Engineered discoidin domain from factor VIII binds αvβ3 integrin with antibody-like affinity

Thesis by
Gene Kym

In Partial Fulfillment of the Requirements for the degree
of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California
2014
(Defended April 23rd, 2014)

ACKNOWLEDGEMENTS

I would like to thank my family for encouraging my studies and keeping me focused on long-term goals. I am thankful for the teachers who educated me, the mentors who trained me, my colleagues in student organizations who showed me the power of teamwork, and my friends for teaching me about the world outside the of academia. I am grateful to Steve, the Mayo Lab, and the individuals at Caltech who made my doctoral years a time of exploration, learning, and fun.

ABSTRACT

Alternative scaffolds are non-antibody proteins that can be engineered to bind new targets. They have found useful niches in the therapeutic space due to their smaller size and the ease with which they can be engineered to be bispecific. We sought a new scaffold that could be used for therapeutic ends and chose the C2 discoidin domain of factor VIII, which is well studied and of human origin. Using yeast surface display, we engineered the C2 domain to bind to αvβ3 integrin with a 16 nM affinity while retaining its thermal stability and monomeric nature. We obtained a crystal structure of the engineered domain at 2.1 Å resolution. We have christened this discoidin domain alternative scaffold the “discobody.”

TABLE OF CONTENTS

FIGURES AND TABLES

ABBREVIATIONS

Da	Daltons
Db	Discobody
ECM	Extracellular matrix
E. coli	*Escherichia coli*
FACS	Fluorescence-activated cell sorting
fVIII	Factor VIII
IgG	Immunoglobulin G
K_d	dissociation constant
kDa	kilodalton
k_{off}	dissociation rate
k_{on}	association rate
LB	lysogeny broth, also known as Luria-Bertani medium
mAb	Monoclonal antibody
μM	micromolar
mM	millimolar
nM	nanomolar
PCR	polymerase chain reaction
PS	phosphatidylserine
S. cerevisea	*Saccharomyces cerevisea*

C h a p t e r 1

INTRODUCTION TO ALTERNATIVE SCAFFOLDS

Monoclonal antibodies (mAbs) can be engineered to bind specific molecular targets with high affinity, and were heralded as being "magic bullets" to target cancer. Despite their increasing commercial success, mAbs have had limited therapeutic value for many diseases [1-4]. A richer understanding of how mAbs engage with and utilize the human immune system is beginning to change how mAbs are engineered [5]. Additionally, antibody engineering innovations [6] have lead to novel mAb formats, including bispecifics [7-9], small domains [10-12], and fusions to cytotoxic payloads [13] or nanoparticles [14]. Small domain alternative scaffolds such as fibronectin or DARPins are driving many of these later generation technologies [11,12,15]. While alternative scaffolds have demonstrated value in protein engineering and as reagents, their clinical utility is just beginning to be explored [10,16].

Antibodies

The mAb immunoglobulin G (IgG) scaffold is a "Y" shaped 150 kDa glycoprotein made from a homodimer of heterodimers. Each homodimer half has a 50 kDa heavy chain (HC) made of four immunoglobulin domains, and a 25 kDa light chain (LC) made of two immunoglobulin domains. The HC and LC pair to form the fragment antigen-binding (Fab) region, which is involved in molecular recognition. There are two Fabs that make up the arms of the "Y" shape. The tail end of the mAb is the fragment crystallizable region (Fc), which mediates immune effector functions and the long serum half-life of mAbs (Fig. 1-1).

The Fab region of the mAb is evolved through a series of VDJ recombination gene rearrangements and somatic hypermutation evolution that introduces point mutations for greater stability of the mAb as well as affinity and specificity of antigen binding. Clones are filtered through a negative selection step that removes self-antigens during B-cell development in the immune system. Traditionally, an antigen of interest that has been tethered to or formulated with immune-stimulating haptens is injected into a mouse. The murine immune system recognizes the antigen as foreign, and activates individual B-cell clones harboring a mAb sequence that recognize the foreign target. This clone is affinity matured and expanded, producing large amounts of antigen-specific antibodies. These B-cells can be isolated from the mouse spleen and fused with an immortalized myeloma cell line to produce hybridomas, which are then screened for the mAb of interest. These hybridomas can be multiplied and utilized to make clonal mAbs [17].

mAbs derived from non-human sources cannot effectively be used as human therapeutics due to their immunogenicity. Mouse mAbs form human anti-mouse antibodies (HAMA), which lead to neutralizing antibodies against the drug. Genetic engineering allowed for techniques to make mouse mAbs more "human" by replacing the majority of their sequence with the human homologue, which is less likely to produce a HAMA immune response. This was traditionally achieved through chimeric or humanized antibodies. Chimeric antibodies retain the mouse variable domain (F_v) grafted onto a human scaffold. The F_v region consists of the variable heavy (V_H) and variable light (V_L) domains. Examples of chimeric antibodies include cetuximab (Erbitux, Bristol-Myers Squibb, Eli Lilly and Company) and rituximab (Rituxan, Biogen Idec, Genentech/Roche). Humanized antibodies have an entirely human sequence except for the mouse antigen binding loops (complementarity determining regions, or CDRs). Trastuzumab (Herceptin, Genentech/Roche) is an example of a humanized mAb [17-19].

Fully human mAbs have been engineered using display technologies or genetically engineered mice with human antibody repertoires [20]. Phage display is the most prominent display method, and utilizes the M13 bacteriophage with antibody libraries fused to the viral p3 coat protein [21]. Yeast display is another useful technology that expresses antibody libraries on the yeast surface as Aga1p-Aga2p yeast mating protein fusions [22]. While larger libraries can be accessed using phage display (10^9-10^{11}) than yeast display (10^7-10^9), the yeast eukaryotic protein processing system may allow for greater library expression [23]. Because display libraries are made from synthetic DNA, including cloned human B-cell sequence repertoires, it is possible to isolate fully human mAbs. Adalimumab (Humira, AbbVie) is an example of a fully human mAb from phage display. A number of genetically engineered mice with completely human B-cell repertoires have been created, including the XenoMouse, which lead to the discovery and commercialization of panitumumab (Vectibix, Amgen) [20,21].

The Fc region of the mAb is critical for serum half-life and effector functions, such as antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), and antibody-dependent cell phagocytosis (ADCP). mAbs usually have serum half-lives of two weeks, allowing for less frequent dosing of the drug. This is because the 150 kDa mAb size is larger than the 60 kDa glomerular filter cutoff which filters out smaller proteins, and the Fc region interaction with the neonatal receptor FcRn. FcRn is present on the surface of endothelial cells, which endocytose mAbs from the serum. Normally, endocytosed proteins are enclosed in vesicles with a pH of 6, and are targeted for lysosomal destruction. However, FcRn proteins in these vesicles bind mAbs via the Fc region at pH 6. FcRn complexed with the mAbs are returned to the surface of the cell and are exposed to serum, where the FcRn-Fc interaction dissociates at neutral pH 7.4. The Fc region also binds four different Fcγ receptors and the first component of complement for immune effector functions. FcγRI, FcγRIIa, and FcγRIIIa are activating receptors, while FcγRIIb is inhibitory. The

four isotyopes of human IgG bind each of these receptors with differing affinities, and therefore have varying effector functions. Fc regions can be engineered to modulate half-life and effector function [24-27].

Several early protein engineering advances demonstrated that the mAb IgG protein scaffold was not the only option for targeted protein therapy (Fig. 1-2). Etanercept (Enbrel, Amgen/Pfizer) is a chimeric tumor necrosis factor (TNF) inhibitor made by fusing the extracellular domain of TNF receptor 2 to an IgG1 Fc, essentially replacing the Fab portion of the mAb with a natural receptor (Fig. 1-2A) [28]. Domain antibodies (dAbs, GlaxoSmithKline) are V_H domains engineered for binding and resistance to aggregation via phage display (Fig. 1-2B). It was hoped that dAbs would show better tissue and tumor penetration due to their small size, and that they could easily be tethered together to make bispecific molecules [29,30]. The 10^{th} repeat of fibronectin III (fibronectin or Adnectin, Bristol-Myers Squibb) was the first alternative scaffold to be engineered. Fibronectin is a β-sandwich with three loops clustered on one end that resembles a dAb (Fig. 1-2C). Fibronectin exhibits thermal stability above 80°C and its binder derivatives have been extensively studied [31-38]. Etanercept, dAbs, and fibronectin have influenced subsequent mAb variants and alternative scaffolds for use as therapeutics.

Despite the success of mAbs, they have inherent limitations that preclude certain applications. mAbs are large, multimeric, glycosylated, and bound by disulfides. This requires expression in expensive mammalian cell lines with slow turnaround times between batches, adding significantly to the time and cost of production for clinical trials. Additionally, they are moderately stable, and do not have long shelf lives compared to many smaller alternative scaffold domains [11,12,15,39]. Their homodimeric nature makes bispecific mAbs difficult to produce, and can require complex

engineering to ensure a heterodimeric mAb pair. Single-chain variable fragment (scFv) formats, which string together the V_H and V_L domains via a long glycine and serine rich (Gly-Ser) linker, are not stable [40]. Additionally, the use of Gly-Ser linkers can potentially be immunogenic [8]. While the 150 kDa size of mAbs is good for a long serum half-life, the large molecular weight can prevent deep tissue penetration and tumor uptake when compared with small domains [11,12,15,39]. Additionally, the complex patent situation around mAbs can make freedom to operate analyses difficult, since many alternative scaffolds have clear-cut intellectual property associated with them [41].

Since Shohei Koide's group demonstrated that fibronectin could be engineered to bind new targets [31], other non-antibody proteins have been explored as alternative scaffolds and have enjoyed varying degrees of success. Besides fibronectin, the most prominent examples include DARPins, Avimers, Anticalins, Affibodies, and peptide-based scaffolds [10,15,16]. Here we review the different scaffolds with a focus on protein engineering and clinical applications. Structures of their respective folds are shown in Figure 3, with engineered regions highlighted in red [16,42-49].

Fibronectin

The 10th repeat of the human fibronectin III domain is a 94 amino acid (10 kDa) β-sandwich with seven β-strands composing the core, and three loops clustered on one face that constitute the binding interface (Fig. 1-3A). These loops are the BC, DE, and FG loops, with FG being the largest. The total number of engineerable amino acids is 10 to 24 residues. Due to the high thermal stability of the parent scaffold, with a T_m above 80°C, sequence diversity, loop length variability, and mutations are well tolerated in engineered variants [10]. Fibronectin has been engineered using phage [31], yeast [36], and mRNA display [50]. mRNA display, and a related technique called

ribosome display, are *in vitro* technologies which are most useful for small domains and peptides [51].

Using loop length variability and recursive mutagenesis in yeast display, Hackel *et al.* extended the affinity limit of the fibronectin scaffold to 1 picomolar against hen egg lysozyme [34]. This is the tightest known engineered protein-protein interaction, though some scFvs can bind hydrophobic molecules such as fluorescein with femtomolar affinities [52]. The bottom AB loop of fibronectin and the β-strands have also been engineered to bind targets, demonstrating the further utility of this scaffold [38,53].

CT-322 (Angiocept, Bristol-Myers Squibb) is a polyethylene glycol (PEG)ylated fibronectin engineered against vascular endothelial growth factor receptor 2 (VEGFR-2). It binds both human and mouse receptors with high affinities, allowing for use of mouse xenograft models for preclinical testing. CT-322 blocks the interaction of VEGFR-2 with VEGF-A, VEGF-C, and VEGF-D. CT-322 had a serum half-life of 100 hours, and in clinical trials demonstrated small decreases in tumor volume in 4/34 patients. 24 patients showed stable disease. Additionally, a PEGylated bispecific fibronectin against epidermal growth factor receptor (EGFR) and insulin-like growth factor 1 receptor (IGF1R) showed greater anti-proliferative effects against various human cancer cell lines than either single domain, demonstrating that the scaffold works in bispecific formats [16,54].

Fibronectin domains have been fused to mAbs to make multi-specific molecules. Two different fibronectins engineered to bind different epitopes of EGFR were coupled to an anti-EGFR mAb to form a triepitopic fusion protein, termed Ab-Fn3. Ab-Fn3 induced EGFR surface crosslinking and downregulation, even in cell lines resistant to anti-EGFR mAb monotherapy. EGFR is a target of interest in cancer, and can undergo downstream activating mutations that render the tumor resistant

to current mAb therapeutics [55]. Novel mAb formats such as Ab-Fn3 highlight the potential of therapies based on alternative scaffolds and mAbs. Folds similar to fibronectin have been engineered for use as therapeutics, including a domain from human tenascin C (Tn3, Medimmune) and a fibronectin consensus domain (Centyrin, Janssen Pharmaceuticals) [56,57].

DARPins

DARPins, or designed ankyrin repeat proteins, are alternative scaffolds based on 33 amino acid repeating units. These repeats consist of a helix-turn-helix-β hairpin fold, and were engineered as consensus designs from human and non-human ankyrin repeat sequences. Jawless vertebrates use similar repeat proteins based on leucine-rich repeats for their immune system, demonstrating a natural precedent for using repeat-based non-antibody folds to provide protection against a broad array of foreign antigens. The average DARPin contains five internal repeats with an N- and C-terminal cap to stabilize the protein, and are typically 130-200 amino acids long. DARPins are diversified along seven residues in the helices and some small loops of each subunit (Fig. 1-3B), and are not believed to change conformation when bound to antigen. With less entropy to lose upon binding, picomolar binders can be isolated from a single round of evolution. Additionally, they display extraordinary stability with a T_m of close to 100°C for four-repeat units, with stability generally increasing as more units are added [12,58,59]. DARPins have primarily been engineered via ribosome display [58], but phage display has also been used using the SRP signal sequence. Because of the fast folding rates of DARPins, traditional Sec-dependent phage display leads to cytoplasmic folding of DARPin libraries and precluded translocation to the periplasm, which is required for display [60,61].

DARPins have been utilized in creative applications for therapeutics [16,58,59]. Of particular interest is viral retargeting to tumors for gene therapy. One study used a tetrameric DARPin expressed separately from the virus, with three of the DARPins binding an adenovirus surface protein. The remaining DARPin was used to target HER2 positive cells. Upon infection with the adenovirus-DARPin assembly, the cells showed a significant increase in luciferase transduction. Because the DARPin construct was expressed independently of the virus, complications associated with expressing virus fused to extraneous constructs were avoided [62]. DARPins have also been used to redirect adeno-associated viruses (AAVs) as direct coat-protein fusions [63].

Because DARPins lack cysteines, they express well in bacterial hosts and can be site-specifically conjugated by the addition of individual cysteine residues, which are amenable to various bioconjugate chemistries. DARPins have been conjugated to protamine polymers complexed with small interfering RNAs (siRNAs), and demonstrated delivery of anti-*bcl*-2 siRNA to tumor cells via epithelial cell adhesion molecule (EpCAM) binding DARPins. EpCAM binding DARPins have also been attached to protein toxins such as *Pseudomonas* exotoxin A, and this fusion demonstrated potency against EpCAM-positive tumor cells. As with the Ab-Fn3 fibronectin example mentioned before, DARPins have been fused to form multiepitopic binders against EGFR. These molecules showed activity equivalent to, and in some cases better than, cetuximab, an EGFR targeting mAb. Anti-VEGF DARPins are undergoing phase II clinical trials for age-related macular degeneration (AGN 150998, Allergan) [16,58,59].

Avimers

Avimers, short for “avidity multimers,” are tandem 35 amino acid (4 kDa) alternative scaffolds based on the human A-domain. These A-domain derivates are strung together into a “pearls on a string” format. The A-domain contains three internal disulfide bonds and a stabilizing calcium ion

(Fig. 1-3C). Over 200 different A-domains are found in human proteins, including low-density lipoprotein receptors. 28 positions per subunit are used for diversity generation, similar to the 25-30 residues involved in mAb binding. Avimer libraries were created using exon shuffling. Unlike other domains, where affinity maturation may be required, the first binding A-domain is subsequently tethered to a new library, where a second randomized A-domain separated by a five amino acid linker is generated. In this fashion, up to eight A-domains were engineered into a single Avimer. Avimers bind with subnanomolar affinities and can be expressed in high yields, even with eight domains linked together [41,64].

Avimers have been tested clinically, and were found to have some advantages over their mAb counterparts. mAbs against c-met or CD28 cause adverse effects, possibly because each mAb has two Fab regions and can bind two antigens at once, leading to receptor clustering and activation. Avimers against these targets did not demonstrate negative effects because each Avimer only binds a single antigen, although on multiple epitopes. An anti-IL-6 triepitopic Avimer inhibited IL-6-induced cell proliferation and showed activity in mice. Since Avimers are small proteins, they are rapidly cleared from the serum through the kidneys. This problem was solved by tethering an anti-IgG domain as part of the Avimer to take advantage of the natural serum half-life extension of the IgG-FcRn interaction [16]. Amgen has since acquired Avimer technology for further development.

Anticalins

Anticalins are engineered from the human lipocalin family. Lipocalins are 170 amino acid (20 kDa) β-barrels with eight β-strands that support four binding loops (Fig. 1-3D). They are found in plasma, and secreted/tissue fluids. They do not contain disulfides, have T_ms above 70°C, and can express well in bacterial hosts. They are unique in that they have a relatively large opening in the binding

interface that can accommodate small molecule binding in a cavity made by the binding loops and β-barrel structure. In nature, this cavity is used to bind small molecule vitamins and hormones for transport, storage, and sequestration. Anticalin libraries can be made with 16-24 random residues in the loops as well as select residues in the small molecule binding pocket, allowing for great diversity generation. Due to their unique structure, they can be engineered to bind proteins or small molecules [16,65,66].

Anticalins have been conjugated to fluorescent molecules via site-specific cysteine insertions, or PEGylated for extended half-life. PEGylated Anticalins against VEGF showed anti-angiogenic and anti-tumor activity in mouse xenograft models. PRS-050, a site-specifically PEGylated Anticalin that binds to all splice forms of human VEGF-A as well as rodent orthologues, was tested in a phase I dose escalation study. PRS-050 showed no immunogenicity, with only a single patient (out of 26) showing ADAs after 17 doses. Half-life was approximately six days, and no free VEGF-A was detected after drug treatment, while PRS-050-VEGF-A complexes were detected up to three weeks after dosing [67]. Bispecific Anticalins, called duocalins, have been made to target two antigens at once [16]. $A^{89}Zn$-labeled anti-MET Anticalin was used to visualize MET-positive tumor expression in mouse models via positron-emission tomography (PET) imaging. Tumor uptake was shown to be dose-dependent [68]. Pieris AG is the company developing Anticalins.

Affibodies

Affibodies are derived from protein A, a *Staphylococcus aureus* protein that is used widely in industry to purify mAbs based on protein A's interaction with the Fc domain. Affibodies are engineered from the B-domain of protein A, and have been modified to remove their Fc binding while increasing stability. The structure consists of a 58 amino acid (6 kDa) three-helix bundle without any cysteines. The antigen-binding portion consists of randomized side chains on two of the

three α-helices (Fig. 1-3E). Anti-HER2 Affibodies have been generated with affinities down to 22 picomolar [16,69].

Radiolabelled variants were used to image HER2-positive xenograft tumors with high contrast in mouse models. Anti-HER2 Affibodies fused to albumin binding domains lead to molecules with increased serum half-life due to the long circulation times of albumin, and had a concomitant increase in tumor uptake. Anti-HER2 Affibodies were fused to *Pseudomonas* exotoxin A and killed BT-474 breast cancer cells in tumor xenograft models. EGFR binding Affibodies were also generated for radiolabelled imaging and showed accumulation in tumors. Anti-EGFR Affibodies were tethered via a Gly-Ser linker to anti-HER2 Affibodies to make bivalent molecules, since both HER2 and EGFR are co-expressed on some tumor types [10,16,69].

Affibodies have also been developed as Fc fusions and in large complexes for drug delivery via viruses, nanoparticles, and liposomes. Affibodies are of particular interest in imaging due to their small size and rapid plasma clearance, which can increase contrast ratios. They have been conjugated to fluorescent proteins, dyes, and radiolabels for such applications [69]. Affibody AB is developing the Affibody technology.

Peptides

Small peptide-based binders have been engineered using a variety of scaffolds such as knottins, the Kunitz domains, and naturally occurring toxins. These domains are 30-60 amino acids (3-7 kDa), and contain several internal disulfide bonds. Due to their small size and rigid structures, most engineered peptide-based binders are based on loop grafts or have been evolved to enhance their natural binding activities. Some of these peptide scaffolds have been engineered for specific

selectivity profiles, since many of their parental molecules demonstrate promiscuity in binding partners [10,16,39]. The universal applicability of these small scaffolds to create *de novo* binders has yet to be demonstrated.

Knottins are 30 amino acid (3 kDa) cystine knot proteins with three internal disulfide bonds that confer stability onto the protein, preventing denaturation even in extreme acid or base conditions. Several knottins, including the human Agouti-related protein (Fig. 1-3F) and the *Ecballium elaterium* trypsin inhibitor were engineered with RGD-based libraries to bind to integrin. Variants with different subtype selectivity profiles against αvβ3, αvβ5, and αvβ1 integrin were engineered with affinities down to 10 nanomolar. These peptides were selected for in yeast display, and binders were synthesized chemically for testing [70,71]. The molecules have been labeled for PET imaging in early cancer detection [72,73]. ω-conotoxins are another class of disulfide-rich peptides, in this case derived from the cone snail *Conus magus* [74]. SNX-111 (Ziconotide, Prialt) is a Food and Drug Administration (FDA) approved therapeutic for chronic pain that blocks N-type voltage-gated calcium channels [75].

Kunitz domains are peptide-based protease inhibitors that are 60 amino acids (7 kDa) in size. They inhibit proteases, including serine proteases such as trypsin [16]. Kunitz domains are also disulfide rich, with three internal disulfides that hold together three loops involved in binding. The Alzheimer's amyloid β-protein precursor Kunitz domain (Fig. 1-3G) was engineered with phage display to selectively inhibit human plasma kallikrein. Libraries with random residues near the binding loop were selected against immobilized kallikrein, and variants with inhibition constants (K_i) down to 75 picomolar were isolated [76-78]. Other Kunitz domain engineering work based on the human lipoprotein associated coagulation inhibitor lead to the identification of ecallantide (Kalbitor, Dyax), an FDA approved therapeutic for hereditary angioedema [79].

Next generation alternative scaffolds

Alternative scaffolds have been engineered to bind a variety of targets, and have been linked together to make multi-specific molecules. Exciting next generation applications of alternative scaffolds are beginning to emerge, including mAb fusions targeting up to five ligands at once, small bicyclic peptides with chemically conjugated hydrophobic cores, and engineered super ligands that are tailor-made to modulate specific biological pathways.

The aforementioned triepitopic Ab-Fn3 against EGFR demonstrated receptor clustering and efficacy against monotherapy resistant cell lines. This concept has also been demonstrated against five different antigens at once using mAbs fused to small alternative scaffolds. These molecules are known as zybodies (Zyngenia). By tethering Affibodies, knottins, or peptides (referred to as "modular recognition domains," or "MRDs" in this study) with varying degrees of structure to the N- and C-termini of HER2 or EGFR targeting mAbs, zybodies showed better efficacy than trastuzumab (anti-HER2) in an angiogenesis-dependent xenograft tumor model [80].

