Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

doi:10.1016/j.jmb.2011.10.003 J. Mol. Biol. (2012) 415, 175–192

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Computational Design of High-Affinity Epitope Scaffolds by Backbone Grafting of a Linear Epitope

Mihai L. Azoitei[1], Yih-En Andrew Ban[1], Jean-Philippe Julien[2], Steve Bryson[2,3], Alexandria Schroeter[1], Oleksandr Kalyuzhniy[1], Justin R. Porter[1,4], Yumiko Adachi[1], David Baker[1], Emil F. Pai[2,3,5] and William R. Schief[1,6,7]*

[1]*Department of Biochemistry, University of Washington, Seattle, WA 98195, USA*
[2]*Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5G 2L7*
[3]*Campbell Family Cancer Research Institute, Ontario Cancer Institute/Princess Margaret Hospital, Toronto, ON, Canada*
[4]*Program in Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA*
[5]*Departments of Medical Biophysics and Molecular Genetics, University of Toronto, Toronto, ON, Canada M5G 2L7*
[6]*IAVI Neutralizing Antibody Center, The Scripps Research Institute, La Jolla, CA 92037, USA*
[7]*Department of Immunology and Microbial Sciences, The Scripps Research Institute, La Jolla, CA 92037, USA*

Received 29 April 2011;
received in revised form
1 October 2011;
accepted 4 October 2011
Available online
31 October 2011

Edited by I. Wilson

Keywords:
protein grafting;
flexible backbone design;
epitope scaffold;
immunogen design

Computational grafting of functional motifs onto scaffold proteins is a promising way to engineer novel proteins with pre-specified functionalities. Typically, protein grafting involves the transplantation of protein side chains from a functional motif onto structurally homologous regions of scaffold proteins. Using this approach, we previously transplanted the human immunodeficiency virus 2F5 and 4E10 epitopes onto heterologous proteins to design novel "epitope-scaffold" antigens. However, side-chain grafting is limited by the availability of scaffolds with compatible backbone for a given epitope structure and offers no route to modify backbone structure to improve mimicry or binding affinity. To address this, we report here a new and more aggressive computational method—backbone grafting of linear motifs—that transplants the backbone and side chains of linear functional motifs onto scaffold proteins. To test this method, we first used side-chain grafting to design new 2F5 epitope scaffolds with improved biophysical characteristics. We then independently transplanted the 2F5 epitope onto three of the same parent scaffolds using the newly developed backbone grafting procedure. Crystal structures of side-chain and backbone grafting designs showed close agreement with both the computational models and the desired epitope structure. In two cases, backbone grafting scaffolds bound antibody 2F5 with 30- and 9-fold higher affinity than corresponding side-chain grafting designs. These results demonstrate that flexible backbone methods for epitope grafting can significantly improve binding affinities over those achieved by fixed

**Corresponding author*. E-mail address: schief@scripps.edu.

Present addresses: Y.-E. A. Ban, Arzeda Corporation, Seattle, WA 98102, USA; J.-P. Julien, Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA.

Abbreviations used: mAb, monoclonal antibody; PDB, Protein Data Bank; SPR, surface plasmon resonance; CCD, cyclic coordinate descent; EDTA, ethylenediaminetetraacetic acid; HBS, Hepes-buffered saline.

backbone methods alone. Backbone grafting of linear motifs is a general method to transplant functional motifs when backbone remodeling of the target scaffold is necessary.

Introduction

The study of protein structures and interactions and their relation to function is essential to understanding biological systems. Developing methods to predict and manipulate protein structures and interactions tests our knowledge of the underlying physics and provides opportunities for diverse applications.

Protein grafting, the transplantation of functional sites onto scaffold proteins, attempts to alter the interactions and functions of proteins via their structures. Protein grafting has been applied to design inhibitors,[1] to stabilize structures of functionally important transient states,[2,3] to increase the specificity and affinity of binding interactions,[4] to introduce catalytic activities into proteins,[5–8] to dissect the interactions of binding partners,[9] to improve vaccine epitope presentation,[10–12] and to present functional sites on proteins with desirable characteristics such as reduced size or high stability.[13,14] Typically, protein grafting relies on identifying scaffolds with local backbone structure similar to the functional motif to be transplanted, followed by altering the scaffold sequence to accommodate the respective motif. Therefore, the ability to graft a motif onto a scaffold is limited by the complexity of the motif and the availability of scaffolds with suitable structure.

We recently described a general computational method to design novel antigens, termed epitope scaffolds, in which linear epitopes are transplanted to scaffold proteins for structural stabilization and immune presentation.[15,16] Unlike peptide antigens that can usually adopt multiple conformations in solution, epitope scaffolds are engineered to preferentially stabilize a particular epitope conformation—usually the conformation of the epitope when it is bound to a neutralizing antibody. As shown for the 2F5 and 4E10 monoclonal antibodies (mAbs) that neutralize diverse strains of the human immunodeficiency virus, epitope scaffolds bind their target antibody with high affinity and are able to elicit structure-specific immune responses.[15,16] In other studies, immunogens designed by epitope transplantation elicited structure-specific responses targeting a snake toxin,[10] the yeast transcription factor GCN4,[11] and the severe acute respiratory syndrome coronavirus S glycoprotein.[12] Epitope scaffolds can thus serve as vaccine candidates or as reagents to map different sera specificities. In addition, epitope scaffolds are a good platform to test the ability of computational design to manipulate protein structure and function.

In previous work,[15,16] we engineered epitope scaffolds by epitope side-chain grafting. In this procedure (Fig. 1), the side chains of a linear epitope were transplanted to other proteins that had exposed backbone segments conformationally similar to the backbone of the antibody-bound epitope. Despite our previous success with side-chain grafting, this method has two significant limitations: first, it relies on the existence of scaffolds with exposed backbone structurally similar to a given epitope, and second, when such scaffolds can be identified, the displayed epitope conformation differs slightly from the antibody-bound epitope conformation due to the imperfect match between the backbone of the native scaffold and that of the epitope (RMSD is typically ~1 Å); this in turn might affect the affinity of the designed antigen–antibody interaction. Generally, these limitations are present in all protein grafting approaches that rely on side-chain transplantation onto a preexisting scaffold backbone and may restrict the diversity or complexity of functional motifs that can be grafted and the affinity of the engineered interactions.

To address these limitations, here we describe and test a new computational protocol for grafting linear functional motifs—epitope backbone grafting. Rather than transplanting the side chains of a functional motif onto a preexisting scaffold backbone, backbone grafting replaces the native backbone of a candidate scaffold with the desired backbone conformation of a functional motif (Fig. 1). Thus, epitope backbone grafting imposes the conformation of a given epitope onto a scaffold and integrates that epitope conformation into the scaffold through backbone remodeling and sequence design in regions flanking the epitope. Previous studies[17–20] successfully modeled protein backbone regions with atomic accuracy using computational approaches. However, none of these studies attempted to maintain a predefined sequence or structure during backbone remodeling but rather iterated between sequence design and structure optimization steps to identify the most favorable solutions energetically. In contrast, backbone grafting of linear motifs maintains the sequence and structure of a functional motif while building novel interactions to integrate the transplanted motif with the scaffold.

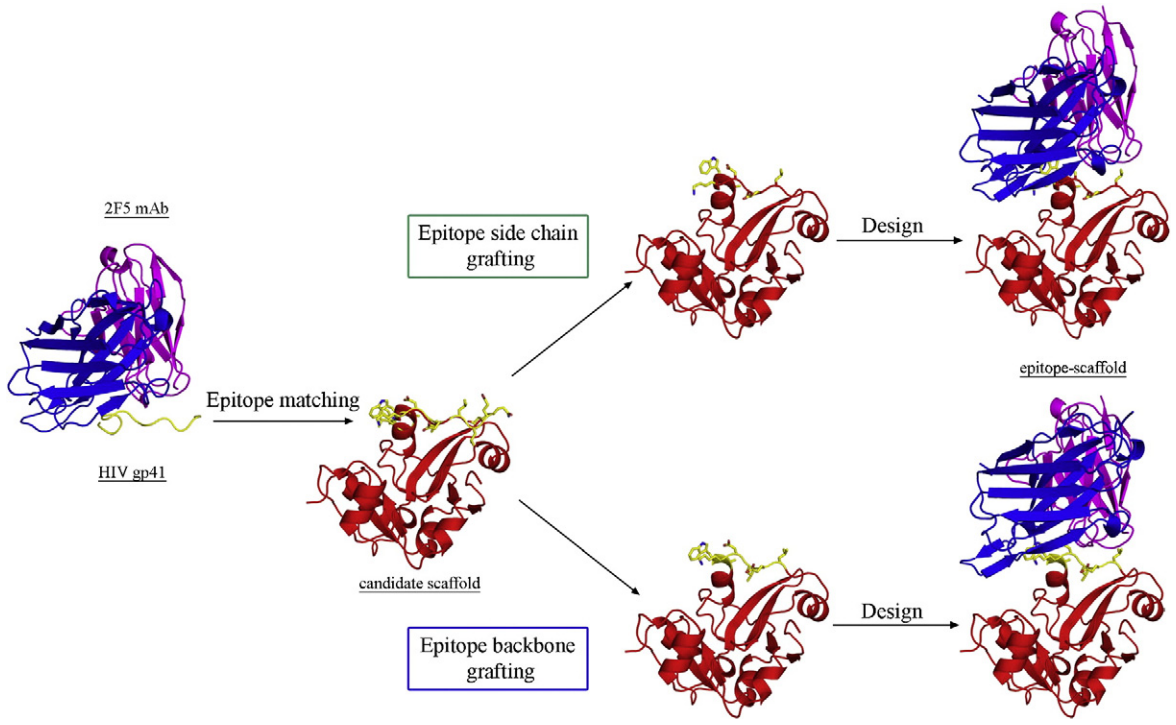

Fig. 1. Main stages of epitope side-chain and backbone grafting protocols. Given a starting antibody–epitope complex, candidate scaffolds (red) are identified in the matching stage for epitope (yellow) transplantation. In epitope side-chain grafting, the residue identity of the candidate scaffold is altered to match the epitope sequence. In epitope backbone grafting, the epitope backbone conformation replaces the native backbone region of the candidate scaffold; the 2F5 mAb is shown in blue (heavy chain) and magenta (light chain).

As a first test of backbone grafting, we transplanted the epitope of 2F5, a broadly neutralizing antibody against human immunodeficiency virus type 1, onto three scaffolds that were also able to accommodate the epitope sequence through side-chain grafting (Fig. 1). Besides evaluating backbone grafting, this approach compared side by side the two protocols for the design of high-affinity antigens. Further, grafting the 2F5 epitope backbone on scaffolds that were able to accommodate the epitope sequence allowed us to evaluate backbone grafting outside potential confounding factors, such as intrinsic scaffold stability or solubility.

