 (2001).
54. Predescu, C., Lerer, A. K., Lippert, R. A., Towles, B., Grossman, J.P., Dirks, R. M., & Shaw, D. E. The *u*-series: A separable decomposition for electrostatics computation with improved accuracy. *J. Chem. Phys.* **152**(8), 084113–1–084113–12 (2020).
55. Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **87**(5), 1117–1157 (1996).
56. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**(3), 1695–1697 (1985).
57. Lippert, R. A. et al. Accurate and Efficient Integration for Molecular Dynamics Simulations at Constant Temperature and Pressure. *J. Chem. Phys.* **139**(16), 164106 (2013).
58. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**(3), 1990–2001 (1992).
59. Shaw, D. E., Grossman, J.P., Bank, J. A., Batson, B., Butts, J. A., Chao, J. C., Deneroff, M. M., Dror, R. O., Even, A., Fenton, C. H., Forte, A., Gagliardo, J., Gill, G., Greskamp, B., Ho, C. R., Ierardi, D. J., Iserovich, L., Kuskin, J. S., Larson, R. H., Layman, T., Lee, L.-S., Lerer, A. K., Li, C., Killebrew, D., Mackenzie, K. M., Mok, S.Y.-H., Moraes, M. A., Mueller, R., Nociolo, L. J., Peticolas, J. L., Quan, T., Ramot, D., Salmon, J. K., Scarpazza, D. P., Schafer, U. B., Siddique, N., Snyder, C. W., Spengler, J., Tang, P. T. P., Theobald, M., Toma, H., Towles, B., Vitale, B., Wang, S. C., & Young, C. Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *Proc. International Conference for High Performance Computing, Networking, Storage and Analysis* (SC14), 41–53 (2014).

Acknowledgements

We thank David Borhani and Michael Eastwood for helpful discussions and critical readings of the manuscript, Jordan Caleb and Kevin Yuh for preparing the molecular graphics, Jordan Caleb for preparing the supplementary movie, and Jessica McGillen and Eric Martens for editorial assistance.

Author contributions

Conceptualization: H.X., D.E.S.; Experimental design: H.X., Q.W.; Data collection: H.X., T.P., Q.W.; Formal analysis: H.X., T.P.; Manuscript writing/editing: H.X., D.E.S.; Supervision: H.X., D.E.S.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-94295-5>.

Correspondence and requests for materials should be addressed to H.X. or D.E.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.