Published in final edited form as:

Proteins. 2014 September ; 82(9): 1884–1895. doi:10.1002/prot.24542.

Effects of protein engineering and rational mutagenesis on crystal lattice of single chain antibody fragments

Sibel Kalyoncu¹, Jeongmin Hyun², Jennifer C. Pai², Jennifer L. Johnson¹, Kevin Entzminger², Avni Jain², David P. Heaner Jr.¹, Ivan A. Morales¹, Thomas M. Truskett², Jennifer A. Maynard^{2,*}, and Raquel L. Lieberman^{1,*}

¹School of Chemistry & Biochemistry, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, GA 30332-0400

²McKetta Department of Chemical Engineering, University of Texas at Austin, MC0400, 1 University Station, Austin, TX 78712

Abstract

Protein crystallization is dependent upon, and sensitive to, the intermolecular contacts that assist in ordering proteins into a three dimensional lattice. Here we used protein engineering and mutagenesis to affect the crystallization of single chain antibody fragments (scFvs) that recognize the EE epitope (EYMPME) with high affinity. These hypercrystallizable scFvs are under development to assist difficult proteins, such as membrane proteins, in forming crystals, by acting as crystallization chaperones. Guided by analyses of intermolecular crystal lattice contacts, two second-generation anti-EE scFvs were produced, which bind to proteins with installed EE tags. Surprisingly, although non-complementarity determining region (CDR) lattice residues from the parent scFv framework remained unchanged through the processes of protein engineering and rational design, crystal lattices of the derivative scFvs differ. Comparison of energy calculations and the experimentally-determined lattice interactions for this basis set provides insight into the complexity of the forces driving crystal lattice choice and demonstrates the availability of multiple well-ordered surface features in our scFvs capable of forming versatile crystal contacts.

Keywords

protein crystallography; crystal packing; intermolecular contacts; short peptide epitope; scFv affinity; protein interaction

Introduction

Macromolecular crystallization requires the association of weak, but specific, intermolecular interactions into a repeated lattice array, and is often the rate limiting step to structure

*To whom correspondence may be addressed: Jennifer A. Maynard, maynard@che.utexas.edu or Raquel L. Lieberman, raquel.lieberman@chemistry.gatech.edu.

Supporting Information

Additional supporting information includes Table S1 and Table S2, which are related to Table II, and Figure S1 which is related to Figure 3.

determination by X-ray crystallography.¹ Although crystallization and structure determination of proteins is successful in many cases, numerous techniques have been implemented in which target proteins are modified to enhance crystallizability. Proteins with increased solubility and stability are less prone to interactions leading to undesirable off-pathways of aggregation or precipitation.^{2,3} Biophysical properties can be augmented through the use of libraries⁴ and/or optimization of target constructs to remove flexible regions,⁵ which reduces the probability of unproductive contacts, and truncation of loops that reduce conformational heterogeneity. A newer strategy, Surface Entropy Reduction (SER), can optimize a protein for the entropic cost of crystal packing by replacing flexible amino acids (Glu, Gln, Lys) with alanines.⁶⁻⁸

An emerging method is co-crystallization chaperones that assist in the crystallization of a target protein, particularly useful for membrane proteins and other proteins that present few surface polar residues to form crystal contacts.⁹ The addition of a chaperone protein can sequester aggregation-prone hydrophobic regions and can serve to immobilize flexible regions.^{10,11} Structure determination of a G-protein coupled receptor (GPCR) using a fused T4 lysozyme in place of a flexible loop and a nanobody chaperone specific for the GPCR demonstrates the utility of this approach.¹² To date, successful crystallization chaperones have been target-specific, limiting the generality of the chaperone method, as a new chaperone must be identified for each new target. Chaperone-assisted crystallization has the potential to become a general platform, however, if chaperones can be engineered for high affinity to a specific short peptide epitope that can be easily installed into any target protein of interest by site-directed mutagenesis (SDM). We believe that an ideal chaperone toolbox would include multiple hypercrystallizable chaperones, each with specificity for a discrete epitope and a predictable crystallization lattice to accommodate client proteins of various sizes. Exquisite control over crystallization may not be easily achievable, but for our set of engineered crystallization chaperones under development, we hypothesized that systematically exploring crystallization propensities may reveal governing principles and lead to improved chaperones.

Here, we used a combination of protein design, crystallization, structure determination and computational analysis to investigate the extent to which protein crystallization may be controlled by tuning the energetics of intermolecular contacts within a highly conserved (sequence identity ~85%) set of murine single chain antibody fragments (scFvs) that are candidate crystallization chaperones. We previously engineered the parent anti-hexahistidine (His₆) scFv 3D5^{13,14} to enhance its biophysical properties and then converted its affinity to that of the hexapeptide EYMPME (EE-tag),¹⁴ (sequences in Supporting Table S1, S2). This process inadvertently resulted in a new crystal lattice employing variable heavy chain (V_H) CDR residues in a crystal contact. In this study, we systematically altered specific amino acid residues mediating these new crystal contacts in the anti-EE scFv in an attempt to restore the original 3D5 lattice. The experimentally prepared variants retain some crystal contact interfaces of the original parent 3D5 antibody, and in some cases even use the same residues for contacts; yet, lattices were altered (Figure 1). Our best second-generation anti-EE scFv described here binds to internally EE-tagged maltose binding protein (MBP) and forms a solution complex with the *E. coli* β -barrel membrane protein called intimin. Taken

together, our results underscore the challenges of directing a particular lattice in hypercrystallizable proteins such as this family of scFvs, but suggest that this plasticity could be an advantage for their use as crystallization chaperones.