In the first part of this study, we designed second-generation 2F5 epitope scaffolds using the side-chain grafting method. Crystallographic and mutagenesis studies identified the ^{657}EQELLELDKWASLW^{670} fragment of gp41 as the 2F5 binding site, with the central "DKW" residues critical for epitope binding and virus neutralization.[21–23] The first-generation side-chain grafting 2F5 epitope scaffolds[15] displayed epitope ranges corresponding to gp41 residues 659–669, and the design models mimicked the target epitope conformation with backbone RMSD values between 1.2 and 1.7 Å. The side-chain grafting epitope scaffolds described here displayed shorter stretches of the epitope region centered on the DKW motif (gp41 ranges 661–667, 662–667, or 661–666). Given that a shorter epitope range was transplanted in this study, it was possible to identify new scaffolds that mimicked the epitope backbone more closely than first-generation designs (backbone RMSD between 0.5 and 0.8 Å). These second-generation epitope scaffolds bound 2F5 mAb with a range of affinities.

In the second part of the study, we performed backbone grafting on three parent scaffolds designed in the first part by side-chain grafting. To test the new backbone grafting protocol, we selected parent scaffolds for which the side-chain grafting method had produced epitope scaffolds with a range of dissociation constants (K_d values) for binding to mAb 2F5 (K_d values from 40 nM to >2 μM). We present biophysical characterization of 10 backbone grafting designs based on three different scaffolds and crystallographic comparison of two pairs of unliganded scaffolds designed by side-chain grafting and backbone grafting. Scaffolds designed by epitope backbone grafting bound 2F5 at least as well as equivalent side-chain grafting scaffolds. Remarkably, for two out of the three cases tested, backbone grafting resulted in scaffolds that bound 2F5 30- and 9-fold better than corresponding side-chain grafting designs, while neither

method produced high-affinity designs in the third case. In the two successful test cases, crystal structures of unliganded epitope scaffolds validated the computational models for both side-chain and backbone grafting and showed that the designed epitope scaffolds closely mimicked the 2F5-bound conformation of the epitope.

These results indicate that backbone grafting may be generally superior to side-chain grafting for generating high-affinity epitope scaffolds, even in cases where side-chain grafting can be applied. Given that epitope backbone grafting can alter both the sequence and the backbone of a scaffold to accommodate a desired functional motif, this procedure may also be useful in cases where no suitable scaffold candidates exist for side-chain grafting.

Results

Computational design of epitope scaffolds by side-chain grafting

Epitope scaffolds were identified from a subset of the Protein Data Bank (PDB)[24] and computationally designed using the side-chain grafting protocol as previously described.[16] Briefly, candidate scaffolds with exposed backbone similar to the 2F5 epitope were selected in the matching stage, the epitope side chains were transferred onto the scaffolds, and additional computational design steps ensured that the epitope scaffolds could successfully interact with the mAb. Compared to previous 2F5 epitope scaffolding work, here we focused on transplanting shorter ranges of the 2F5 epitope (LExDKWA, ExDKWA, and LExDKW) to identify epitope scaffolds with better mimicry of the epitope core. The transplanted epitope ranges contribute at least 60% of the buried surface area and include 10 out of the 13 hydrogen bonds and salt bridges present at the interface of the 2F5–gp41 complex.[21] Given the large contributions of the selected epitope subranges to the overall binding interface, we reasoned that their stabilization through scaffolding would be sufficient to generate high-affinity epitope scaffolds.

Twenty-eight epitope scaffolds based on 28 different parent proteins were initially designed to accommodate the 2F5 epitope, and the resulting epitope scaffolds were predicted to productively interact with the antibody (see Fig. 2 for models of representative epitope scaffolds). On each of these

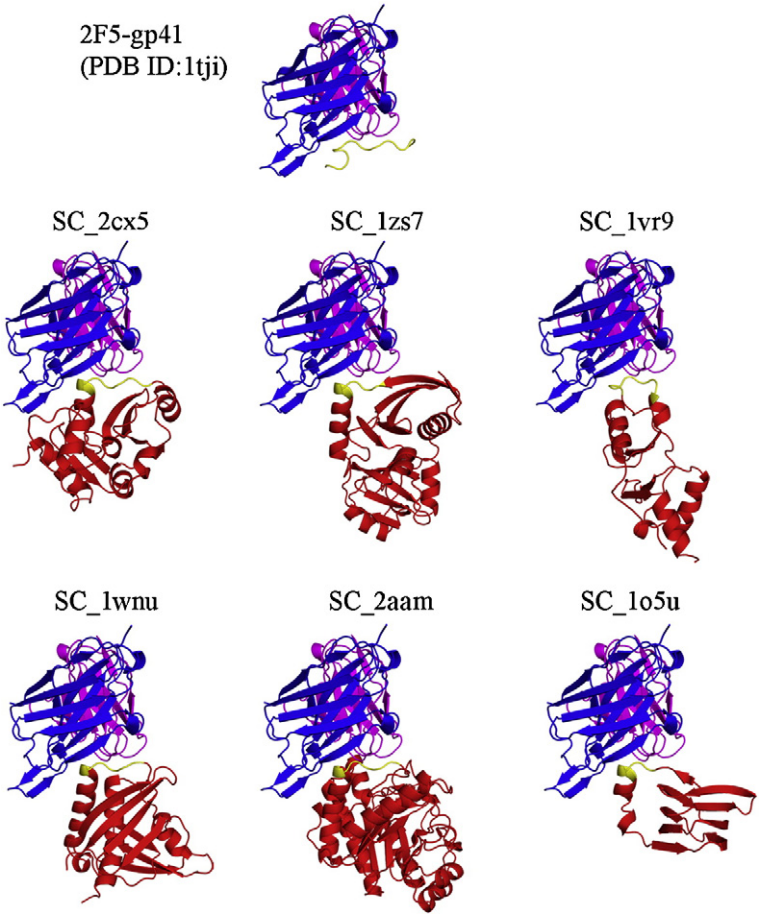

Fig. 2. Models of representative epitope scaffolds designed by side-chain grafting and chosen for experimental characterization. The scaffolds are shown in red with the epitope in yellow; the 2F5 mAb is in blue (heavy chain) and magenta (light chain).

epitope scaffolds, two to five additional variants were subsequently designed using Rosetta in an attempt to optimize protein solubility and stability. Modifications included removing unpaired cysteines, removing exposed hydrophobics, trimming domains, and creating negatively or positively charged variants by redesigning surface residues.[25] Overall, 98 epitope scaffolds were selected for experimental characterization. These epitope scaffolds differed from the parent proteins by 15 mutations on average and were predicted to present epitope backbone conformations within 1 Å RMSD of the 2F5-bound gp41 (Table S1).

Biophysical characterization of side-chain grafting epitope scaffolds

Epitope scaffolds were expressed in *Escherichia coli* and purified by immobilized metal affinity chromatography followed by size-exclusion chromatography. Of the 98 computational designs, 35 (36%), based on 15 different parent proteins, were soluble and purifiable. In comparison, none of the first-generation 2F5 epitope scaffolds were soluble when expressed in *E. coli*.[15] In an initial screen, 2F5 binding was assessed by surface plasmon resonance (SPR), with epitope scaffolds flowed at high concentrations (>5 μM) over 2F5 IgG captured on an amine-linked anti-human IgG CM5 chip. For proteins that bound to 2F5 in this screen, the thermal stability, solution multimeric state, and K_d for 2F5 were subsequently measured (Figs. 3 and 4; Figs. S1 and S2). Table 1 summarizes the biophysical data for seven representative epitope scaffolds with a K_d lower than 1 μM. Four of these seven epitope scaffolds were design variants that included mutations to enhance solubility or stability as mentioned above—design variants without the additional mutations were either insoluble or unstable during purification. In these cases, the additional mutations were introduced outside the region interacting with antibody and were as follows: SC_2cx0 and SC_1o5u were resurfaced for net negative charge by designing all the exposed surface residues not contacting the antibody to negatively charged or neutral polar residues (total charges were −16 and

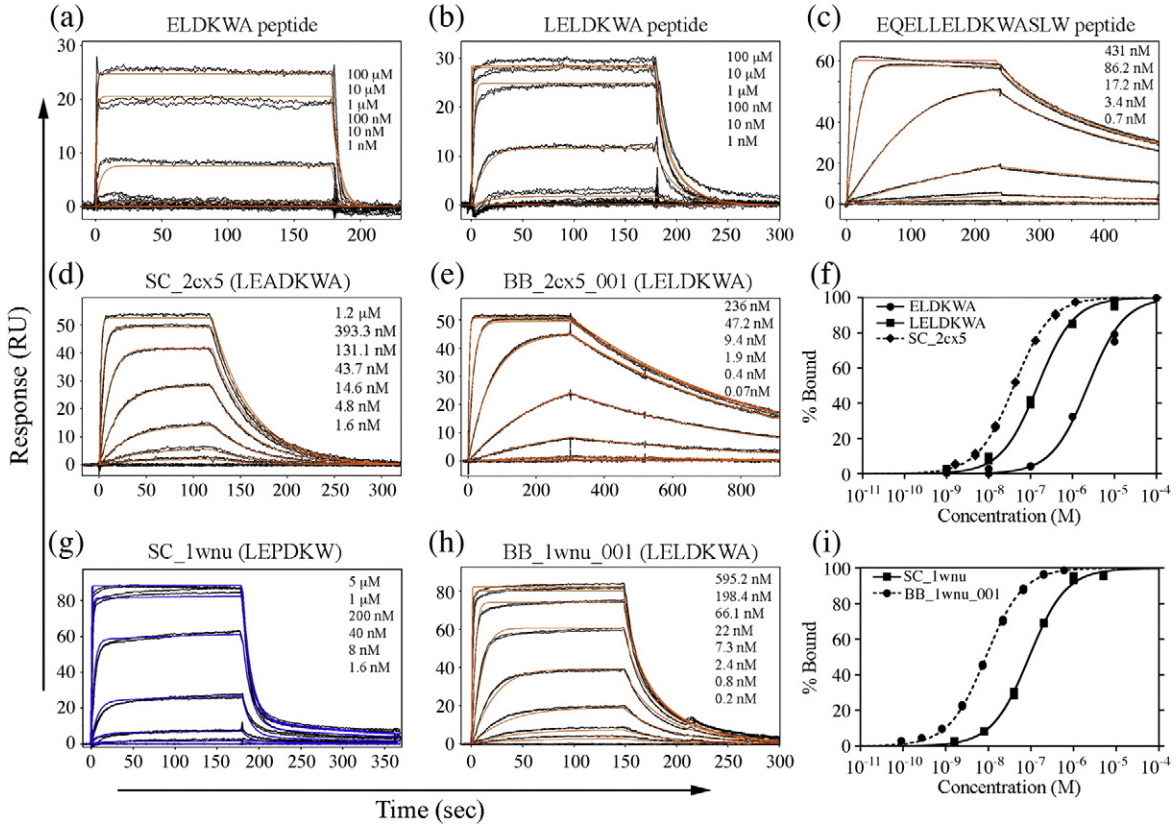

Fig. 3. SPR analysis of the 2F5 mAb interactions of selected epitope scaffolds and epitope peptides. (a–c) Binding responses of epitope peptides of different lengths. (d and g) 2F5 binding of corresponding epitope scaffolds designed by side-chain and backbone grafting (e and h). (f and i) Equilibrium binding analysis and affinity comparison of epitope peptides and epitope scaffolds designed by the two different computational protocols; the collected data are shown in black and the model fits are shown in red (1:1 Langmuir binding) and blue (conformational change model).

