Conversion of scFv peptide-binding specificity for crystal chaperone development

Jennifer C.Pai¹, Jeffrey A.Culver², Jason E.Drury², Rakesh S.Motani³, Raquel L.Lieberman^{2,4} and Jennifer A.Maynard^{1,4}

¹Department of Chemical Engineering, University of Texas at Austin, MC0400, 1 University Station, Austin TX 78712, USA, ²School of Chemistry and Biochemistry and Institute for Bioscience and Bioengineering, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, GA 30332, USA and ³Department of Microbial Engineering, University of Minnesota, Minneapolis, MN 55455, USA

⁴To whom correspondence should be addressed. maynard@che.utexas.edu (J.A.M); raquel.lieberman@chemistry.gatech.edu (R.L.L.)

Received September 3, 2010; revised November 18, 2010; accepted November 26, 2010

Edited by James Marks

In spite of advances in protein expression and purification over the last decade, many proteins remain recalcitrant to structure determination by X-ray crystallography. One emerging tactic to obtain high-quality protein crystals for structure determination, particularly in the case of membrane proteins, involves co-crystallization with a protein-specific antibody fragment. Here, we report the development of new recombinant single-chain antibody fragments (scFv) capable of binding a specific epitope that can be introduced into internal loops of client proteins. The previously crystallized hexa-histidine-specific 3D5 scFv antibody was modified in the complementary determining region and by random mutagenesis, in conjunction with phage display, to yield scFvs with new biochemical characteristics and binding specificity. Selected variants include those specific for the hexa-histidine peptide with increased expression, solubility (up to 16.6 mg/ml) and sub-micromolar affinity, and those with new specificity for the EE hexa-peptide (EYMPME) and nanomolar affinity. Complexes of one such chaperone with model proteins harboring either an internal or a terminal EE tag were isolated by gel filtration. The 3.1 Å resolution structure of this chaperone reveals a binding surface complementary to the EE peptide and a \sim 52 Å channel in the crystal lattice. Notably, in spite of 85% sequence identity, and nearly identical crystallization conditions, the engineered scFv crystallizes in a different space group than the parent 3D5 scFv, and utilizes two new crystal contacts. These engineered scFvs represent a new class of chaperones that may eliminate the need for de novo identification of candidate chaperones from large antibody libraries.

Keywords: antibody/binding affinity/co-crystallization/ protein complex/protein engineering

Introduction

Even though the numbers of protein databank entries continue to increase, numerous proteins are rejected from the pipeline leading to structure determination. Specifically, there is a need for strategies to overcome the crystallization limitation, especially for membrane proteins and proteins with inherent conformational variability. A number of strategies to improve the likelihood of growing crystals of so-called 'difficult' proteins have emerged over the last decade. Beyond improvements in recombinant expression and protein purification that enable more expansive crystallization trials, these techniques either involve modifying the protein to be crystallized in a way that improves its properties for crystallization, or introducing a second protein, a crystallization chaperone, to provide the crystal lattice. The former category includes random mutagenesis and homolog shuffling (Pedelacq et al., 2002; Yang et al., 2003; Keenan et al., 2005), limited proteolysis to generate a compact, stable protein entity (Wernimont and Edwards, 2009), the identification of ligands to optimally stabilize a particular conformation of the protein (Vedadi et al., 2006), modification of the protein surface to reduce entropy (Derewenda, 2004; Cooper et al., 2007; Goldschmidt et al., 2007) and protein symmetrization by cross-linking (Banatao et al., 2006), among others.

The chaperone category involves the formation of a specific complex between a client protein and a soluble protein that provides hydrophilic residues to form crystal contacts and thus increases the chances of growing wellordered, highly diffracting crystals of the complex. Since the first report of a crystallization chaperone used to determine the HIV capsid protein structure (Prongay et al., 1990), efforts have focused on generating complexes between membrane proteins, which suffer from particularly unfavorable surface properties for crystal formation. Non-covalent complexes of target membrane proteins with tailored antibody fragments (Kovari et al., 1995; Ostermeier et al., 1995; Hunte et al., 2000; Zhou et al., 2001; Stura et al., 2002; Rasmussen et al., 2007; Uysal et al., 2009), affibodies (Warke and Monmany, 2007), V_{HH} camelid domains (Tereshko et al., 2008) and designed ankyrin repeat proteins (DARPins, Huber et al., 2007; Sennhauser and Grutter, 2008; Milovnik et al., 2009) have been reported. In general, crystallization chaperones recognize native membrane protein sequence, and require identification of a new chaperone for each protein of interest. Fusion to or insertion of a chaperone into a flexible loop has also been described (Prive et al., 1994; Byrne et al., 2000; Hunte et al., 2000; Cherezov et al., 2007; Rasmussen et al., 2007; Rosenbaum et al., 2007). The location of the fusion protein is key, as long linkers confer flexibility typically detrimental to crystallization (Byrne et al., 2000). Ideally, the chaperone should not interfere with activity or function of the client protein of interest. Nevertheless, in principle, any stable soluble protein tethered to or with high affinity for the membrane protein of interest could be used in co-crystallization experiments.

Here, we describe the first steps in development of a generalizable approach to chaperoning crystal growth: antibody fragments that can be used as a co-crystallization chaperone for any protein in which a short peptide sequence, the EYMPME epitope (EE), is inserted. We selected the hexahistidine-specific (His₆) 3D5 single-chain antibody fragments (scFv) as the framework for protein engineering because it does not employ complementary determining region (CDR) residues in major crystal contacts and the CDRs face a wide channel that could accommodate a client protein (Lindner et al., 1997; Muller et al., 1998; Kaufmann et al., 2002). We hypothesized that these CDRs could be modified to recognize a new peptide epitope, and a peptide or client protein could bind, without compromising existing crystal contacts. Nevertheless, 3D5 possesses several shortcomings that require optimization for use as a crystallization chaperone including low affinity $(K_d \sim 1 \,\mu M)$ for only extreme C-terminal histidines (Kaufmann et al., 2002), pH sensitive binding (Muller et al., 1998), relatively poor expression in Escherichia coli, and limited solubility (Kaufmann et al., 2002). Lastly, terminal His₆ tags, which are commonly used for protein purification, are not always removed before crystallization and can degrade over time due to low-level protease contamination. These features limit the broader application of the His₆ tag as a receptor for a crystallization chaperone and motivate conversion of 3D5 to new peptide specificity. The EE peptide (EYMPME) was chosen for (i) its short length, (ii) the presence of tyrosines to form hydrophobic interactions and hydrogen bonds that commonly dominate protein-protein binding energetics (Fellouse et al., 2007), (iii) charged residues to form electrostatic interactions. (iv) the presence of a proline to restrict conformational diversity and (v) the availability of high affinity commercial antibodies binding these peptides (Grussenmeyer et al., 1985; Prickett et al., 1989) (Covance, Sigma). Indeed, our optimized scFv exhibits enhanced crystallization propensity, including elevated solubility, stability, affinity and the ability to bind internal peptide sequences. This engineered, peptidebinding scFv represents a new class of crystallization chaperones that may eliminate the need for de novo identification of candidate chaperones from large antibody libraries.