Materials and Methods

Molecular biology, expression and purification of scFv chaperones

The two initial crystal chaperones with enhanced biophysical properties derived from parent 3D5 scFv,¹³ namely, the anti-His₆ 3D5/His₆₈₃ and anti-EE 3D5/EE₄₈, were expressed and purified as described previously.¹⁴ As described in Results, anti-EE variants investigated in this study target specific crystal contact residues: 3D5/EE₄₈.A harbors the heavy chain (V_H) amino acid changes S30T and S32A, and 3D5/EE₄₈.K, harbors mutations S30T and S32K. Amino acid residues are numbered according to the Kabat system and sequence information for all scFv variants is presented in Supporting Table S1 and Table S2. ScFvs 3D5/EE₄₈.K and 3D5/EE₄₈.A were generated by site-directed mutagenesis (SDM, Quickchange II, Agilent Technologies). Primers for 3D5/EE₄₈.A scFv variant: forward: 5'-atgggtgtg aactgggtt aaacagagt ccagg-3', reverse: 5'-cctataagt gactggtgac gaccatacc cacacttg-3'. SDM primers for 3D5/EE₄₈.K scFv variant: forward: 5'-atgggtgtg aactgggtt aaacagagt ccagg-3', reverse: 5'-cctataagt gactggtga tcccatc ccacacttg-3'. Sequences were verified by DNA sequencing (MWG Operon), and proteins were expressed and purified as described previously.¹⁴

Molecular biology, expression and purification of proteins presenting the EE epitope for complexation with scFv chaperones

Peptide epitopes were incorporated into proteins of interest via SDM (Quickchange II, Agilent Technologies) and verified by DNA sequencing (MWG Operon). The EE-tagged MBP (MBP-KEE) used in this study presents the six residue EE tag in the context of a native surface exposed loop. The EE tag was placed immediately after Lys 171 in MBP via SDM (forward primer: 5'-ggttatgcg tcaaggaa tacatgcc atggaggac attaaagac gtggcgctg g-3', reverse primer: 5'-gaacgcata acccccgctc agcagcaat cagcgcca ggtgaagta cg-3'). MBP-KEE was expressed and purified as previously described for the corresponding C-terminal EE-tagged construct¹⁴. *E. coli* intimin was selected as a test membrane protein,¹⁵ with the expression plasmid generously provided by Dr. Susan K. Buchanan (NIH). The EE epitope was incorporated into an extramembraneous loop in wild type intimin¹⁵ between residues 314-321 via SDM (forward primer: 5'-cggtactt ccgtatgag tggttggca tgaatacat gcccatgga agattacga tgaacgcc ggcaaatgg ctttgatat tcg-3' reverse primer: 5'-cgaatatca aagccattt gccggcgct tcactgtaa tcttcctg ggcattgat tcattccaa ccaactcata cggaagtag ccg-3'). EE-tagged intimin (intimin-EE) was expressed and purified as previously described for wild type intimin (WT-intimin).¹⁵

Biophysical characterization

Protein purity and size were assessed by standard reducing sodium-dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).¹⁶ Qualitative analysis of the oligomeric state distribution in solution at equilibrium (scFv monomer-to-dimer ratio) was determined by calculating the area under each elution peak from size exclusion chromatography

(Superdex 75 pg, GE Healthcare) using Unicorn software (GE Healthcare). Protein solubility was determined by quantifying the concentration of soluble protein after concentration of the protein to ~20 mg/mL, incubation for four days at 4 °C, and centrifugation to pellet insoluble material. Thermal stability was evaluated by differential scanning fluorimetry (DSF).¹⁷ Briefly, purified protein (20 µl of 200µM) or buffer blank were mixed with Sypro Orange (1 µl of a 1:1000 dilution; Molecular Probes), heated in a Real Time PCR machine (Viia™7; Applied Biosystems) at increments of 0.96°C/min from 25°C to 90°C and analyzed with Viia™7 software (Applied Biosystems) for the melting temperature (T_m), the midpoint of unfolding. Reported values are averages of at least two independent samples.

Complexation and binding assays

Kinetic binding assays were performed with a BIAcore 3000 (GE Healthcare) instrument using an MBP-KEE as a client for 3D5/EE_48, 3D5/EE_48.A and 3D5/EE_48.K variants. Protein ligands were coupled to CM5 chips using NHS-EDC chemistry to a level of ~500 RU. Responses due to sample refractive index changes and non-specific binding were corrected using signal from a flow cell coupled with a control MBP lacking the EE-tag.¹⁸ Purified scFv proteins were injected in a duplicate dilution series from 3 to 0.1875 µM at a flow rate of 50 µl/min to minimize mass transport effects. Surface regeneration was performed after each run with a one minute injection of 2M MgCl₂. The association rate constant (k_{on}), dissociation rate constant (k_{off}), and equilibrium dissociation constant (K_d ; $K_d = k_{off}/k_{on}$) were calculated assuming a Langmuir 1:1 binding model with BIAevaluation software. Only data sets with $\chi^2 < 0.5$ were used.

The (intimin-EE)-(3D5/EE_48.K) interaction was evaluated using an Åkta FPLC instrument equipped with a Superose 12 GL 10/300 SEC column (GE Healthcare) equilibrated with 50 mM Tris pH 7.5, 200 mM NaCl, 0.01% Na Azide, 0.05% n-dodecyl-β-D-maltoside (DDM, Anatrace), at 4 °C. Wild type intimin or intimin-EE was mixed in a 1:1 molar ratio with 3D5/EE_48.K and incubated for 3 hours on ice before injection. Elution fractions for each peak were analyzed with reducing SDS-PAGE.

Protein Crystallization

Initial crystallization conditions for 3D5/His_683 were identified by sparse matrix screening (Hampton Research HR2-139 and HR2-138). For 3D5/EE_48.A and 3D5/EE_48.K, we initially used conditions that resulted in diffraction quality crystals of the parent 3D5/EE_48 scFv, but included several optimization steps to increase the diffraction quality of the second generation scFvs. Variables optimized include protein concentration, protein:drop ratio, temperature, buffer concentration, salt, and polyethylene glycol (PEG) molecular mass in the mother liquor. Ultimately, the best crystals of 3D5/His_683 (8mg/ml in 10 mM Hepes pH 7.5, 150 mM NaCl, HBS) were grown at room temperature with 0.2M KI, 0.001M Guanidinium HCl, 18% PEG 8000 (Hampton Research). These crystals appeared in 2-3 days and grew to a maximal size of 120-150 µm within 1 week. Crystals of 3D5/EE_48.K (7.5mg/ml in HBS) were grown at 4°C with 0.1M Tris pH 8.5, 0.2M Li₂(SO₄), 3% 6-aminohexanoic acid, 24% PEG 8000; crystals appeared in 4-5 days and grew to a maximal size of 30-40 µm within 2 weeks. Crystals of 3D5/EE_48.A (7.5mg/ml in HBS) were grown

at 4°C with 0.1M BisTris pH 6.5, 0.2M Mg(OAc)₂, 21% PEG 8000, crystals appeared in 4-5 days and grew to a maximal size of 20-30 µm within 3-4 weeks. All crystals were grown utilizing a 1:1 reservoir:protein ratio.