Small-molecule conjugated bicyclic peptides were engineered using a modified phage display to create rigid cyclic peptides with synthetic cores for added stability. Peptide libraries with diverse loops interspersed by three fixed cysteine residues were designed. These cysteine residues were conjugated to tris-(bromomethyl)benzene (TBMB), an aromatic ring with three thiol-reactive groups. To reduce background binding to the phage itself, a cysteine free D1-D2 domain of the p3 fusion protein was utilized. Conjugated phage showed reduced infectivity, possibly due to background conjugation of the TBMB molecule to surface lysines. By randomizing two six-residue loops between the three cysteines, bicyclic peptides against kallikrein were engineered with K_is of

1.5 nanomolar. The bicyclic peptides can be synthesized chemically [81], and Bicycle Therapeutics is developing the technology for clinical use.

Interleukins (IL) are immune signaling molecules that bind to receptor complexes consisting of different subunits. Affinity towards the receptor depends on which subunits interact with the bound IL. IL-2 is one such molecule that exerts immune stimulating effects. The IL-2 receptor consists of α, β, and γ subunits that all form interactions with the IL-2 ligand. IL-2 can bind to the αβγ complex with high affinity, or the βγ complex alone with a lower affinity. IL-2 variants were made to bind either the αβγ or βγ complexes with increased affinity, allowing selective IL-2 activation on different receptor assemblies. Certain cell types, such as leukocytes, that are targets of IL-2 activation can be specifically modulated by the βγ binding IL-2 variant. Using this strategy, background cells which express the αβγ complexes, such as lung endothelial cells, can be avoided [82].

Factor VIII

Factor VIII (fVIII) is a blood-clotting protein that is encoded by the *F8* gene. Defects in the *F8* gene cause hemophilia A, which is usually inherited as an X-linked recessive gene, but can also arise from spontaneous mutation. fVIII is a 2332 amino acid (330 kDa) protein that consists of six domains (A1-A2-B-A3-C1-C2) that are expressed as a single polypeptide from vascular and tubular endothelium, as well as sinusoidal cells of the liver. The six-domain fVIII protein is proteolytically cleaved and the B domain is removed, leaving a heterodimer of A1-A2 that is complexed with A3-C1-C2. This heterodimer circulates in the serum with an additional protein called the von Willebrand factor (VWF). Upon injury, fVIII is further processed to fVIIIa, a heterotrimer of A1, A2, and A3-C1-C2, which is released from VWF and binds to platelet membranes. The C2 domain is responsible for binding to phosphatidylserine (PS) rich membranes on activated platelet cells.

fVIIIa binding to membranes catalyzes the interaction of factors IXa and X (fIXa, fX), which leads to further downstream processing that ultimately results in coagulation [49,83-88].

Patients with hemophilia A are missing or lack sufficient quantities of fVIII, and they require infusions of fVIII derived from natural or recombinant sources. fVIII can be isolated from donated blood (Hemofil M and Monarc-M, Baxter; Monoclate-P, ZLB Behring). Naturally derived fVIII requires extensive purification due to cases of contaminating HIV or hepatitis in the 1980s. Recombinant fVIII is also available (Advate and Recombinate, Baxter; Helixate FS and Kogenate FS, Bayer; ReFacto, Wyeth), and avoids the problems of serum virus contamination. Hemophiliac patients who lack natural fVIII and receive fVIII infusions sometimes form ADAs against the foreign protein. Additionally, fVIII has a short serum half-life at 8-12 hours, requiring many injections [89-91]. Newer fVIII technologies are being developed to overcome these shortcomings. mAbs have long serum half-lives due to their Fc regions, which interact with FcRn receptors. Fc fusion proteins have been made for a variety of proteins. fVIII was fused to an Fc to make rVIIIFc, which demonstrated a 1.54-1.7 fold longer $t_{1/2}$ in human patients, potentially requiring less frequent dosing [92]. PEGylated fVIII variants demonstrated extended serum half-life and reduced ADA formation in mouse models [93]. Gene therapy can be used to introduce a gene encoding for functional fVIII, and has shown promise in a variety of animal models [94]. A particularly creative approach was to replace fVIII with a bispecific mAb that can bind both fIXa and fX, replacing fVIII's natural function of clustering these two proteins. This approach successfully catalyzed the downstream coagulation processes necessary for blood clotting. Additionally, no ADAs against the mAb were formed, and the mAb demonstrated a serum half-life of two weeks [95]. This was tested in a nonhuman primate model, and may allow for innovative therapeutics for hemophiliacs.

The C2 domain of fVIII is of critical importance. Hemophilia A can also be acquired through autoimmune antibodies against the C2 domain, and ADAs that form against the C2 domain of fVIII drugs can lead to decreased efficacy [89]. The HAMSTeRS database (Hemophilia A Mutation, Structure, Test, and Resource Site) is a compilation of fVIII genetic mutations [96]. The C2 domain has been co-crystallized with a variety of inhibitory mAbs. Additionally, alanine scanning mutagenesis, which tests the importance of certain residues for binding, has been performed on the C2 domain. These mutants were tested for binding to a panel of inhibitory mAbs, helping identify which residues can be engineered to avoid B-cell epitopes for future drug development [97].

Discoidin Domain Family

The C2 domain is part of the discoidin domain family, a 150 amino acid (18 kDa) domain that mediates many biological processes, including cellular adhesion, vasculogenesis, angiogenesis, migration, and development. DDR1 and DDR2 are receptor tyrosine kinases that contain the discoidin domain and are involved in extracellular matrix remodeling, cell adhesion, and proliferation. Neurexins and neuropilins also contain discoidin domains, and are important in nervous system development. Discoidin domains bind to a variety of cognate ligands such as galactose, collagen, growth factors, phospholipids, and other non-charged lipids. Discoidin domains are found as subunits of larger proteins, and can be arrayed in a repeated fashion or found as single units within the protein. The domain was first identified in the amoeba *Dictyostelium discoideum*, and was described as a lectin that bound galactose with high affinity (10^{-6} to 10^{-8} M K_d). Since then, genome sequencing has identified over 100 eukaryotic and 300 prokaryotic proteins that contain the discoidin domain [98-101].

Over 20 X-ray and NMR structures of the discoidin domain, either alone or in its parental protein, have been solved to date. The structure consists of eight β-strands that form a distorted β-barrel. The

β-strands support three conformationally flexible spikes on one end that make up the binding interface (Fig. 1-3H). While the core β-barrel structure is fairly conserved within the discoidin domain family, the three binding spikes demonstrate diversity as result of the range of ligands that the family can bind to. However, each individual discoidin domain is only known to bind to a single cognate ligand [101]. The fVIII discoidin domain contains hydrophobic residues at the ends of spikes one and three. These intercalate into phospholipid membranes, and have positively charged residues at the base of the spikes that interact with the negatively charged phospholipid head groups [83,86]. In the case of the *Dictyostelium discoideum* discoidin domain, the spikes specifically coordinate binding to galactose residues to provide specific binding. The N- and C-termini of the discoidin domain are adjacent to one another, and bound by a disulfide in eukaryotic discoidin domains. The discoidin domain is believed to have originated from a single gene. Its modular architecture facilitates evolution into new proteins via domain shuffling [98-101].

Discobody Alternative Scaffold

The C2 domain exhibits several useful properties that make it attractive as an alternative scaffold. It is small, monomeric, and of human origin [87]. Additionally, the PS binding spikes have been mutated to no longer interact with the cognate PS membranes or VWF factor [83]. The large body of knowledge related to C2 mutations and fVIII engineering as a therapeutic are useful for protein engineers trying to manipulate C2 function for novel purposes [86,88,97]. Towards this end, we engineered the C2 domain of fVIII to bind to a new protein target - αvβ3 integrin. αvβ3 integrin is a cell surface protein involved in cell adhesion and migration during development, and is overexpressed in a number of tumors, where its dysregulation is associated with angiogenesis and metastasis [102,103].

Our work is the first demonstration that the discoidin domain scaffold can be engineered for non-cognate binding purposes. We call engineered discoidin domains “discobodies.” Engineering the C2 domain of fVIII may also illustrate ways to further engineer fVIII for next generation hemophilia A therapeutics.

Figures

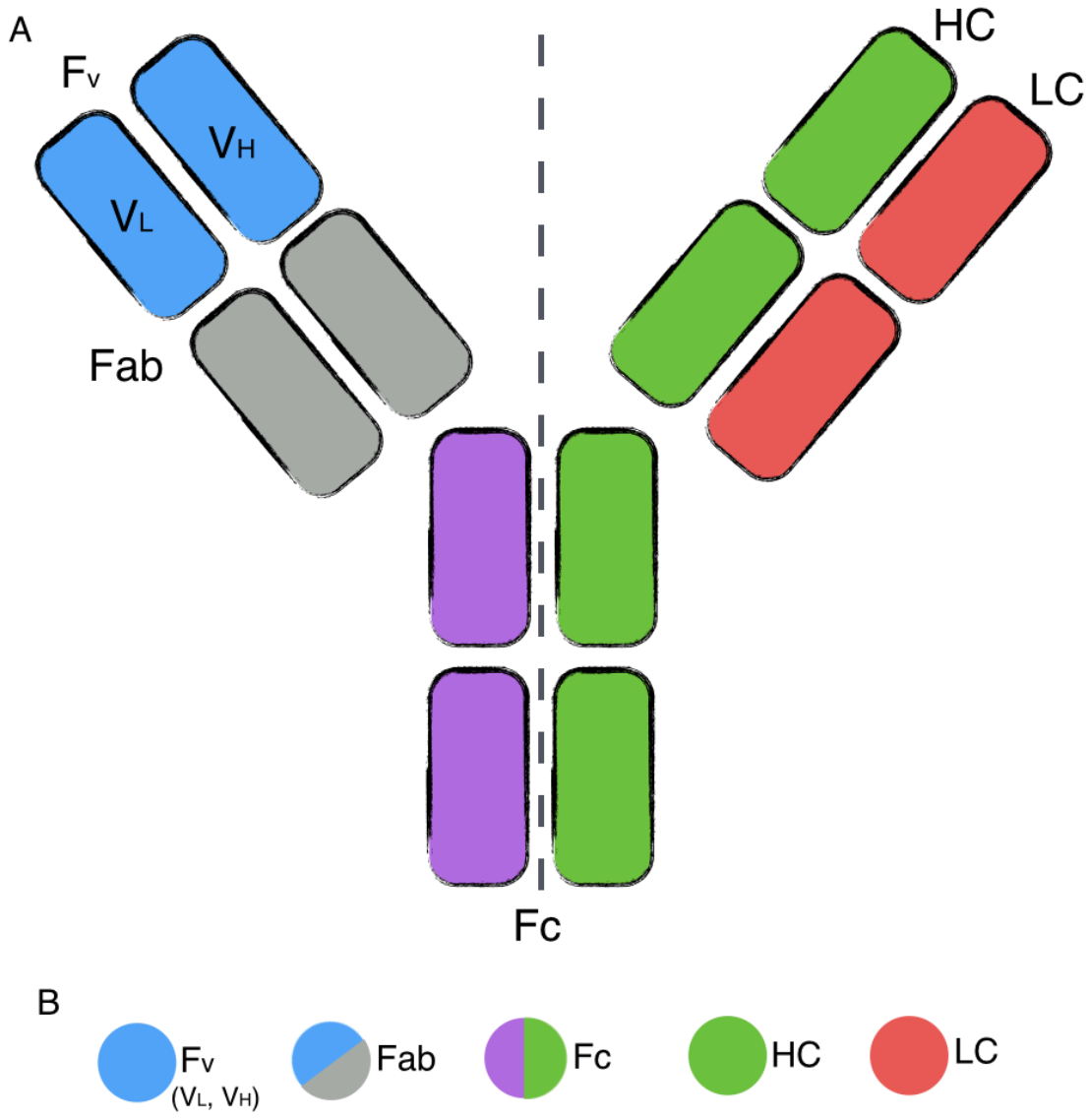


Fig. 1-1. Domain schematic of a mAb. (A) Upper left: F_v in light blue, with V_L and V_H domains labeled. F_v is part of the Fab region shown in light blue and grey. There are two Fab fragments on the upper left and upper right, only the left one is annotated in this figure. Bottom: Fc region shown in purple and green. Upper right: HC and LC are shown in green and red, respectively. There are two HCs and LCs per mAb, only the right side is annotated in this figure. Regions of the mAb with overlapping nomenclature and are colored and labeled differently on each side. The axis of symmetry between each homodimer of heterodimers is shown as a dashed line bisecting the Fc region. (B) Legend summarizing the color schematics.

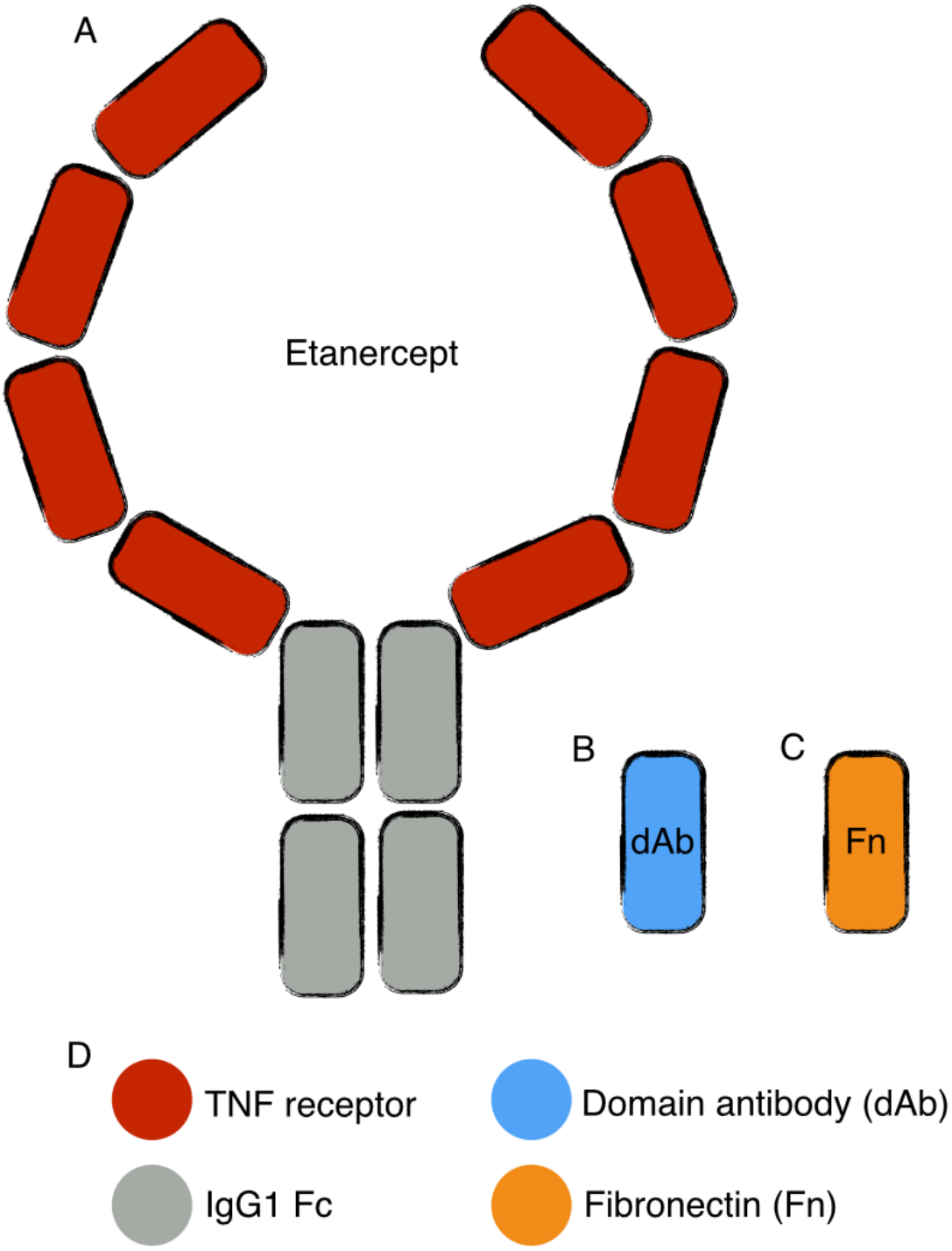

A
Etanercept
B
dAb
C
Fn
D
TNF receptor
Domain antibody (dAb)
IgG1 Fc
Fibronectin (Fn)

Fig. 1-2. Early mAb alternatives. (A) Etanercept, or Enbrel, a chimeric protein consisting of the extracellular domain of TNF receptor 2 in red fused to the Fc region of an IgG1 in grey. (B) Domain antibody, or dAb, shown in light blue. (C) Fibronectin domain shown in orange. (D) Legend summarizing colors and labels.

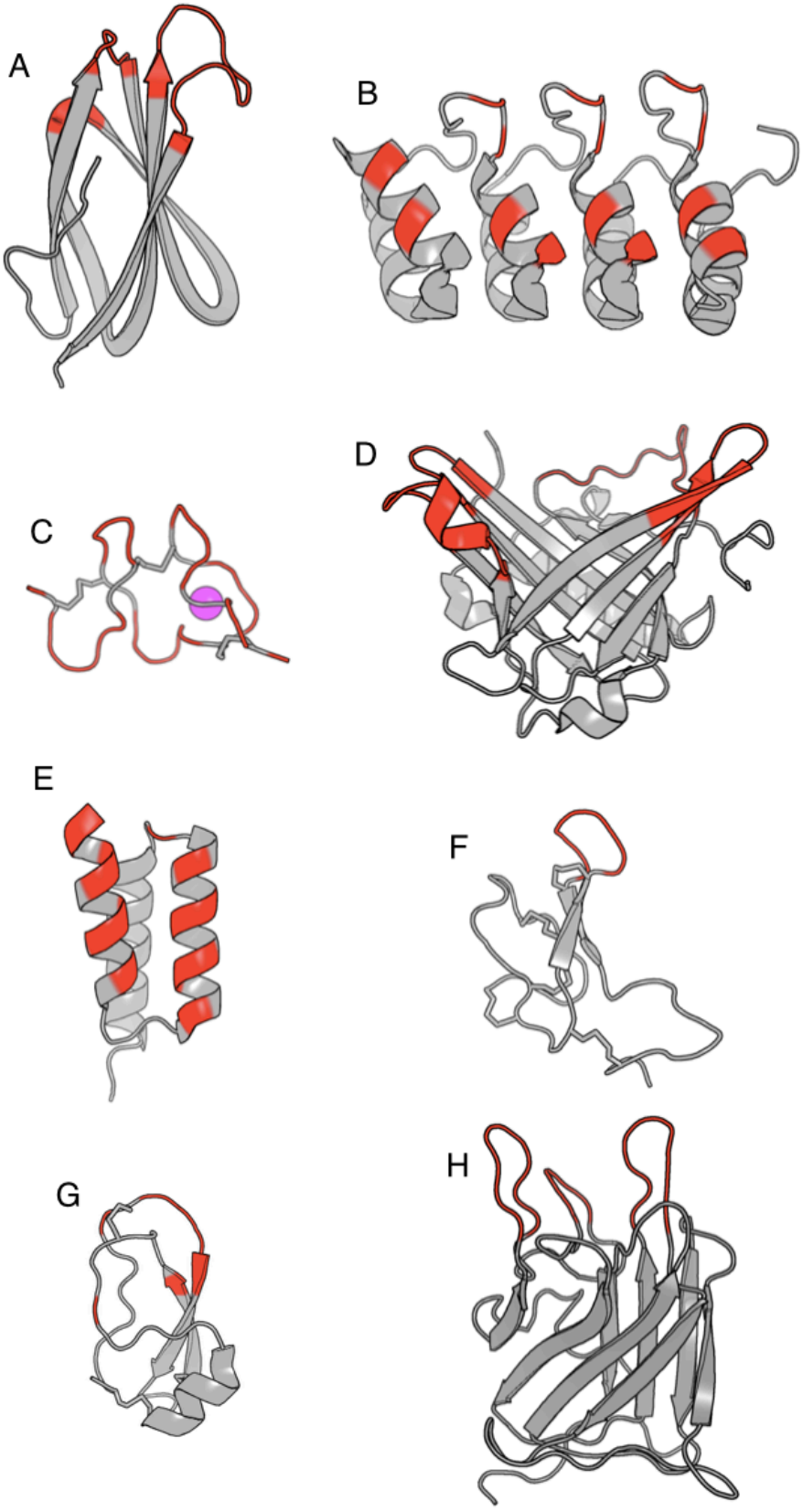

A
B
C
D
E
F
G
H

Fig. 1-3. Alternative scaffold structures in grey with engineerable binding interfaces highlighted in red. (A) Fibronectin, (B) DARPin, (C) Avimer subunit with stabilizing calcium ion shown in pink, (D) Anticalin, (E) Affibody, (F) knottin, (G) Kunitz domain, and (H) the discoidin domain (modified form PDB IDs 1FNA, 2V5Q, 1AJJ, 3BX7, 3MZW, 1HYK, 1KTH, and 2R7E, respectively).

Chapter 2

DISCOBODY ENGINEERING AND CHARACTERIZATION

Monoclonal antibodies (mAbs) are an important part of the biotechnology industry and have demonstrated value as therapeutics, as research reagents, and in diagnostic assays [1-3]. mAbs can be engineered to bind specific molecular targets with high affinity. However, this ability is not limited to mAbs, as in recent years alternative scaffolds based on fibronectin, protein A, ankyrin repeats, lipocalins, and other proteins have been engineered to bind various antigens. Alternative scaffolds can be engineered to have binding affinities and specificities characteristic of mAbs. Many of these scaffolds have also demonstrated superior biochemical and biophysical properties, such as high bacterial expression levels and thermal stability [15]. These attributes are important for reagents and diagnostic tools, where low material cost and high thermal stability (to facilitate storage and use in real-world conditions) are priorities [1,104,105].

The therapeutic utility of alternative scaffolds is assessed with different criteria [11,106]. Immunogenicity is a critical issue for any protein therapeutic. Because immunogenicity remains difficult to predict *a priori* [106], it is important to mitigate this risk by starting with a human protein [11]. The first mAbs were of non-human origin and, despite successes in laboratory settings, were not suitable as human drugs. They were too immunogenic for repeated dosing and required humanization to make them useful as a therapeutic [107]. In contrast, scaffolds derived from human proteins, such as fibronectin [31], have elicited less immunogenic responses in patients. A Phase I trial of CT-322, a fibronectin engineered to bind VEGFR-2, only showed anti-drug antibodies (ADAs) against the engineered binding loops and not the scaffold itself. Importantly, the ADAs did not affect plasma levels of CT-322 or a VEGF-A biomarker, and were

not associated with any adverse events [54]. It remains to be seen whether alternative scaffolds built from consensus designs [combinations of homologous sequences that were fully [57] or partially [108] derived from human proteins] will form ADAs upon repeated intravenous injections [106]. Other alternative scaffolds from non-human sources, such as camelid nanobodies, have been humanized prior to clinical use [109].

Alternative scaffolds have been touted for their small size, which could lead to enhanced tumor penetration [11,15,105,110]. Additionally, monomeric scaffolds can be strung in tandem to form bispecifics that target two different antigens, or recognize two different sites on the same antigen and thus enhance binding. Bispecifics have been created using fibronectin [37], camelid nanobodies [111], DARPins [112,113], and more [15]. Another use of alternative scaffolds is to tether them to mAbs to form multi-specific molecules. Various proteins have been appended to mAbs including fibronectin [55], peptides, knottins, and affibodies [80]. The majority of these domains have N- and C-termini on opposing ends of the molecule (Fig. 2-1A, C, D, and F). This architecture requires a long and potentially immunogenic linker [8] when connecting tandem domains (Supplementary Fig. 2-S1A).