Materials and methods

Molecular biology and expression of proteins presenting EE and His₆ peptides

Antibody binding sites (peptide sequences) were incorporated into proteins of interest via site-directed mutagenesis with mutagenic oligonucleotides (Integrated DNA Technologies). To generate a ligand with a C-terminal EE tag, maltose binding protein (MBP) was amplified from the E.coli genome, appended with an EE tag and cloned into the pAK400 vector (Krebber et al., 1997), with or without a stop codon before the vector-encoded His6 tag to generate the MBP-EE and MBP-EE-His₆ ligands. The variant without a His₆ tag was used during phage screening and panning; both variants as well as a third in which MBP was cloned directly into pAK400 (MBP-His₆) were used in Western analysis. To account for steric accessibility, we generated ligands with

varying numbers of internal EE tags by introducing tandem repeats of the EE coding sequence into the flexible linker connecting the heavy and light chains of unrelated scFv proteins. A single repeat was introduced into the DO11.10 scFv gene to generate scFv-EE₁, two repeats into the 14B7 scFv gene to generate scFv-EE₂ (Maynard et al., 2002) and three repeats into a non-native scFv consisting of the 3D5 light chain and 14B7 heavy chain to generate scFv-EE₃. These proteins also contain C-terminal, vector-encoded His₆ tags to facilitate purification. The original 14B7 scFv with only a C-terminal His₆ tag was used as a hexa-histidine-tagged ligand. All ligand proteins were expressed from pAK400 in E.coli BL21 in 250 ml cultures of TB media, induced with 1 mM IPTG for 3-5 h at 25°C before cell harvesting and periplasmic fractionation via osmotic shock as previously described (Maynard et al., 2005). Recombinant antibodybased ligands were purified via sequential immobilized Ni²⁺ affinity chromatography and size exclusion chromatography (SEC), using HEPES-buffered saline (HBS; 10 mM HEPES, 150 mM NaCl, pH 7.4). The MBP-based ligands were purified using an amylose affinity column and eluted with maltose-containing buffer (200 mM HEPES, 150 mM NaCl, 10 mM maltose, pH 7.4), prior to SEC.

Library generation by CDR and random mutagenesis

The 3D5 gene was generated by total gene synthesis from published amino acid sequence (Genscript), including previously identified solubilizing mutations (referred to here as 3D5) (Kaufmann et al., 2002) and cloned via SfiI-SfiI ligation into the pMoPac24 phage display vector (Hayhurst et al., 2003), which introduces a C-myc tag at the C-terminus of the displayed protein. To identify clones with desired peptide specificity, two CDR libraries were generated by modified Kunkel mutagenesis under error-prone conditions: HCDR3 only and HCDR2 + HCDR3 simultaneously. Randomized codons were selected to retain anti-peptide binding capabilities as described previously (Cobaugh et al., 2008). For generation of EE-specific antibodies, amino acid sequencing by Edman degradation (ICMB Protein Microanalysis Facility, University of Texas at Austin) of proteolytic fragments of the commercially available GluGlu antibody (Covance) identified a candidate HCDR1 sequence. The HCDR2 loop included 13 amino acid modifications (theoretical size, 9×10^{14}), while HCDR3 loop length was set at seven residues to reflect observed diversity in antipeptide antibodies (theoretical diversity, 6.4×10^8). The 3' H2 oligonucleotide sequence is CACGGTGAGTGTGGCC **CTSNNCTTVNHMNYSBNGTTATASNNSNNSNNSNNAY** YSNNARDMYDAATSNNTCCGATCCACTCCAGACC. while the 3' H3 oligonucleotide sequence is CCTTGACCC CAGTAATCCATAGCSNNSNNSNNGCTSNNSNNSNNSN NABNTGCACAGTAGTATACG.

A third library was generated to introduce random mutations into a pool of scFvs selected for EE-specificity. Here, 3D5 variants with desired epitope specificity were subjected to error-prone PCR with Mutazyme II DNA Polymerase (Stratagene) using flanking primers (5' scback and 3' scforlong, IDT; Krebber *et al.*, 1997) and exponential amplification according to the manufacturer's instructions. Briefly, reaction mixtures were heated at 95°C for 4 min, followed by 25 cycles of incubation at 95°C for 30 s, 52°C for 30 s and 72°C for 1 min to introduce a predicted 3-4

mutations per 1000 bp. Gel-purified PCR products (Qiagen) were used in a modified Kunkel mutagenesis step to produce two libraries. Library mutation rate and diversity was assessed by plasmid DNA sequencing at the University of Texas Core Facility using primer 5' pAKpel (Krebber *et al.*, 1997) (IDT).

Selection and screening by phage display

M13 phage monovalently displaying scFv–gpIII fusions were prepared as previously described (Krebber *et al.*, 1997). After precipitation with 1/4-volume PEG-2.5 M NaCl and resuspension in PBS, the phage concentration was quantified by absorbance: virions/ml = $[(A269-A320)\times6\times10^{16}/(#bases per virion)]$.

For panning, 10^{12} plaque forming units (pfu) were added to blocked ELISA wells (Costar) coated with either anti-C-myc antibody (9E10, Sigma), His₆ or EE presenting ligand. After equilibration, and washing with PBS-0.05% Tween, bound phage were eluted with 0.1 M glycine-HCl pH 2.2, transferred to a new tube and neutralized with 2 M Tris, pH 7.0. Phage were then amplified in *E.coli* in preparation for the next panning round or used to infect *E.coli* and plated to isolate single clones. Panning involved two cycles, each consisting of three selection rounds: one with immobilized anti-c-myc antibody to enrich for full-length scFv and remove variants with primer-encoded stop codons or frameshifts, followed by two rounds with peptide-tagged ligands.

Individual phage clones were analyzed by phage ELISA to confirm enrichment of peptide-specific clones and to screen candidates for biophysical characterization. Phage from single clones were produced in 200 µl in sterile 96-well plates (Costar). The plates were centrifuged and supernatant transferred to coated and blocked ELISA wells (Costar). After washing with PBS-0.05% Tween, bound phage were detected by anti-M13-HRP (1:2500, GE Healthcare) with tetramethylbenzidine (TMB, Sigma) substrate and the resulting absorbance at 450 nm recorded. Binding of each clone to immobilized anti-C-myc antibody (1 µg/ml), peptide ligand (MBP-EE or 14B7-His₆ at $4 \mu g/ml$) and blocked wells (5% milk) was monitored. Clones with a high ratio of peptide ligand to anti-C-myc signal, indicating high peptide-binding specific activity, were further characterized. To rank the relative affinities of these high-activity variants, phage were produced from 100 ml cultures and the concentration of phage particles was quantified by absorbance prior to ELISA analysis. The phage concentration resulting in 50% of the maximum ELISA signal (EC50) was compared to select final candidates for expression and characterization as soluble scFv protein.