Data collection, structure determination and refinement

Crystals were mounted in nylon loops and flash-cooled with liquid nitrogen prior to synchrotron X-ray data collection. Glycerol was used as a cryo-protectant by supplementing the reservoir conditions with 30% (v/v) (3D5/His_683), 20% (3D5/EE_48.A) and 25% (3D5/EE_48.K). Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID at the Advanced Photon Source at Argonne National Laboratory (Darien, IL). Crystallographic data sets were processed with HKL2000¹⁹ and unmerged data were exported for structure determination by molecular replacement using Phaser.²⁰ For 3D5/His_683 the search model was the 3D5 scFv (protein data bank, PDB ID 1KTR) whereas for 3D5/EE_48.A and 3D5/EE_48.K, the 3D5/EE_48 monomer structure (PDB ID 3NN8) was used. The atomic models were fit to the respective electron density maps using Coot,²¹ and then iteratively refined with Phenix,²² taking advantage of parameterization specifically designed for low resolution structures.²³ POLYGON²⁴ evaluation of model qualities indicate they are in the acceptable range.^{24,25} Crystallographic statistics are presented in Table I and representative electron density relevant to crystal contact engineering is shown in Supporting Figure S1. Atomic coordinates and structure factors for 3D5/His_683, 3D5/EE_48.A and 3D5/EE_48.K have been deposited at the PDB with accession codes 4NKO, 4NKD, and 4NKM, respectively.

Computational analyses

PDBe Protein Interfaces, Surfaces and Assemblies (PISA)²⁶ was used to rank and analyze crystal lattice contacts by surface area and energy, as well as catalog critical amino acids in crystal contacts based on their formation of polar interactions. After excluding peptide binding (ID4 and ID5 of 3D5) and the native heavy-light chain interface (ID1 of all scFv variants) within the scFv monomer, identified by the Complexation Significance Score (CSS), the top three interfaces were considered major crystal contacts and used for further analysis.

To assess the overall thermodynamic stability of each scFv in its native or the 3D5 lattice, the three refined anti-EE scFv monomer structures were fit into the P3₂21 spacegroup using secondary-structure matching (SSM)²⁷ in Coot. The overall free energy of each lattice was assessed by summing the energetic contributions of major crystal contacts, identified by PISA as those with a surface area >50 Å² or $G > 0.84$ kJ/mol and excluding biological interfaces. The free energy contributions of each contact are calculated as $G_{\text{contact}} = G_{\text{interface}} + G_{\text{H-bonds}} + G_{\text{salt bridges}}$, with $G_{\text{interface}}$ = the solvation free energy gain upon formation of the interface, calculated as difference in total solvation energies of isolated and interfacing structures, and the contributions of each hydrogen bond and salt bridge assessed at 2.09 and 1.25 kJ/mol respectively. The total free energy of lattice formation per scFv was obtained by dividing the energies by the number of monomers in each asymmetric unit.

Results

Analysis of first generation scFv chaperones

The anti-His₆ 3D5 scFv framework was selected as the initial basis for our crystallization chaperone toolbox since it is one of the few scFvs that does not use CDRs in major crystal contacts, instead directing these towards a large solvent cavity that could accommodate a membrane protein (Figure 1a).^{13,14} Anti-His₆ 3D5/His₆₈₃ varies from the parent 3D5 by 13 amino acid residues in the V_H CDR3 loop (Supporting Table S2) with favorable biophysical characteristics compared to 3D5 (see biophysical characterization below).¹⁴ 3D5/EE₄₈ was engineered by further manipulating 3D5/His₆₈₃ V_H CDRs to convert its specificity from anti-His₆ to anti-EE. 3D5/EE₄₈ was previously crystallized under identical conditions as 3D5, but used the cubic space group F23 as opposed to the P3₂21 space group used by 3D5.¹⁴ Lattice analysis reveals that 3D5 (Figure 1a) and 3D5/EE₄₈ (Figure 1b) both present large solvent channels, but are in distinct arrangements.

Here we add limited comparison the structure of 3D5/His₆₈₃, which readily grew crystals in a variety of conditions. These crystals exhibit significant pseudomeroheredral twinning as identified by Xtriage within Phenix.²⁸ Dozens of crystals were tested for space group assignment and reduced twin fraction, but no differences were detected. The final reported 3D5/His₆₈₃ crystals grew in conditions reminiscent of, but different from, 3D5. No crystals were obtained in the original 3D5 conditions. The best diffracting crystals grew with a higher protein concentration and growth temperature, in a cocktail containing no buffer and a very low concentration of the denaturant guanidinium hydrochloride (see Materials and Methods). The 3D5/His₆₈₃ structure was solved in space group C222₁ (Table 1). In this lattice, a pair of three nearly identical scFv molecules in the correct asymmetric unit is related by a two-fold rotation, but the structure could not be refined in a hexagonal lattice (P6₅22 or P6₁22 with one molecule in the asymmetric unit). Lattice analysis reveals that the 3D5/His₆₈₃ lattice does not harbor an analogous large channel like 3D5 or 3D5/EE₄₈ to accommodate a client protein (Figure 1c). Due to the combination of the crystal properties and disadvantages of targeting a pH-sensitive, His₆ tag for chaperoning, 3D5/His₆₈₃ was not considered further in this study or for our crystallization chaperone toolbox.

Though they crystallize in different space groups, structures of the scFv monomers 3D5, 3D5/EE₄₈, and 3D5/His₆₈₃ are nearly superimposable (overall root mean squared deviation is ~ 1 Å), indicating there are no lattice-induced deformations (Figure 2a). PISA crystal contact analysis (Table II) reveals one common interface among the three structures (Figure 2b). This contact relies on polar interactions between C-terminal Lys and Arg residues within variable light chain (V_L) and other nearby residues of the V_H domain, which sometimes also involves a Gly from the linker region, depending on the variant (Table II, Figure 2c). In 3D5, this contact is ranked third (ID3) and involves a surface area of 305 Å². It is likewise ID3 in 3D5/His₆₈₃ and ID4 in 3D5/EE₄₈. Unlike 3D5, which utilizes predominantly V_H-V_L protein contacts (Figure 3a), those in 3D5/EE₄₈ involve residues in symmetry related V_H domains, whereas for 3D5/His₆₈₃ there is a V_H-V_H contact and a V_L-V_L contact (Tab II).