Despite the promising results with existing scaffolds, new alternative scaffolds tailored for use as bispecific therapeutics are still needed. Here, we present an alternative scaffold based on the C2 discoidin domain of human factor VIII (fVIII). We call this discoidin domain-based scaffold the "discobody," or "Db." Discobodies are of human origin and have favorable N- and C-termini geometries that may facilitate bispecific tethering. The N- and C-termini are adjacent to one another and are linked by a disulfide bond (Fig. 2-1B and E). Tandem domains facing opposite directions to maximize binding to two targets should not require a substantial linker, as is needed for other alternative scaffolds (Supplementary Fig. 2-S1B).

fVIII is a blood coagulation protein encoded by the *F8* gene, which is involved in the formation of blood clots after injury. Mutation to the *F8* gene results in the medical condition hemophilia A, requiring hemophiliacs to receive injections of functional fVIII to aid blood clotting [85,114]. The C2 domain of fVIII is an 18 kDa distorted β-barrel that contains three loops, or "spikes," on its binding interface. Spikes 1 and 3 extend the farthest and contain hydrophobic residues at their ends. The native function of these spikes is to insert into phosphatidylserine (PS)-rich platelet membranes to initiate the coagulation cascade [83,87]. An important feature of our chosen scaffold is that it is based on the β-barrel topology. β-barrel folds are similar to the β-sandwich structures of the canonical mAb immunoglobulin and fibronectin scaffolds. Much like the mAb immunoglobulin and fibronectin folds, the discoidin β-barrel has loops clustered on one end [101] that may be engineered to bind specific antigens (Fig. 2-1).

Due to its importance in fVIII function, the C2 domain has been extensively studied [84]. The domain can be expressed independently of fVIII, and positions on the spikes have been identified that abolish wild-type PS recognition, thereby providing a template for future engineering efforts [83,87]. The discoidin domain is a ~150 amino acid domain that is found in over 100 eukaryotic and 300 prokaryotic proteins and is often involved in cellular adhesion and development. The discoidin family demonstrates high sequence diversity in the spike regions, with different discoidins able to bind to a variety of different cognate lipids, carbohydrates, and proteins [101]. This inherent ability to encode natural diversity in one scaffold means the fold can tolerate different sequences and should therefore be amenable to the production of synthetic libraries. Its small size, monomeric nature, tolerance to mutation, and extensive biochemical characterization all make the C2 domain an attractive protein for alternative scaffold design [12].

We chose to validate the fVIII C2 scaffold by engineering a binder against αvβ3 integrin, a cell surface protein and a target of interest in oncology [102]. Additionally, αvβ3 is a well-studied protein and is known to interact with the linear Gly-Arg-Gly-Asp-Ser (GRGDS) peptide, exhibiting a 740 nM IC_{50} in extracellular matrix (ECM) competition assays [115]. The Arg-Gly-Asp (RGD) motif is found in ECM proteins such as fibronectin and is used by αvβ3 integrin-expressing cells as footholds to navigate the ECM for cellular motility [116]. αvβ3 integrin is overexpressed in a number of cancers [103], and the RGD-αvβ3 interaction contributes to tumor growth, angiogenesis, and metastasis. Blocking this interaction could provide therapeutic value [102]. We sought to build an RGD-based library of discobody variants and select for binders against αvβ3 integrin.

Materials and Methods

Reagents and strains

αvβ3 integrin was purchased from R&D Systems. Yeast display *Saccharomyces cerevisea* strain EBY100, yeast display plasmid pPNL6, yeast secretion strain YVH10, and yeast secretion plasmid pPNL9 were obtained from Pacific Northwest National Laboratory. All yeast display protocols, buffers, and reagents were used as previously described [22,70]. Oligonucleotides were obtained from Integrated DNA Technologies, and PCR assembly primers were designed using DNAWorks from NIH's Helix Systems [117]. KOD Hot Start Polymerase from Novagen was used to PCR assemble gene inserts. The inserts had flanking 20-40 bp overlaps with the desired cloning site in their respective, linearized plasmids. These were assembled into the appropriate plasmids using either yeast homologous recombination or Gibson cloning into TOP10 *Escherichia coli* cells. Gibson cloning used reagents previously described [118], and constructs cloned into TOP10 competent cells were subsequently mini-prepped and transformed into yeast or BL21 (DE3) *E. coli* for protein expression. Bacterial expression used the pET11a vector with

Strep-Tag II cloned onto the C-terminus. Purification resins included HIS-select HF Nickel Affinity Gel from Sigma and StrepTactin Sepharose High Performance from GE Healthcare Life Sciences. Unless otherwise stated, all chemicals were from Sigma, and all *E. coli* strains from Life Technologies. Soluble protein expression was performed in 2.5 L Ultra Yield Flasks from Thomson. Protein was concentrated in Amicon 10,000 Da MWCO centrifugal filters from EMD Millipore. DNA extraction was performed with kits from Qiagen.

Loop grafts, library construction, flow cytometry, and sorting

RGD-based integrin binding loops from six different engineered knottin peptides [70,71] and from wild-type fibronectin [43] were grafted into the spike 1 region of the discobody starting template (4-Ala C2 discoidin domain of fVIII). Yeast displayed protein was expressed at 20ºC and 250 RPM for 16 hours. Displayed protein was incubated with αvβ3. Flow cytometry was performed on a BD Biosciences FACSCalibur, and data was analyzed with FlowJo from Tree Star. 50,000 cell counts were collected for loop graft binding experiments. All yeast display experiments were performed in Integrin Binding Buffer (IBB) as previously described [70]. Libraries were constructed by high efficiency yeast electroporation [22]. FACS was performed on the MoFlo XDP instrument from Beckman Coulter using polypropylene BD Falcon FACS tubes. Sorted sequences were represented using WebLogo [119].

Expression and characterization of soluble discobodies

Discobodies were expressed both in YVH10 yeast secretion culture and BL21(DE3) *E. coli.* YVH10 transformed with a discobody construct in pPNL9 with a TEV-cleavable 6HIS tag were grown overnight in 4L of SDCAA–Ura. This was spun down and transferred to YPGR media (1% yeast extract, 2% bacto peptone, 2% galactose, 2% raffinose, 0.1% dextrose; denoted in w/v, yeast media components from BD Biosciences, sugars from Sigma) with penicillin-streptomycin

from Life Technologies. This culture was expressed at 20ºC, 200 RPM, for 48 hours. Pellets were spun down and discarded twice, and the supernatant was ammonium sulfate (VWR) precipitated at 80% salt. Precipitate was resuspended in YVH10 binding/wash buffer (300 mM NaCl, 20 mM sodium phosphate, pH 7.8, 0.05% tween 20, 2.5% glycerol, 10 mM imidazole) and loaded onto HIS-select resin. After washing with the same buffer, samples were eluted in YVH10 elution buffer (300 mM NaCl, 20 mM sodium phosphate, pH 7.8, 0.05% tween 20, 2.5% glycerol, 200 mM imidazole), concentrated, and run over an analytical Superdex-75 column from Amersham Pharmacia on the ÄKTA FPLC system. Sample volume was 0.5 mL, run at 0.5 mL/min in discobody storage buffer (200 mM NaCl, 20 mM Tris, pH 7.4). Typical yields were 1 mg/L from yeast secretion culture.

For *E. coli* expression, single colonies of BL21(DE3) containing the discobody construct in pET11a with a TEV-cleavable Strep-Tag II were picked and inoculated in 5 mL LB cultures with 100 ug/mL amplicillin (LB-amp) and grown overnight at 37ºC, 250 RPM. The next day, we inoculated 4 L of LB-amp per clone, at 1 L per flask, with a 1:333 dilution of overnight culture and grew it at 37ºC until it reached an OD_{600} of 0.5. Cultures were inoculated with 1 mM IPTG and grown at 16ºC, 200 RPM, overnight to express protein. The following day, 4 L of the discobody clone were spun down and pooled, after which the pellet was split between two 50 mL conical Falcon tubes. The concentrated pellets were frozen at -20ºC then thawed at room temperature. Each 50 mL tube corresponded to 2 L worth of pellet and weighed 6.5 g. We added 35 mL of BugBuster Master Mix from EMD Millipore to each 50 mL tube with heavy vortexing, then nutated the mix at room temperature for 30 minutes. This was then spun for 45 minutes at 15,000 rcf, 4ºC on an angled rotor to separate the pellet from the discobody-containing supernatant. The resulting 30 mL of supernatant was poured into a new 50 mL tube and mixed with 6 mL of 5X StrepTactin binding/wash buffer (1X is 150 mM NaCl, 1 mM EDTA, 100 mM

Tris pH 8). This mix was run through equilibrated StrepTactin resin in a gravity column, washed once, then eluted in StrepTactin elution buffer (150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin, 100 mM Tris pH 8). Eluate was concentrated and purified by gel filtration using discobody storage buffer as described above. Yields were 1 mg/L.

Binding kinetics in display

Kinetics were performed as described [24]. All incubations were performed while nutating at room temperature. Quenched samples were placed on ice in a 4ºC cold room. All work with cold samples was performed at 4ºC with chilled tips and a dedicated cold room centrifuge. Samples were handled in 1.5 mL Eppendorf tubes, then transferred to a Nunc polypropylene 96-well V-bottom plate prior to addition of secondary antibody for ease of handling and standardization between samples. Competition experiments were performed entirely in 96-well plates. Prior to flow cytometry analysis, samples were spun down and kept as pellets on ice, covered from light. Immediately prior to analysis, individual samples were resuspended in 0.5 mL of IBB and transferred to a polystyrene BD Falcon FACS tube for loading. 20,000 cell counts were collected for each sample in kinetic assays. Analysis was performed in GraphPad Prism version 6.0d for OSX 10.9 from GraphPad Software. K_d was fit to the Michaelis-Menten model, k_{on} calculated association kinetics of two or more concentrations of hot ligand, and k_{off} was determined via one phase exponential decay dissociation kinetics. k_{on} values using one ligand concentration association kinetics were performed by constraining with the calculated k_{off} and 200 nM hot ligand. Competition plots were fit to a five-parameter logistic equation for asymmetric sigmoidal curves, with the discobody competition plot being performed in duplicate.

Crystallization

Initial crystal screening was performed using an Index screen (Hampton Research). Crystallization was conducted in 96-well plate format using a MRC sitting-drop crystallization plate (Hampton Research) at 20ºC; protein drops at 10 mg/mL were mixed with precipitant in a 0.5 μL x 0.5 μL ratio and were equilibrated with 50 μL of precipitant. Four initial hits were observed in 25% PEG3350 in different buffers. Upon crystal optimization, the best diffracting crystal was obtained in 0.1 M Bis-Tris pH 6.5, 10% PEG750 methyl ether (Sigma). Crystals started to appear after four days and grew to maturation in a week with an approximate length of 300 μm.

X-ray data collection and structure determination

Crystals were harvested and soaked in a cryoprotectant solution containing Fomblin Y 16/6 (Sigma). Crystals were frozen and sent to the Stanford Synchrotron Radiation Laboratory (SSRL). The diffraction data were collected at SSRL beamline 12-2 equipped with a PILATUS 6M PAD detector. The diffraction data were indexed, integrated, and scaled using XDS [120]. Initial phase was obtained by molecular replacement using PhaserMR [121] of the scaled experimental data using the previously solved C2 domain of human factor VIII [87] (PDB ID: 1D7P). Model building was done using Coot [122], and the structure was refined using Phenix.Refine [123]. Composite omit map was calculated using CCP4 suite [124].

Results

The search process

We set out to find a new scaffold by searching the RCSB protein data bank [125] for a protein that matched the following biophysical characteristics: monomeric, 1–50 kDa, and 25–100% β-sheet. We hoped that a β-barrel scaffold would allow for a binding interface predicated on loops linking β-strands, much like the immunoglobulin β-sandwich fold [126]. In this case, we could

adopt many of the lessons learned from engineering loop-based binders and apply them to our scaffold, possibly also allowing loop grafting from known mAbs or other binding proteins. In addition to these biochemical search terms, we searched for “blood” or “serum” as keywords, and limited the search to human proteins. We expected that this constrained search would find a scaffold that behaved well as an intravenously injectable human therapeutic. An initial hit of the C2 discoidin domain of factor V, another blood coagulation factor, stood out to us. Upon further review of the discoidin domains involved in coagulation, we settled on the C2 domain of factor VIII, a very similar protein that has been more extensively characterized [83,84,86].

Display of template discoidin scaffold, loop grafting, and binding to integrin

The C2 domain of fVIII contains hydrophobic residues on the most solvent-exposed tips of spikes 1 and 3. Previous studies [83] have shown that mutating these residues (Met-Phe of spike 1 and Leu-Leu of spike 3) to Ala-Ala on each spike to produce a four-point mutant (called 4-Ala) results in a 35-fold reduction in binding to PS membranes and a 91% reduction in specific activity in the activated partial thromboplastin time assay, one way of measuring fVIII activity. The 4-Ala mutant C2 domain expresses in yeast surface display [22] format, so we used this approach to determine binding to our target protein, αvβ3 integrin, as measured by flow cytometry. As shown in Fig. 2-2A, 4-Ala shows no inherent binding to αvβ3 integrin and was therefore chosen as our starting scaffold for engineering.

To assess whether our discobody scaffold could bind a new target, we grafted a number of loops from αvβ3 binding proteins into the spike 1 region of the 4-Ala template, replacing the FTNAAAT spike 1 sequence. These sequences were taken from engineered integrin binding peptides (Fig. 2-2B–G) as well as fibronectin (Fig. 2-2H), which naturally binds αvβ3 [43,70,71,116]. Of our seven loop grafts, four bound αvβ3 (Fig. 2-2C, E, G, and H). Based on

these data, we proceeded to build an RGD-containing degenerate codon library into spike 1 of 4-Ala.

Library generation and FACS selection

We mutated the three most solvent-exposed amino acids at the tip of spike 1 to RGD, without changing its length, to ensure that the RGD motif was available to guide binding. Because both the position of RGD in the loop and its flanking amino acids significantly affect binding affinity [70,115,127], we randomly mutated two residues on each side of the motif using NNS degenerate codons. The FTNAAAT sequence of spike 1 was replaced with XXRGDXX, where "X" represents the NNS codon. NNS encodes all twenty amino acids as well as a single stop codon [128]. The theoretical library size was 1.6×10^5, and we were able to obtain ~1000-fold coverage, as we attained a transformation efficiency of 10^8 into the EBY100 yeast display strain.

We expressed a library size of 10^9 and performed five sequential fluorescence-activated cell sorting (FACS) sorts with decreasing concentrations of αvβ3 integrin. Sorting against 200 nM for the first two sorts, 100 nM for the third and fourth sorts, and 10 nM for the final sort, we sorted the top 1-5% of each population for both antigen binding and surface display levels. We monitored each round by sequencing 10 clones each time. By round five, the XXRGDXX library converged upon a dominant clone with the loop sequence ACRGDTC (Fig. 2-3). This engineered variant (Eng-Db) was selected for further characterization and analysis.

Binding and characterization of engineered discobody

We measured binding kinetics of Eng-Db in yeast surface display format using soluble αvβ3 integrin. Prior studies comparing yeast display measurements against other measurement formats show a strong correlation between display and conventional assays [33]. We examined k_{on} by

incubating aliquots of the displayed discobody with 20, 50, and 100 nM αvβ3 integrin and quenching on ice at time points up to four hours. Quenching was performed by washing the sample in cold buffer and keeping it on ice until analysis. The k_{on} was fit to 1.3 ± 0.2 (10^3 $M^{-1}s^{-1}$) (Fig. 2-4A, Supplementary Fig. 2-S2). This slow association may be due to the conformational change that αvβ3 integrin undergoes when binding to RGD-containing partners [129-131]. We measured k_{off} by incubating yeast displaying Eng-Db with 200 nM antigen for four hours, washing aliquots at different time points, and leaving them with 1 uM soluble Eng-Db competitor at room temperature for intervals up to 26 hours to allow antigen dissociation. We calculated a k_{off} of 2.0 ± 0.1 (10^{-5} s^{-1}) and a $t_{1/2}$ of 9.6 hours (Fig. 2-4B, Supplementary Fig. 2-S3). Using this k_{off} value, we measured the k_{on} again using a single concentration of 200 nM αvβ3 integrin and obtained a k_{on} of 2.1 ± 0.4 (10^3 $M^{-1}s^{-1}$), which is similar to the value obtained using different antigen concentrations (Supplementary Fig. 2-S4). We also performed an equilibrium titration to determine K_d at equilibrium. Yeast displaying Eng-Db were incubated with varying concentrations of αvβ3 integrin (0–250 nM) for 24 hours. We measured a binding affinity of 16.2 ± 1.7 nM (Fig. 2-4C, Supplementary Fig. 2-S5). Binding measurements are summarized in Table I.

To confirm that the discobodies function outside of display format, we expressed soluble 4-Ala and Eng-Db in both *E. coli* and *S. cerevisiae* hosts. Thermal stability was assessed using the Thermofluor denaturation assay [132], which yielded T_ms of 55ºC and 61ºC for the 4-Ala and Eng-Db, respectively (Supplementary Fig. 2-S6). Both proteins show similar elution profiles via gel filtration and appear to be monomeric by comparison to FPLC standards (Supplementary Fig. 2-S7). Using these purified proteins, we performed competition assays to see if either protein could prevent binding of displayed Eng-Db to αvβ3 integrin. Soluble Eng-Db prevented binding

in a dose-dependent manner, but soluble 4-Ala did not, indicating that Eng-Db functions outside of display (Fig. 2-5A, Supplementary Fig. 2-S8).

Integrin αvβ3 naturally binds to the ECM protein fibronectin [116]. To see if our protein competes for the same binding site, we displayed fibronectin in yeast surface display and incubated it with αvβ3 integrin in the presence of soluble 4-Ala or Eng-Db. As before, Eng-Db blocked binding in a dose-dependent manner and 4-Ala did not, indicating that Eng-Db competes with the fibronectin binding site (Fig. 2-5B, Supplementary Fig. 2-S9).

X-ray crystal structure of engineered discobody

To gain further insight into Eng-Db, we obtained its crystal structure at 2.1 Å resolution (Fig. 2-6A, Supplementary Table I). The engineered variant looks very similar to previously solved wild-type fVIII C2 domain structures [87,127]. The C2 domain's membrane-binding spikes exhibit some conformational heterogeneity, as illustrated in a structural overlay (Fig. 2-6C). Eng-Db's loops fall within the observed range. The RGD tripeptide motif is positioned at the most solvent exposed tip, as we had planned for during library design. Note the presence of an additional disulfide bond within spike 1, which may account for the 6ºC increase in Tm (Fig. 2-6A).

Discussion

Discoidin domain as an alternative scaffold

The C2 discoidin domain of factor VIII has many attributes that make it a good alternative scaffold. Structurally, it is a β-barrel in the immunoglobulin superfamily that utilizes loops on one end for binding its cognate target (Fig. 2-1). The discoidin domain family binds many different targets in nature, and the β-barrel framework allows for diversity in the binding loops that in turn

allows for library construction [101]. Because the binding interface is made up of loops, similar to those of antibodies and engineered fibronectins [31,126], it is possible to transfer loops from preexisting binders onto the discobody scaffold, as has been demonstrated in our loop graft experiments (Fig. 2-2).

Discobodies begin fully human, but loop mutations still need to be chosen with care to avoid introducing immunogenic sequences [106]. Certain hemophiliac patients who receive fVIII injections develop ADAs [89], but it seems unlikely that this would happen in an individual with endogenous fVIII due to negative selection during immune system development (personal communication, Gary Gilbert).

αvβ3 integrin and RGD-containing peptides

To minimize the number of variables while validating a new scaffold, we chose the well understood system of αvβ3 integrin and the RGD motif. Additionally, αvβ3 integrin is over-expressed in a number of cancers and is a drug target currently being pursued by the pharmaceutical industry [102].

Our structure allows us to examine the presentation of the RGD tripeptide motif within Eng-Db. The distance between Cβ atoms of Arg and Asp has been found to influence binding to αvβ3 integrin [115]. In that study, cyclic peptides with $Arg_{C\beta}$-$Asp_{C\beta}$ distances of 6.7 Å or less were optimal, with the lowest IC_{50} values (~50 nM) in an integrin-ECM competition assay. Our measured Cβ-Cβ distance is 8.3 Å (Fig. 2-6B) but still provides a K_d in the low nanomolar range. While the assay systems and measurements are different, it is interesting to think about our results in the context of these studies.

αvβ3 bound to a cyclic Arg-Gly-Asp-{D-Phe}- {*N-methyl*-Val-} pentapeptide has been crystalized [130] (PDB ID: 1L5G). The cyclic pentapeptide from this structure can be overlaid with spike 1 of Eng-Db (Supplementary Fig. 2-S10A), and the backbone RMSD of the respective Arg-Gly-Asp sequences is only 0.3 Å (Supplementary Fig. 2-S10B), demonstrating that the two binding sites may be very similar. In Eng-Db, the Arg and Asp sidechains are directed away from each other, much like in the target-bound peptide. Although the side chains differ slightly starting from their respective Cβs (Supplementary Fig. 2-S10A), their conformation may change upon binding to αvβ3 integrin.

Future work for a universal discobody scaffold

To truly be a universal alternative scaffold, discobodies will need to be engineered to bind a variety of targets. Future work will include building general libraries that are not RGD-focused. Our structure indicates that the C2 domain of fVIII can tolerate extensive mutations in its spikes and still retain its native fold (Fig. 2-6C). Additional work will be required to understand which residues can be engineered [38], if the spikes can tolerate length diversity [32,34], what range of antigens discobodies can bind [35,36], and what the affinity limits of this scaffold are [34]. Realizing the potential of bispecific discobodies and discobody-mAb fusions will require further study. Although the T_m of our template discobody (4-Ala) is 55ºC, it may be advantageous to engineer a more thermostable variant [133]; this would allow for greater diversity in library generation given that mutations often decrease thermal stability [34]. These efforts, however, will have to avoid introducing mutations that induce immunogenicity, which can arise from a single, erroneously placed point mutation [106].

Conclusions

Alternative scaffolds have found an important niche as complements to mAbs and are also being used to enhance binding as mAb fusions. Although many alternative scaffolds have entered the clinic, each has its strengths and weaknesses, requiring careful examination of the desired goals prior to scaffold selection. We saw the need for a new alternative scaffold that is tailored from the outset for use as a human therapeutic with bispecific capabilities. Discobodies are derived from a human protein and have a useful geometry with adjacent N- and C-termini, allowing for potential bispecific fusions with minimal linkers. Further work is needed to explore the potential of the discobody scaffold in protein engineering and drug development.