To confirm conversion of specificity, phage displaying the 3D5/EE_48 or commercial anti-His-HRP (Invitrogen) were used to probe a western blot containing host proteins presenting the EE (MBP-EE, scFv-EE), His₆ (scFv-His₆) or both peptides (scFv-EE₂). Phage displaying 3D5 were not used as divalent display is required, due to low affinity His₆ recognition (Kaufmann *et al.*, 2002). A 15% SDS–PAGE gel was loaded with 10 μ l cellular lysate from cells expressing ligand, and electrophoresed prior to transfer to PVDF membrane. After blocking with 5% non-fat milk in PBS, the blot was incubated with 8×10^{10} virions/ml fresh 3D5/EE_48 scFv displaying phage for 1 h at room temperature, washed

three times with HBS-0.05% Tween and incubated with anti-M13-HRP secondary antibody (1:5000). Signal was developed with SuperSignal West Extended Duration substrate (Pierce), and the resulting image captured on Kodak film. The blot was stripped with mild stripping buffer (200 mM glycine pH 2.2 with 0.1% w/v SDS and 1% Tween-20), re-blocked and re-probed with commercial anti-His-HRP (1:5000, Invitrogen) to detect His₆-containing ligands. The blot was stripped a second time, blocked and probed with anti-MBP-HRP (1:2500, Invitrogen) to confirm similar loading between wells.

Chaperone protein expression, purification and complexation

The parent 3D5 and scFv genes selected from phage display experiments were subcloned into the SfiI–SfiI site of pAK400 for scFv expression (Krebber *et al.*, 1997), or pMoPac54, to produce an scAb (an scFv appended with a human kappa constant domain as a convenient detection handle; Hayhurst *et al.*, 2003). Protein was secreted into the bacterial periplasm of *E.coli* strain BL21, isolated by osmotic shock and purified by immobilized Ni²⁺ affinity chromatography and SEC using a Superdex S75 column (GE Healthcare), as previously described (Maynard *et al.*, 2005). The Superdex S75 column was calibrated using a Low Molecular Mass gel filtration calibration kit (GE Healthcare).

Protein purity and size were characterized by SDS-PAGE under reducing or non-reducing conditions (Sambrook, 2001). Protein solubility was determined by concentrating the protein to ~ 20 mg/ml, incubating for 4 days at 4°C, centrifuging for 10 min at high speed to pellet insoluble particles and quantifying the concentration of protein remaining soluble. Stability was assessed as the mid-point for thermal unfolding, using a fluorescence assay (Lavinder et al., 2009). Purified protein (20 μ l at 280 μ g/ml) or buffer blank and Sypro Orange (1 µl of a 1:1000 dilution; Molecular Probes) were heated in a Real Time PCR instrument (7900HT Fast Real-Time PCR System; Applied Biosystems) from 20°C to 85°C in increments of 0.5°C and analyzed with SDS.2 (Applied Biosciences). The scFv monomer-to-dimer ratio was determined from SEC traces by calculating the area under the curve for each peak with Unicorn software (GE Healthcare). Protein concentration was assessed by micro-BCA assay with a BSA standard curve and buffer blank (Pierce). To facilitate direct comparisons, all 3D5 and variant characterization values reported here were performed with these methods and specific values may differ slightly from those previously reported (Kaufmann et al., 2002).

Complex formation between 3D5/EE_48 and two ligands, scFv-EE₃ and MBP-EE (see above), was assessed by SEC. Equimolar volumes of purified 3D5/EE_48 and either scFv-EE₃ or MBP-EE ($\sim 1 \mu$ M each) were combined and allowed to incubate on ice for 6 min followed by separation using an analytical Superdex S75 column (GE Healthcare) equilibrated with 10 mM HEPES, 150 mM NaCl, pH 7.4. Fractions of interest were concentrated and characterized by SDS–PAGE. Control experiments applied the same quantity of each species alone as used in complexation experiments. The Superdex 75 column was calibrated using a Low Molecular Mass gel filtration calibration kit (GE Healthcare) supplemented with cross-linked albumin (Sigma).

Determination of chaperone-peptide binding affinity

Direct ELISA with purified scFv protein was performed in two orientations: scFv as an immobilized capture molecule or scAb detection of immobilized EE-tagged protein. For the former, wells were coated overnight at 4° C with 50 µg/ml scFv variant in PBS, prior to blocking with 5% milk in PBS. MBP-EE was serially diluted (1:2) from an initial concentration of 100 µg/ml, followed by washing and detection with anti-MBP-HRP (1:2500). For the inverse configuration, plates were coated with EE-tagged proteins (4 µg/ml) followed by 1:2 serial dilutions of scAb protein from 200 µg/ ml. In this case, detection was achieved with anti-human-kappa-HRP (1:2500; Sigma) and TMB substrate. To assess the pH sensitivity of the binding interaction, ELISAs were performed in which the scFv-ligand interaction proceeds at pHs ranging from 6.0 to 8.0, in 0.5 increments. To rank the relative affinities, the EC_{50} concentrations were compared.

Kinetic binding assays were performed with proteins bearing C-terminal His₆ or EE-tags and internal EE-tags to quantify scFv peptide selectivity using a BIAcore 3000 (GE Healthcare). Peptide-binding scFv or protein ligands were coupled to CM5 chips using NHS-EDC chemistry to a level of \sim 500 RU. The signal from a flow cell coupled with a control scFv (14B7-His₆; Maynard et al., 2002) was used to correct for non-specific binding to the matrix, while control scFv injections corrected for changes in sample refractive index. Soluble protein ligands were injected in a duplicate dilution series from 8 to 0.1875 μ M at a flow rate of 50 μ l/ min to minimize mass transport effects. 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 $\chi^2 < 0.6$ were used.

Protein crystallization

3D5/EE_48 was crystallized by the sitting drop vapor diffusion method at 4°C. Conditions were optimized based on those reported for 3D5 (Kaufmann *et al.*, 2002). One to two microliters of protein solution in HBS buffer at 3.8 mg/ml chilled on ice were mixed with 1 μ l sample of reservoir solution containing 0.1 M Mes (pH 6.4), 0.1 M magnesium acetate and 20–24% (w/v) PEG 8000. Crystals of rectangular or triangular shape appeared in 4 days and grew to a maximal size of 40–60 μ m within 4 weeks. The largest crystals grew when the reservoir to protein ratio was 1:1.33–1.66.

Data collection, structure determination and refinement

Crystals were harvested at 4°C and cryocooled using a solution consisting of 85.5% (v/v) reservoir solution and 14.5% (v/v) ethylene glycol. Crystallographic data were collected using a wavelength of 1 Å at the GM/CA-CAT beamline (Darien, IL) equipped with a 5 μ m mini-beam setup. Data were processed with XDS (Kabsch, 1993) and Scala (CCP4, 1994). The structure of 3D5/EE_48 was solved by molecular replacement with Molrep (CCP4, 1994) using a polyalanine search model derived from parent 3D5 asymmetric unit (PDB ID 1KTR) from which all non-protein atoms and loop residues were removed. All four 3D5/EE_48 scFv monomers

present in the asymmetric unit were identified from Molrep. The atomic model was fit to the respective electron density map using Coot (Emsley and Cowtan, 2004), and then iteratively refined with Refmac (CCP4, 1994). After several initial rounds of refinement using tight non-crystallographic symmetry restraints, refinement including Translation/ Libration/Screw motions and medium non-crystallographic restraints was conducted. Of the 947 residues present in the asymmetric unit, 99.2% are in most favored and additional allowed regions of the Ramachandran plot. The final model has been deposited in the PDB (PDB ID 3NN8). Figures were generated using Pymol (The PyMOL Molecular Graphics System, Version 0.99rc6, Schrödinger, LLC). Electrostatic surface potentials were calculated using APBS (Baker et al., 2001) and visualized using Python Molecular Viewer 1.5.4 (Sanner, 1999). Computational peptide docking was conducted with ClusPro (Comeau et al., 2004).