Rational design of second-generation scFv candidate chaperones to control the crystal lattice

On the basis of the disparate crystal contacts found among 3D5 and the two first-generation scFvs 3D5/His_683 and 3D5/EE_48, residues were identified which, upon substitution, might restore the desirable 3D5 lattice (Figure 1a) via the original contacts (Figure 3a). Residues considered were those involved in apparent hydrogen bonding and/or salt bridge interactions spanning the intermolecular interfaces of 3D5, 3D5/His_683 and 3D5/EE_48, but were not involved in the common crystal contact described above (Table II, Table III, and Supporting Table S2). Five V_H residues, Ser 30, Ser 32, Lys 57, Arg 58, and Ser 65, were evaluated for their contributions to the respective crystal contact for each protein, and extent of accessible and buried surface areas (Table III). The top two residues, V_H Ser 30 and V_H Ser 32, found in 3D5/EE_48 ID3 and not predicted to interact with peptide,¹⁴ were then selected for mutation. In 3D5/EE_48, Ser 30 occurs one residue before the start of V_H CDR1 and participates interacts with Thr 77 and Tyr 79 of a neighboring V_H domain via the Ser 30 side chain and main chain, respectively (Figure 3b). In this position in the structures of 3D5 and 3D5/His_683 there is a Thr, which is not involved in a crystal contact (Table II and Supporting Table S2). V_H Asp 72 and V_H Lys 73 are present in the sequences of 3D5 and 3D5/EE_48, and are involved in the 3D5 contact. Thus, by making the conservative S30T substitution, we expected to remove the 3D5/EE_48-specific Ser 30 interactions in favor of the desired V_H-V_L contact. In addition, in 3D5/EE_48, Ser 32 participates in hydrogen bonding interactions with the main chain carbonyl of Lys 19 from a neighboring V_H domain (Figure 3b). We reasoned that replacement of Ser 32 with Ala would remove the side chain interaction, or substitution with Lys would disrupt the side chain interaction as well as increase entropy, making this contact less likely to form. Reversion to the Tyr 32 found in 3D5 and 3D5/His_683 did not seem prudent because while Tyr 32 does not form a crystal contact in the 3D5 structure, it does participate in a crystal contact in 3D5/His_683 and Tyr residues commonly mediate antibody-ligand interactions (Table II and Supporting Table S2).²⁹ In sum, our second-generation scFv chaperone candidates are the double mutants 3D5/EE_48.A (S30T, S32A) and 3D5/EE_48.K (S30T, S32K).

Biophysical properties of 3D5/EE_48.A and 3D5/EE_48.K

As analyzed by size exclusion chromatography (SEC), the scFv variants 3D5/His_683 initially purify as a mixture of a monomer and dimer but to differing extents (Figure 4a). The anti-EE scFvs exhibit a higher monomer:dimer ratio compared to anti-His 3D5/His_683 and 3D5¹⁴ (Table IV), and the monomeric protein is stable once purified. Solubility, thermal stability, and yield of purified protein are also important practical considerations for crystallization. Based on such biophysical properties (Table IV), 3D5/EE_48 and 3D5/EE_48.K are the most promising anti-EE scFv crystallization chaperones; although 3D5/EE_48.A exhibited improved stability over 3D5/EE_48, its yield and solubility were limited.

Structure and lattice analysis of second-generation scFv chaperones (3D5/EE_48.A and 3D5/EE_48.K)

3D5/EE_48.A crystallizes in a cocktail nearly identical to that of 3D5 and 3D5/EE_48, but produces smaller crystals (20-30 μm in the largest dimension) than 3D5 (70 μm)¹³ and 3D5/

EE_48 (40-60 μm).¹⁴ As with 3D5/EE_48 and 3D5/His_683, the scFvs are all readily superimposable and diffract to medium resolution (Figure 2a). Crystals of 3D5/EE_48.A belong to the same space group as 3D5/EE_48, namely, F23, with a similar 50Å solvent channel present in the lattice (Figure 1d). Interestingly, while V_H Ser 30 of 3D5/EE_48 and V_H Thr 30 of 3D5/EE_48.A are within essentially the same contacts in the two structures (Figure 3b,c), V_H Ala 32 of 3D5/EE_48.A now torques the crystal contact to form an apparent interaction with V_H Ser 21 via the main chain of V_H Ala 32 (Figure 3c, Supporting Figure S1a), instead of with V_H Lys 19 as seen for V_H Ser 32 in 3D5/EE_48 (Figure 3b). Notably, V_H Ser 21 is not a contact residue in 3D5/EE_48 (Figure 3b). Comparison of the second common contact (ID2 in 3D5/EE_48 and ID3 of 3D5/EE_48.A, Figure 3d, Supporting Figure S1b) also reveals an apparent slippage between residues interacting across the contact. The only residue to form a contact at this interface in 3D5/EE_48.A is V_H Ser 65, which interacts with a symmetry-related V_H Ser 65, compared to four additional contacts in the same interface of 3D5/EE_48 (Figure 3d, Supporting Figure S1b). Thus, the two amino acid substitutions were not sufficient to alter the crystal lattice but did change the lattice contacts.

By comparison, crystallization of 3D5/EE_48.K was successful under different conditions; in spite of considerable effort, no crystals were obtained in the condition used for 3D5. For 3D5/EE_48.K, the combination of higher protein concentration with a cocktail composed of monovalent rather than divalent salt, and higher pH, promoted crystal growth. The crystals belong to space group C2 (Figure 1e), and like those of 3D5/His_683 (Figure 1c), do not harbor a wide channel within the lattice. In the 3D5/EE_48.K structure, there are three main crystal contacts: (1) a hybrid interface (ID2, Table II, Figure 3e, Supporting Figure S1c) between the top two contacts seen in 3D5/EE_48 and ID2 in 3D5 (Table II, Figure 3a,b,d left panel, Supporting Figure S1b) (2) the common V_H-V_L interface (ID3) seen among all structures (Table II, Figure 2b for overlay, Figure 2c details of 3D5 and 3D5/EE_48), and (3) a new crystal interface (ID4) involving V_H CDR3 and V_L CDR1 (Figure 3f, Supporting Figure S1d) proximal to ID2. Residues participating in 3D5/EE_48.K ID4 are located within V_H CDR3, the most important loop for peptide recognition, and V_L CDR1 (Figure 3f, Supporting Figure S1d). One ID4 residue, V_L Asn 30, participates in the desirable common contact in the 3D5 lattice (Figure 3a). In the hybrid contact of 3D5/EE_48.K ID2, V_H Thr 30, one of the two residues altered in second generation variants due to its lack of involvement in contacts in the 3D5 or 3D5/His_683 lattices, participates in stabilizing interactions with the carbonyl oxygens of V_H Lys 64 and V_H Ser 65, both of which are residues within V_H CDR2 (Figure 3e, Supporting Figure S1c, Supporting Table S2). In turn, V_H Ser 65 is stabilized by V_H Ser 28, an interaction not observed in 3D5/EE_48 or 3D5/EE_48.A (Figure 3d, Supporting Figure S1b). Residue 28 is an invariant serine within the anti-EE scFvs, a threonine in 3D5 and 3D5/His_683, and did not previously participate in crystal contacts (Supporting Table S2). Tyr 59 from V_H CDR2, which is unchanged in the structures under discussion and not previously in a crystal contact, forms an interaction with the carbonyl group of V_H Ser 74 in a neighboring molecule (Figure 3e, Supporting Figure S1c); V_H Ser 74 is a contact residue in 3D5 (Figure 3a) but not in 3D5/EE_48 or 3D5/EE_48.A (Figure 3b,c, Supporting Figure S1a). Another commonality between the desired 3D5 contact and ID2 of 3D5/EE_48.K is the involvement of the residue at position 57 of