Figures and Tables

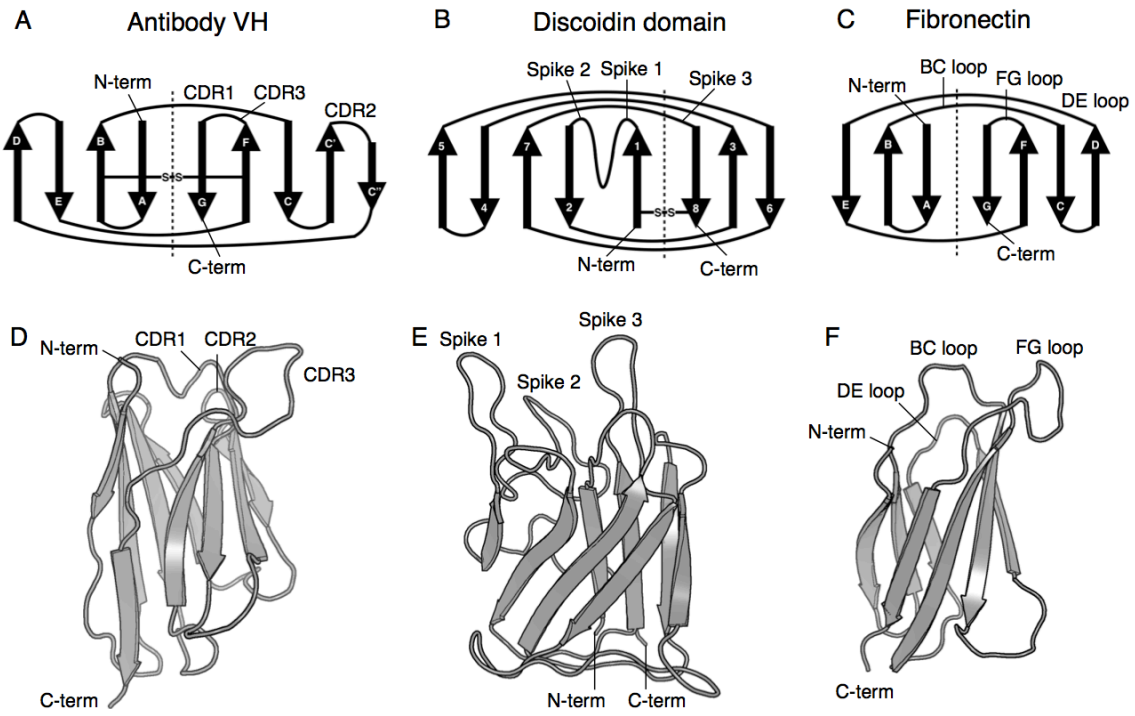


Fig. 2-1. Topology diagrams comparing (A) the antibody immunoglobulin fold, (B) the discoidin domain, and (C) fibronectin. Structures of (D) the antibody VH domain of trastuzumab, (E) the C2 discoidin domain of factor VIII, and (F) the tenth repeat of fibronectin III. These correspond to PDB ID's (D) 1N8Z, (E) 2R7E, and (F) 1FNA. N- and C-termini as well as binding loops are annotated.

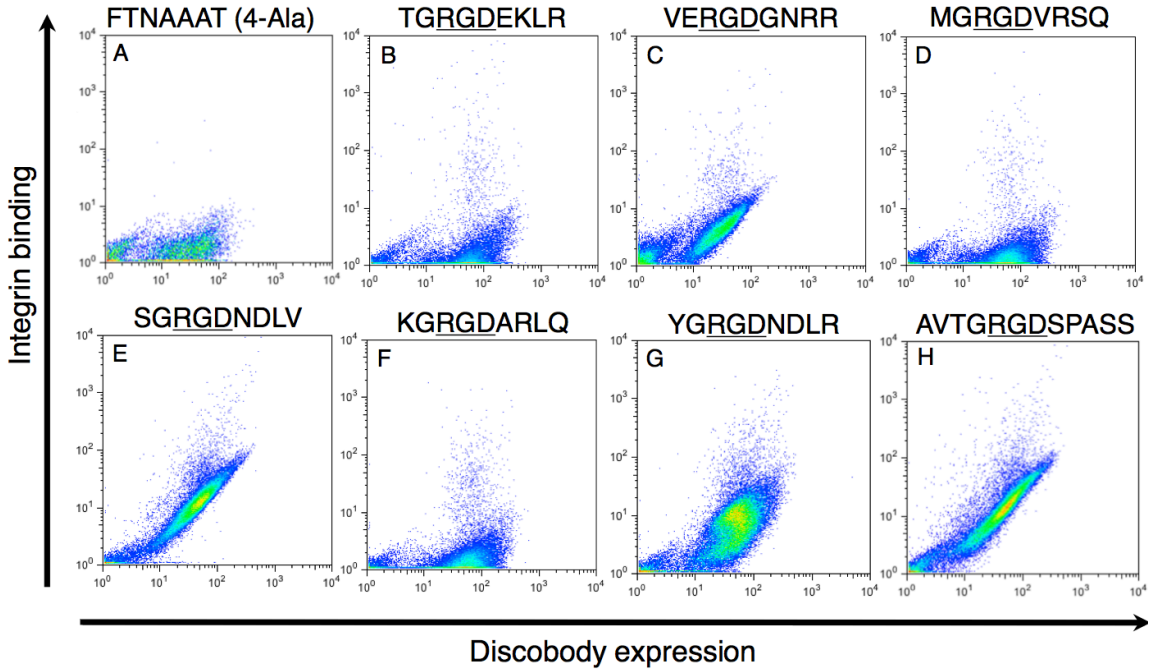


Fig. 2-2. Flow cytometry plots showing binding of 100 nM integrin to (A) 4-Ala and various loop grafts from known integrin-binding proteins, including (B-G) engineered knottins and (H) wild-type fibronectin. C, E, G, and H demonstrated binding to integrin.

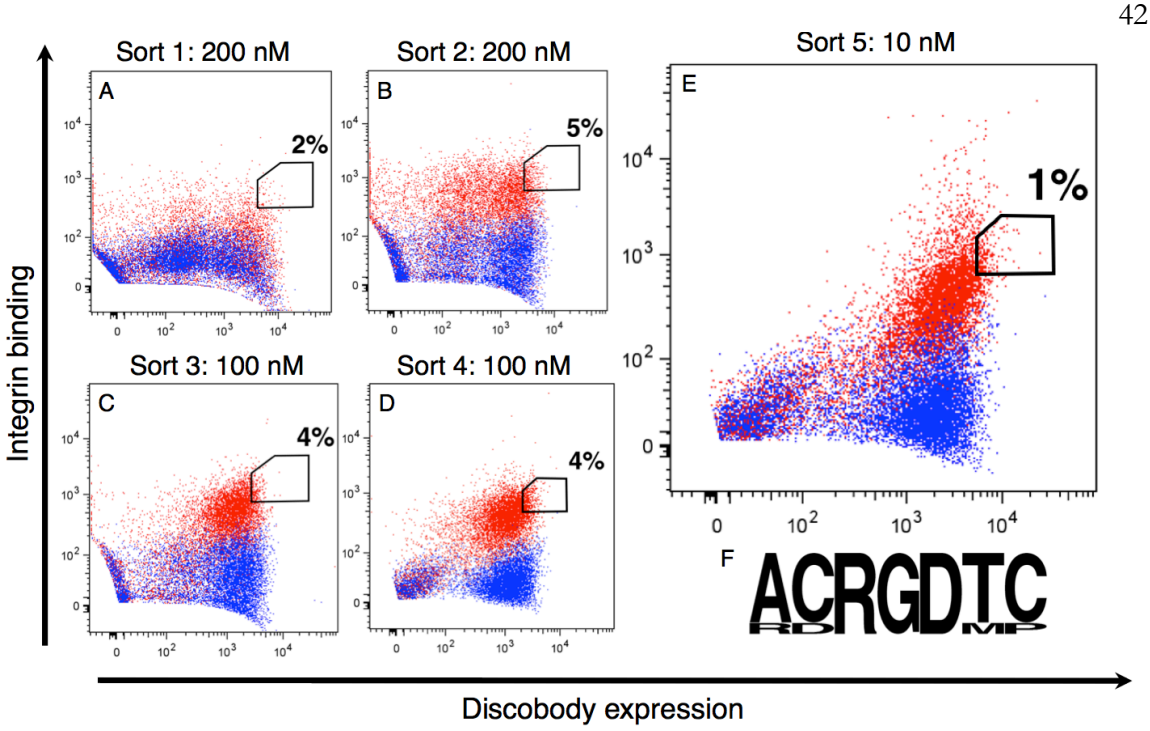


Fig. 2-3. FACS plots for sequential sorts (A-E) of the library against decreasing concentrations of integrin. Red is the library with integrin, blue is without integrin and shows the level of background binding to secondary reagents. Sort gates and percentages of the library sorted are shown. 7/10 sequenced clones after sort five showed the (F) loop sequence ACRGDTC.

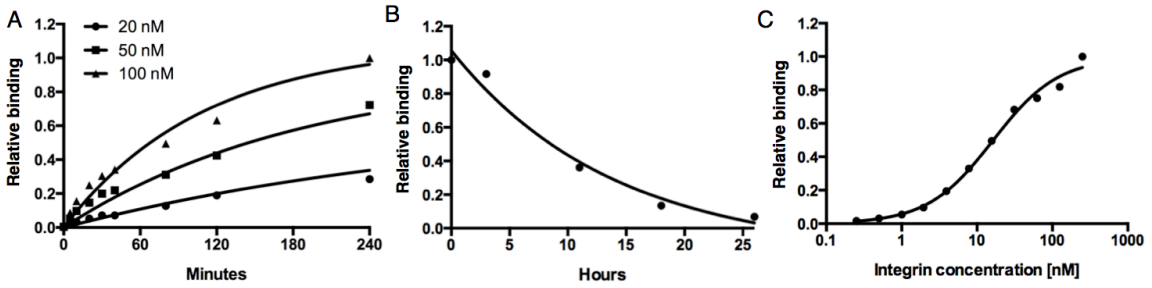


Fig. 2-4. Binding characterization of Eng-Db. (A) On rate measured with different concentrations of integrin. (B) Off rate measured with 200 nM integrin. (C) Equilibrium titration of displayed Eng-Db against integrin.

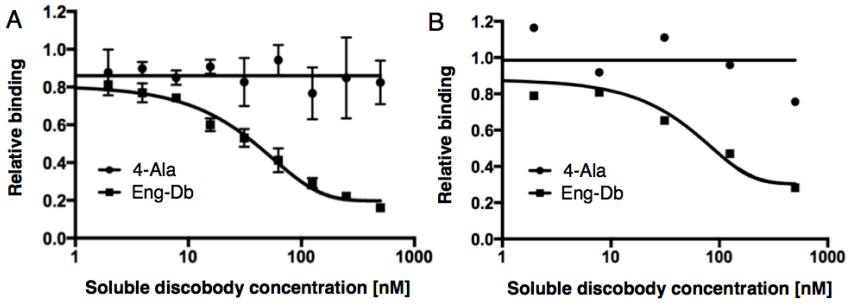


Fig. 2-5. (A) Binding of displayed Eng-Db to 125 nM integrin with competing soluble 4-Ala or Eng-Db, performed in duplicate. (B) Binding of displayed fibronectin to 125 nM integrin with competing soluble 4-Ala or Eng-Db.

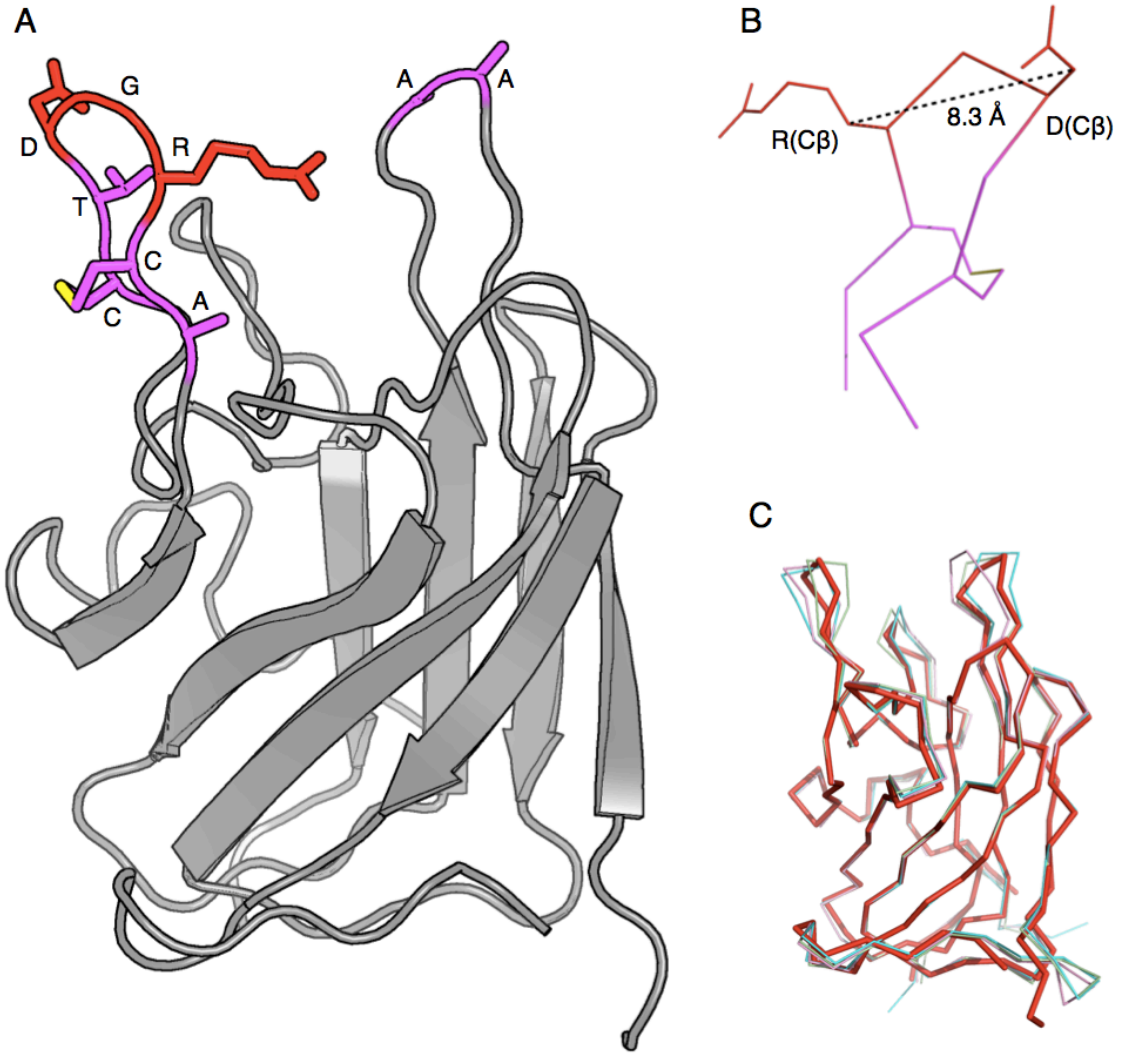
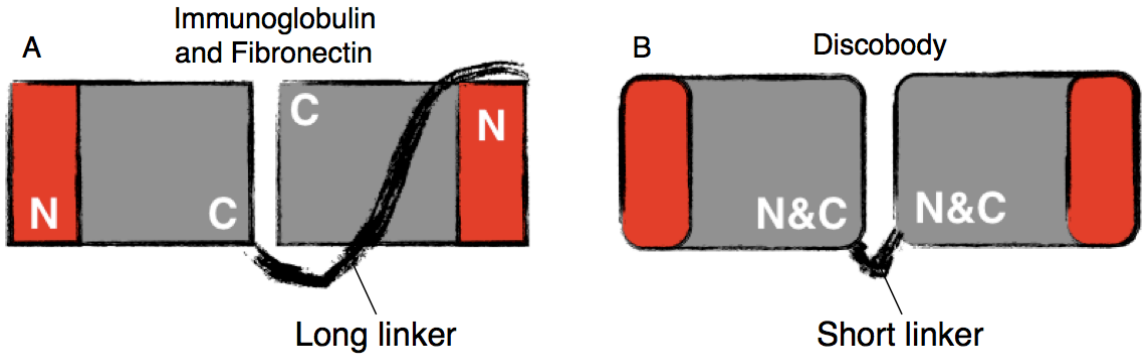
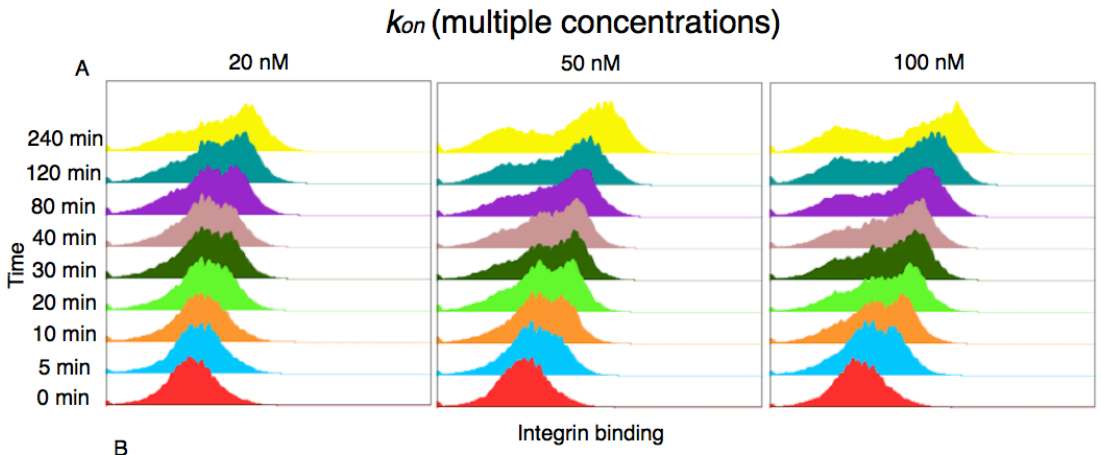


Fig. 2-6. (A) Cartoon representation of the Eng-Db structure with stick side chains in regions of interest; residues differing from wild type are shown in magenta, RGD is shown in red, the introduced disulfide bond in yellow, and mutations are labeled. (B) Line representation of spike 1 with the Cβ-Cβ distance between Arg and Asp indicated with a dotted line. (C) Ribbon overlay of the Eng-Db structure in red compared to wild-type structures of the C2 discoidin domain from fVIII: 1D7P in cyan, 3HNY in light green, and 3HOB in light pink; the Eng-Db ribbon is thicker for improved visualization.



Supplementary Fig. 2-S1. Orientations of N- and C-termini of discobodies allow for shorter linkers. Bispecific domain schematics of (A) the mAb single-chain variable fragment and fibronectin compared to (B) the discobody. Binding interfaces are shown in red.



B

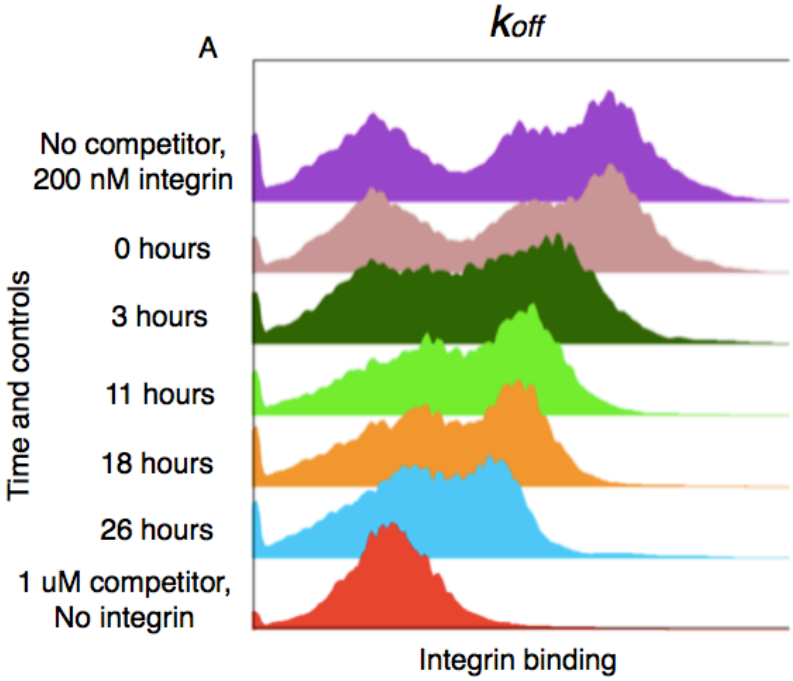
Minutes	20 nM	50 nM	100 nM
0.0	0.000000	0.004065045	0.01626018
5.0	0.0230353	0.05284559	0.08672093
10.0	0.03387534	0.09756096	0.1558266
20.0	0.05284559	0.1463415	0.2493225
30.0	0.07181571	0.199187	0.304878
40.0	0.07181571	0.2195122	0.3414634
80.0	0.1273713	0.3116531	0.4945799
120.0	0.1897019	0.4241193	0.6314363
240.0	0.2859079	0.7222223	1.000000

C

	Nonlin fit	A	B	C	D
		20 nM	50 nM	100 nM	Global (shared)
		Y	Y	Y	Y
1	Association kinetics - Two or more conc. of hc				
2	Best-fit values				
3	Koff	0.001496	0.001496	0.001496	0.001496
4	Kon	76218	76218	76218	76218
5	Hotnm	= 20.00	= 50.00	= 100.0	
6	Bmax	1.296	1.296	1.296	1.296
7	Kd	1.963e-008	1.963e-008	1.963e-008	1.963e-008
8	Std. Error				
9	Koff	0.001181	0.001181	0.001181	0.001181
10	Kon	13037	13037	13037	13037
11	Bmax	0.2127	0.2127	0.2127	0.2127
12	Kd	1.767e-008	1.767e-008	1.767e-008	1.767e-008
13	95% Confidence Intervals				
14	Koff	-0.0009408 to 0.003933	-0.0009408 to 0.003933	-0.0009408 to 0.003933	-0.0009408 to 0.003933
15	Kon	49311 to 103124	49311 to 103124	49311 to 103124	49311 to 103124
16	Bmax	0.8572 to 1.735	0.8572 to 1.735	0.8572 to 1.735	0.8572 to 1.735
17	Kd	-1.684e-008 to 5.610e-008	-1.684e-008 to 5.610e-008	-1.684e-008 to 5.610e-008	-1.684e-008 to 5.610e-008
18	Goodness of Fit				
19	Degrees of Freedom				24
20	R square	0.9446	0.9628	0.9651	0.9710
21	Absolute Sum of Squares	0.003724	0.01464	0.02625	0.04461
22	Sy.x				0.04311
23	Constraints				
24	Koff	Koff is shared	Koff is shared	Koff is shared	
25	Kon	Kon is shared	Kon is shared	Kon is shared	
26	Hotnm	Hotnm = 20.00	Hotnm = 50.00	Hotnm = 100.0	
27	Bmax	Bmax is shared	Bmax is shared	Bmax is shared	
28					
29	Number of points				
30	Analyzed	9	9	9	

Supplementary Fig. 2-S2. (A) Flow cytometry diagrams of displayed Eng-Db binding to 20, 50, or 100 nM integrin over time. (B) Mean fluorescence intensity values of each population normalized between 0-1 for the lowest and highest signal. (C) Prism calculations for k_{on} with two or more concentrations of hot ligand.