Results

Selection of 3D5 variants

We identified an scFv scaffold to use as a starting point for engineering peptide-binding chaperones by examining a family of structurally characterized antibody fragments binding small molecules (PDB IDs including 1KTR, 1MAJ, 2CJU, 1DLF, 2UUD, 1DSF, 1WZ1, 1N4X, 2G60) that share a highly conserved variable light chain (V_L from the murine V κ 1 germline, >90% identity) and, if crystallized, a major crystal contact. One member of this family, the His₆-specific 3D5 scFv, had previously been displayed on M13 bacteriophage (Lindner et al., 1997). We hypothesized that we could enhance and/or convert scFv peptide specificity while retaining the favorable crystallization characteristics of 3D5 through CDR and random mutagenesis, coupled with a phage display selection strategy in which peptide binding affinity, solubility, stability and expression level are used as proxy variables for crystallizability. Similar scaffolding approaches have been effective for antibody humanization and thermodynamic stabilization (Baca et al., 1997; Jung and Pluckthun, 1997).

To increase the versatility of our crystallization chaperones, we sought to identify variants with affinity for either the His₆ or EE (sequence: EYMPME) hexa-peptides. The chemical diversity of the EE peptide would be expected to enhance binding interactions while the inclusion of a proline would limit conformational entropy (Reiersen and Rees, 2001). In order to engineer scFvs with the desired peptide specificity (His₆ or EE), three libraries with randomized CDRs were generated by methods previously described (Cobaugh *et al.*, 2008). Since the heavy chain (V_H) typically dominates ligand interactions (Xu and Davis, 2000), V_H CDRs 2 and 3 (HCDR2 and HCDR3) were randomized to convert peptide specificity while retaining the desirable crystallization properties of 3D5. The three libraries of scFv HCDR variants (actual library size $\sim 10^7$ each) were monovalently displayed on the surface of M13 phage via fusion to coat protein gpIII, and scFv variants were selected for ligand binding specificity using a series of panning cycles. First, full-length scFvs, which present a C-terminal C-myc epitope, were enriched from prematurely truncated variants resulting from oligonucleotide-encoded stop codons via

Table I. Comparison of 3D5 scFv CDR H3 regions																				
Kabat #	90	91	92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	101	102	103	104	105
3D5	Υ	Y	С	Е	S	Q	S	G								А	Υ	W	G	Q
3D5/His_683	Υ	Y	С	А	А	S	S	Р	Y	S	Μ	R	А	А	М	D	Υ	W	G	Q
3D5/His_67	Υ	Y	С	Е	R	А	R			S	Р	R	А	А	М	D	Υ	W	G	Q
3D5/EE_48	Y	Y	С	А	R	R	G	G	S	S	Н	Y	Y	А	М	D	Y	W	G	Q

immunoprecipitation. Next, eluted phage were amplified in *E.coli*, and variants with desired peptide specificity selected via phage binding to an immobilized host protein presenting either the His₆ or EE peptide. One host protein, the 14B7 scFv with a terminal His₆ peptide (scFv-His₆), was employed for selection of hexa-histidine-specific variants. Two ligand proteins were used to select for EE peptide binders: MBP with a singe C-terminal EE tag (MBP-EE), and another scFv with two internal tandem EE tags to allow for steric accessibility within the Gly–Ser linkage between V_H and V_L immunoglobulin domains (scFv-EE₂).

The amplification and selection procedure was repeated twice, using different immobilized host proteins during each cycle to ensure selection for peptide, as opposed to host protein, specificity. Next, weakly peptide-reactive phage were pooled and subjected to random mutagenesis to yield the libraries, one based on EE-specific scFv. Sequencing of 20 individual clones from each library comprising $\sim 10^7$ unique members confirmed library diversity and the anticipated $\sim 0.5\%$ mutagenic rate. An additional three rounds of phage selection yielded the pool of EE peptide-specific scFv variants.

After screening several hundred clones by monoclonal phage ELISA followed by phage titration ELISA to rank clones by binding affinity, two His₆ (denoted 3D5/His_#) and six EE-specific (denoted 3D5/EE #) scFv variants with unique sequences were identified (Table I). Of these, two clones, 3D5/His_683 and 3D5/EE_48, provided the highest specific binding activity (measured as the ratio of peptide tag/anti-c-myc ELISA signal). Western blot analysis provided a clear verification of peptide specificity: 3D5/EE_48 displayed on phage bound host proteins with either internal or C-terminal EE peptides, but not those with only C-terminal His₆ peptides (Fig. 1a). These scFv variants were then expressed and purified as soluble protein (Fig. 1b and c), and characterized for binding activity by ELISA and surface plasmon resonance (SPR; Fig. 2a-f) analysis and for enhanced biophysical properties (see below).

Characterization of 3D5/His variants

The selected variants 3D5/His_67 and 3D5/His_683 differ from each other and 3D5 in the HCDR2 (3D5/His_67) and HCDR3 (3D5/His_67 and 683). These variants harbor longer CDR3 lengths with several amino acid differences (Table I). In our hands, 3D5/His_683 expressed nearly 3-fold better in *E.coli* than 3D5 (8.5 versus 3.1 mg/l culture; Table II), exhibits enhanced scFv solubility (estimated as 16.6 versus 2.3 mg/ml, respectively) and modestly improved affinity (K_d 808 versus 4700 nM). At concentrations relevant to crystallization (~4 mg/ml), 3D5/His_683 elutes from a gel filtration column as a mixture of a monomer and dimer (Fig. 1c). In contrast, 3D5/His_67 expressed at lower levels, but exhibited similar affinity for His₆ (K_d 760 nM). One of the two key



Fig. 1. Characterization of selected scFv variants. (a) Western blot detection of peptide tagged ligands by anti-His₆-HRP, $3D5/EE_48$ displayed on M13 phage and anti-MBP-HRP on the same blot. Lanes are 1, MBP; 2, MBP-EE; 3, MBP-EE-His₆; 4, scFv-His₆. (b) Size and purity of scFv chaperones by SDS–PAGE after initial purification. (c) Size analysis of purified scFv by SEC. Molecular weight standards elute on S75 at 7.4 ml (void volume; dextran 2000); 8.9 ml (132 kDa albumin dimer); 9.6 ml (75 kDa conalbumin); 10.9 ml (44 kDa ovalbumin); and 12.7 ml (29 kDa carbonic anhydrase).

residues in HCDR3 that stabilize the bound His₆ in the 3D5 crystal structure, Glu93 or Ser 96, is retained in each variant (Table I), yet these variants and others we tested all exhibited micoromolar affinity for His₆ (K_d 3–4 μ M). Thus, even though these variants possess rather different HCDR3s than