V_H, which is a Thr in 3D5 and Lys in 3D5/EE_48.K (Figure 3a,e, Supporting Figure S1c, Supporting Table S2). Finally, an apparent side-chain Lys-Asp salt bridge is present in both lattices (V_L K50-V_H D72 of 3D5 and V_H K23-V_H D56 of 3D5/EE_48.K), although specific residues are different (Figure 3a,e, Supporting Figure S1c).

Computational energetics analysis

To date, we have not crystallized our engineered anti-EE scFv variants in the 3D5 crystal lattice of P3₂21 even though the residues mediating these interactions are conserved. To gain insight into why this may be the case, we superimposed each engineered scFv onto the 3D5 lattice and computed the free energies for each lattice (Table V). When compared to the experimentally-obtained lattices on a per-asymmetric unit basis, P3₂21 is consistently less favorable, indicating that thermodynamics do contribute to lattice choice. However, when compared per scFv molecule, free energy values corresponding to the experimental lattices are either very similar (3D5/EE_48 and 3D5/EE_48.K) or somewhat less favorable (3D5/EE_48.A) than P3₂21 (Table V).

Binding and complexation with EE-tagged soluble and membrane proteins

Important to our design of second-generation scFv crystallization chaperones is tight binding to client EE-tagged proteins. As proof of concept, we initially tested scFv binding to MBP-KEE, a molecule that presents the EE peptide in a native surface-exposed loop (see Materials and Methods), by ELISA (data not shown) and SPR (Figure 5a). In spite of amino acid changes in and near V_H CDR1, both 3D5/EE_48.A and 3D5/EE_48.K retain strong binding affinity for MBP-KEE in the nanomolar range, similar to 3D5/EE_48 (K_ds of 100, 147 and 166 nM, respectively; Figure 5a, Table IV). Of note, the affinities measured for 3D5/EE_48 binding to the EE epitope when constrained by presentation in an internal loop is significantly tighter than when the peptide is present as a C-terminal extension with considerable conformational flexibility (~700 nM).¹⁴

Due to its higher solubility and expression level, 3D5/EE_48.K was assessed for complexation with a proof-of-concept membrane protein. Guided by published mutagenesis studies evaluating intimin topology and supported by examination of the crystal structure,¹⁵ the EE-tag was introduced into an extramembraneous loop using just one round of site-directed-mutagenesis (SDM). Expression and purification of intimin-EE proceeded as for wild-type. Co-elution of 3D5/EE_48.K and intimin-EE is seen by gel filtration combined with SDS-PAGE analysis, concomitant with a reduction in the height of the elution peak corresponding to scFv when compared to the negative control (Figure 5b).

Discussion

We attempted to combine our knowledge of the crystallization propensity of 3D5, 3D5/His_683 and 3D5/EE_48 to redirect second-generation anti-EE scFvs to the 3D5 lattice, which uses largely non-CDR residues for contacts and whose lattice includes a wide ~75 Å diameter channel. Side chain entropy has been shown to be the principal negative correlate of crystallization, with charged residues the least likely to participate in crystal contacts.³⁰ In the case of proteins that do not crystallize in their native state, mutation of key high-entropy

side chains has been highly successful in identifying crystallizable variants.⁸ Thus, we used similar considerations of contact residue contributions to buried and accessible surface area, as well as to surface entropy. The variants 3D5/EE_48.A and 3D5/EE_48.K retained potent EE binding affinity but crystallized in different, non-3D5 lattices. In the case of 3D5/EE_48.A, the lattice of 3D5/EE_48 was retained with minor modifications in crystal contacts. Our result for 3D5/EE_48.A is consistent with general observations from SER, namely that the substitution of Ser for Ala removes the polar side chain interaction and leads instead to a stereochemically constrained backbone interaction that likely contributes to hydrophobic stabilization in the water-depleted crystal contact environment.⁸ By contrast, our attempt to use SER underpinnings to kill a contact by introducing a lysine residue led to an unexpected result. The offending crystal contact from 3D5/EE_48 was altered, but a new hybrid contact was generated that incorporated some, but not all, of the desired parent contacts.

Our study demonstrates that the disruption of just one intermolecular contact can be sufficient to modify the lattice if more than one protein-protein interaction is available for crystallization. Such crystallization variability has been explored in classically studied proteins like lysozyme³¹⁻³⁴ and ribonuclease A,³⁵ but few studies strategically attempt mutagenesis to alter the crystal lattice to a desired one starting from first principles. In one study, structures of *A. cellulolyticus* coh1 mutants were solved as part of an effort to generate a protein-based lattice capable of templating nanocomposite materials or as an immobilized catalyst.³⁶ In this case, surface lysine residues were eliminated by replacing them with alanine or tryptophan. Although not a typical residue employed in SER,⁸ the latter mutant resulted in a crystal lattice with higher porosity and solvent content than the parent, with the Trp playing a key contact role.³⁶ For coh1, the presence of a surface-exposed Trp residue did not cause problems in terms of protein yield or solubility. In a second example, an alternative lattice was a serendipitous outcome of attempts to revert the oligomeric state of sweet potato β -amylase from monomer to the tetramer found in another ortholog.³⁷ Similar to what we report here with our scFvs, residues comprising the original sweet potato β -amylase lattice contacts were not mutated, yet another lattice was adopted.