A



B

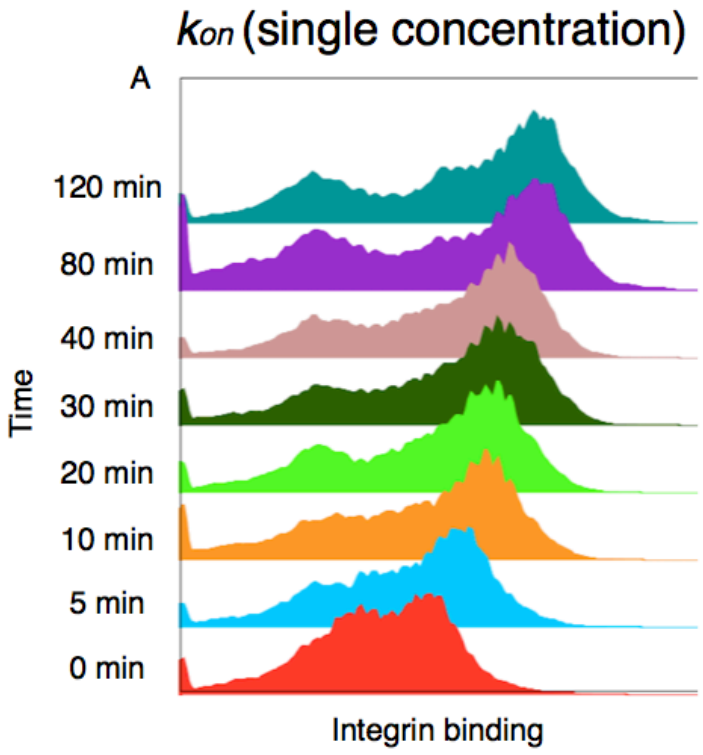
Hours	Binding
0.0	1.000000
3.0	0.916037
11.0	0.3618808
18.0	0.1343409
26.0	0.06884971

C

	Nonlin fit	A
		Control
		Y
1	Dissociation - One phase exponential decay	
2	Best-fit values	
3	Y0	1.056
4	NS	-0.1539
5	K	0.07192
6	HalfLife	9.637
7	Std. Error	
8	Y0	0.08136
9	NS	0.2733
10	K	0.03593
11	95% Confidence Intervals	
12	Y0	0.7056 to 1.406
13	NS	-1.330 to 1.022
14	K	0.0 to 0.2265
15	HalfLife	3.060 to +infinity
16	Goodness of Fit	
17	Degrees of Freedom	2
18	R square	0.9785
19	Absolute Sum of Squares	0.01639
20	Sy.x	0.09053
21	Constraints	
22	K	K > 0.0
23		
24	Number of points	
25	Analyzed	5

Supplementary Fig. 2-S3. (A) Flow cytometry diagrams of displayed Eng-Db dissociating from 200 nM integrin in the presence of excess soluble Eng-Db competitor over time. (B) Mean fluorescence intensity values of each population normalized to 1 for the highest signal. (C) Prism calculations for k_{off} with one phase exponential decay dissociation kinetics.

A



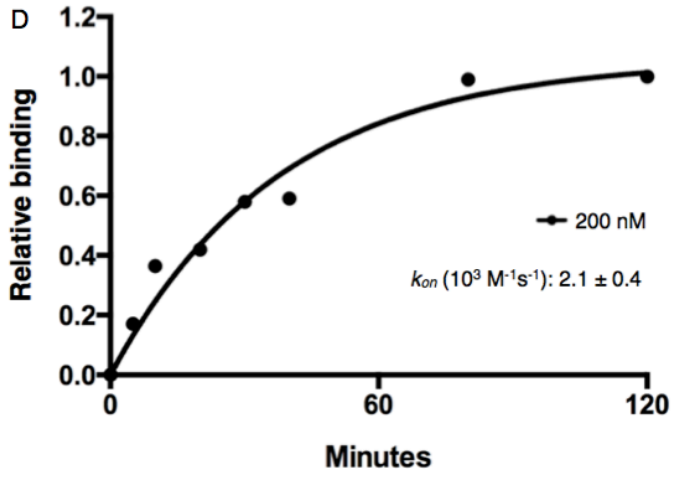
B

Minutes	200 nM
0.0	0.000000
5.0	0.1702586
10.0	0.3642241
20.0	0.419181
30.0	0.5797414
40.0	0.5905172
80.0	0.9892241
120.0	1.000000

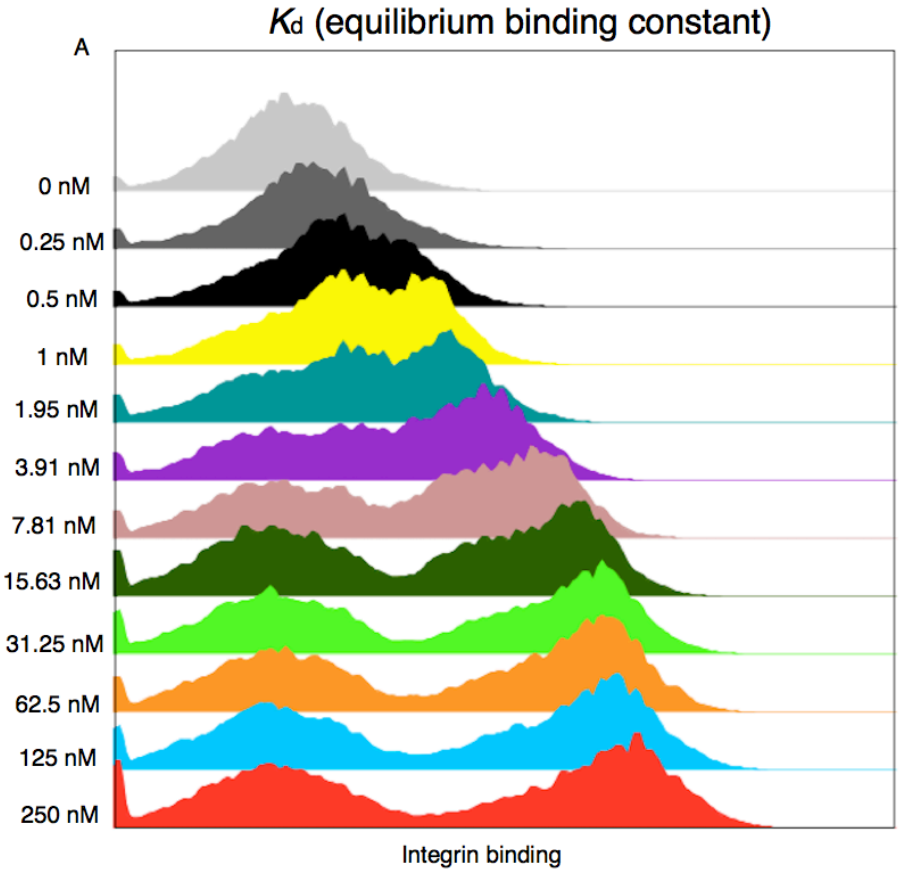
C

	Nonlin fit	A
		200 nM
		Y
1	Association kinetics - One conc. of hot	
2	Best-fit values	
3	Koff	= 0.001199
4	Kon	126710
5	Hotnm	= 200.0
6	Bmax	1.106
7	Kd	= 9.463e-009
8	Std. Error	
9	Kon	23605
10	Bmax	0.09315
11	95% Confidence Intervals	
12	Kon	68950 to 184469
13	Bmax	0.8784 to 1.334
14	Goodness of Fit	
15	Degrees of Freedom	6
16	R square	0.9668
17	Absolute Sum of Squares	0.02945
18	Sy.x	0.07006
19	Constraints	
20	Koff	Koff = 0.001199
21	Hotnm	Hotnm = 200.0
22		
23	Number of points	
24	Analyzed	8

D



Supplementary Fig. 2-S4. (A) Flow cytometry diagrams of displayed Eng-Db binding to 200 nM integrin over time. (B) Mean fluorescence intensity values of each population normalized between 0-1 for the lowest and highest signal. (C) Prism calculations for k_{on} with single ligand concentration association kinetics, using k_{off} values derived from Figure 4B. (D) Curve fit from calculations



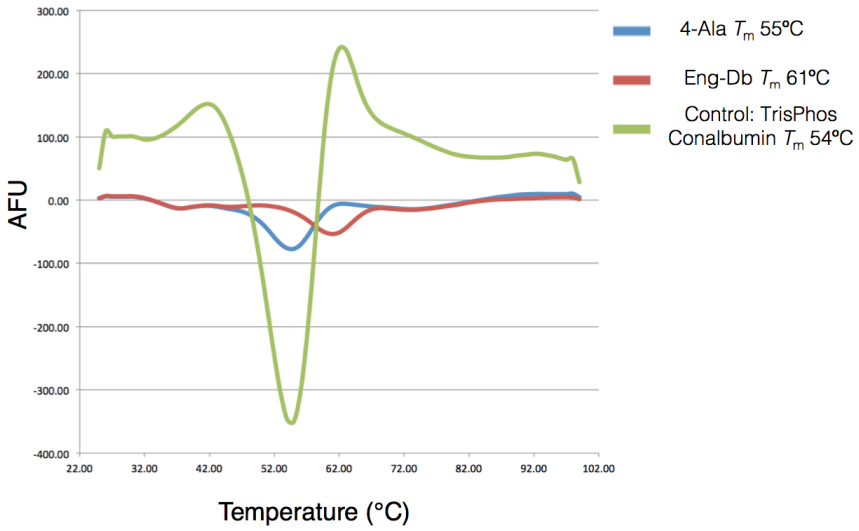
B

Integrin [nM]	Binding
0.000000	0.000000
0.250000	0.0164306
0.500000	0.03172804
1.000000	0.05439093
1.953125	0.09688386
3.906250	0.1943343
7.812500	0.3297451
15.625000	0.4957507
31.250000	0.6827195
62.500000	0.7507082
125.000000	0.8186969
250.000000	1.000000

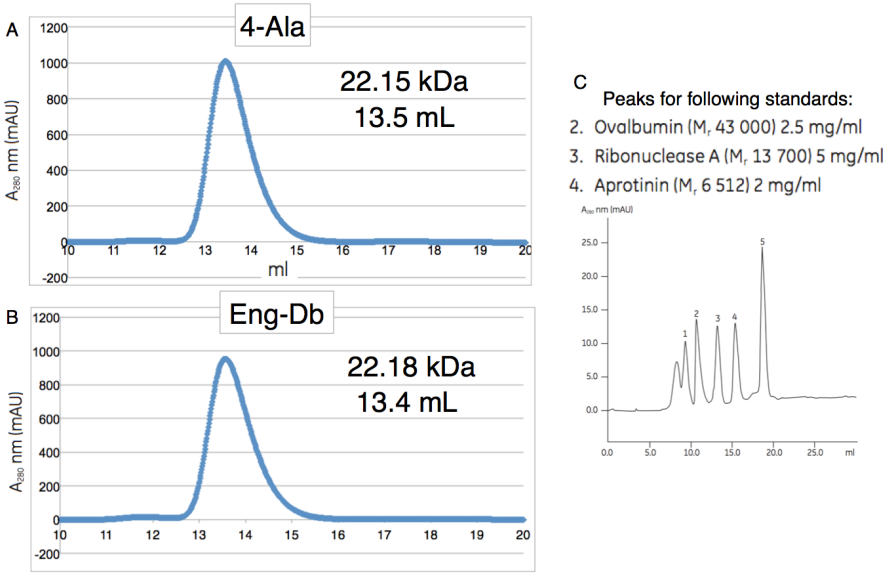
C

	Nonlin fit	A
		Data Set-A
		Y
1	Michaelis-Menten	
2	Best-fit values	
3	Vmax	0.9947
4	Km	16.23
5	Std. Error	
6	Vmax	0.02860
7	Km	1.688
8	95% Confidence Intervals	
9	Vmax	0.9309 to 1.058
10	Km	12.47 to 19.99
11	Goodness of Fit	
12	Degrees of Freedom	10
13	R square	0.9926
14	Absolute Sum of Squares	0.01068
15	Sy.x	0.03268
16	Constraints	
17	Km	Km > 0.0
18		
19	Number of points	
20	Analyzed	12

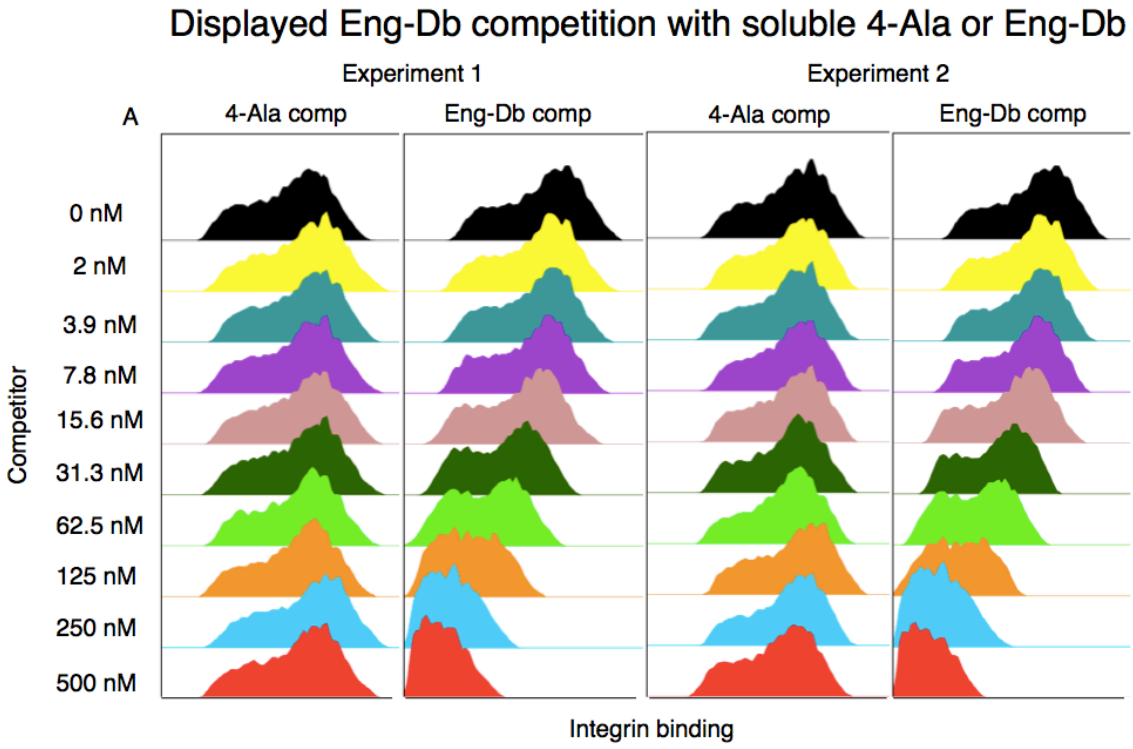
Supplementary Fig. 2-S5. (A) Flow cytometry diagrams of displayed Eng-Db binding to 0-250 nM integrin at equilibrium. (B) Mean fluorescence intensity values of each population normalized from 0-1 for 0-250 nM. (C) Prism calculations for K_d fit to the Michaelis-Menten model.



Supplementary Fig. 2-S6. Melting temperature (T_m) as assessed by the Thermofluor denaturation assay shows that Eng-Db is 6ºC more thermostable than 4-Ala.



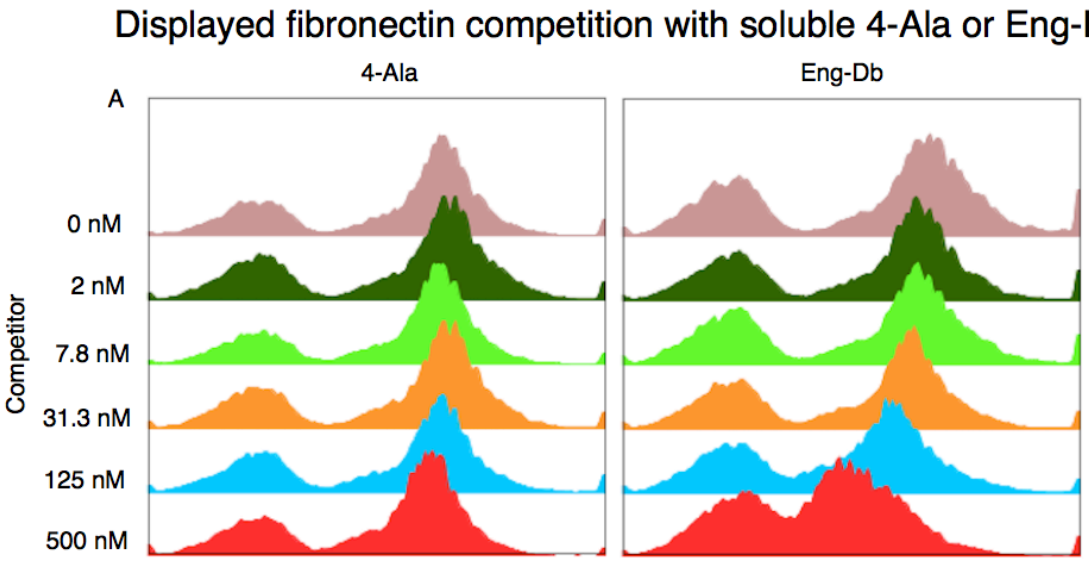
Supplementary Fig. 2-S7. Gel filtration profiles of soluble (A) 4-Ala and (B) Eng-Db. (C) Standards from the AKTA manual as comparison.



B

Soluble Db [nM]	4-Ala (1)	Eng-Db (1)	4-Ala (2)	Eng-Db (2)
0.0000	0.7059322	0.7067797	1.000000	0.9037383
1.9531	0.7915254	0.820339	0.9626168	0.8056075
3.9062	0.8728814	0.7338983	0.9224299	0.8056075
7.8125	0.8220423	0.7440678	0.8766355	0.7401869
15.6250	0.8813559	0.5762712	0.9345794	0.624299
31.2500	0.7364407	0.5644068	0.9158878	0.4971963
62.5000	1.000000	0.4567797	0.8859813	0.3682243
125.0000	0.6694915	0.3101695	0.8635514	0.2719626
250.0000	0.6974576	0.2364407	1.000000	0.2056075
500.0000	0.7432203	0.1508475	0.9065421	0.171028

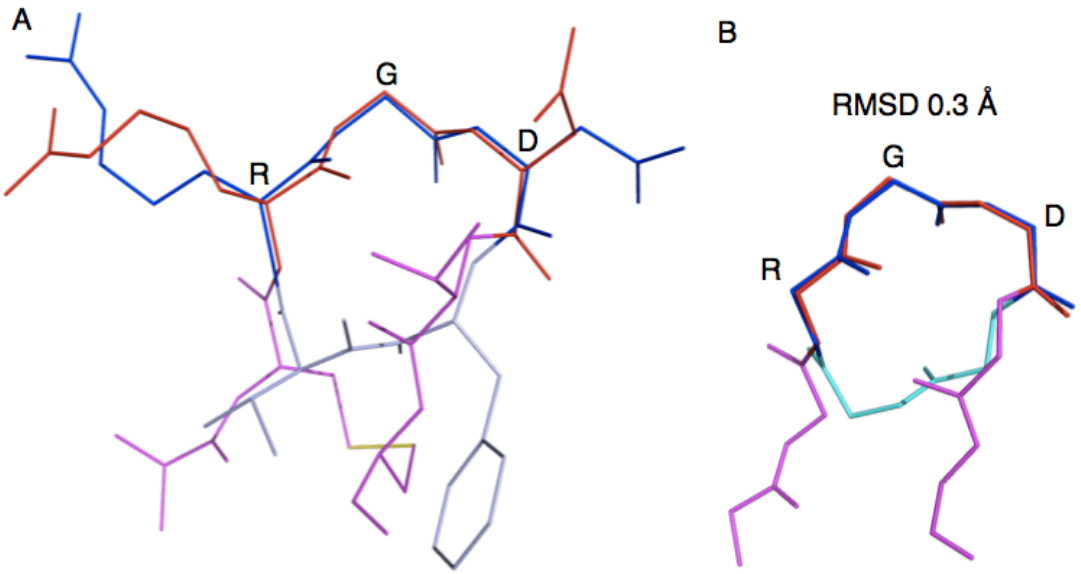
Supplementary Fig. 2-S8. (A) Flow cytometry diagrams of displayed Eng-Db binding to 125 nM integrin in the presence of soluble competing 4-Ala or Eng-Db, performed in duplicate. (B) Mean fluorescence intensity values of each population normalized to 1 for the highest signal.



B

Soluble Db [nM]	4-Ala	Engineered
0.000000	1.000000	1.000000
1.953125	1.164865	0.7891683
7.812500	0.9189189	0.8085107
31.250000	1.110811	0.6537718
125.000000	0.9594595	0.4700193
500.000000	0.7567568	0.2823985

Supplementary Fig. 2-S9. (A) Flow cytometry diagrams of displayed fibronectin binding to 125 nM integrin in the presence of soluble competing 4-Ala or Eng-Db. (B) Mean fluorescence intensity values of each population normalized to 1 for the highest signal.



Supplementary Fig. 2-S10. (A) Line representations of spike 1 of Eng-Db (magenta) overlaid with that of the Arg-Gly-Asp-{D-Phe}- {*N-methyl*-Val-} integrin-binding cylic pentapeptide from the 1L5G αvβ3-RGD-Mn structure (cyan). RGD from spike 1 of Eng-Db is shown in red and the disulfide is in yellow; RGD from the pentapeptide is shown in blue. (B) Line representation of RGD backbones of spike 1 of Eng-Db (magenta) and pentapeptide (cyan); RMSD value for the RGD backbone only is 0.3 Å.

Table 2-I. Eng-Db binding and thermal stability characterization

K_d (nM)	k_{on} (10^3 M^{-1} s^{-1})	k_{off} (10^{-5} s^{-1})	$t_{1/2}$ (hours)	T_m (°C)
16.2 ± 1.7	1.3 ± 0.2	2.0 ± 0.1	9.6	66

Supplementary Table 2-I. Data Collection and Refinement Statistics

Data collection statistics	
PDB ID	4PT6
Space group	P1
Unit cell	
a, b, c (Å)	32.45, 38.45, 48.84
α, β, g (°)	98.56, 91.77, 111.03
Wavelength (Å)	0.9795
Resolution (Å)	31.15 - 2.10
No. of unique reflections	11234
Mean (I/σ_I)	15.4 (7.3)
Completeness	88.6 (90.3)
Average multiplicity	3.8 (3.8)
R_{merge} (%)	8.22 (14.9)
Structure refinement statistics	
Resolution	31.15 - 2.10
Average B factor ($Å^2$)	45.9
R_{work}/R_{free} (%)	19.88/25.20
RMS deviations	
Bond length (Å)	0.006
Bond angle (°)	1.2
Ramachandran	
Favored (%)	97.7
Outliers (%)	0

Rotamer outliers (%)	1.1
MolProbity clashscore	9.27
MolProbity overall score	1.61

A p p e n d i x A

DETAILED DISCOBODY DATA

Here we present extensive data related to the Discobody engineering project that supports the main thesis body. All methods were detailed in the preceding chapters, and will not be covered here. Due to the supplementary nature of the section, most information will be presented in figure format with corresponding legends.