Fig. 2. Peptide-scFv binding kinetics. Top row, analysis of 3D5 and 3D5/His_683 binding affinity and specificity His₆ by SPR. Binding partners were injected in duplicate in concentrations ranging from 8000 to 125 nM. (a) 3D5 scFv recognition of immobilized scFv-His₆ ligand; (b) binding of scFv-His₆ ligand to immobilized 3D5/His_683; (c) 3D5 recognition of control MBP-EE ligand. Bottom row, analysis of 3D5/EE_48 affinity and specificity for EE-tagged proteins by SPR. Binding of (d) a single, internal EE peptide presented by scFv-EE₁, (e) a control scFv lacking an EE tag, scFv-His₆ (buffer spikes erased for clarity); and (f) a C-terminal tag presented by MBP-EE-His₆ to immobilized 3D5/EE_48 scFv ligands. Ligands were injected at concentrations ranging from 6000 to 475 nM for MBP-EE, 2000 to 125 nM for the scFv ligands. Cartoons illustrate the orientations and identity of binding partners.

3D5 and more favorable biochemical properties, their affinity for His_6 is not substantially improved over that of 3D5. These results suggest that the His_6 binding site is well organized for peptide binding, or that HCDR3 may contribute fewer productive interactions than expected. For these reasons, plus concerns regarding the utility of His_6 as a peptide ligand, these clones were not pursued further.

Characterization of 3D5/EE_48

The lead 3D5/EE scFv candidate, 3D5/EE_48, retains 85% amino acid identity relative to 3D5. In addition to novel HCDR sequences (Table I), two key amino acid changes in the $V_{\rm H}$ framework identified during random mutagenesis (E6Q and S74T) were instrumental in improving scFv expression and affinity. The impact of Glu versus Gln at position 6 has been previously described (Kipriyanov *et al.*, 1997; de Haard *et al.*, 1998).

The 3D5/EE_48 scFv displayed no detectable binding affinity for C-terminal His₆ tags and instead is able to bind both C-terminal and internal EE-tags with similar affinities, K_{d} 389 and 212 nM, respectively (Figs 1a and 2c-e; Tables II and III). A terminal EE tag followed by a His tag was recognized with higher affinity than a naked EE tag (K_d 389 versus 767 nM), perhaps due to protease protection and reduced entropy with an additional C-terminal extension. Varying the pH in 0.5 increments from 6.0 to 8.0 or increasing the number of internal EE tag repeats from two to three had no detectable effect on affinity as measured by ELISA (data not shown) and SPR (K_d 25–30 nM; Table III). An increase in affinity was observed for ligands harboring multiple EE repeats versus single repeats, likely due to re-binding effects, as the measured on-rates are similar but the off-rates are slower (Table III). The use of 3D5/EE 48 displaying phage as detection reagents in ELISA (data not shown) and western blot (Fig. 1a) demonstrated specific binding of EE but not His₆-tagged ligands. Expression levels

Table II	Biophysical	characteristics	of 3D5	scEv variants

	3D5	3D5/His_683	3D5/EE_48
Expression level (mg/l culture)	$\begin{array}{c} 3.1 \\ 2.3 \\ 46.5 \pm 0.5 \\ 100 \\ 50 \\ 4700 \ (0.08) \\ \text{ND}^{\text{b}} \end{array}$	8.5	2.1
Solubility (mg/ml)		16.6	12.8
Melting temperature (°C)		53.6 \pm 0.0	47.2 \pm 0.3
% AA identity		96.2	85
% Monomeric protein ^a		62	81
K_d (nM), 14B7-His ₆ (χ^2)		808 (0.17)	ND ^b
K_d (nM), MBP-EE (χ^2)		ND ^b	767 (0.03)

^a% monomer measured upon initial purification.

^bND, not detected.

Table III.	Characterization	of 3D5/EE_	48 scFv	binding	kinetics	by SP	R
------------	------------------	------------	---------	---------	----------	-------	---

					•	
Ligand	$k_{\rm on}~(1/{\rm Ms})$	SD (%)	$k_{\rm off} (1/{ m s})$	SD (%)	$K_{\rm d}~({\rm nM})$	χ^2
MBP-EE-His MBP-EE scFv-EE ₁ scFv-EE ₂ scFv-EE ₃	3.5E + 4 1.5E + 5 2.4E + 5 3.9E + 5 3.6E + 5	47 62 32 30 37	1.2E-2 7.3 E-2 4.7E-2 9.0 E-3 9.2 E-3	2 24 21 34 41	389 767 212 25.5 29.6	0.028 0.030 0.010 0.586 0.598

of 3D5/EE_48 (2.1 mg/l culture) are similar to that observed for 3D5 (3.1 mg/l culture), but the solubility increased from 2.3 to >12 mg/ml. In addition, 3D5/EE_48 is initially purified as a predominantly monomeric species (\sim 80% of total eluted protein; Table II) and retains this monomeric state when concentrated up to at least 3 mg/ml (Fig. 1c). This contrasts with the lower initial ratio of monomeric to dimeric protein (Table II) and slow conversion of purified monomer to dimer observed for 3D5 and 3D5/His_683 under similar conditions (Fig. 1c). The melting temperature of 3D5/EE_48 is almost identical to 3D5, indicating similar thermal stabilities (Table II). Overall, 3D5/EE_48 exhibits similar or enhanced biophysical properties when compared with 3D5 in terms of affinity, expression level, solubility, stability and homogeneity of oligomerization state.

Complexation of 3D5/EE_48 with EE-tagged proteins

The ability to isolate complexes of 3D5/EE 48 and client proteins expressing the EE tag was assessed next using SEC. Equimolar concentrations ($\sim 1 \mu M$) of 3D5/EE_48 and the client protein were combined and fractionated using an analytical gel filtration column. Fractions corresponding to the eluted peaks were analyzed by SDS-PAGE and compared with control runs with isolated binding partners. One client protein, scFv-EE₃ used originally for selections (see above), elutes as a dimer, and runs slightly higher than its expected molecular mass by SDS-PAGE, likely due to an extended conformation of the individual V_H and V_L domains within the monomer (Fig. 3a). Complexation with 3D5/ EE_48 results in a single elution peak that corresponds to a molecular mass consistent with a heterotetramer, i.e. an scFv-EE₃ dimer with two bound 3D5/EE 48 monomers (Fig. 3a). The second client protein tested was MBP-EE, which harbors only a C-terminal EE tag. Although MBP-EE by itself elutes as a monomer, complexation with 3D5/ EE 48 yields two higher molecular weight complexes with molecular masses consistent with a heterodimer and heterotetramer (Fig. 3b). Given the lack of dimerization precedent for MBP, the heterotetramer could arise from a domain



Fig. 3. Isolation of 3D5/EE_48 complexes with EE-tagged client proteins by SEC. (a) 3D5/EE_48 incubated with scFv-EE₃ ligand. Elution peak 1 corresponds closely to the expected retention volume of a scFv-EE₃ homodimer complexed with two 3D5/EE_48 molecules as described in text. (b) 3D5/EE_48 incubated with MBP-EE. Elution peaks 1 and 2 correspond to heterotetramer and heterodimer complexes as described in text. For (a) and (b), SDS-PAGE analysis of elution peaks confirms protein and complex identities. In addition, individual elution profiles are overlayed. Molecular weight standards as in Fig. 1.

swapped 3D5/EE_48 dimer in which two distinct binding sites for MBP-EE are presented. Domain swapping has been proposed as a mode for 3D5 dimerization (Kaufmann *et al.*, 2002).