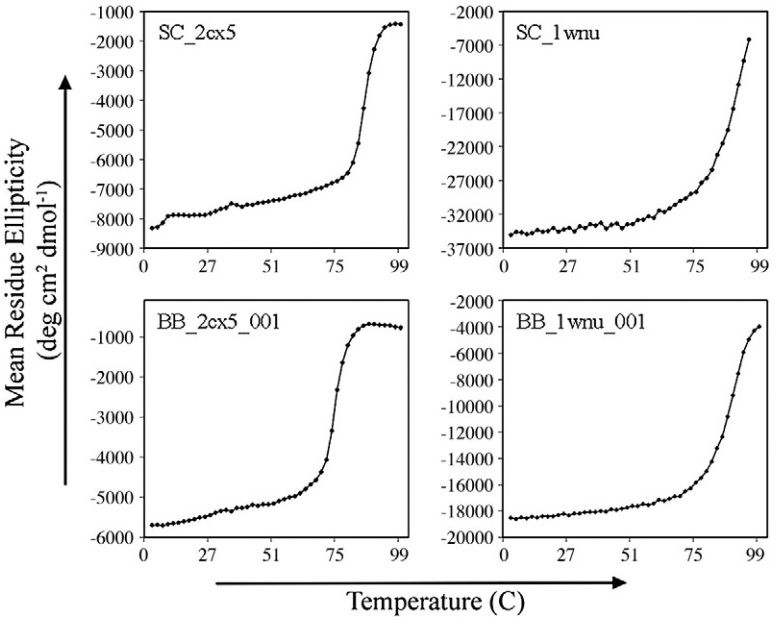

Fig. 4. Thermal stability of corresponding epitope scaffolds designed by side-chain and backbone grafting.

−10, respectively); one unpaired cysteine was removed from SC_2aam; four nonpolar surface residues on SC_1zs7 were changed to polar.

Epitope scaffolds were thermally stable with melting temperatures greater than 80 °C as measured by circular dichroism (CD) temperature melt analysis (Fig. 4 and Fig. S2). Although most of the parent proteins for these scaffolds originated from thermophilic organisms, the high thermal stability is still remarkable given that approximately 10% of the total residues were mutated to create the epitope scaffolds. Static light scattering analysis showed that except for SC_2cx0 (mixture of monomer and trimer species), all epitope scaffolds were monomeric in solution (data not shown). The parent protein for SC_1vr9 is a natural dimer, but the dimerization interface was intentionally disrupted during the design process, which resulted in SC_1vr9 being a monomer in solution.

To provide controls for the interaction of 2F5 with side-chain grafting epitope scaffolds, we first used SPR to measure the dissociation and rate constants of three gp41 peptides of different lengths centered at the "^{664}DKW^{666}" core (Table 1 and Fig. 3). Peptides were injected over 2F5 captured on a CM5 chip with amine-coupled anti-human IgG. The full-length 2F5 epitope (^{657}EQELLELDKWASLW^{670}) bound the antibody with high affinity (K_d = 4.1 nM) in agreement with previously published values.[26,27] Shorter epitope peptides, however, showed considerably reduced affinity for the antibody. ^{661}LELDKWA^{667} bound 2F5 with a K_d of 144 nM while ^{662}ELDKWA^{667}

Table 1. Biophysical characterization of side-chain grafting epitope scaffolds

	Epitope sequence	K_d kinetic (nM)	K_d equilibrium (nM)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	Oligomeric state	T_M (°C)
2F5 peptide	EQELLELDKWASLW	4.1	—	1.45×10^{6}	6.06×10^{-3}	—	—
SC_2cx0[a]	LELDKWA	6.6	—	2.29×10^{5}	1.52×10^{-3}	Monomer/trimer	ND
SC_2cx5	LEADKWA	40.2	38.3	5.87×10^{5}	2.36×10^{-2}	Monomer	84
SC_1vr9	ELDKWA	71.3	66.3	2.83×10^{6}	2.02×10^{-1}	Monomer	>90
SC_1zs7	ELDKWA	75.2	74.9	5.40×10^{5}	4.06×10^{-2}	Monomer	>90
SC_1wnu[b]	LEPDKW	1: 109.4 2: 3.8×10^{9}	91	1: 9.04×10^{5} 2: 1.26×10^{-3}	1: 9.89×10^{-2} 2: 4.80×10^{-3}	Monomer	>90
SC_2aam	LEADKW	148.6	145.2	1.05×10^{6}	1.56×10^{-1}	Monomer	>90
SC_1o5u	EPDKWA	357.4	354	6.77×10^{5}	2.42×10^{-2}	ND	ND
2F5 peptide	LELDKWA	144.2	142	7.0×10^{5}	1.01×10^{-1}	—	—
2F5 peptide	ELDKWA	2.3×10^{3}	2.3×10^{3}	8.01×10^{4}	1.85×10^{-1}	—	—

ND, experiment was not performed.

For all reported SPR data, standard error is ≤±5 of the last significant figure.

[a] Reported K_d is affected by avidity since epitope scaffold is not monomeric.

[b] SPR data for SC_1wnu were fit using a conformational change model; the second on-rate has units of s^{-1}.

showed a K_d of 2.3 μM, corresponding to 36-fold and 564-fold lower affinity than the full-length epitope peptide, respectively. Decreases in k_{on} and increases in k_{off} contributed to the reduction in affinity. Based on the crystal structure of the 2F5–gp41 epitope complex (PDB ID: 1tji),[21] the shorter peptides make two fewer hydrogen bonds ($Gln_{658}O$ to $Gln_{L27}N2$ of 2F5; $Glu_{659}O2$ to $Ala_{1L}N$ of 2F5) and bury less surface area than the full-length epitope peptide when interacting with 2F5. Indeed, ^{662}ELDKWA^{667} and ^{661}LELDKWA^{667} peptides present 62% and 74% of the total binding surface of the ^{657}EQELLELDKWASLW^{670} peptide, respectively.[21]

Determination of 2F5 binding parameters for epitope scaffolds was done as for the peptides above. Epitope scaffolds bound 2F5 with higher affinity than corresponding epitope peptides, consistent with the assumption that epitope stabilization in the antibody-bound state results in better 2F5 binding. SC_1vr9 (K_d = 71.3 nM) and SC_1zs7 (K_d = 75.2 nM), both of which present the ELDKWA portion of the epitope, have a 30-fold higher affinity for 2F5 than the ^{662}ELDKWA^{667} peptide. SC_2cx5 (K_d = 38.3 nM), which presents the LELDKWA portion of the epitope, binds 2F5 fourfold tighter than the ^{661}LELDKWA^{667} peptide (Fig. 3). With one exception (SC_1vr9a), the binding improvements of epitope scaffolds over corresponding epitope peptides were due to slower dissociation rates (k_{off}).

Computational design of epitope scaffolds by backbone grafting

Backbone grafting (Fig. 5) was employed to transplant the 2F5 epitope onto three parent proteins used for side-chain grafting. Two of the side-chain grafting designs bound to 2F5 with a K_d lower than 100 nM (SC_2cx5, SC_1wnu), while one (SC_1zko) bound 2F5 weakly, with a K_d >2 μM (Table 1). In the matching stage for backbone grafting, all possible alignments were tested for the N or C terminus of the epitope over the whole structure of the three parent scaffolds (PDB IDs: 2cx5, 1wnu, and 1zko). In this alignment procedure, backbone atoms on one end of the epitope were superimposed onto the scaffold backbone, and RMSD values were measured between the other end of the epitope and proximal positions on the scaffold. Epitope alignments were selected if the non-superimposed end was closer than 3 Å from a scaffold backbone residue and if the alignments satisfied additional filters that assessed potential clashes between the epitope and the scaffold and between the antibody and the epitope presented in the scaffold context. The resulting matches were restricted further, for the purpose of comparing backbone grafting scaffolds directly with side-chain grafting scaffolds. Only backbone matches that placed the epitope at the same scaffold position as in the side-chain grafting designs were chosen for subsequent computational design. The native scaffold backbone corresponding to the epitope was removed and replaced by the epitope itself per the alignment, resulting in scaffolds anchoring the epitope at one end and leaving an open gap at the opposing end (Fig. 5). Backbone steric clash filters were then assessed to check the suitability of the scaffold to accommodate the epitope shape.