A major challenge in directing a particular crystal lattice is that beyond a general preference for amino acids with low entropic contribution,^{6,8} physicochemical characteristics common to protein crystal contacts, are not obvious.^{32,38} In our system, one possible explanation for the observed lattice alterations is the length and flexibility of surface loops. Both 3D5/His_683 and 3D5/EE_48 have V_H CDR3 loops that are seven residue longer, while the 3D5/EE_48 V_H CDR1 is two residues longer, than in 3D5. Notably, while 3D5 exhibits the typical surface area for crystal packing of 100-500 Å²³² with a corresponding lattice free energy per asymmetric unit on the low end of published values (G^0 = ~-25-40 kJ/mol),^{1,39} those for our engineered anti-EE variants are higher in terms of contact surface areas, and are more energetically favorable per asymmetric unit. Viewed on a per-scFv basis, however, the free energies are roughly comparable; suggesting that lattice choice may be influenced by the crystallization pathway. Specifically, the additional residues in our EE variants may direct crystal nucleation through repulsive or attractive events on the pathway to crystallization, such as by releasing key bound water molecules.³⁹ The pathway may also be

influenced by the composition of the mother liquor,^{33,35} or the presence/absence of a ligand/binding partner.

Ideally, our anti-EE chaperones would template a porous three dimensional lattice using contacts remote from CDRs, in a manner suitable to immobilize an EE-tagged client protein for structure determination at a higher resolution than a crystal of the membrane protein alone. Although controlling the crystallization lattice by rational design may not be achievable, our scFv platform holds promise for crystallization chaperoning. As evidenced by this study, our hypercrystallizable scFv molecules have multiple well-ordered surface patches available for crystal contacts, including those utilized in 3D5 that would not interfere with epitope binding; indeed the presence of a bound ligand may direct lattice formation to that used by 3D5. In addition, our protein engineering efforts have led to a high affinity binding site, capable of complexation with a purified, detergent-solubilized, EE-tagged membrane protein. Engineering of alternative parent scFv and other antibody fragment scaffolds are currently being considered to improve the chaperone crystal diffraction limit. It is our hope that with our approach, a single (or a few) chaperone(s) can be engineered for optimal crystallizability, and then used to co-crystallize a wide variety of client proteins simply by installing the hexapeptide epitope, the placement of which the researcher has complete control. Co-crystallization trials with intimin and other tagged membrane proteins are underway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the National Institutes of Health (#DK91357, RLL and #GM095638, JAM) and the Welch Foundation (#F-1696, TMT; #F-1767, JAM). Partial student support is acknowledged from Georgia Tech Senior Biophysics Training program (SK, JLJ), Department of Education Graduate Assistance in Areas of National Need (GAANN) Grant #P200A090307 (JLJ), and Petit Undergraduate Scholars program (DPH, IAM).

References

1. Price WN, Chen Y, Handelman SK, Neely H, Manor P, Karlin R, Nair R, Liu JF, Baran M, Everett J, Tong SCN, Forouhar F, Swaminathan SS, Acton T, Xiao R, Luft JR, Lauricella A, DeTitta GT, Rost B, Montelione GT, Hunt JF. Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat Biotechnol.* 2009; 27(1):51–57. [PubMed: 19079241]
2. Dale GE, Oefner C, D'Arcy A. The protein as a variable in protein crystallization. *J Struct Biol.* 2003; 142(1):88–97. [PubMed: 12718922]
3. Mcelroy HE, Sisson GW, Schoettlin WE, Aust RM, Villafranca JE. Studies on Engineering Crystallizability by Mutation of Surface Residues of Human Thymidylate Synthase. *J Cryst Growth.* 1992; 122(1-4):265–272.
4. Pedelacq JD, Piltch E, Liong EC, Berendzen J, Kim CY, Rho BS, Park MS, Terwilliger TC, Waldo GS. Engineering soluble proteins for structural genomics. *Nat Biotechnol.* 2002; 20(9):927–932. [PubMed: 12205510]
5. Wernimont A, Edwards A. In situ proteolysis to generate crystals for structure determination: an update. *PLoS One.* 2009; 4(4):e5094. [PubMed: 19352432]

6. Cooper DR, Boczek T, Grelewska K, Pinkowska M, Sikorska M, Zawadzki M, Derewenda Z. Protein crystallization by surface entropy reduction: optimization of the SER strategy. *Acta Crystallogr D*. 2007; 63:636–645. [PubMed: 17452789]
7. Derewenda ZS. Rational protein crystallization by mutational surface engineering. *Structure*. 2004; 12(4):529–535. [PubMed: 15062076]
8. Derewenda ZS. It's all in the crystals. *Acta Crystallogr*. 2011; D67(Pt 4):243–248.
9. Lieberman RL, Culver JA, Entzminger KC, Pai JC, Maynard JA. Crystallization chaperone strategies for membrane proteins. *Methods*. 2011; 55(4):293–302. [PubMed: 21854852]
10. Hunte C, Michel H. Crystallisation of membrane proteins mediated by antibody fragments. *Curr Opin Struct Biol*. 2002; 12(4):503–508. [PubMed: 12163074]
11. Koide S. Engineering of recombinant crystallization chaperones. *Curr Opin Struct Biol*. 2009; 19(4):449–457. [PubMed: 19477632]
12. Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*. 2007; 450(7168):383–387. [PubMed: 17952055]
13. Kaufmann M, Lindner P, Honegger A, Blank K, Tschopp M, Capitani G, Pluckthun A, Grutter MG. Crystal structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen. *J Mol Biol*. 2002; 318(1):135–147. [PubMed: 12054774]
14. Pai JC, Culver JA, Drury JE, Motani RS, Lieberman RL, Maynard JA. Conversion of scFv peptide-binding specificity for crystal chaperone development. *Protein Eng Des Sel*. 2011; 24(5):419–428. [PubMed: 21217145]
15. Fairman JW, Dautin N, Wojtowicz D, Liu W, Noinaj N, Barnard TJ, Udho E, Przytycka TM, Cherezov V, Buchanan SK. Crystal Structures of the Outer Membrane Domain of Intimin and Invasin from Enterohemorrhagic E. coli and Enteropathogenic Y pseudotuberculosis. *Structure*. 2012; 20(7):1233–1243. [PubMed: 22658748]
16. Sambrook, J.; Russell, DW. *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2001.
17. Lavinder JJ, Hari SB, Sullivan BJ, Magliery TJ. High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. *J Am Chem Soc*. 2009; 131(11):3794–3795. [PubMed: 19292479]
18. Maynard JA, Maassen CBM, Leppla SH, Brasky K, Patterson JL, Iverson BL, Georgiou G. Protection against anthrax toxin by recombinant antibody fragments correlates with antigen affinity. *Nat Biotechnol*. 2002; 20(6):597–601. [PubMed: 12042864]
19. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol*. 1997; 276:307–326.
20. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40:658–674. [PubMed: 19461840]
21. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D*. 2010; 66:486–501. [PubMed: 20383002]
22. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D*. 2012; 68:352–367. [PubMed: 22505256]
23. Headd JJ, Echols N, Afonine PV, Grosse-Kunstleve RW, Chen VB, Moriarty NW, Richardson DC, Richardson JS, Adams PD. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr D*. 2012; 68:381–390. [PubMed: 22505258]
24. Urzhumtseva L, Afonine PV, Adams PD, Urzhumtsev A. Crystallographic model quality at a glance. *Acta Crystallogr D*. 2009; 65:297–300. [PubMed: 19237753]
25. Brown EN, Ramaswamy S. Quality of protein crystal structures. *Acta Crystallogr D*. 2007; 63:941–950. [PubMed: 17704562]
26. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007; 372(3):774–797. [PubMed: 17681537]

27. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D*. 2004; 60:2256–2268. [PubMed: 15572779]
28. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D*. 2010; 66:213–221. [PubMed: 20124702]
29. Koide S, Sidhu SS. The Importance of Being Tyrosine: Lessons in Molecular Recognition from Minimalist Synthetic Binding Proteins. *Acs Chem Biol*. 2009; 4(5):325–334. [PubMed: 19298050]
30. Cieslik M, Derewenda ZS. The role of entropy and polarity in intermolecular contacts in protein crystals. *Acta Crystallogr D*. 2009; 65:500–509. [PubMed: 19390155]
31. Janin J, Rodier F. Protein-protein interaction at crystal contacts. *Proteins*. 1995; 23(4):580–587. [PubMed: 8749854]
32. Carugo O, Argos P. Protein-protein crystal-packing contacts. *Protein Sci*. 1997; 6(10):2261–2263. [PubMed: 9336849]
33. Zhang XJ, Wozniak JA, Matthews BW. Protein Flexibility and Adaptability Seen in 25 Crystal Forms of T4 Lysozyme. *J Mol Biol*. 1995; 250(4):527–552. [PubMed: 7616572]
34. Janin J. Specific versus non-specific contacts in protein crystals. *Nat Struct Biol*. 1997; 4(12):973–974. [PubMed: 9406542]
35. Crosio MP, Janin J, Jullien M. Crystal Packing in 6 Crystal Forms of Pancreatic Ribonuclease. *J Mol Biol*. 1992; 228(1):243–251. [PubMed: 1447785]
36. Wine Y, Cohen-Hadar N, Lamed R, Freeman A, Frolov F. Modification of Protein Crystal Packing by Systematic Mutations of Surface Residues: Implications on Biotemplating and Crystal Porosity. *Biotechnol Bioeng*. 2009; 104(3):444–457. [PubMed: 19575413]
37. Kang YN, Adachi M, Mikami B, Utsumi S. Change in the crystal packing of soybean beta-amylase mutants substituted at a few surface amino acid residues. *Protein Eng*. 2003; 16(11):809–817. [PubMed: 14631070]
38. Bahadur RP, Chakrabarti P, Rodier F, Janin J. A dissection of specific and non-specific protein - Protein interfaces. *J Mol Biol*. 2004; 336(4):943–955. [PubMed: 15095871]
39. Vekilov PG, Feeling-Taylor AR, Yau ST, Petsev D. Solvent entropy contribution to the free energy of protein crystallization. *Acta Crystallogr D*. 2002; 58:1611–1616. [PubMed: 12351872]

Figure 1. Crystal lattices of scFv variants described in this manuscript. (a) 3D5, (b) 3D5/EE_48, (c) 3D5/His_683, (d) 3D5/EE_48.A, (e) 3D5/EE_48.K. Lines indicate solvent channels, with diameters listed.

Figure 2.

Common features among scFv variants. (a) Superposition of 3D5, 3D5/His₆₈₃, 3D5/EE₄₈, 3D5/EE₄₈.A, 3D5/EE₄₈.K monomers. (b) Common V_H-V_L crystal contact interface of aforementioned five structures (V_L in grey, V_H in black). (c) Close up view of contact in (b) for 3D5 (PDB code 1KTR) and 3D5/EE₄₈ (PDB code 3NN8). L: light chain, H: heavy chain.

Figure 3.

Outcome of rational design efforts. (a) Desired V_H - V_L lattice contact in 3D5 (ID2). (b) 3D5/EE_48 contact selected for mutagenesis (ID3) (c) Interface of 3D5/EE_48.A ID2 similar to 3D5/EE_48 ID3. (d) Comparison of 3D5/EE_48 ID2 (left) and 3D5/EE_48.A ID3 (right). (e) Interface ID2 formed in 3D5/EE_48.K (f) Interface ID4 formed in 3D5/EE_48.K. L: light chain, H: heavy chain. Note that 3D5/EE_48.K interfaces are based on 3.7 Å resolution structural data. See Supporting Figure S1 for corresponding electron density maps.

Figure 4.

Biophysical characterization of second-generation scFv variants. (a) Oligomeric state (dimer:monomer) analysis of purified scFvs by size exclusion chromatography on Superdex 75 column. (b) Size and purity analysis of scFvs by SDS-PAGE.

Figure 5.

Binding studies with EE-tagged client proteins. **(a)** BIAcore SPR binding analysis of second-generation scFv variants to MBP-KEE or a control MBP surface. **(b)** Overlay of intimin-EE-3D5/EE_48.K (black) with WT-intimin-3D5/EE_48.K (red) upon mixing and fractionation by SEC, and corresponding SDS-PAGE analysis.

Table I

Crystallographic data collection and refinement statistics.