Structure of 3D5/EE_48

Crystals of 3D5/EE_48, grown under conditions used to crystallize 3D5 (Kaufmann et al., 2002), appeared within 4 days and continued to grow over several weeks. The structure of 3D5/EE 48 was solved by molecular replacement using a search model derived from the 3D5 coordinates (see Materials and methods, Table IV). Although the crystals were grown under similar conditions, and the proteins share a high level of sequence identity, the two scFvs do not crystallize in the same manner (Fig. 4). First, whereas 3D5 crystals belong to a trigonal space group (P3₂21), crystals of 3D5/EE_48 belong to a cubic space group (F23). The asymmetric unit of 3D5/EE 48 contains four molecules whereas 3D5 contains one $V_H - V_L$ pair (Fig. 4a). In addition, in spite of the fact that no amino acid changes occurred in the major 3D5 crystal contact, the contact is not preserved in the 3D5/ EE_48 lattice. Whereas the crystal lattice of 3D5 is built by alternating V_H/V_L subunits from neighboring molecules, that of 3D5/EE_48 relies primarily on HCDR residues from adjacent molecules (see arrows, Fig. 4b). The second largest contact in the 3D5 crystal lattice (305 $Å^2$) has become the largest crystal contact (560 $Å^2$) in the 3D5/EE_48 lattice with several additional hydrogen bonds formed at this interface (see shaded area, Fig. 4b). Finally, as a consequence of lattice changes, 3D5/EE_48 crystals consist of 66% solvent with a channel \sim 52 Å wide while 3D5 crystals consist of 77% solvent and a channel \sim 70 Å wide (Fig. 4c).

The overall structure of the 3D5/EE_48 scFv remains very similar to that of the parent 3D5 (average rmsd \sim 0.55 Å for main chain atoms in V_L and \sim 1 Å for V_H domains); however, changes observed in the CDR regions reconfigure the peptide binding region to accommodate an EE-tag

Table IV. Data collection and refinement statistics

	3D5/EE_48
Data collection	
Space group	F23
Cell dimensions	
a = b = c (Å)	266.64
$\alpha = \beta = \gamma$ (°)	90
Resolution (Å)*	154.3-3.1 (3.18-3.10)
R _{sym}	12.4 (34.2)
I/oI	5.6 (2.1)
Completeness (%) ⁺	99.6 (99.6)
Redundancy	5.6
Refinement	
Resolution (Å)	154.3-3.1
No. of reflections	26932
$R_{\rm work}/R_{\rm free}$	18.3/23.7
No. of molecules	
Protein residues	947
<i>B</i> -factor protein (A^2)	40.8
R.m.s deviations (rmsd)	
Bond lengths (Å)	0.016
Bond angles (°)	1.871
- · · · ·	

*Highest resolution shell is shown in parentheses.

+5% of reflections were selected for R_{free}.



Fig. 4. Comparison of 3D5/EE_48 (top) and 3D5 (bottom) crystal lattices. (a) Asymmetric units. (b) Crystal contacts. The preserved contact common to both lattice networks shaded grey. New crystal contact comparison depicted in arrows. (c) Lattice structure with size of central cavity indicated. For both structures, lighter hue indicates the V_L chain while the darker hue indicates V_H . CDRs for 3D5/EE_48 are depicted in red and those for 3D5 are yellow.



Fig. 5. Analysis of 3D5/EE_48 structure. (a) Superimposition of 3D5/EE_48 and 3D5 with CDRs labeled. Amino acid changes discussed in the text are represented as ball-and-stick. LCDR1, LCDR2 and LCDR3 indicate V_L CDR loops 1, 2 and 3, respectively. HCDR1, HCDR2 and HCDR3 indicate V_H CDR loops 1, 2 3, respectively. (b) Calculated electrostatic surface potential of the binding site [-15 kT (red) to + 15 kT (blue)]. (c) Modeled peptide binding interaction of EE peptide with 3D5/EE_48. Color scheme identical to Fig. 4.

(Fig. 5). In the V_L CDR1 loop (LCDR1) of 3D5/EE_48, slight movement in residues His 27d-Asn 30 may be influenced by the presence of neighboring Leu 93 in the V_L CDR3 loop (LCDR3), instead of the corresponding His residue at this position in 3D5 (Fig. 5a). Another substitution in LCDR3 of 3D5/EE_48, introduction of a Pro 96 for the Phe in this position in 3D5 appears to open up the peptidebinding groove to accommodate longer peptides, and, in particular, may allow internal peptides to be recognized (Fig. 5a). Compared with the LCDRs, HCDRs are more divergent both in sequence and in structure. In 3D5/EE 48, the beta-hairpin in HCDR2 as a whole shifts closer to HCDR1. HCDR3 differs primarily in its longer length, which significantly alters the shape of the peptide-binding region when compared with 3D5. The binding surface near the interface of the heavy and light chains forms a pronounced tri-lobed hydrophobic pocket (Fig. 5b and c). The electrostatic surface potential reflects a charge distribution complementary to that of the peptide in this region (Fig. 5b).

We turned to computational docking to model EE-peptide binding to 3D5/EE_48 (Fig. 5c) because no crystals of adequate size for structure determination containing both 3D5/ EE_48 and EE-peptide have been grown to date, and soaking with the commercial EE-peptide (Covance) has not yielded crystals with bound peptide. The EE peptide is predicted to bind in an orientation in which the central proline (Pro 4) introduces a kink, allowing peptide residues Tyr 2 and Met 3, to reach into the hydrophobic binding pocket. In this working model, $V_{\rm H}$ residue His 50 appears to stabilize peptide Tyr 2 through hydrophobic interactions, while V_H residue Arg 95 forms key polar interactions with multiple peptide side chains (Tyr 2, Met 3, Glu 6). Peptide residues Glu 1 and Glu 6 stabilize this binding mode through surface electrostatic interactions, and hydrogen bonding interactions between the peptide backbone amide and carbonyl stabilize the peptide conformation. In the case of a terminal EE tag, the C-terminus may compete for the Glu 6 side chain interactions, and/or more flexibility of the tag may destabilize the peptide backbone interactions. In this docked model, peptide residue Met 5 does not appear to be directly recognized by 3D5/EE_48. Notably, V_L residues predicted to form interactions with peptide are conserved between 3D5 and 3D5/EE_48, while V_H residues contributing to peptide interactions, such as Arg 95, were altered during engineering, underscoring the role of the V_H in peptide recognition.