To close the gap, we employed multiple cycles of a loop closure procedure utilizing a low-resolution scoring function, fragment insertion,[28] cyclic coordinate descent (CCD),[29] and Monte Carlo sampling (Fig. 5). In this stage, the amino acid identity of the scaffold and epitope–antibody complex was changed to alanine for efficient sampling and scoring. During loop closure, the values for the dihedral angles of four scaffold residues and one epitope residue were allowed to vary on each side of the grafted epitope while the backbone of the remaining residues was held fixed. Thus, the conformation of the epitope itself was held fixed except for the ϕ and ψ angles for single residues at the epitope termini. Throughout the process, the antibody was kept in fixed rigid-body orientation relative to the epitope. Once a satisfactory closure was achieved, a backbone refinement step was employed to correct problematic backbone conformations modeled in the previous stage. The full sequence information

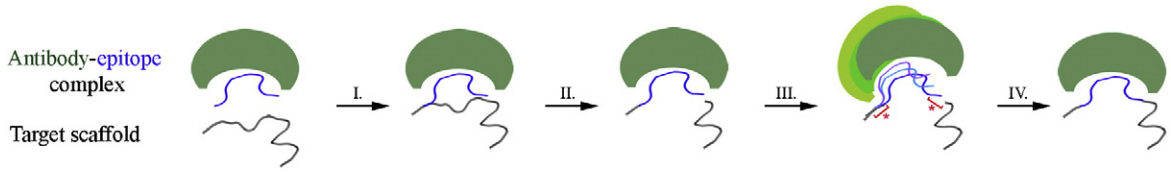

Fig. 5. Stages of epitope backbone grafting. (I) The epitope is aligned on the target scaffold. (II) The native scaffold backbone corresponding to the epitope is deleted, resulting in a disconnected polypeptide chain. (III) To integrate the epitope with the scaffold, novel backbone regions are modeled between the epitope termini and the scaffold (red stars). (IV) Final closure of the chain sets the rigid-body orientation of the epitope and the antibody relative to the scaffold; sequence design ensures the stabilization of the epitope conformation and the productive interaction of the antibody with the resulting epitope scaffold.

was then recovered and the region of the scaffold adjacent to the binding interface was subjected to a round of sequence design in the presence of the antibody using a high-resolution energy function. Scaffold residues were designed if they were located within 4 Å of either the grafted epitope, defined as intra-design residues, or the antibody, defined as inter-design residues. During the automated design process, intra-design residues were allowed to change to any amino acid except cysteine, while inter-design residues were changed to either serine, alanine, glycine, or threonine. This stage of the procedure ensured that the grafted backbone interacted favorably with the rest of the scaffold and that the antibody could productively bind the resulting epitope scaffold only through epitope-mediated contacts. Finally, all-atom loop refinement[30,31] was employed to refine the conformation of the backbone regions that were allowed to change during the process. On the resulting epitope scaffolds, there is a very slight deviation (backbone RMSD <0.2 Å) between the modeled conformation of the epitope and that of the native antibody-bound epitope. This is because the conformations of the N- and C-terminal epitope residues were altered during the loop closure stage to connect the epitope with the scaffold. Outside these peripheral residues, the epitope conformation on the scaffold perfectly matches that of the antibody-bound epitope.

This procedure modeled different backbone conformations of the scaffold regions that anchored the epitope, resulting in different antibody binding angles between epitope scaffolds designed by side-chain and backbone grafting (Fig. 6a and b). The different backbone conformations in the epitope region led to the design of different sequences in the backbone grafting designs compared to the side-chain grafting designs (Fig. 6c). The sequence changes in epitope backbone grafting designs were deemed to be beneficial by stabilizing the epitope conformation, by favoring the epitope backbone conformation in the particular scaffold region, by transplanting additional epitope residues, or by improving the overall characteristics of the resulting epitope scaffold. A total of 5000 backbone grafting variants were designed on each of the three scaffolds. Designs were filtered and ranked according to the following criteria: (1) the energy of the scaffold alone and in complex with the antibody, (2) the theoretical binding energy ($E_{complex} - E_{partner1} - E_{partner2}$), (3) the final values of the backbone dihedral angles moved during the procedure, and (4) the overall core packing of the scaffold. For each of the three scaffolds, the backbone grafting method generated at least 10 models that scored better on all the criteria than the corresponding side-chain grafting designs. Three or four design variants based on each of the parent scaffolds were selected for experimental characterization based on sequence and structure diversity. Following visual inspection of the selected models, one additional round of human-guided computational sequence design was performed to correct the following defects: solvent-exposed hydrophobic groups, buried hydrogen donor or acceptor groups, and non-epitope-mediated contacts to the antibody. This stage corrects for any imperfections due to incomplete sampling and inaccuracies in the Rosetta energy function by manually restricting the set of amino acids allowed at a given position during the computational procedure—for example, only polar residues are allowed at solvent-exposed position on the scaffold.

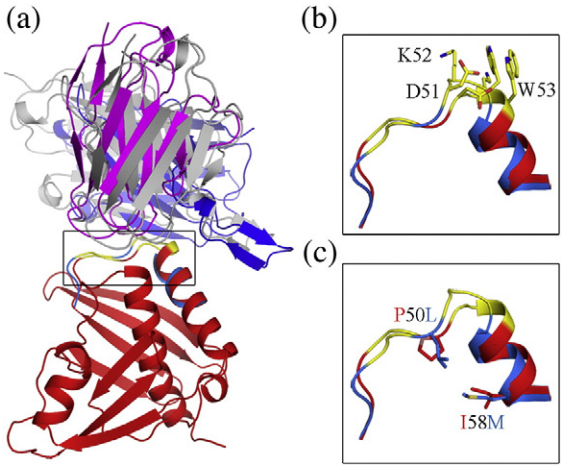

Fig. 6. Comparison of epitope-scaffold models generated by side-chain and backbone grafting based on the same native protein (PDB ID: 1wnu). (a) Alignment of side-chain grafting (red) and the backbone grafting models (blue) in complex with 2F5 (shown in gray when interacting with the backbone grafting design and in blue and magenta in complex with the side-chain grafting design). Changes in the orientation of the epitope region (yellow) relative to the scaffold alter the binding orientation of 2F5. (b) Detailed view (black box) of the different epitope conformations. The central DKW epitope residues are shown in stick representation. (c) Changes in the conformation of the scaffold backbone in the epitope region allow the addition of extra epitope residues (P50L) as well as the engineering of stabilizing interaction between the epitope and the scaffold (I58M) in the backbone grafting model.

Epitope backbone grafting scaffolds bind 2F5 with high affinity

Following the computational procedure, the epitope regions on backbone grafting design models were different by 1.8 Å (backbone RMSD) on average from the epitope regions on corresponding side-chain grafting design models. To optimize the sequence of these conformationally different regions, we introduced, on average, eight sequence

changes at the design stage between backbone grafting epitope scaffolds and equivalent side-chain grafting designs (Fig. S3). Importantly, backbone grafting epitope scaffolds based on SC_2cx5 and SC_1wnu showed significantly higher 2F5 affinity than corresponding side-chain grafting designs (Table 2).

Of the three 2cx5 backbone grafting variants characterized, the 2F5 affinity of BB_2cx5_001 was 30 times higher than that of the side-chain grafting design SC_2cx5 (K_d = 1.4 nM *versus* 40.2 nM) (Fig. 4). Two other backbone grafting variants, BB_2cx5_002 and BB_2cx5_003, showed modest, yet significant, affinity improvements (fourfold and twofold, respectively). BB_2cx5_001 differs from SC_2cx5 by seven mutations that are mapped onto the respective crystal structures in Fig. 7d. The mutations were introduced for the following reasons: (1) to avoid clashes between the predicted rigid-body orientation of the antibody and the scaffold (E56S and D98A), (2) to accommodate the backbone changes required to connect the N terminus of the epitope with the scaffold (G81T and P82D), (3) to remove clashes between the scaffold and the remodeled regions connecting the epitope (V51A and T76A), and (4) to add stabilizing interactions between the epitope and the scaffold (A86L). Of the sequence changes on BB_2cx5_001, five were introduced in the automatic stage of backbone grafting and two of them (D82 and D98) were introduced after visual inspection in the human-guided computational design stage, to remove solvent-exposed hydrophobic residues on the surface of the scaffold.

Of these seven mutations, all the backbone grafting BB_2cx5 variants share A86L and D98A, while the other five changes are unique to BB_2cx5_001. L86 is equivalent to L663 of the native epitope sequence (^{661}LELDKWA^{667}) that contributes 6% of the total epitope buried surface in the 2F5–epitope complex.[21,22] In the side-chain graft SC_2cx5, the leucine at position 86 would have caused a clash with the rest of the scaffold, due to the preexisting scaffold backbone conformation where the epitope side chains were transplanted. However, during backbone grafting, the orientation of the epitope relative to the scaffold was changed such that L86 could be accommodated on BB_2cx5 variants. The higher affinity of BB_2cx5 designs could thus be partially explained due to extra contacts between L86 and 2F5 mAb. However, other changes must have also contributed to the improved affinity, because L86 is common to all the BB_2cx5 variants but BB_2cx5_001 binds 2F5 with fivefold higher affinity than BB_2cx5_002.

Three of the four experimentally characterized backbone grafting designs based on SC_1wnu were stable monodisperse proteins that bound 2F5 more tightly than SC_1wnu. Both backbone grafting designs BB_1wnu_001 and BB_1wnu_002 had K_d values of 9 nM, a 10-fold improvement over the side-chain grafting design with K_d = 91 nM (Tables 1 and 2; Fig. 4i). As in the 2cx5 case, small sequence differences were responsible for the enhanced 2F5 binding. Four residue changes relative to SC_1wnu were common to all the BB_1wnu variants (T22A, D26A, P50L, and G54A) and two additional mutations were particular to BB_1wnu_001 (I55W and I58M) (Fig. S3). Figure 7h shows the locations of these mutations on the crystal structures of SC_1wnu and BB_1wnu_001. Three of these mutations (T22A, I55W, and I58M) add interactions between the epitope and the scaffold, and one (D26A) removes nonspecific contacts between the epitope scaffold and the predicted 2F5 orientation upon binding. Similar to the 2cx5 case, the new rigid-body orientation of the epitope relative to the scaffold in the BB_1wnu designs allowed the transplantation of additional gp41 residues (L663 and A667) from the native 2F5 epitope to the backbone grafting scaffolds (mutations P50L and G54A). Together, these two

Table 2. Biophysical characterization of backbone grafting epitope scaffolds

Backbone grafting design[a]	Parent SC design	K_d kinetic (nM)	K_d equilibrium (nM)	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	Epitope RMSD[b] (Å)	Mutations[b]	Oligomeric state	T_M (°C)
BB_2cx5_001	SC_2cx5	1.4	—	1.31×10^{6}	1.79×10^{-3}	1.4	7	Monomer	75
BB_2cx5_002[c]	SC_2cx5	1:7.7 2:8×10^{9}	—	1: 1.21×10^{6} 2: 2.81×10^{-3}	1: 9.33×10^{-3} 2: 2.24×10^{-3}	1.4	4	ND	ND
BB_2cx5_003	SC_2cx5	15.5	15.6	1.11×10^{6}	1.73×10^{-2}	1.6	5	Monomer	ND
BB_1wnu_001	SC_1wnu	8.3	8.7	8.8×10^{6}	7.3×10^{-2}	2.1	6	Monomer	>90
BB_1wnu_002	SC_1wnu	8.9	9.4	1.34×10^{7}	1.19×10^{-1}	2.5	8	ND	ND
BB_1wnu_003	SC_1wnu	29	28	2.0×10^{6}	5.8×10^{-2}	2.0	6	Monomer	ND

ND, experiment was not performed.

For all reported SPR data, standard error is ≤±5 of the last significant figure.

[a] Three backbone grafting epitope scaffolds based on the parent protein 1zko showed weak binding (K_d > 2 μM) and were not further characterized.

[b] Epitope RMSD and mutations values were calculated in reference to the corresponding epitope side-chain (SC) grafting design.

[c] SPR data for BB_2cx5_002 were fit using a conformational change model; the second on-rate has units of s^{-1}.