	3D5/His_683 (4NKO)	3D5/EE_48.K (4NKM)	3D5/EE_48.A (4NKD)
Data collection			
Space group	C222 ₁	C2	F23
Cell dimensions			
a, b, c (Å)	60.6, 104.9, 284.4	103.3, 92.3, 142.7	275.7, 275.7, 275.7
α, β, γ (°)	90, 90, 90	90, 110.9, 90	90, 90, 90
Resolution (Å)^a	42.2 - 3.49 (3.62 - 3.49)	46.1 - 3.71 (3.84 - 3.71)	38.6 - 3.30 (3.42 - 3.30)
Rmerge^a	0.092 (0.49)	0.10 (0.22)	0.11 (0.30)
Rmeas	0.11	0.13	0.12
I/σI^a	12.0 (3.0)	11.9 (4.3)	20.2 (7.0)
Completeness (%)^a	93.1 (95.2)	84.8 (82.2)	100.0 (100.00)
Multiplicity^a	3.3 (3.5)	2.0 (1.9)	7.6 (7.6)
Refinement			
Resolution (Å)	42.2 - 3.49	46.1 - 3.71	38.6-3.30
No. of reflections	11089	11363	26045
R_{work}/R_{free}	0.308/0.344	0.264/0.282	0.195/0.231
No. of molecules			
Protein residues	708	944	944
B-factor protein (Å²)	90.4	84.2	50.0
R.m.s deviations (rmsd)			
Bond lengths (Å)	0.004	0.003	0.003
Bond angles (°)	1.25	0.87	0.94
Ramachandran			
Favored (%)	92	95	97
Outliers (%)	1.6	0.75	1.0

^aValues for the highest resolution shell are in parenthesis.

PISA analysis of all scFv variants described in this study. In all variants, ID1 is the native V_L - V_H binding interface within the scFv monomer. Residues are listed according to Kabat numbering. See also Supporting Table S1 and Table S2.

Table II

scFv	PISA Interface	Interface area ($\text{\AA}^2$)	G (kJ/mole)	V_H Contacts that formed H-bonds/salt bridges	V_L Contacts that formed H-bonds/salt bridges
3D5	ID2	467	11.3	T57, V71, D72, K73, S74	N30, K50, N53, R54
	ID3	305	-5.0	D11, Y91, Q105, T108	K107, R108
	ID6	160	10.0	n/a	Q42, R108
3D5/His_683	ID2	465	-2.1	D31, Y32, N52, S99	n/a
	ID3	353	-12.1	Q6, D11, Y91, Q105, G106, T108	L106, K107, R108, G109 (linker), G110 (linker)
	ID4	185	-2.5	n/a	S65, S67, D70
First-generation	ID2	465	3.3	K57, R58, S65, T68	n/a
	ID3	305	-16.3	P9, K19, S30, S32, D72, K73, S75, T77, Y79, S98	n/a
	ID4	393	-10.5	Q6, D11, Y91, Q105, T108	L106, K107, R108, G110 (linker)
	ID2	409	-15.5	P9, S21, T30, A32, T77, Y79, S98	n/a
3D5/EE_48.A	ID3	407	2.1	S65	n/a
	ID4	364	-12.6	Q6, D11, Q105, T108	K107, R108, G110 (linker)
	ID2	484	-1.3	K23, S28, T30, D56, K57, Y59, K64, S65, S74, R82A	n/a
3D5/EE_48.K ^a	ID3	356	-8.4	Q6, S7, D11, Q105, T108	K107, R108, G109 (linker), G110 (linker)
	ID4	226	-9.2	S98, Y100B	N28, N30

^aBased on 3.7 Å resolution data.

Table III
Comparison of average solvent accessible surface areas of selected crystal contact residues.

Rank	Residue	Kabat #	PISA Interface	ASA ^a	BSA ^b	% buried area ^c
1	Ser	30	ID3/7	46.1	44.7	96.9
	Ser	30	ID3/8	47.1	45.1	95.7
	<i>Average</i>			46.6	44.9	96.3
2	Ser	32	ID3/7	88.7	81.8	92.2
	Ser	32	ID3/8	82.0	76.0	92.6
	<i>Average</i>			85.4	78.9	92.4
3	Ser	65	ID2/5	91.8	83.5	91.0
	Ser	65	ID2/5	92.4	86.3	93.4
	Ser	65	ID2/6	94.1	86.5	92.0
4	Ser	65	ID2/6	82.3	73.3	89.1
	<i>Average</i>			90.1	82.4	91.4
5	Lys	57	ID2/5	68.9	55.4	80.4
	Lys	57	ID2/5	79.3	59.9	75.4
	Lys	57	ID2/6	69.6	50.6	72.8
5	<i>Average</i>			72.6	55.3	76.2
	Arg	58	ID2/5	151.6	68.0	44.8

^a Solvent-accessible surface area in Å², average values are in bold.
^b Solvent-accessible surface area that is buried upon interface formation, in Å², average values are in bold.
^c Buried area percentage (BSA/ASA*100), average values are in bold.

Table IV

Biophysical characteristics of scFv variants.

	3D5/His_683 ^a	3D5/EE_48 ^a	3D5/EE_48.A	3D5/EE_48.K
Expression level (mg/l culture)	8.5 ^a	2.1 ^a	0.5 ± 0.2	3.6 ± 0.4
Solubility (mg/ml)	16.6 ^a	12.8 ^a	3.0 ± 0.03	6.8 ± 0.32
Melting temperature (°C)	53.6 ± 0.0	47.2 ± 0.3	51.8 ± 0.1	51.3 ± 0.2
% Monomeric protein	62 ^a	81 ^a	73.5 ± 6.5	76.1 ± 2.11
Kd (nM), MBP-KEE (χ^2)	ND	166 (0.031)	100 (0.029)	147 (0.020)

^aValues from Pai, J. C., Culver, J. A., Drury, J. E., Motani, R. S., Lieberman, R. L., and Maynard, J. A. (2011) Conversion of scFv peptide-binding specificity for crystal chaperone development, *Protein Eng Des Sel* 24, 419-428.

Table V

Free energy analysis of scFv variants in differing crystal lattices

scFv	Crystal Lattice ^a	G ⁰ (kJ/mol asymmetric unit)	# scFv per asymmetric unit	G ⁰ (kJ/mol scFv)
3D5	P 3₂21	-18.8	1	-18.8
3D5/EE_48	F23	-70.3	4	-17.6
	P 3 ₂ 21	-15.9	1	-15.9
3D5/EE_48.A	F23	-59.8	4	-15.0
	P 3 ₂ 21	-29.7	1	-29.7
3D5/EE_48.K	C121	-109.6	4	-27.4
	P 3 ₂ 21	-25.5	1	-25.5

^a Experimentally observed lattice in bold.