Discussion

Crystallization chaperones proposed are to aid co-crystallization by several distinct mechanisms, including immobilizing flexible regions, concealing exposed hydrophobic regions and providing polar surfaces capable of forming lattice contacts (Zhou et al., 2001; Hunte and Michel, 2002). To date, most co-crystal structures have employed antibody fragments because the molecular requirements for ligand binding are well understood and their hypervariable regions can be modified to recognize nearly any epitope of interest (Chothia and Lesk, 1987; Cobaugh et al., 2008). Typically, antibodies that recognize specific epitopes on unmodified target proteins are identified through traditional hybridoma screening or library selection techniques (Rothlisberger et al., 2004; Huber et al., 2007; Uysal et al., 2009; Veesler et al., 2009) where there is minimal control over the epitope recognized. Moreover, the identification and optimization of a chaperone tailored to each client protein of interest is an expensive and timeconsuming process. An attractive alternative is the use of commercially available purified monoclonal antibodies for common epitopes, such as commercially available anti-His antibodies. Unfortunately, the hybridomas secreting these antibodies are not available and the cost to purchase purified antibody sufficient for use in crystallization trials is prohibitive. Moreover, without the gene sequence available, the biophysical properties and format of the antibody (e.g. scFv, scAb, Fab) cannot be readily altered. Finally, even in the case where sequences are known and genes for the corresponding antibody fragments can be synthesized for recombinant expression, antibody fragments often express with relatively low yield in E.coli and lack suitable solubility and stability profiles.

The engineered scFv chaperone approach complements non-antibody formats that have been developed to allow modular recognition of a specific binding partner (e.g. DARPin, V_{HH}) (Huber *et al.*, 2007; Warke and Monmany, 2007; Sennhauser and Grutter, 2008; Tereshko *et al.*, 2008). Whereas these alternative frameworks express in very high levels (up to 200 mg/l in the bacterial cytoplasm for DARPins) and possess a stable structure, a potential disadvantage is their small size ~15 kDa, which limits the hydrophilic surface area available for generating protein–protein crystal contacts. In contrast, scFvs are nearly twice as large, and can be readily converted to a ~50 kDa Fab format to accommodate larger client proteins with a larger hydrophobic surface area.

Our engineered scFv chaperones, derived from the previously crystallized 3D5 scFv framework and binding short His₆ or EE peptide sequences, overcome several of the aforementioned limitations of antibody fragments and represent a potentially generalizable solution to the production of high affinity protein complexes for crystallization of difficult proteins. We overcame the affinity, pH sensitivity and solubility limitations specific to 3D5 by employing a two-step protein engineering process of randomizing the HCDR 2 and 3, followed by random mutagenesis of the selected scFvs. This selection scheme does not directly select for the ability to crystallize, as there is no clear biophysical correlate of crystallization propensity, but can select for 'well-behaved' proteins, as evidenced by the increased expression levels, solubility and peptide binding affinity of our characterized variants. The initial library design focused on the HCDRs because these can be sufficient to confer high affinity and specificity (Rader *et al.*, 1998; Sidhu and Weiss, 2004) and in the 3D5 family of antibody fragments, the V_L domains are highly conserved.

After limited success in improving the biochemical characteristics of His₆-specific scFvs, we converted 3D5 to EE epitope specificity. The 3D5/EE 48 scFv is expressed in high yield in E.coli, is highly soluble, is predominantly monomer, and is readily crystallized. The affinity of 3D5/ EE_48 for internal EE-tags (K_d 212 for single, 26 nM for multiple peptide insertions) likely reflects the combined effects of restricted conformational variability due to the presence of a proline in the EE-peptide, as well as the ability of the remaining peptide residues to participate in hydrogen bonding, electrostatic or hydrophobic interactions. Combined with the general reduced entropic costs of binding an internal peptide, this scenario represents a desirable binding configuration for crystallization chaperone and tagged client protein. Indeed, complexes of 3D5/EE 48 with host proteins are sufficiently tight to withstand separation by SEC. In the context of a co-crystallization experiment, a modest 5 mg/ml concentration of a 30 kDa scFv chaperone protein equates to 170 µM, which is nearly 1000-fold above the measured equilibrium dissociation constant and will drive complex formation within the crystallization drop.

Unexpectedly, the crystal lattice of 3D5, whose open framework and limited use of CDRs in crystal contacts was an initial design criterion, was not preserved in 3D5/EE_48. Although the use of CDR residues in crystal contacts appears to render 3D5/EE 48 not ideal for co-crystallization, none of the residues participating in the major crystal contact of 3D5 has in fact been altered. Thus, it should be possible for 3D5/ EE 48 to revert back the 3D5 lattice framework when most CDR residues are participating in a complex and CDR-based crystal contacts are no longer accessible. We are optimistic about the prospect of the ability of 3D5/EE 48 to promote crystallization of 'difficult' proteins, either by mediating formation of crystal contacts (as observed for KcsA; Zhou et al., 2001) or by immobilizing flexible loops (as observed for GPCRs; Rasmussen et al., 2007; Milovnik et al., 2009). Our current efforts are focused on further engineering of 3D5/EE_48 to render the CDR crystal contacts less favorable than those found in 3D5. We are also co-crystallizing 3D5/ EE 48 with MBP and candidate membrane proteins in which the EE peptide has been installed into an accessible but functionally silent location. In the long term, we plan to extend our approach to generate 3D5-based scFvs or Fabs that recognize other peptide sequences, leading to a toolbox of peptide binding crystallization chaperones with homotypic crystal contacts that could be used to crystallize any protein of interest.

Accession numbers

The coordinates of the 3D5/EE_48 structure are deposited in the Protein Data Bank under PDB ID 3NN8.

Acknowledgements

We thank Benjamin Roy for the scFv-EE1 construct.

Funding

This work was supported by the National Institutes of Health (grant number AI066239 to J.A.M.), the Packard Foundation (grant number 29098 to J.A.M.), the National Science Foundation (grant number 0845445 to R.L.L.) and the American Federation for Aging Research (R.L.L.). GM/ CA-CAT has been funded in whole or in part with Federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104). Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. DE-AC02–06CH11357.