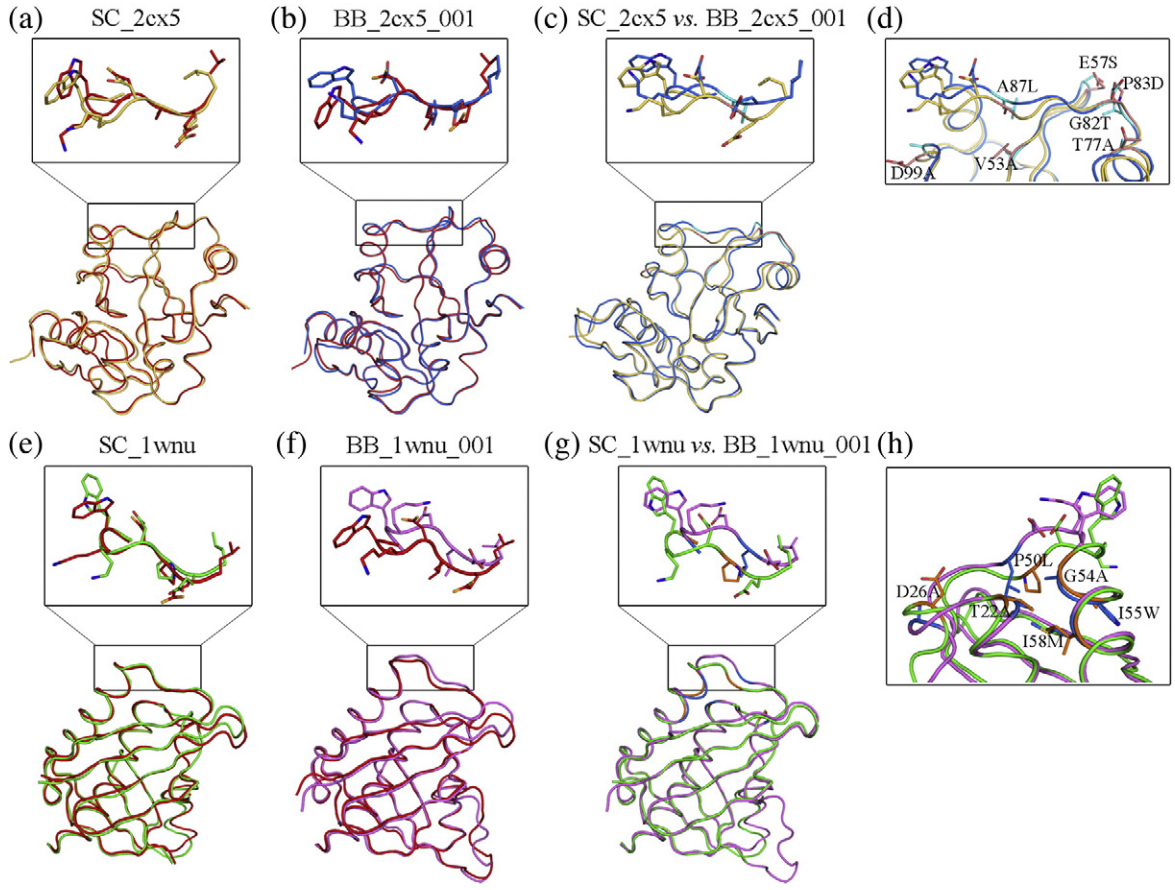

Fig. 7. Crystal structure analysis of epitope scaffolds and comparison with the computational models. Epitope regions are shown in detailed view. (a) Alignment of the SC_2cx5 crystal structure (yellow) and the computational model (red). (b) Alignment of the BB_2cx5_001 crystal structure (blue) and the computational model (red). (c) Alignment of the SC_2cx5 (yellow) and BB_2cx5_001 (blue) structures. (d) Mutated residues between SC_2cx5 (yellow) and BB_2cx5_001 mapped on the respective structures; the mutated residues and the DKW epitope motif are shown as sticks; SC_2cx5 residues are shown in salmon and BB_2cx5_001 residues are shown in cyan. (e) Alignment of the SC_1wnu crystal structure (green) and the computational model (red). (f) Alignment of the BB_1wnu_001 crystal structure (magenta) and the computational model (red). (g) Alignment of the SC_1wnu (green) and BB_1wnu_001 (magenta) structures. (h) Mutated residues between SC_1wnu (green) and BB_1wnu_001 (magenta) mapped on the respective structures; the mutated residues and the DKW epitope motif are shown as sticks; SC_1wnu residues are shown in orange and BB_2cx5_001 residues are shown in blue.

residues contribute 11% of the buried surface in the 2F5–gp41 complex (PDB ID: 1tji); this could explain the increased affinity of BB_1wnu designs over SC_1wnu.

None of the three backbone grafting designs based on the 1zko parent protein showed significant 2F5 binding. The 1zko side-chain grafting design had similarly low 2F5 affinity (K_d >2 μM). These results together suggest either that the epitope sequence is incompatible with the structure of the native 1zko protein or that there are unaccounted clashes in the modeling of the antibody–scaffold interaction in both SC and BB designs. Backbone grafting was thus unable to "rescue" this particular epitope scaffold, but backbone grafting performed no worse than side-chain grafting. Overall, of the 10 backbone grafting designs, 9 were purifiable and 6 bound 2F5 better than the corresponding side-chain grafting designs.

Crystallographic analysis of epitope scaffolds

To assess the computational methods and the resulting designs at the atomic level, we determined high-resolution crystal structures of two pairs of epitope scaffolds designed by side-chain and backbone grafting (SC_2cx5 and BB_2cx5_001; SC_1wnu and BB_1wnu_001). Crystallization conditions and X-ray diffraction and refinement statistics are listed in Table 3.

In all four crystal structures, crystal packing contacts are made by epitope residues with symmetry-related

Table 3. Statistics for data collection and model refinement

	SC_2cx5	BB_2cx5_001	SC_1wnu	BB_1wnu_001
Crystal data				
Crystal growth	0.1 M Hepes, pH 7.5, 2.0 M ammonium sulfate, and 4% polyethylene glycol 400	0.2 M sodium cacodylate, pH 6.5, 0.2 M ammonium sulphatesulfate, and 30% polyethylene glycol 8000	0.1 M Tris, pH 8.0, and 2.8 M NaCl	0.1 M Tris, pH 8.0, and 2.8 M NaCl
Space group	*P*4(3)	*C*2	*I*4	*C*2
Unit cell axes: *a*, *b*, *c* (Å)	70.0, 70.0, 131.4	61.6, 80.4, 67.3	103.65, 103.65, 85.67	93.8, 36.1, 46.3
Unit cell angles: α, β, γ (°)	90.0, 90.0, 90.0	90.0, 111.9, 90.0	90.0, 90.0, 90.0	90.0, 99.4, 90.0
$\langle I \rangle / \langle \sigma(I) \rangle$	19.0 (5.4)	12.8 (3.5)	26.7 (5.2)	25.0 (5.9)
Completeness (%)	99.0 (97.9)	98.0 (99.5)	99.8 (100.0)	97.7 (89.0)
Measured reflections	167,626	43,282	62,749	26,983
Unique reflections	27,812	14,246	11,260	15,316
R_{sym} (%)	6.9 (38.0)	7.4 (33.9)	6.4 (36.5)	2.2 (13.2)
Refinement				
Resolution (Å)	17.0–2.3	17.0–2.25	73.3–2.80	17.0–1.80
Number of monomers in the asymmetric unit	4	2	2	1
Number of protein atoms/solvent molecules	4540/449	2277/176	2230/0	1202/88
RMSD bond length (Å)/RMSD bond angle (°)	0.008/1.56	0.008/1.51	0.008/1.11	0.0069/1.6
Number of reflections used for R_{work}/R_{free}	26,144/1385	12,879/665	10,772/538	13,181/691
Ramachandran plot: preferred region/outliers (%)	93.1/1.5	96.4/0.7	89.8/4.7	95.3/1.3
R_{cryst}/R_{free} (%)	25.2/29.4	21.7/23.6	28.0/31.2	21.9/23.7
Average overall *B*-value for protein/solvent (Å^2)	28.5/35.9	23.1/35.7	73.7/—	25.1/38.5

molecules, particularly the W of the ^{664}DKW^{666} motif (Fig. S4). It is not surprising that this region takes part in mediating crystal packing interactions, as the epitope was purposely grafted onto the surface of the scaffold for eventual functional presentation. The present analysis of the epitope conformation was performed taking these artifactual molecular environments into consideration.

The computational models of SC_2cx5 and BB_2cx5 were in close agreement with the determined structures, with overall backbone (N, C$^\alpha$, C, and O) RMSD values of 0.62 Å and 0.59 Å, respectively. Within the overall backbone alignment, the SC_2cx5 model had a backbone RMSD of 0.71 Å to the crystal structure over the grafted epitope region. Over the 15-residue segment that was structurally modified to accommodate the epitope by backbone grafting, BB_2cx5_001 had an RMSD of 1.13 Å. The RMSD value was slightly higher (1.56 Å) for the seven-residue epitope region alone (LELDKWA) (Fig. 7a and b). Consistent with the high affinities for 2F5, the epitope conformation on both of these scaffolds was similar to that of the 2F5-bound structure of the gp41 peptide described in 1tji (backbone RMSD values of 0.61 Å for SC_2cx5 and 0.66 Å for BB_2cx5_001). Further, modeling the interaction of these molecules with 2F5 showed that both scaffolds could productively interact with the antibody (not shown).

The epitope backbone had an RMSD of 1.38 Å between the crystal structures of SC_2cx5 and BB_2cx5_001 (Fig. 7c). The A86L mutation identified by backbone grafting provided additional contacts between 2F5 and BB_2cx5 and was probably responsible for most of the change in epitope conformation between the two structures (Fig. 7d). When modeled computationally on the crystal structure of SC_2cx5, a leucine at position 86 clashed with the scaffold. In the BB_2cx5_001 crystal structure, the new orientation of the epitope backbone allowed L86 to make stabilizing interactions with A53 and L95 of the scaffold. Firm conclusions about the epitope conformation in solution based on the crystallographic analysis of the SC_2cx5 and BB_2cx5_001 structures cannot be made, given that the conformation of the epitope region may be influenced by significant contacts with symmetry-related molecules in both crystals (Fig. S4).