Figures and Tables

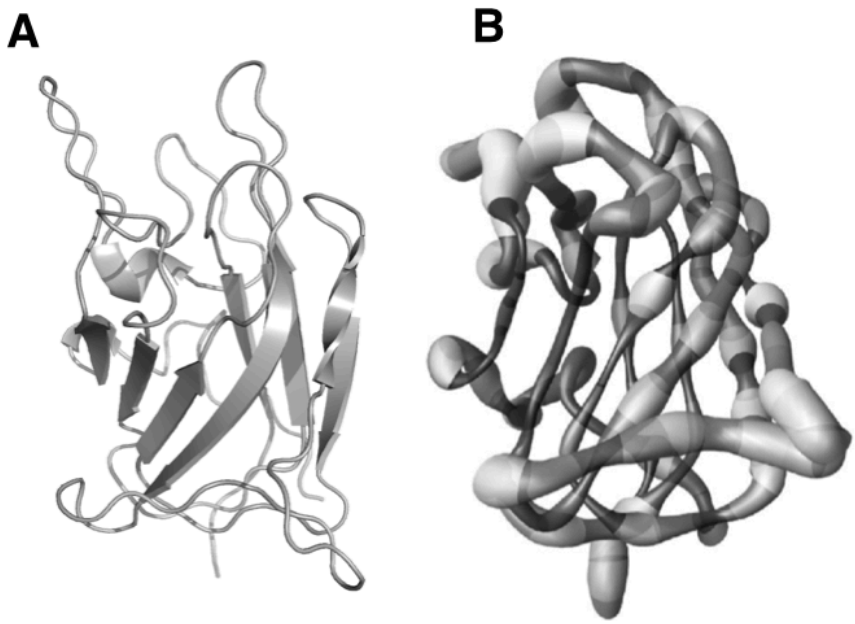


Fig. A-1. Discoidin domain structural information. (A) The C2 domain of factor V (PDB ID: 1CZT) was the first resulting hit from our PDB database search. (B) A structural overlay of discoidin domain family members from Kiedzierska, et al, illustrates the sequence diversity within the discoidin family, with diverse regions indicated as thick and conserved regions as thin. Loops on top are attractive for including sequence diversity, such as libraries.

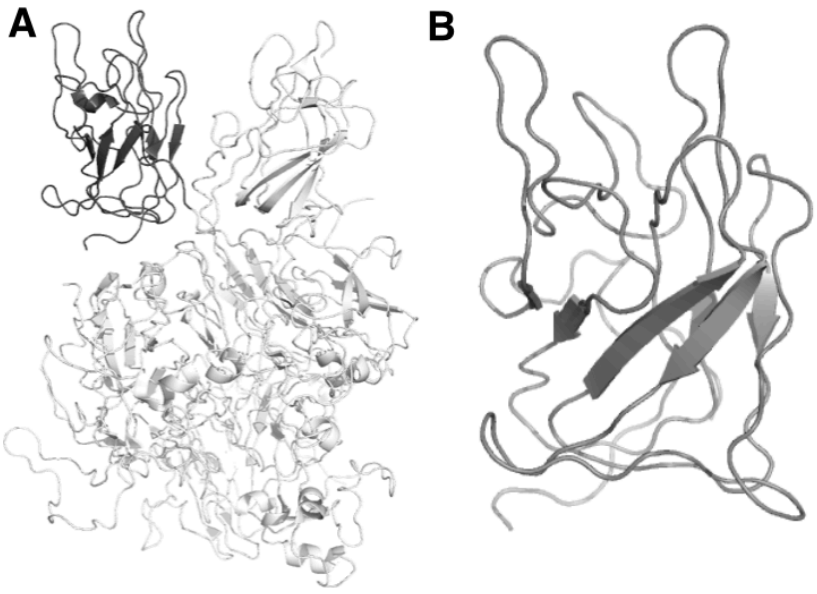


Fig. A-2. Factor VIII. (A) The complete fVIII structure (PDB ID 2R7E) with the C2 domain highlighted in dark grey. (B) The isolated C2 domain from the same structure.

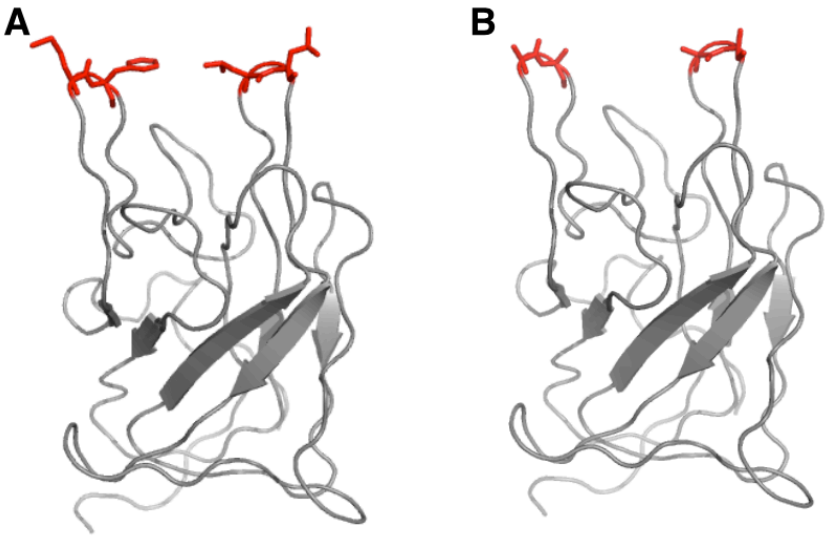


Fig. A-3. Hydrophobic residues mediate lipid binding. (A) The wild type fVIII C2 discoidin domain has Met-Phe and Leu-Leu, from left to right. (B) Our starting 4-Ala mutant has these four residues mutated to Ala-Ala and Ala-Ala. Mutations performed in Pymol.

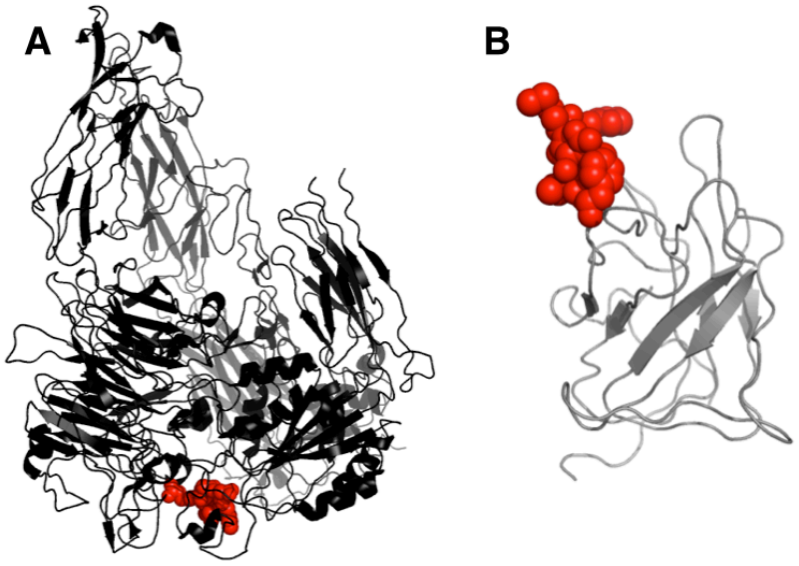


Fig. A-4. αvβ3 integrin and the RGD motif. (A) αvβ3 integrin in black bound to an RGD peptide in red. (B) The discobody with the position of an inserted RGD noted in red.

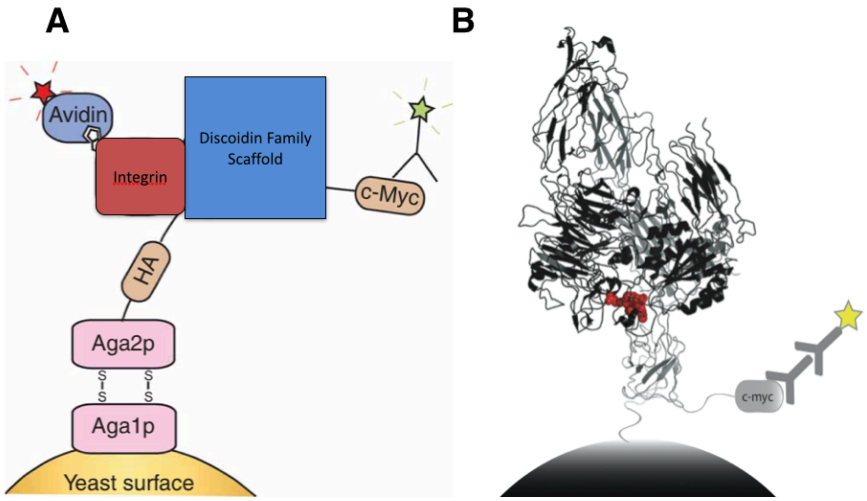


Fig. A-5. Yeast surface display schematic. (A) Diagram of yeast surface display constructs with biotinylated integrin, adapted from Chao, et al. Antigen can either be biotinylated or stained with a specific antibody. (B) The layout of the discobody displayed on the yeast surface with bound integrin.

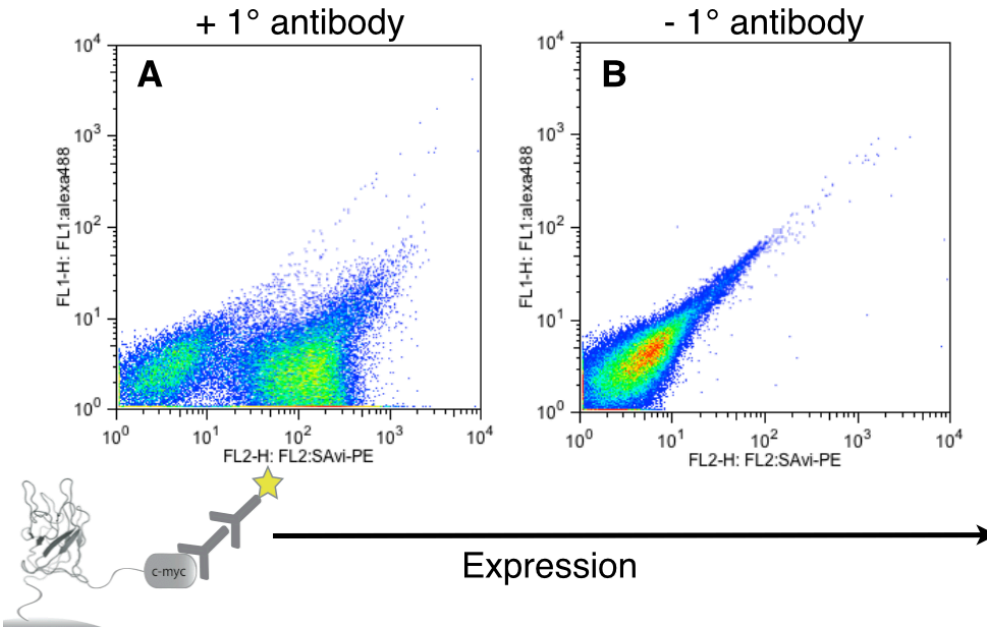


Fig. A-6. Flow cytometry plots demonstrate surface expression of 4-Ala discobody. (A) Display constructs with primary anti-cmyc antibody and secondary with fluorophore. (B) Secondary antibody only.

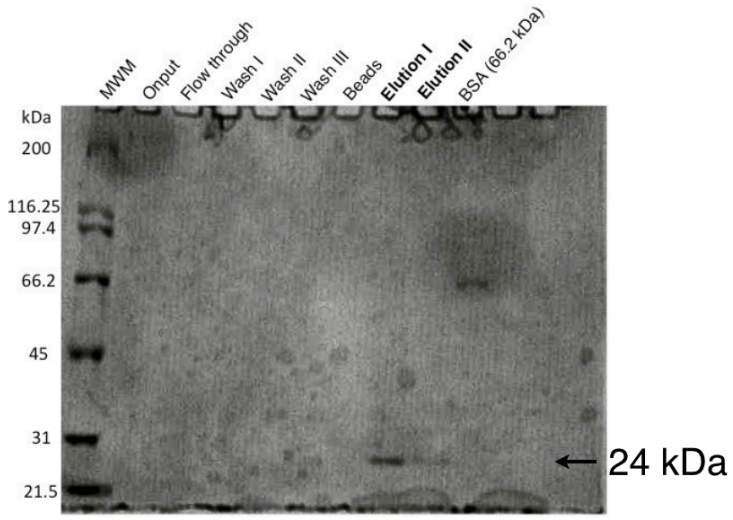


Fig. A-7. Coomassie stain of soluble expression and purification from yeast of 4-Ala demonstrates that the starting template can be expressed outside of display format.

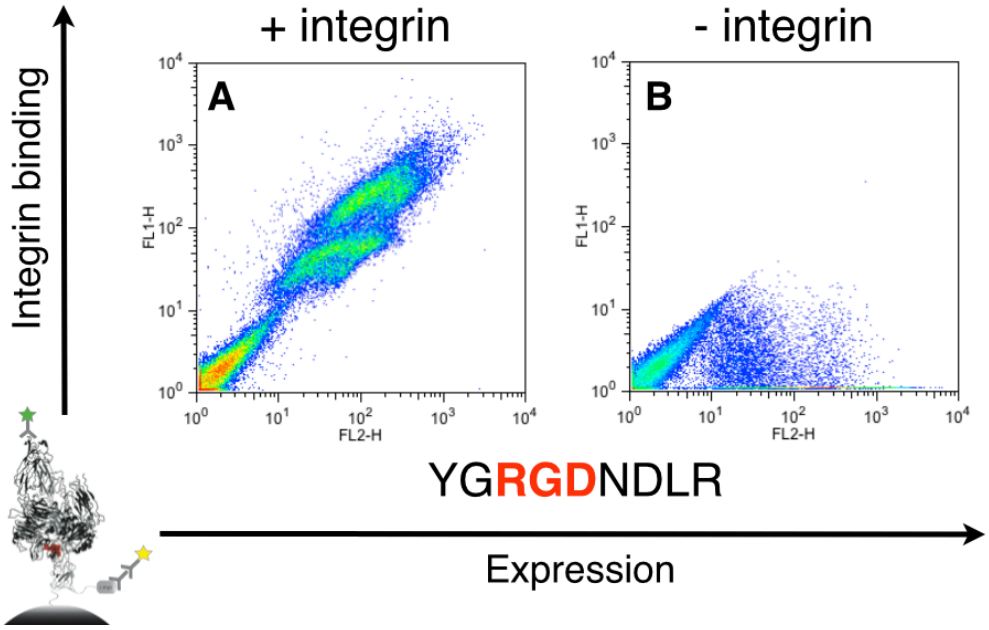


Fig. A-8. Flow cytometry plots of integrin binding AgRP control peptides from Silverman, et al, in display format. (A) With antigen, and (B) without antigen. Sequence of the binding loop is shown with RGD highlighted in red.

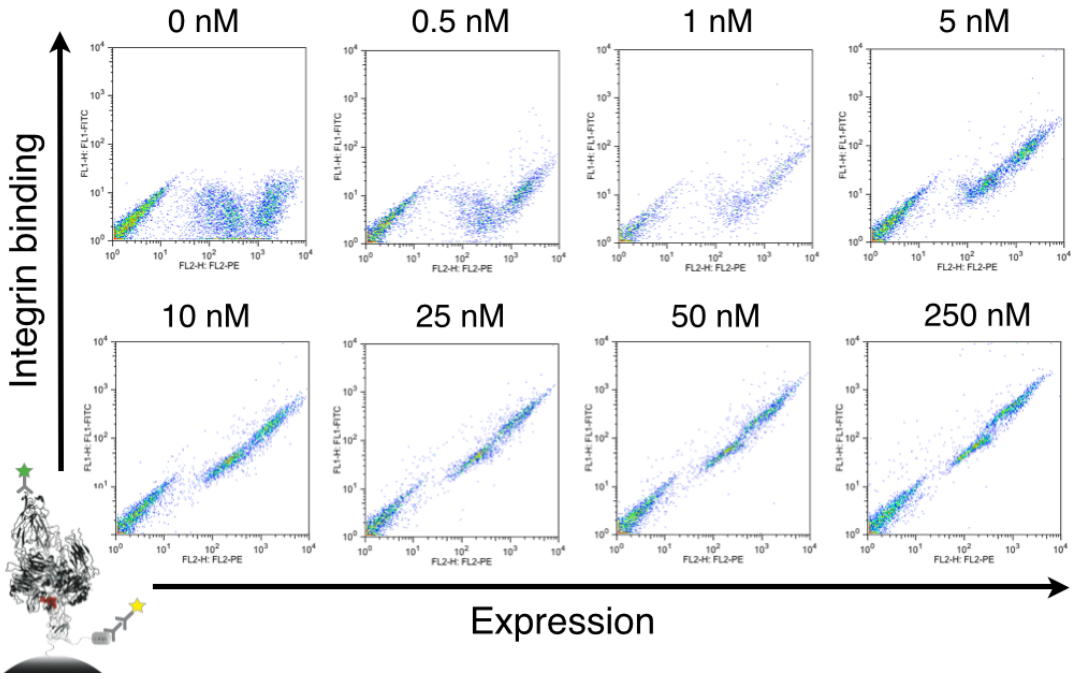


Fig. A-9. Affinity titration of control AgRP peptide against αvβ3 integrin, data are flow cytometry plots for displayed peptide. Incubation was performed for 1 hour at room temperature.

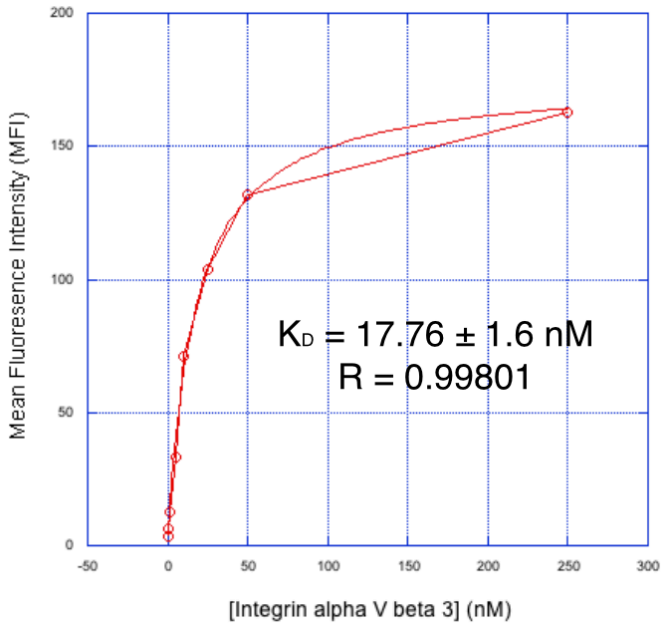


Fig. A-10. Curve fit for AgRP and αvβ3 integrin based on data from Figure A-9.

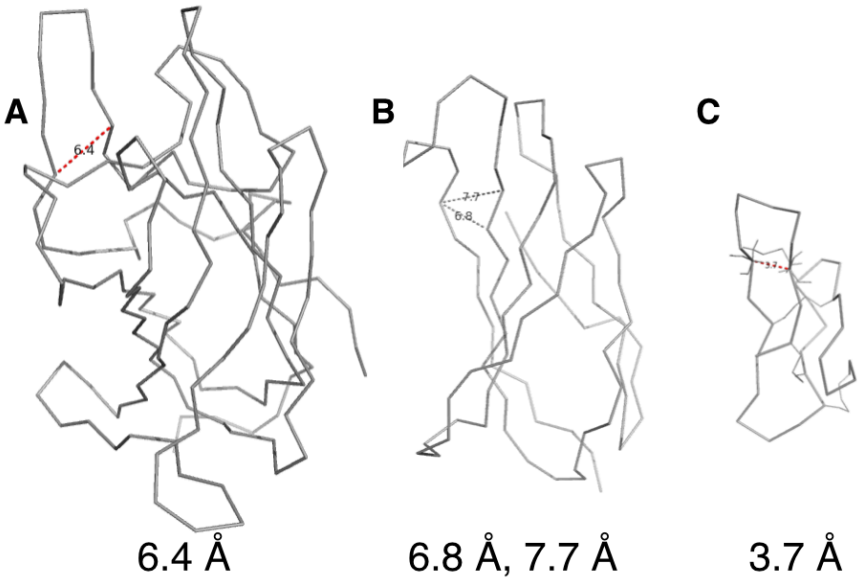


Fig. A-11. Ribbon diagrams measuring backbone distances for loop grafting from onto (A) the discoidin domain. Loops containing RGD from (B) Fibronectin and (C) AgRP peptide were grafted onto the 4-Ala scaffold. Measured distances indicated below in angstroms. PDB IDs (A) 2R7E, (B) 1FNA, and (C) 1HYK.

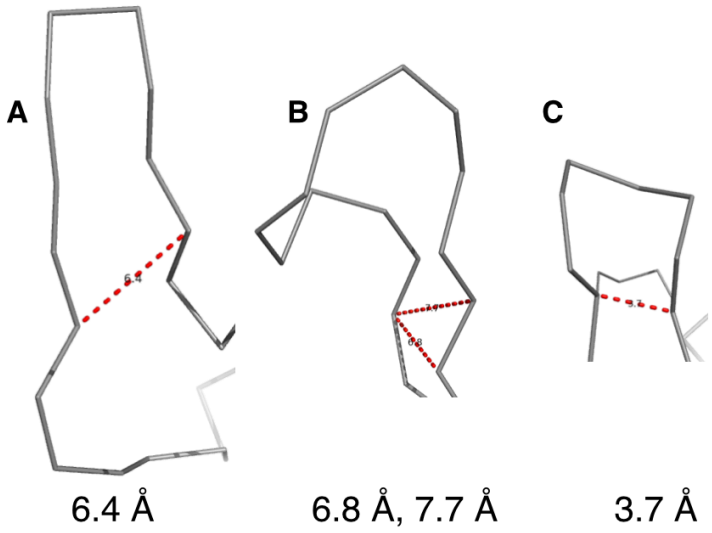


Fig. A-12. Close-up ribbon diagrams from Figure A-11.

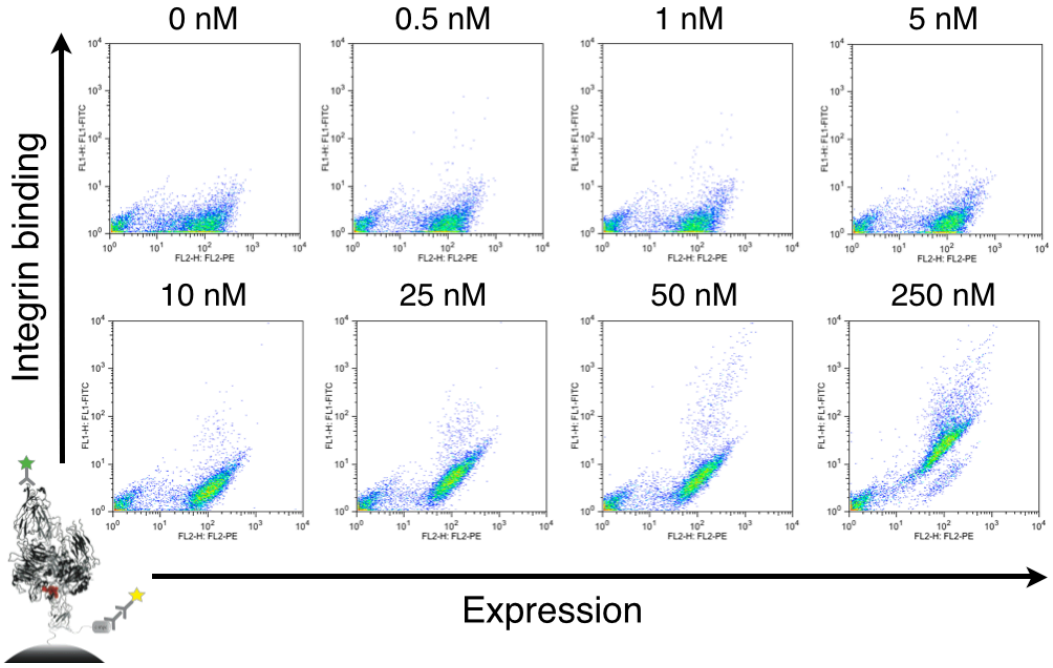


Fig. A-13. Affinity titration of 4-Ala with the SGRGDNDLV loop graft against αvβ3 integrin via flow cytometry of displayed protein.