References

- Baca, M., Presta, L.G., O'Connor, S.J. and Wells, J.A. (1997) J. Biol. Chem., 272, 10678–10684.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001) Proc. Natl Acad. Sci. USA, 98, 10037–10041.
- Banatao, D.R., Cascio, D, Crowley, C.S, Fleissner, M.R., Tienson, H.L. and Yeates, T.O. (2006) *Proc. Natl Acad. Sci. USA*, **103**, 16230–16235.
- Byrne, B., Abramson, J., Jansson, M., Holmgren, E. and Iwata, S. (2000) *Biochim. Biophys. Acta*, **1459**, 449–455.
- Collaborative Computational Project. (1994) Acta Cryst., **D50**, 760–763.
- Cherezov, V., Rosenbaum, D.M., Hanson, M.A., et al. (2007) Science, 318, 1258–1265.
- Chothia, C. and Lesk, A.M. (1987) J. Mol. Biol, 196, 901-917.
- Cobaugh,C.W., Almagro,J.C., Pogson,M., Iverson,B. and Georgiou,G. (2008) J. Mol. Biol., **378**, 622–633.
- Comeau,S.R., Gatchell,D.W., Vajda,S. and Camacho,C.J. (2004) Nucleic Acids Res., 32, W96–99.
- Cooper, D.R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M. and Derewenda, Z. (2007) *Acta Crystallogr. D: Biol. Crystallogr.*, **63**, 636–645.
- de Haard,H.J., Kazemier,B., van der Bent,A., Oudshoorn,P., Boender,P., van Gemen,B., Arends,J.W. and Hoogenboom,H.R. (1998) *Protein Eng.*, **11**, 1267–1276.
- Derewenda, Z.S. (2004) Structure (Camb), 12, 529-535.
- Emsley, P. and Cowtan, K. (2004) Acta Crystallogr. D: Biol. Crystallogr., 60, 2126–2132.
- Fellouse, F.A., Esaki, K., Birtalan, S., et al. (2007) J. Mol. Biol., 373, 924–940.
- Goldschmidt,L., Cooper,D.R., Derewenda,Z.S. and Eisenberg,D. (2007) *Protein Sci.*, **16**, 1569–1576.
- Grussenmeyer, T., Scheidtmann, K.H., Hutchinson, M.A., Eckhart, W. and Walter, G. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 7952–7954.
- Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B.L. and Georgiou, G. (2003) J. Immunol. Methods, 276, 185–196.
- Huber, T., Steiner, D., Rothlisberger, D. and Pluckthun, A. (2007) J. Struct. Biol., 159, 206-221.
- Hunte, C. and Michel, H. (2002) Curr. Opin. Struct. Biol., 12, 503-508
- Hunte, C., Koepke, J., Lange, C., Rossmanith, T. and Michel, H. (2000) *Structure Fold Des.*, **8**, 669–684.
- Jung, S. and Pluckthun, A. (1997) Protein Eng., 10, 959-966.
- Kabsch, W. (1993) J. Appl. Cryst., 26, 795-800.
- Kaufmann,M., Lindner,P., Honegger,A., Blank,K., Tschopp,M., Capitani,G., Pluckthun,A. and Grutter,M.G. (2002) J. Mol. Biol., 318, 135–147.
- Keenan, R.J., Siehl, D.L., Gorton, R. and Castle, L.A. (2005) Proc. Natl Acad. Sci. USA, 102, 8887–8892.
- Kipriyanov,S.M., Moldenhauer,G., Martin,A.C., Kupriyanova,O.A. and Little,M. (1997) *Protein Eng.*, 10, 445–453.

- Kovari,L.C., Momany,C. and Rossmann,M.G. (1995) *Structure*, **3**, 1291–1293.
- Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H.R. and Pluckthun, A. (1997) J. Immunol Methods, 201, 35–55.
- Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C., Honegger, A., Klinger, B., Mocikat, R. and Pluckthun, A. (1997) *Biotechniques*, **22**, 140–149.
- Lavinder, J.J., Hari, S.B., Sullivan, B.J. and Magliery, T.J. (2009) J. Am. Chem. Soc., 131, 3794–3795.
- Maynard, J.A., Maassen, C.B., Leppla, S.H., Brasky, K., Patterson, J.L., Iverson, B.L. and Georgiou, G. (2002) *Nat. Biotechnol.*, **20**, 597–601.
- Maynard, J., Adams, E.J., Krogsgaard, M.K., Petersson, K., Liu, C. and Garcia, K.C. (2005) J. Immunol. Methods, 306, 51–67.
- Milovnik, P., Ferrari, D., Sarkar, C.A. and Pluckthun, A. (2009) Protein Eng. Des. Sel., 22, 357–366.
- Muller, K.M., Arndt, K.M., Bauer, K. and Pluckthun, A. (1998) Anal. Biochem., 259, 54–61.
- Ostermeier, C., Essen, L.O. and Michel, H. (1995) Proteins, 21, 74-77.
- Pedelacq,J.D., Piltch,E., Liong,E.C., Berendzen,J., Kim,C.Y., Rho,B.S., Park,M.S., Terwilliger,T.C. and Waldo,G.S. (2002) *Nat. Biotechnol.*, **20**, 927–932.
- Prickett,K.S., Amberg,D.C. and Hopp,T.P. (1989) *Biotechniques*, 7, 580–589.
- Prive,G.G., Yerner,G.E., Weitzman,C., Zen,K.H., Eisenberg,D. and Kaback,H.R. (1994) *Acta Cryst.*, **D50**, 375–379.
- Prongay,A.J., Smith,T.J., Rossmann,M.G., Ehrlich,L.S., Carter,C.A. and McClure,J. (1990) Proc. Natl Acad. Sci. USA, 87, 9980–9984.
- Rader, C., Cheresh, D.A. and Barbas, C.F., 3rd (1998) Proc. Natl Acad. Sci. USA, 95, 8910–8915.
- Rasmussen, S.G., Choi, H.J., Rosenbaum, D.M., et al. (2007) Nature, 450, 383-387.
- Reiersen, H. and Rees, A.R. (2001) Trends Biochem. Sci., 26, 679-684.
- Rosenbaum, D.M., Cherezov, V., Hanson, M.A., et al. (2007) Science, 318, 1266–1273.
- Rothlisberger, D., Pos, K.M. and Pluckthun, A. (2004) FEBS Lett, 564, 340–348.
- Sambrook, J. (2001) *Molecular Cloning*, 3rd edn, Cold Spring Harbor Laboratory Press, New York.
- Sanner, F. (1999) J. Mol. Graph Model, 17, 57-61.
- Sennhauser, G. and Grutter, M.G. (2008) Structure, 16, 1443–1453.
- Sidhu,S. and Weiss,G. (2004) In Clackson,T. and Lowman,H.B. (eds), *Phage Display: A Practical Approach*. Oxford, UK: Oxford University Press.
- Stura, E.A., Graille, M., Housden, N.G. and Gore, M.G. (2002) Acta Crystallogr. D: Biol. Crystallogr., **58**, 1744–1748.
- Tereshko, V., Uysal, S., Koide, A., Margalef, K., Koide, S. and Kossiakoff, A.A. (2008) *Protein Sci.*, **17**, 1175–1187.
- Uysal, S., Vasquez, V., Tereshko, V., Esaki, K., Fellouse, F.A., Sidhu, S.S., Koide, S., Perozo, E. and Kossiakoff, A. (2009) *Proc. Natl Acad. Sci. USA*, **106**, 6644–6649.
- Vedadi, M., Niesen, F.H., Allali-Hassani, A., et al. (2006) Proc. Natl Acad. Sci. USA, 103, 15835–15840.
- Veesler, D., Dreier, B., Blangy, S., et al. (2009) J. Biol. Chem., 284, 30718-30726.
- Warke, A. and Monmany, C. (2007) Cryst. Growth Des., 7, 2219-2225.
- Wernimont, A. and Edwards, A. (2009) PLoS ONE, 4, e5094.
- Xu,J.L. and Davis,M.M. (2000) Immunity, 13, 37-45.
- Yang,J.K., Park,M.S., Waldo,G.S. and Suh,S.W. (2003) Proc. Natl Acad. Sci. USA, 100, 455–460.
- Zhou,Y.F., Morais-Cabral,J.H., Kaufman,A. and MacKinnon,R. (2001) *Nature*, **414**, 43–48.