The structures of SC_1wnu and BB_1wnu_001 were in close agreement with the computational models, with overall backbone RMSD values of 0.9 Å (Fig. 7e and f). Over the remodeled epitope region (LELDKWA), the backbone RMSD values were 1.1 Å for SC_1wnu and 2 Å for BB_1wnu_001, but the conformation of this region may be affected by extensive contacts between the epitope and symmetry-related molecules present in the crystal

structure (Fig. S4). The backbone conformation of a 15-residue segment was modified to transplant the 2F5 epitope to BB_1wnu_001, and over this region, the computational models showed an RMSD of 1.1 Å to the crystal structure. The epitope conformation presented on SC_1wnu accurately mimicked the epitope conformation induced by 2F5 upon gp41 peptide binding (backbone RMSD values of 0.9 Å) and could interact with 2F5 when the antibody–scaffold complex was modeled (not shown). In the case of BB_1wnub, although the conformation of the epitope is very similar to the conformation of 2F5-bound gp41 (RMSD = 0.5 Å), modeling the interaction of 2F5 with the determined structure resulted in significant backbone clashes between the antibody and regions of the scaffold that were not modified during the computational procedure. Given the high affinity of the interaction between 2F5 and BB_1wnu_001 (K_d = 9 nM), these clashes suggested that the solution conformation of the epitope is different from the one present in the crystal structure, which is affected by crystal packing or that there is enough flexibility in the epitope region to allow for conformational changes upon 2F5 binding.

The structures of BB_1wnu and SC_1wnu were very similar (backbone RMSD of 1 Å), but there were considerable differences in the epitope region (RMSD of 2.2 Å) (Fig. 7g and h). The DKW motif in SC_1wnu was located between a proline and a glycine that were mutated to leucine and alanine, respectively, in BB_1wnu (P50L and G54A mutations in Fig. 7h). These two residues are equivalent to L663 and A667 of the 2F5 epitope on gp41 and contribute additional contacts to the antibody. Furthermore, in BB_1wnu, these two residues interact with M58, which is mutated from an isoleucine in SC_1wnu.

Overall, the crystallographic analysis validated the computational procedure and indicated that the designed epitope scaffolds present the 2F5-bound conformation of the epitope on their surfaces.

Discussion

Epitope backbone *versus* side-chain grafting

Here, we described a new computational method, backbone grafting of linear motifs, and tested the method by engineering novel 2F5 epitope scaffolds. The epitope scaffolds designed by backbone grafting bound the antibody with higher affinity than gp41-derived peptides corresponding to the scaffolded epitope segments. Further, backbone grafting scaffolds had higher affinity than comparable side-chain grafting scaffolds based on the same parent scaffolds, in two of three cases. In the third case (1zko), neither the side-chain grafting nor the backbone grafting scaffolds had significant affinity for the 2F5 antibody; thus, backbone grafting was unable to rescue a failed scaffold in that case, but there was no evidence that backbone grafting was less successful than side-chain grafting. These results show that backbone grafting of linear epitopes is a promising approach to design high-affinity antigens in the future.

In the cases for which backbone grafting produced higher-affinity scaffolds compared to side-chain grafting, structural analysis of unliganded epitope scaffolds showed good agreement between the structures and the computational models and allows us to speculate on the reasons for the binding improvements. Small changes in the backbone conformation allowed the transplantation of additional epitope residues for backbone grafting scaffolds, which in turn provided more contacts to the antibody. It is also possible that the designed alterations to the scaffold backbone made attaining the ideal epitope conformation upon antibody binding more favorable energetically. Crystal structures of the scaffolds in complex with 2F5 would greatly contribute to confirming these hypotheses; however, diffraction-quality crystals of such complexes so far remain elusive.

Compared to epitope side-chain grafting, backbone grafting is considerably less dependent on the native backbone of a given scaffold. In the matching stage of backbone grafting, candidate scaffolds are selected based on their ability to accommodate only the terminal residues of the epitope as well as the ability to sterically accommodate the epitope shape. This is in contrast to the matching stage of epitope side-chain grafting in which candidate scaffolds must structurally match *all* the backbone residues of a given motif. However, integrating and stabilizing the desired conformation of an epitope on a scaffold are more challenging in the case of backbone grafting. While epitope stabilization is intrinsic to the side-chain grafting procedure, backbone grafting relies on the ability of the computational protocol to accurately engineer and model novel protein backbone segments and to identify sequence changes on the scaffold to stabilize the desired epitope conformation. Therefore, for the transplantation of a given epitope, side-chain grafting may be more efficient if appropriate candidate scaffolds can be identified, while backbone grafting would be required in cases when no suitable scaffolds exist for side-chain grafting. However, a key finding of this study is that backbone grafting scaffolds can achieve higher affinities than side-chain grafting scaffolds in some cases. This demonstrates that flexible backbone modeling can improve the design of high-affinity epitope scaffolds, even in cases where side-chain grafting is possible.

Transplantation of the 2F5 epitope

The epitope scaffolds described here were soluble and expressed in large quantities in *E. coli* cells without requiring refolding or production in mammalian cells, unlike the first-generation side-chain grafting designs previously reported for the 2F5 epitope.[15,26] These improvements were probably due to additional design steps in this study, such as surface redesign.[25] The 2F5 affinities of the epitope scaffolds in this study are comparable to the first-generation 2F5 epitope scaffolds, although the scaffolds in this study present less of the 2F5 epitope on gp41. The 2F5 affinities of the first-generation epitope scaffolds were initially determined by Ofek *et al.* by SPR with the epitope scaffold immobilized on the sensor chip and 2F5 Fab injected as analyte (measured K_d values between 0.6 and 18.8 nM).[15] In a subsequent study by Guenaga *et al.*,[26] these interactions were measured in the same format employed here—2F5 IgG immobilized on the chip and the epitope scaffold injected as analyte—and with one exception (ES2, $K_d = 1.1$ nM), the reported 2F5 affinities were lower than those determined by Ofek *et al.* (K_d values between 41 and 85 nM). Furthermore, as noted in the follow-up study by Guenaga *et al.*, the actual K_d values for the 1:1 interaction between 2F5 and some of the first-generation epitope scaffolds might be even higher, due to the possible oligomerization of epitope scaffolds that may have led to avidity gains in the measurements. No solution multimeric state or thermal stability data were reported by those two studies on the first-generation epitope scaffolds. The epitope scaffolds described in this study were shown to have high melting temperatures and to be monomeric in solution, and the determined dissociation constants represent true measurements of the 1:1 interaction between 2F5 and the epitope scaffolds (with the noted exception of SC_2cx0a). Overall, 10 of the epitope scaffolds discussed here bound 2F5 with K_d values between 1 nM and 100 nM. These epitope scaffolds present between 62% (^{660}ELDKWA^{667}) and 74% (^{661}LELDKWA^{667}) of the 2F5 binding area on gp41, while the first-generation epitope scaffolds displayed 87% (^{659}ExLELDKWAxL^{669}) of the binding area and allowed one additional hydrogen bond across the interface (gp41 $Gln_{658}O$–2F5 $Gln_{L27}N^{\varepsilon 2}$). Given the similar affinities, this suggests increased epitope stabilization and/or better conformational mimicry on the epitope scaffolds designed in this study. To further support this, the kinetic association rates (k_{on}) of the epitope scaffolds discussed here (values between 5×10^5 and 10^7 $M^{-1}s^{-1}$) are at least 1 order of magnitude faster than the k_{on} values of the first-generation epitope scaffolds, except for ES2 ($k_{on} = 10^6$ $M^{-1}s^{-1}$).[15] The epitope scaffold with the highest affinity in this study, BB_2cx5_001 ($K_d = 1.4$ nM), was engineered by backbone grafting. Its 2F5 affinity is in the same range as that of the longer, full-length gp41epitope ($K_d = 4.5$ nM) and that of the best first-generation 2F5 epitope scaffold (ES2, $K_d = 1.1$ nM), which also presented a longer stretch of the epitope.

Future applications of epitope backbone grafting

Transplantation of functional motifs onto heterologous proteins is a general approach to engineer novel molecules, such as antigens, protein inhibitors, or enzymes. The majority of previous protein grafting studies relied on grafting the side chains of a given functional motif onto regions of high structural homology of candidate scaffold proteins. These approaches are restricted by the existence of structurally characterized scaffolds with exposed backbone similar in conformation to a particular motif. To avoid this limitation, numerous efforts involved the grafting of motifs located on common secondary-structure elements, such as single α-helices and polyproline type II helices,[2,32–35] α-helical coiled coils,[3,9,11,12,36] and β-sheets,[10,13,14,37] to analogous regions on heterologous proteins. In contrast, backbone grafting of linear motifs requires reduced structural similarity between a functional motif and candidate scaffolds, since only the termini of the motif, rather than the entire backbone length, have to be structurally matched with a scaffold for transplantation. Furthermore, backbone grafting has the added ability to alter the backbone conformation of a candidate scaffold and to engineer interactions between the transplanted motif and the scaffold such that the desired structural conformation of the motif is stabilized. Therefore, backbone grafting may allow transplantation of structurally diverse linear functional motifs when no suitable scaffolds exist for side-chain grafting. Although the new protocol presented here is capable in principle of aggressive scaffold remodeling and design, the results presented here only validate this method for motif transplantation into structurally similar regions of target scaffolds. Further studies will assess the ability of this computational method to graft functional motifs into scaffold regions with low structural homology.

Materials and Methods

Epitope side-chain grafting

Scaffold matching for side-chain grafting

To identify potential scaffolds, we aligned subranges of the 2F5 epitope centered on the core sequence ^{664}DKW^{666} against proteins from a curated version of the PDB, containing 13,337 monomeric proteins crystallized without

ligands at a resolution better than 3 Å, as previously described[16] (Fig. S5). RMSD values were computed between the backbone of the epitope and that of the candidate protein scaffold (Table S1). An RMSD cutoff of 0.1 Å per residue was used to select potential scaffolds, which were further filtered based on potential backbone clashes with the 2F5 mAb.

Epitope side-chain grafting

The design stage was composed of two steps: (1) a binding interface optimization and (2) a sequence design step for epitope accommodation and removal of extraneous interfacial interactions. In the binding interface optimization, the native rotamers of epitope side chains were transplanted to variants of the candidate scaffolds in which the identity of all other scaffold residues were initially changed to glycine. Maintaining the full side-chain detail of the epitope residues and the antibody, both side-chain and rigid-body degrees of freedom at the antibody–scaffold binding interface were simultaneously minimized.[38] The temporary usage of glycine during this step for scaffold residues noncritical for binding ensures that extraneous or glancing interfacial interactions from the original scaffold residues do not exist and allows for a computational speedup during the minimization. Subsequently for sequence design, the native side chains of the scaffold outside the epitope region were recovered and the scaffold was further designed in the context of its complex with 2F5 mAb. Residues within 4 Å of the epitope were considered "intra-design" positions and all residues of the scaffold within 4 Å of the antibody were defined as "inter-design" positions. During the automated design process, inter-design residues were changed to either serine, alanine, glycine, or threonine, while the intra-design residues were allowed to change to any amino acid except cysteine. A Rosetta energy-based $\Delta\Delta G$ was calculated between the epitope scaffolds and the antibody, and scaffolds that showed unfavorable interactions were discarded. Finally, a step of human-guided design was performed to optimize the scaffolds further, such as removing scaffold–mAb contacts outside the epitope, reverting scaffold mutations not affecting the scaffold–mAb interaction back to the native scaffold residues, and eliminating unpaired cysteines and undesired functional sites.