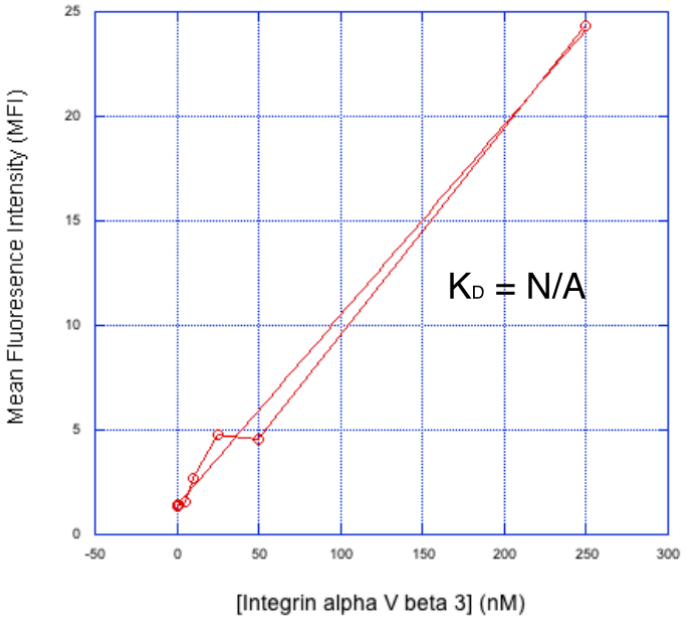


Fig. A-14. Curve fit of Figure A-13 does not show a measureable affinity.

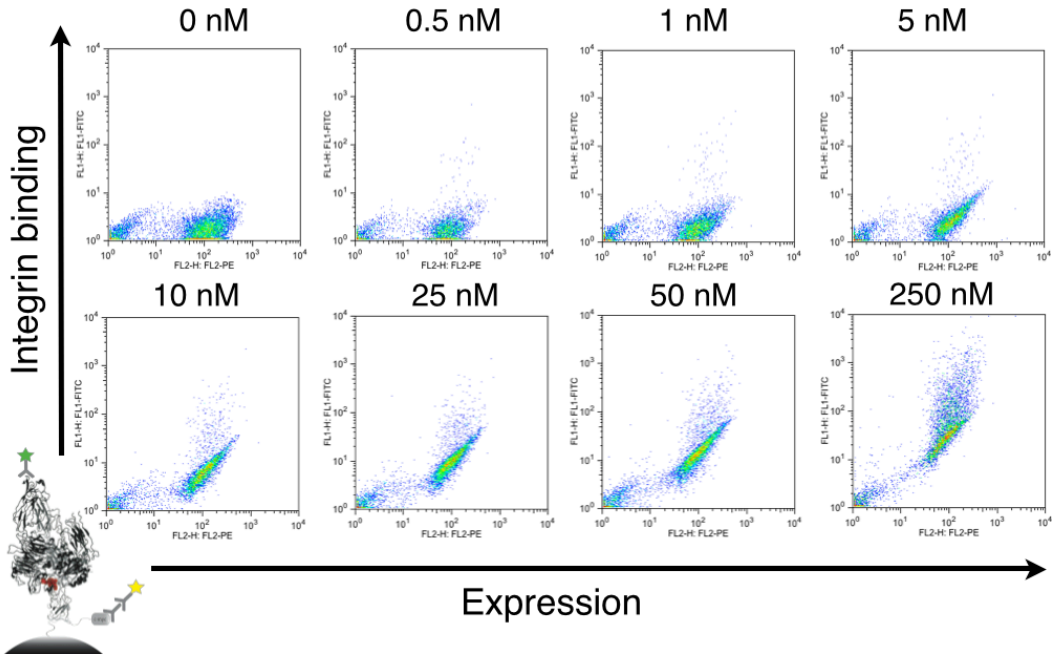


Fig. A-15. Affinity titration of 4-Ala with the AVTGRGDSPASS loop graft against αvβ3 integrin via flow cytometry of displayed protein.

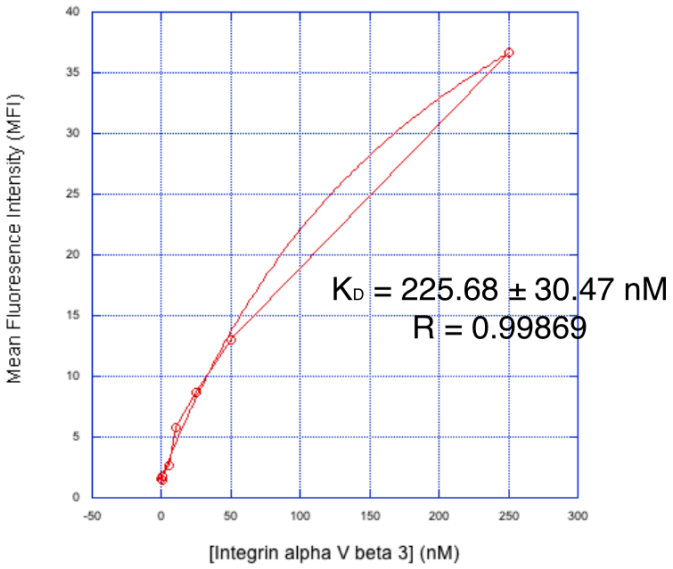


Fig. A-16. Curve fit of Figure A-15 demonstrates an affinity of 225.68 ± 30.47 nM at equilibrium.

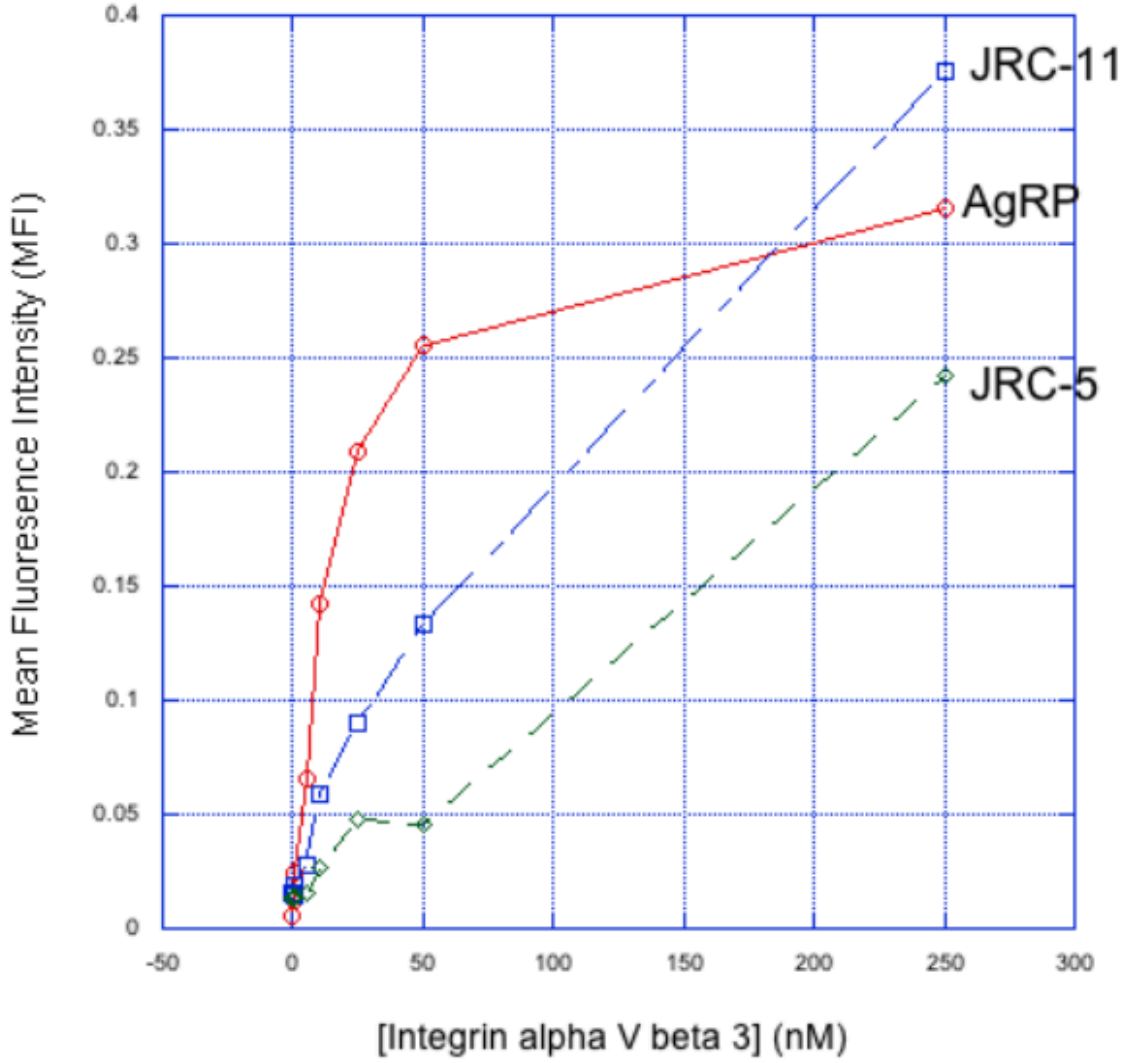


Fig. A-17. Binding curves normalized by expression for AgRP (Figure A-10), JRC-5 (SGRGDNDLV, Figure A-14), and JRC-11 (AVTGRGDSPASS, Figure A-16).

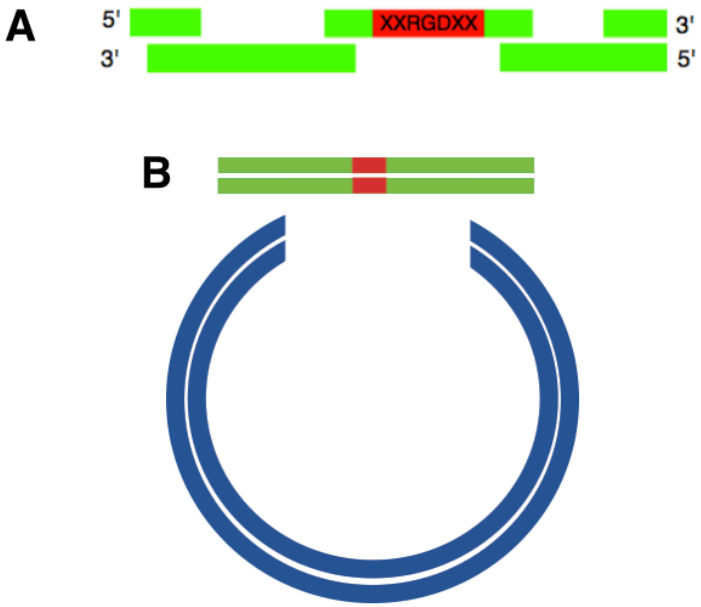


Fig. A-18. Library construction schematics. (A) Overlap PCR oligos for constructing the degenerate codon insert in green, RGD and flanking diverse regions shown in red. (B) Insert for cloning into blue pPNL6 yeast display vector via yeast homologous recombination.

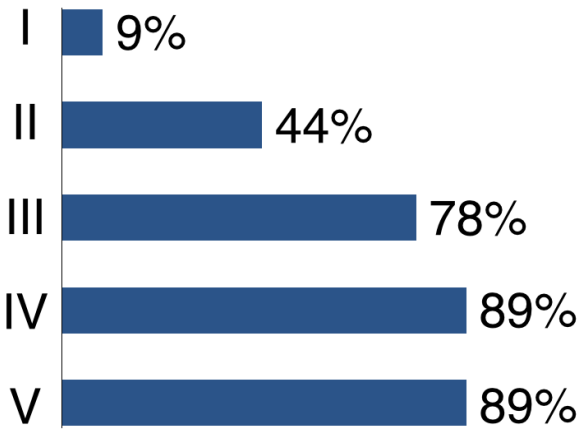


Fig. A-19. Percent lift of library population per FACS sorting round. Library "lift" is defined as the percentage of the displayed population that is lifted above baseline. Lift peaks after sort IV, indicating the selected population has plateaued.

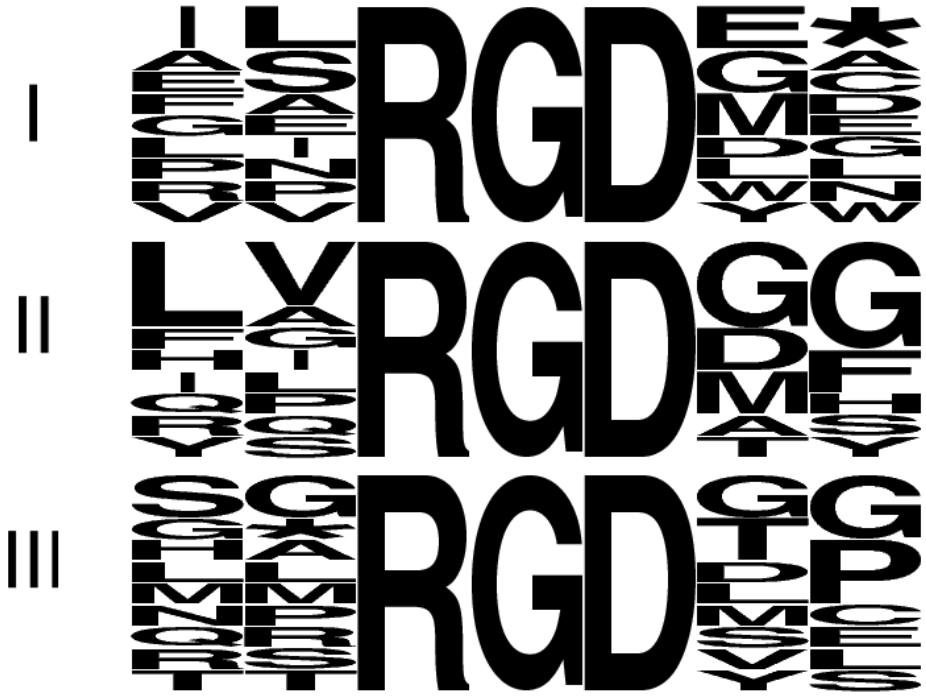


Fig. A-20. Sequence diversity of each FACS sorting round of the library. Round I-III demonstrate no convergence.

Fig. A-21. Sequences converge in sort round V. (A) VL, a minor population, converges upon RPRGDIE. (B) VR, the major population, converges upon ACRGDTC.

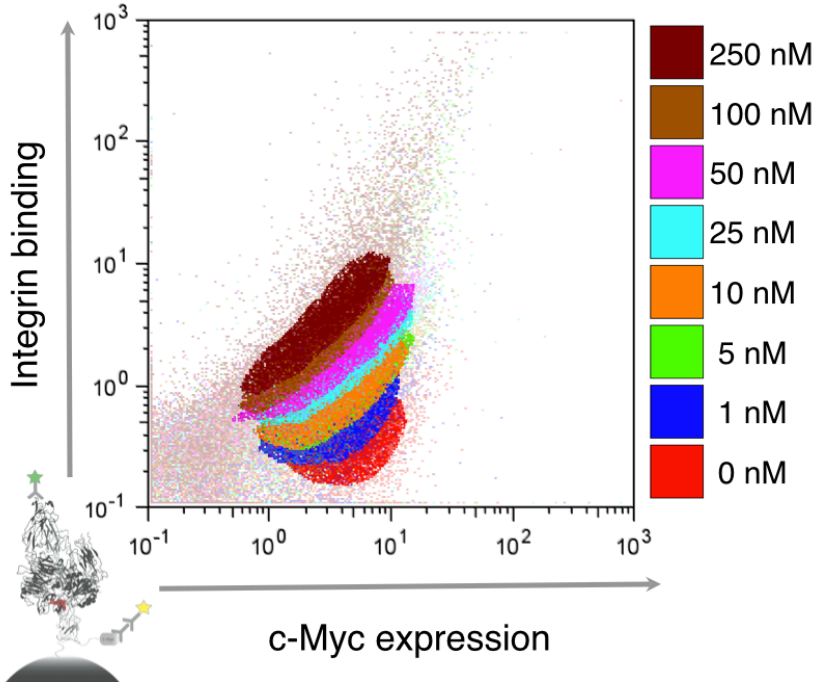


Fig. A-22. Overlaid flow cytometry plots for affinity titration of RPRGDIE in display format. αvβ3 integrin was incubated with displayed RPRGDIE discobody for 30 minutes at room temperature and the geometric mean of the center of each population was used to calculate lift.

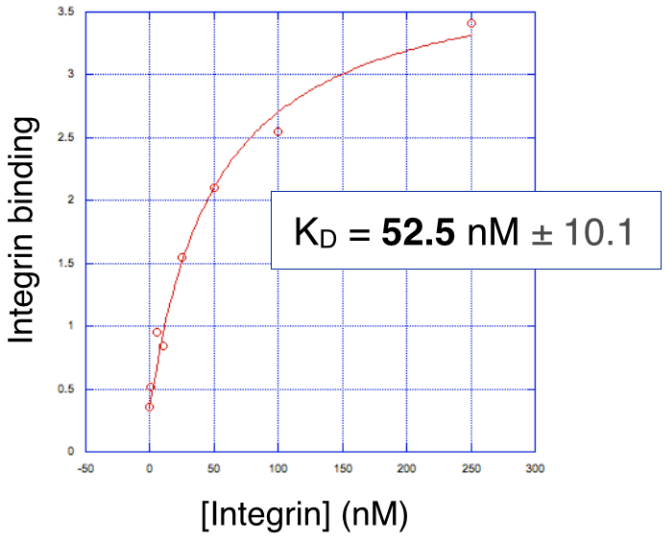


Fig. A-23. Curve fit for RPRGDIE data from Figure A-22.

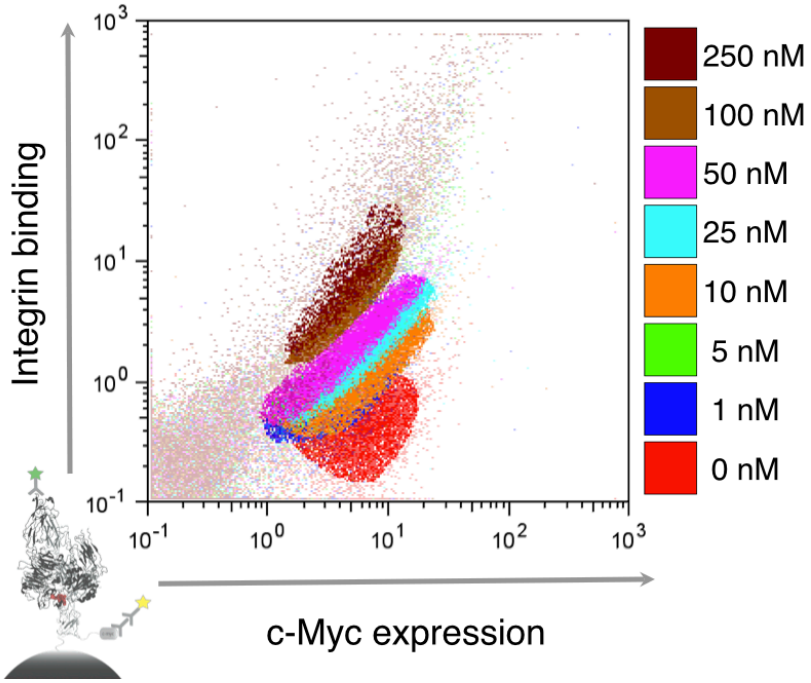


Fig. A-24. Overlaid flow cytometry plots for affinity titration of ACRGDTC in display format. αvβ3 integrin was incubated with displayed ACRGDTC discobody for 30 minutes at room temperature and the geometric mean of the center of each population was used to calculate lift. Later titrations were performed for much longer equilibrium times due to the slow on-rate.

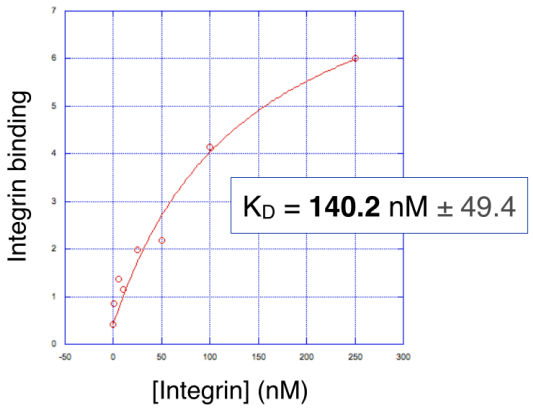


Fig. A-25. Curve fit for ACRGDTC data from Figure A-24. Later titrations with much longer equilibration times for the slow on-rate showed affinities an order of magnitude better.

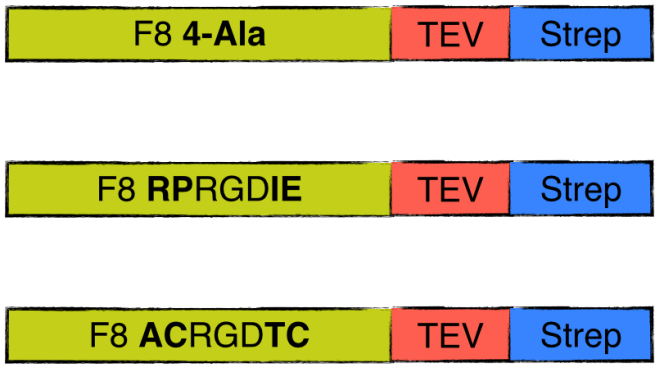


Fig. A-26. Constructs for bacterial expression. Discobody constructs are separated by a TEV cleavable linker to the Strep-II tag.

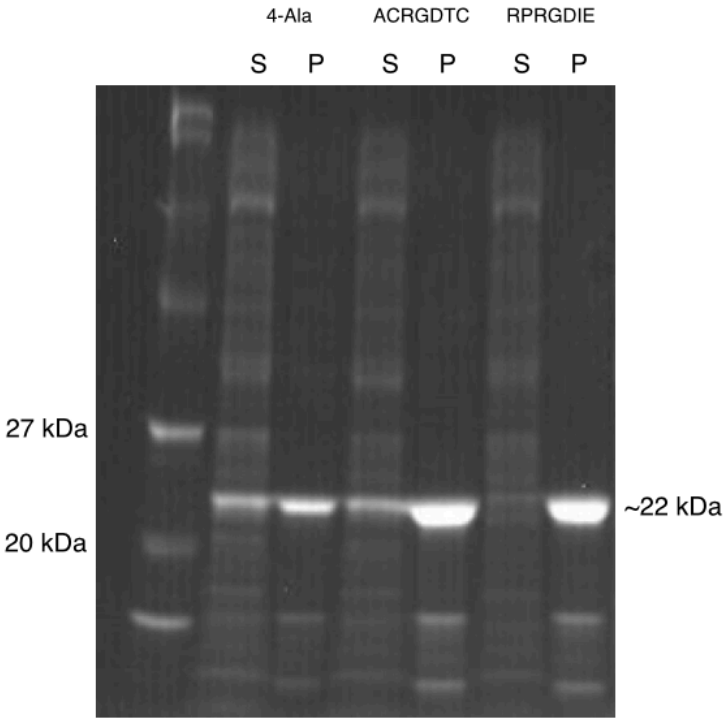


Fig. A-27. UV image of protein gel for soluble discobody clones. "S" is supernatant and "P" is pellet fraction. 4-Ala and ACRGDTC show expression in both fractions, RPRGDIE is primarily in the insoluble pellet.