Epitope backbone grafting

Epitope backbone grafting is a computational protocol that (a) identifies scaffolds that can accommodate the backbone conformation of a given linear epitope, (b) replaces corresponding native scaffold backbone with the desired backbone conformation of an epitope, (c) integrates the epitope with the native scaffold such that the desired epitope conformation is stabilized, and (d) ensures that the antibody can interact with the resulting epitope scaffold. The computational procedure is divided into two stages—matching and design—as discussed below.

Scaffold matching for epitope backbone grafting

Two types of alignments, "N2C" and "C2N", are used to identify scaffolds for epitope backbone grafting (Fig. S5). In N2C alignment, the backbone atoms (N, C, and C^{α}) of the N terminus residue of the epitope were aligned on candidate scaffolds and RMSD values were measured between the C terminus of the epitope and proximal residues on the scaffold. In C2N alignment, the backbone atoms (N, C, and C^{α}) of the C terminus residue of the epitope was aligned on candidate scaffolds and RMSD values were measured between the N terminus of the epitope and proximal residues on the scaffold. Scaffolds with RMSD values below 3 Å were considered a preliminary match. For each match, the native scaffold backbone corresponding to the epitope was removed and replaced by the epitope itself per the alignment. This resulted in scaffolds anchoring the epitope at one end and leaving an open gap at the opposing end. Two steric clash filters using the Lennard–Jones component of the Rosetta energy function were then applied as an elementary assessment of the ability of the scaffold to accommodate the epitope shape. Only matches passing both filters continued onward to the grafting stage. The first "intra-clash" filter evaluated the clash of the newly placed epitope backbone against the existing scaffold backbone. The presence of steric clash beyond a threshold indicates that the resulting epitope-scaffold backbone may self-intersect in an irrecoverable manner and that the match should be removed from consideration. For this study, the threshold for intra-clash was 1000 Rosetta energy units. The second "inter-clash" filter evaluated the clash of the antibody backbone, as placed in the match-derived rigid-body orientation relative to the epitope, against the existing scaffold backbone. Matches presenting an excessive amount of inter-clash beyond a threshold (a value of 3000 Rosetta energy units in this study) are indicative of anti-complementarity at the interface between antibody and scaffold and were removed from further consideration.

Epitope backbone grafting design

A standard Rosetta procedure for loop closure utilizing a low-resolution scoring function, fragment insertion,[28] CCD,[29] and Monte Carlo sampling was employed to close the gaps in N2C and C2N matches. In this stage, the amino acid identity of the scaffold and epitope–antibody complex was temporarily changed to alanine. The usage of alanine allows for a relatively neutral sampling of backbone dihedrals in Ramachandran space during loop closure, as we do not know prior to sequence design what the optimal sequence of the moveable residues should be. In addition, it ensures that extraneous or glancing interfacial interactions from the original scaffold do not exist, and the reduced number of atoms allows for a computational speedup. During loop closure, the values for the dihedral angles of four scaffold residues and one epitope residue were allowed to vary on each side of the grafted epitope, while the backbone of the remaining residues was held fixed. For each mobile residue, two sets of 200 backbone fragments, one set each of fragment length 3 and fragment length 1, were selected randomly based upon the preexisting secondary-structure type (loop, helix, or sheet) present in the scaffold and epitope. Throughout the process, the antibody was kept in fixed rigid-body orientation relative to the epitope, allowing the antibody to swing with the epitope as the epitope moved during loop closure. The polypeptide

chain was considered closed when the chainbreak score[30] was less than 0.2. Additionally, the Ramachandran score[28] for any residue with modified backbone dihedrals was required to be less than 1.5 to ensure that no backbone dihedrals significantly violated the allowed space of Ramachandran angles. Due to the number of moveable residues (10), each loop closure attempt consisted of 1000 total cycles of the procedure. A total of five loop closures were attempted, and the three best-scoring closed epitope scaffolds were allowed to move to the next stage of the procedure.

For epitope scaffolds that achieved satisfactory closure, backbone refinement utilizing a high-resolution energy function was employed to catch problematic conformations of the backbone regions that were allowed to change. Since the prior loop closure procedure was performed using a low-resolution scoring function, issues with the backbone such as improper clashes could have been introduced. Backbone refinement therefore consisted of utilizing all-atom loop refinement while the scaffold was still all-alanine in sequence[30] to allow for adjustments based primarily on the backbone and not side-chain interactions. Problematic backbones with large clashes were expected to fail to re-obtain proper loop closure geometry during loop refinement. Subsequently, the same closure conditions employed during the prior loop closure procedure were checked again, and any epitope scaffolds failing those conditions were removed from further consideration. Full-sequence and side-chain information of both epitope scaffold and antibody were then recovered, and a full-atom design stage was employed to ensure that the grafted backbone interacted favorably with the rest of the scaffold and that the antibody could bind to the resulting epitope scaffold. This design stage used the identical inter- and intra-residue classification as presented in epitope side-chain grafting. Finally, all-atom loop refinement was performed to refine the regions of the backbone that were allowed to change during the grafting process, allowing the epitope scaffold to settle into a final state.[30] Resulting epitope scaffolds were ranked based on their predicted ability to bind 2F5, their protein core interactions, the number of buried hydrogen donors or acceptors, and the values of the dihedral angles in the regions allowed to move during loop closure.

Expression and purification of epitope scaffolds

Plasmids encoding individual computational designs were synthesized with a 6×His-terminal tag in a pet29b+ vector (GenScript) and subsequently transformed in Arctic Express™ *E. coli* cells (Stratagene). Single colonies were grown overnight in 50 mL of LB with kanamycin (1 mg/mL). Starter culture (10 mL) was expanded into 1 L of LB culture with kanamycin (1 mg/mL) and grown at 37 °C to an absorbance of 0.5 (~3 h). Protein expression was induced overnight at 12 °C by adding 500 μL of 1 M IPTG (Sigma-Aldrich).

Cells were harvested by spinning at 6000 rpm for 20 min, resuspended in start buffer (160 mM imidazole, 4 M NaCl, and 20 mM Na_2PHO_4, pH 7.4) with one tablet of ethylenediaminetetraacetic acid (EDTA)-free Protease Inhibitor (Roche), and frozen at −20 °C until ready for processing. Ten milliliters of 10× Bugbuster (Novagen), 50 μL of Benzoase Nuclease (Novagen), and 1.7 μL of rLysozyme (Novagen) were added to the resuspended cells, which were then lysed by rocking at room temperature for 20 min. The cell lysate was spun for 20 min at 11,000 rpm, and the resulting supernatant was filtered (0.22 μm, Millipore) and rocked on an orbital shaker at 4 °C with 5 mL of Ni^{2+} Sepharose 6 Fast Flow beads (GE Healthcare) for 1 h. The beads were pelleted and washed three times with 30 mL of wash buffer (50 mM imidazole, 500 mM NaCl, and 20 mM Na_2PHO_4, pH 7.4) by spinning at 500*g* for 10 min. The Ni beads in the wash buffer were then applied to a sterile column, and the protein was eluted with 20 mL of elution buffer (250 mM imidazole, 500 mM NaCl, and 20 mM sodium phosphate, pH 7.4).

Proteins were further purified using size-exclusion chromatography on a Superdex 75 26/60 column (GE Healthcare) in Hepes-buffered saline (HBS) buffer (10 mM Hepes, 150 mM NaCl, and 3 mM EDTA, pH 7.4). The collected fractions were analyzed on a 4–12% SDS denaturing gel (Invitrogen). Positive fractions were combined, and the final concentration was determined by measuring the UV absorption signal at 280 nm (Nanodrop 1000). The overall yield was 5–15 mg of pure protein per liter of culture.

Biophysical characterization of epitope scaffolds

CD spectra

The stability of the proteins was measured by CD on an Aviv 62A DS spectrometer by far-UV (200–260 nm) wavelength scan followed by thermal melting. Two hundred fifty microliters of each protein, at a concentration between 20 and 50 μM, in either phosphate-buffered saline or HBS, was analyzed in a 1-mm cuvette. Following the UV scan, the following wavelengths were selected to monitor the change in ellipticity during temperature denaturation experiments: SC_2cx5, 220 nm; SC_1zs7, 220 nm; SC_1wnu, 220 nm; SC_1vr9, 224 nm; BB_2cx5_001, 218 nm; BB_1wnu_001, 219 nm. Experiments were carried over a temperature range from 5 to 99 °C, with 2 °C increments every 3 min, and the resulting data were converted to mean residue ellipticity and fitted to a two-state model.

Light scattering

The oligomerization state of the proteins was determined by static light scattering (miniDAWN TREOS, Wyatt) coupled in-line to HPLC (Agilent, 1200 series). One hundred twenty microliters of each protein sample at 2–5 mg/mL was run in phosphate-buffered saline, and the resulting data were analyzed using the ASTRA software (Wyatt).

2F5 mAb affinity

Binding to 2F5 mAb was measured by SPR (Biacore 2000) using the human antibody capture kit (GE Healthcare). We immobilized 7000–9000 response units of mouse anti-human IgG on a CM5 chip via amine coupling following the manufacturer's instructions (GE Healthcare). For kinetic analysis, 200–500 response units of 2F5 IgG

was captured, and after 1 min of surface stabilization, increasing concentrations of epitope scaffolds were injected in duplicates at 50–100 μL/min for 1–5 min. Between protein runs, two injections of 10 μL 3 M $MgCl_2$ at a flow rate of 10 μL/min were used to dissociate 2F5 IgG from the surface. Experiments were conducted in HBS-EP buffer [GE Healthcare, 0.01 M Hepes, pH 7.4, 0.15 M NaCl, 3 mM EDTA, and 0.005% (v/v) Surfactant P20]. One flow cell with no 2F5 IgG captured was used as a reference cell, and buffer-only injections were used for blank subtraction. Data analysis was done using Scrubber 2.0 (BioLogic Software). For kinetic analysis, biosensor data were globally fit to a mass transport limited 1:1 Langmuir binding model. The data sets for two proteins (SC_1wnu and BB_2cx5_002) were fitted with a two-stage interaction model (1: A+B↔AB; 2: AB↔AB*). For equilibrium analysis, each data set was fitted to a single-site interaction model based on the response value at equilibrium, the concentration of the protein, and the maximum response obtained when all binding sites are occupied (R_{max}).

Crystal structure determination

Crystals were grown using the hanging drop vapor diffusion method with the protein concentration at 8–10 mg/mL. The starting volume of the drops was 2 μL: 1 μL of protein solution mixed with 1 μL of reservoir solution. Details of the crystallization conditions are given in Table 3. Before flash-freezing in boiling nitrogen, crystals were soaked for a few seconds in reservoir solution supplemented with 20% glycerol. X-ray diffraction data for SC_2cx5, BB_2cx5_001, and BB_1wnu_001 were collected on a Rigaku MicroMax-007 HF generator run at 40 kV and 30 mA and equipped with Rigaku VariMax multi-layer optics and a mar345 image plate detector. X-ray diffraction data for SC_1wnu were collected at the Canadian Light Source, Beamline 08ID-1, using a MarMosaic 225 CCD detector.

Raw diffraction data were reduced using either XDS[39] or HKL2000.[40] Phases for all structures were calculated by molecular replacement using the program Phaser[41] and using the PDB entry of the parent scaffold protein as the initial search model. Experimental electron density maps were calculated with CNS[42] and interpreted using the program Coot.[43] The resulting model was refined with CNS. Ramachandran plot statistics were calculated using Coot.[43] Solvent molecules were picked with the CNS algorithm using the following input parameters: (1) a peak over 3.5 σ in the difference electron density map and (2) at hydrogen bond distance from neighboring donor or acceptor atoms. Statistical data for data collection and model refinement are given in Table 3. Figures were prepared using the program PyMOL.

Accession numbers

The crystal structures discussed in this article were deposited in the PDB with the following codes: 3RIJ (SC_2cx5), 3RI0 (BB_2cx5_001), 3RHU (SC_1wnu), and 3RFN (BB_1wnu_001).

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.10.003

Acknowledgements

We would like to thank Annie Cunningham and Emelie Szabo for their assistance with the crystallization efforts. We would also like to thank Bruno E. Correia and Sergey Menis for comments on the manuscript. This work was supported by a grant from the International AIDS Vaccine Initiative Neutralizing Antibody Consortium and by the Canada Research Chairs Program. The data set for the SC_1wnu crystal was collected at the Canadian Light Source, which is supported by the Natural Sciences and Engineering Research Council of Canada, National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada, and the University of Saskatchewan.

References

1. Kritzer, J. A., Zutshi, R., Cheah, M., Ran, F. A., Webman, R., Wongjirad, T. M. & Schepartz, A. (2006). Miniature protein inhibitors of the p53–hDM2 interaction. *ChemBioChem*, **7**, 29–31.
2. Cobos, E. S., Pisabarro, M. T., Vega, M. C., Lacroix, E., Serrano, L., Ruiz-Sanz, J. & Martinez, J. C. (2004). A miniprotein scaffold used to assemble the polyproline II binding epitope recognized by SH3 domains. *J. Mol. Biol.* **342**, 355–365.
3. Sia, S. K. & Kim, P. S. (2003). Protein grafting of an HIV-1-inhibiting epitope. *Proc. Natl Acad. Sci. USA*, **100**, 9756–9761.
4. Chin, J. W. & Schepartz, A. (2001). Concerted evolution of structure and function in a miniature protein. *J. Am. Chem. Soc.* **123**, 2929–2930.
5. Zanghellini, A., Jiang, L., Wollacott, A. M., Cheng, G., Meiler, J., Althoff, E. A. *et al.* (2006). New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci.* **15**, 2785–2794.
6. Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Rothlisberger, D., Zanghellini, A. *et al.* (2008). De novo computational design of retro-aldol enzymes. *Science*, **319**, 1387–1391.
7. Rothlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J. *et al.* (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, **453**, 190–195.
8. Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., St Clair, J. L. *et al.* (2010). Computational design of an enzyme catalyst for a stereoselective bimolecular Diels–Alder reaction. *Science*, **329**, 309–313.
9. Li, Y., Kaur, H. & Oakley, M. G. (2008). Probing the recognition properties of the antiparallel coiled coil motif from PKN by protein grafting. *Biochemistry*, **47**, 13564–13572.
10. Drakopoulou, E., Zinn-Justin, S., Guenneugues, M., Gilqin, B., Menez, A. & Vita, C. (1996). Changing the structural context of a functional beta-hairpin. Synthesis and characterization of a chimera containing

the curaremimetic loop of a snake toxin in the scorpion alpha/beta scaffold. *J. Biol. Chem.* **271**, 11979–11987.
11. Lu, S. M. & Hodges, R. S. (2002). A de novo designed template for generating conformation-specific antibodies that recognize alpha-helices in proteins. *J. Biol. Chem.* **277**, 23515–23524.
12. Tripet, B., Kao, D. J., Jeffers, S. A., Holmes, K. V. & Hodges, R. S. (2006). Template-based coiled-coil antigens elicit neutralizing antibodies to the SARS-coronavirus. *J. Struct. Biol.* **155**, 176–194.
13. Vita, C., Roumestand, C., Toma, F. & Menez, A. (1995). Scorpion toxins as natural scaffolds for protein engineering. *Proc. Natl Acad. Sci. USA*, **92**, 6404–6408.
14. Rajagopal, S., Meyer, S. C., Goldman, A., Zhou, M. & Ghosh, I. (2006). A minimalist approach toward protein recognition by epitope transfer from functionally evolved beta-sheet surfaces. *J. Am. Chem. Soc.* **128**, 14356–14363.
15. Ofek, G., Guenaga, F. J., Schief, W. R., Skinner, J., Baker, D., Wyatt, R. & Kwong, P. D. (2010). Elicitation of structure-specific antibodies by epitope scaffolds. *Proc. Natl Acad. Sci. USA*, **107**, 17880–17887.
16. Correia, B. E., Ban, Y. E., Holmes, M. A., Xu, H., Ellingson, K., Kraft, Z. *et al.* (2010). Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure*, **18**, 1116–1126.
17. Hu, X., Wang, H., Ke, H. & Kuhlman, B. (2007). High-resolution design of a protein loop. *Proc. Natl Acad. Sci. USA*, **104**, 17668–17673.
18. Murphy, P. M., Bolduc, J. M., Gallaher, J. L., Stoddard, B. L. & Baker, D. (2009). Alteration of enzyme specificity by computational loop remodeling and design. *Proc. Natl Acad. Sci. USA*, **106**, 9215–9220.
19. Correia, B. E., Ban, Y. E., Friend, D. J., Ellingson, K., Xu, H., Boni, E. *et al.* (2011). Computational protein design using flexible backbone remodeling and resurfacing: case studies in structure-based antigen design. *J. Mol. Biol.* **405**, 284–297.
20. Sammond, D. W., Bosch, D. E., Butterfoss, G. L., Purbeck, C., Machius, M., Siderovski, D. P. & Kuhlman, B. (2011). Computational design of the sequence and structure of a protein-binding peptide. *J. Am. Chem. Soc.* **133**, 4190–4192.
21. Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J. R., Wyatt, R. & Kwong, P. D. (2004). Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J. Virol.* **78**, 10724–10737.
22. Julien, J. P., Bryson, S., Nieva, J. L. & Pai, E. F. (2008). Structural details of HIV-1 recognition by the broadly neutralizing monoclonal antibody 2F5: epitope conformation, antigen-recognition loop mobility, and anion-binding site. *J. Mol. Biol.* **384**, 377–392.
23. Bryson, S., Julien, J. P., Hynes, R. C. & Pai, E. F. (2009). Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. *J. Virol.* **83**, 11862–11875.
24. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
25. Lawrence, M. S., Phillips, K. J. & Liu, D. R. (2007). Supercharging proteins can impart unusual resilience. *J. Am. Chem. Soc.* **129**, 10110–10112.
26. Guenaga, J., Dosenovic, P., Ofek, G., Baker, D., Schief, W. R., Kwong, P. D. *et al.* (2011). Heterologous epitope-scaffold prime:boosting immuno-focuses B cell responses to the HIV-1 gp41 2F5 neutralization determinant. *PLoS One*, **6**, 1–12.
27. Tian, Y., Ramesh, C. V., Ma, X., Naqvi, S., Patel, T., Cenizal, T. *et al.* (2002). Structure–affinity relationships in the gp41 ELDKWA epitope for the HIV-1 neutralizing monoclonal antibody 2F5: effects of side-chain and backbone modifications and conformational constraints. *J. Pept. Res.* **59**, 264–276.
28. Rohl, C. A., Strauss, C. E., Chivian, D. & Baker, D. (2004). Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, **55**, 656–677.
29. Canutescu, A. A. & Dunbrack, R. L., Jr (2003). Cyclic coordinate descent: a robotics algorithm for protein loop closure. *Protein Sci.* **12**, 963–972.
30. Wang, C., Bradley, P. & Baker, D. (2007). Protein–protein docking with backbone flexibility. *J. Mol. Biol.* **373**, 503–519.
31. Qian, B., Raman, S., Das, R., Bradley, P., McCoy, A. J., Read, R. J. & Baker, D. (2007). High-resolution structure prediction and the crystallographic phase problem. *Nature*, **450**, 259–264.
32. Rutledge, S. E., Volkman, H. M. & Schepartz, A. (2003). Molecular recognition of protein surfaces: high affinity ligands for the CBP KIX domain. *J. Am. Chem. Soc.* **125**, 14336–14347.
33. Golemi-Kotra, D., Mahaffy, R., Footer, M. J., Holtzman, J. H., Pollard, T. D., Theriot, J. A. & Schepartz, A. (2004). High affinity, paralog-specific recognition of the Mena EVH1 domain by a miniature protein. *J. Am. Chem. Soc.* **126**, 4–5.
34. Gemperli, A. C., Rutledge, S. E., Maranda, A. & Schepartz, A. (2005). Paralog-selective ligands for bcl-2 proteins. *J. Am. Chem. Soc.* **127**, 1596–1597.
35. Zondlo, N. J. & Schepartz, A. (1999). Highly specific DNA recognition by a designed miniature protein. *J. Am. Chem. Soc.* **121**, 6938–6939.
36. Domingues, H., Cregut, D., Sebald, W., Oschkinat, H. & Serrano, L. (1999). Rational design of a GCN4-derived mimetic of interleukin-4. *Nat. Struct. Biol.* **6**, 652–656.
37. Martin, L., Stricher, F., Misse, D., Sironi, F., Pugniere, M., Barthe, P. *et al.* (2003). Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. *Nat. Biotechnol.* **21**, 71–76.
38. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A. & Baker, D. (2003). Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281–299.
39. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800.
40. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr., Part A*, **276**, 307–326.

41. Mccoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
42. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
43. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.