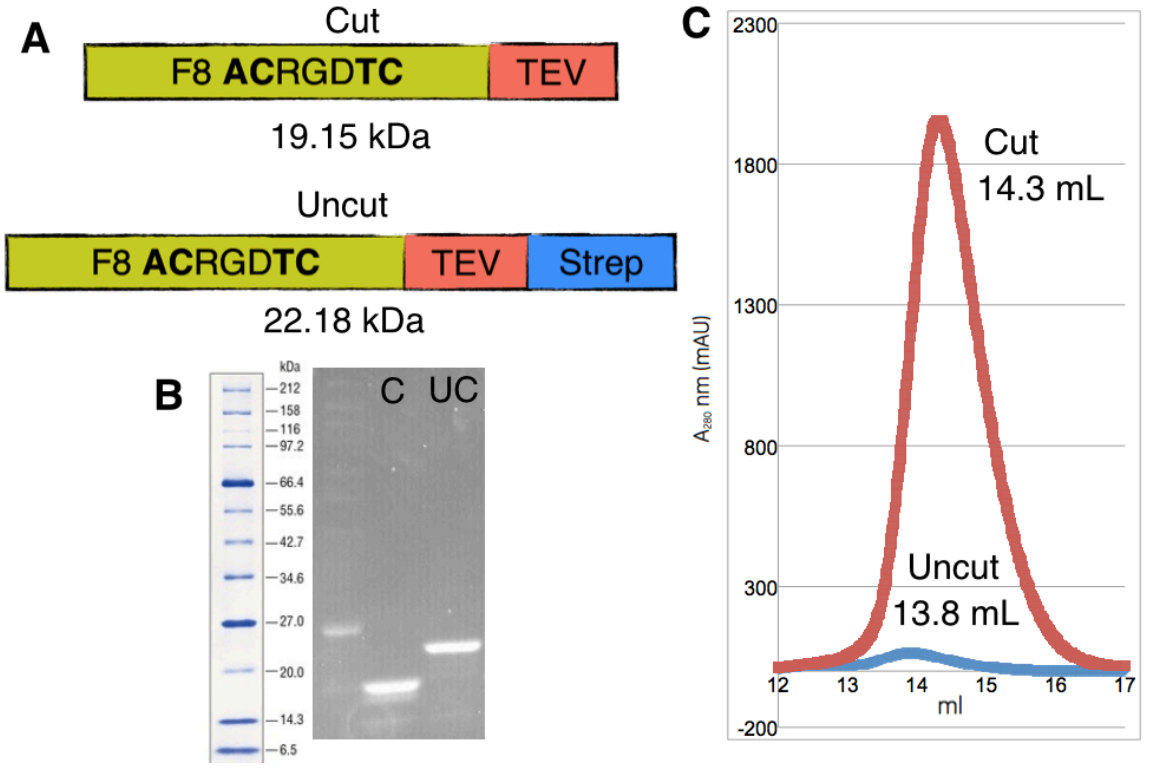


Fig. A-28. TEV cleavage of ACRGDTC discobody (Eng-Db). (A) Cut and uncut schematics. (B) Protein gel of cut and uncut proteins show clear difference in molecular weight, "C" and "UC," respectively. (C) Gel filtration of cut and uncut shows different elution times, the cut peak represents several milligrams of protein, while the uncut peak is 100 micrograms for comparison.

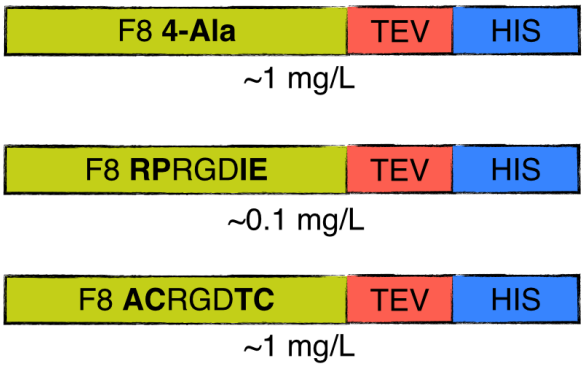


Fig. A-29. Constructs for yeast secretion expression. Discobody constructs are separated by a TEV cleavable linker to a 6xHIS tag. Expression yields are shown.

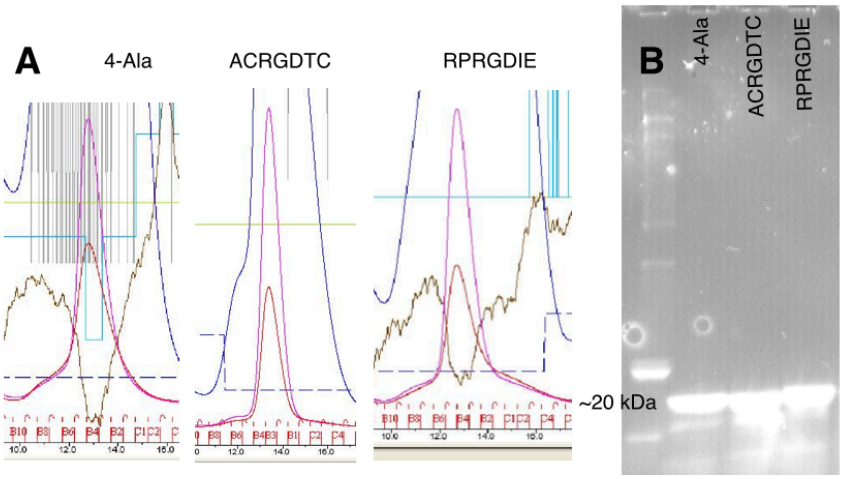


Fig. A-30. Gel filtration and reducing SDS-PAGE gel of soluble discobodies from yeast. (A) Gel filtration data shows monomeric elution profiles for all three constructs, though RPRGDIE is slightly wider at its base. (B) SDS-PAGE gel of three constructs shows expected molecular weight.

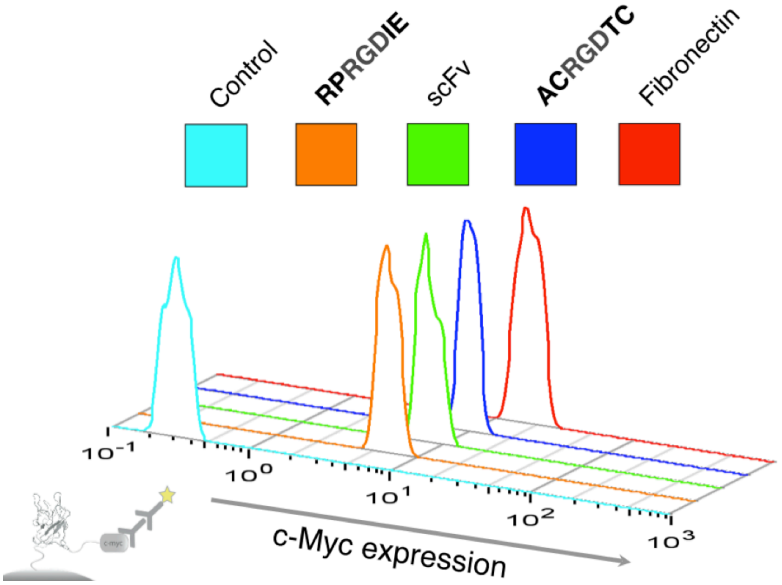


Fig. A-31. Yeast display levels of different discobody clones compared to scFv and fibronectin constructs.

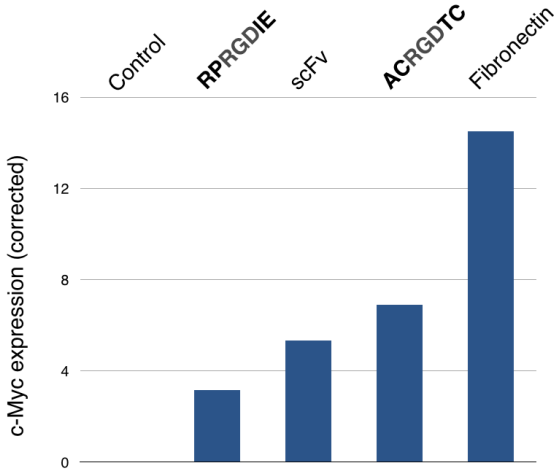


Fig. A-32. Relative expression levels from Figure A-31 plotted. Expression levels correlate with known stabilities of each protein.

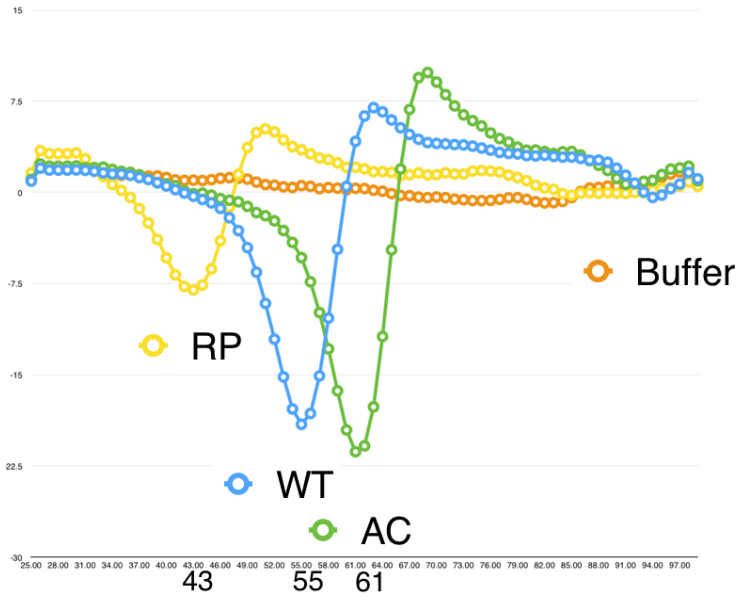


Fig. A-33. Thermal melts of soluble discobody constructs via thermofluor denaturation. "RP" is RPRGDIE, "WT" is 4-Ala, and "AC" is ACRGDTC. RPRGDIE precipitated shortly after expression. 4-Ala and ACRGDTC demonstrated stability for weeks in the cold room. Relative levels of display from Figure A-31 and A-32 correlate with soluble protein thermal stabilities.

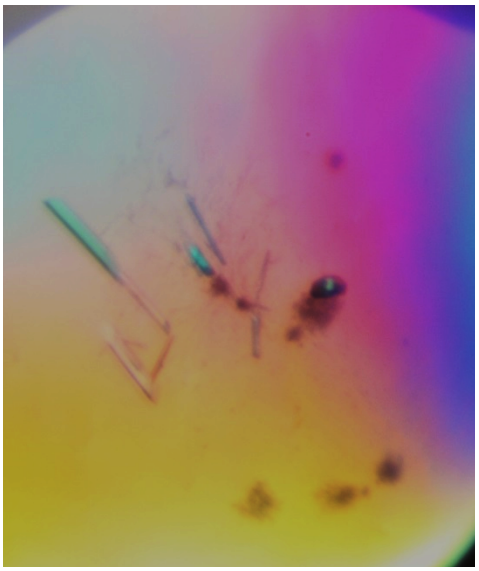


Fig. A-34. Crystal collected for diffraction.

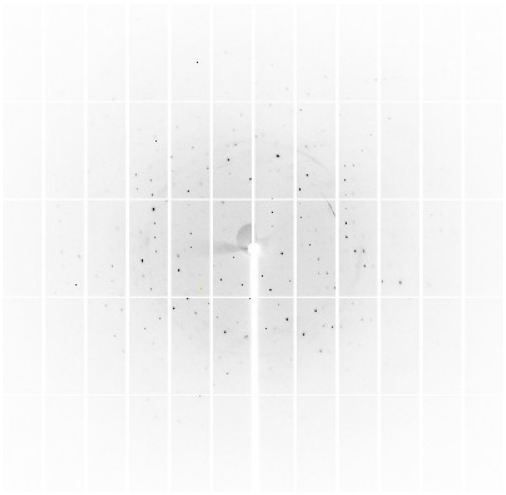


Fig. A-35. Diffraction during pre-data collection.

Table A-1. Sequences of discobody constructs

Wild type sequence of fVIII C2 domain : hydrophobic feet in bold, spikes 1-3 underlined in order. Sequence is residues L2171-Y2332 from 2R7E structure.

LNSCSMPLGMESKAISDAQITASSYFTN**MF**ATWSPSKARLHLQGRSNAWRPQVNNPKEW

LQVDFQKTMKVTGVTTQGVKS**LL**TSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ

DSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

4-Ala mutant : with MF/LL → AA/AA mutations in bold, same underline scheme as above.

LNSCSMPLGMESKAISDAQITASSYFTN**AA**ATWSPSKARLHLQGRSNAWRPQVNNPKEW

LQVDFQKTMKVTGVTTQGVKS**AA**TSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ

DSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

DNA sequence of 4-Ala mutant : XhoI cut site in bold, spike 1 underlined. This construct was used to clone degenerate codon libraries into pPNL6 yeast display vector.

GCTAGCCTAAACAGCTGCTCTATGCCGTTGGGAATGGAATCAAAGGCGATCAGCGAC

GCTCAGATCACTGC**CTC**GAGCTATTTTACCAATGCTGCCGCCACATGGTCCCCTTCCA

AGGCGAGACTGCATTTGCAAGGTAGATCAAACGCGTGGAGACCACAAGTGAACAACC

CGAAAGAGTGGCTACAGGTGGACTTTCAGAAGACCATGAAGGTGACTGGCGTGACGA

CTCAGGGTGTAAAGTCAGCCGCGACCAGCATGTATGTCAAGGAGTTCTTGATCAGCTC

CAGCCAGGACGGGCACCAATGGACGCTTTTTTTCCAGAACGGTAAGGTTAAGGTGTTC

CAGGGAAACCAGGACTCATTTACGCCCGTGGTGAACAGCCTAGATCCCCCGTTGTTGA

CCAGATACTTGAGGATACACCCGCAGTCTTGGGTGCATCAAATTGCCTTGCGTATGGA

GGTATTGGGCTGTGAGGCTCAGGATCTGTATGGTTCCGGTAGC

RGD library schematic : NNS codons denoted with "X" and spikes 1-3 underlined in order.

LNSCSMPLGMESKAISDAQITASSY**XXRGDXX**WSPSKARLHLQGRSNAWRPQVNNPKEW

LQVDFQKTMKVTGVTTQGVKS**AA**TSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ

DSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

Eng-Db sequence : spikes 1-3 underlined in order.

LNSCSMPLGMESKAISDAQITASSY**ACRGDTC**WSPSKARLHLQGRSNAWRPQVNNPKEW

LQVDFQKTMKVTGVTTQGVKS**AA**TSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ

DSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

BIBLIOGRAPHY

1. Borrebaeck CA: **Antibodies in diagnostics - from immunoassays to protein chips.** *Immunol Today* (2000) **21**(8):379-382.

2. Nissim A, Chernajovsky Y: **Historical development of monoclonal antibody therapeutics.** *Handb Exp Pharmacol* (2008) 181):3-18.

3. Reichert JM: **Monoclonal antibodies as innovative therapeutics.** *Curr Pharm Biotechnol* (2008) **9**(6):423-430.

4. Eccles SA: **Monoclonal antibodies targeting cancer: 'Magic bullets' or just the trigger?** *Breast cancer research : BCR* (2001) **3**(2):86-90.

5. Page DB, Postow MA, Callahan MK, Allison JP, Wolchok JD: **Immune modulation in cancer with antibodies.** *Annual review of medicine* (2014) **65**(185-202.

6. Boder ET, Jiang W: **Engineering antibodies for cancer therapy.** *Annual review of chemical and biomolecular engineering* (2011) **2**(53-75.

7. Stork R, Campigna E, Robert B, Muller D, Kontermann RE: **Biodistribution of a bispecific single-chain diabody and its half-life extended derivatives.** *J Biol Chem* (2009) **284**(38):25612-25619.

8. Spiess C, Merchant M, Huang A, Zheng Z, Yang NY, Peng J, Ellerman D, Shatz W, Reilly D, Yansura DG, Scheer JM: **Bispecific antibodies with natural architecture produced by co-culture of bacteria expressing two distinct half-antibodies.** *Nature biotechnology* (2013) **31**(8):753-758.

9. Baeuerle PA, Reinhardt C: **Bispecific t-cell engaging antibodies for cancer therapy.** *Cancer Res* (2009) **69**(12):4941-4944.

10. Stern LA, Case BA, Hackel BJ: **Alternative non-antibody protein scaffolds for molecular imaging of cancer.** *Curr Opin Chem Eng* (2013) **2**(4).

11. Skerra A: **Alternative non-antibody scaffolds for molecular recognition.** *Curr Opin Biotechnol* (2007) **18**(4):295-304.

12. Gebauer M, Skerra A: **Engineered protein scaffolds as next-generation antibody therapeutics.** *Current Opinion in Chemical Biology* (2009) **13**(3):245-255.

13. Sievers EL, Senter PD: **Antibody-drug conjugates in cancer therapy.** *Annual review of medicine* (2013) **64**(15-29.

14. Cardoso MM, Peca IN, Roque AC: **Antibody-conjugated nanoparticles for therapeutic applications.** *Current medicinal chemistry* (2012) **19**(19):3103-3127.

15. Lofblom J, Frejd FY, Stahl S: **Non-immunoglobulin based protein scaffolds.** *Curr Opin Biotechnol* (2011) **22**(6):843-848.

16. Weidle UH, Auer J, Brinkmann U, Georges G, Tiefenthaler G: **The emerging role of new protein scaffold-based agents for treatment of cancer.** *Cancer genomics & proteomics* (2013) **10**(4):155-168.

17. Murphy K, Travers P, Walport M, Janeway C: *Janeway's immunobiology.* Garland Science, New York (2012).

18. Hwang WY, Foote J: **Immunogenicity of engineered antibodies.** *Methods* (2005) **36**(1):3-10.

19. Stern M, Herrmann R: **Overview of monoclonal antibodies in cancer therapy: Present and promise.** *Critical reviews in oncology/hematology* (2005) **54**(1):11-29.

20. Lonberg N: **Fully human antibodies from transgenic mouse and phage display platforms.** *Current opinion in immunology* (2008) **20**(4):450-459.

21. Barbas CF: *Phage display : A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2001).

22. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD: **Isolating and engineering human antibodies using yeast surface display.** *Nat Protoc* (2006) **1**(2):755-768.

23. Bowley DR, Labrijn AF, Zwick MB, Burton DR: **Antigen selection from an hiv-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage.** *Protein engineering, design & selection : PEDS* (2007) **20**(2):81-90.

24. Kontermann R, Dübel S, SpringerLink (Online service): **Antibody engineering.** In: Springer-Verlag Berlin Heidelberg,, Berlin, Heidelberg (2010).

25. Chames P: *Antibody engineering : Methods and protocols.* Humana Press, New York (2012).

26. Lazar GA, Dang W, Karki S, Vafa O, Peng JS, Hyun L, Chan C, Chung HS, Eivazi A, Yoder SC, Vielmetter J *et al*: **Engineered antibody fc variants with enhanced effector function.** *Proceedings of the National Academy of Sciences of the United States of America* (2006) **103**(11):4005-4010.

27. Vaccaro C, Zhou J, Ober RJ, Ward ES: **Engineering the fc region of immunoglobulin g to modulate in vivo antibody levels.** *Nature biotechnology* (2005) **23**(10):1283-1288.

28. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, Smolen J, Emery P, Harriman G, Feldmann M, Lipsky P: **Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: A randomised phase iii trial. Attract study group.** *Lancet* (1999) **354**(9194):1932-1939.

29. Holt LJ, Herring C, Jespers LS, Woolven BP, Tomlinson IM: **Domain antibodies: Proteins for therapy.** *Trends Biotechnol* (2003) **21**(11):484-490.

30. Jespers L, Schon O, Famm K, Winter G: **Aggregation-resistant domain antibodies selected on phage by heat denaturation.** *Nature biotechnology* (2004) **22**(9):1161-1165.

31. Koide A, Bailey CW, Huang X, Koide S: **The fibronectin type iii domain as a scaffold for novel binding proteins.** *Journal of molecular biology* (1998) **284**(4):1141-1151.

32. Batori V, Koide A, Koide S: **Exploring the potential of the monobody scaffold: Effects of loop elongation on the stability of a fibronectin type iii domain.** *Protein Eng* (2002) **15**(12):1015-1020.

33. Lipovsek D, Lippow SM, Hackel BJ, Gregson MW, Cheng P, Kapila A, Wittrup KD: **Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type iii domain and selected by yeast surface display: Molecular convergence with single-domain camelid and shark antibodies.** *Journal of molecular biology* (2007) **368**(4):1024-1041.

34. Hackel BJ, Kapila A, Wittrup KD: **Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling.** *Journal of molecular biology* (2008) **381**(5):1238-1252.

35. Hackel BJ, Ackerman ME, Howland SW, Wittrup KD: **Stability and cdr composition biases enrich binder functionality landscapes.** *Journal of molecular biology* (2010) **401**(1):84-96.

36. Hackel BJ, Wittrup KD: **The full amino acid repertoire is superior to serine/tyrosine for selection of high affinity immunoglobulin g binders from the fibronectin scaffold.** *Protein engineering, design & selection : PEDS* (2010) **23**(4):211-219.

37. Emanuel SL, Engle LJ, Chao G, Zhu RR, Cao C, Lin Z, Yamniuk AP, Hosbach J, Brown J, Fitzpatrick E, Gokemeijer J *et al*: **A fibronectin scaffold approach to bispecific inhibitors of epidermal growth factor receptor and insulin-like growth factor-i receptor.** *MAbs* (2011) **3**(1):38-48.

38. Koide A, Wojcik J, Gilbreth RN, Hoey RJ, Koide S: **Teaching an old scaffold new tricks: Monobodies constructed using alternative surfaces of the fn3 scaffold.** *Journal of molecular biology* (2012) **415**(2):393-405.

39. Banta S, Dooley K, Shur O: **Replacing antibodies: Engineering new binding proteins.** *Annu Rev Biomed Eng* (2013) **15**(93-113.

40. Welschof M, Krauss Jr: *Recombinant antibodies for cancer therapy : Methods and protocols.* Humana Press, Totowa, N.J. (2003).

41. Jeong KJ, Mabry R, Georgiou G: **Avimers hold their own.** *Nature biotechnology* (2005) **23**(12):1493-1494.

42. Arnoux B, Ducruix A, Prange T: **Anisotropic behaviour of the c-terminal kunitz-type domain of the alpha3 chain of human type vi collagen at atomic resolution (0.9 a).** *Acta crystallographica Section D, Biological crystallography* (2002) **58**(Pt 7):1252-1254.

43. Dickinson CD, Veerapandian B, Dai XP, Hamlin RC, Xuong NH, Ruoslahti E, Ely KR: **Crystal structure of the tenth type iii cell adhesion module of human fibronectin.** *Journal of molecular biology* (1994) **236**(4):1079-1092.

44. Schonfeld D, Matschiner G, Chatwell L, Trentmann S, Gille H, Hulsmeyer M, Brown N, Kaye PM, Schlehuber S, Hohlbaum AM, Skerra A: **An engineered lipocalin specific for ctla-4 reveals a combining site with structural and conformational features similar to antibodies.** *Proceedings of the National Academy of Sciences of the United States of America* (2009) **106**(20):8198-8203.

45. Fass D, Blacklow S, Kim PS, Berger JM: **Molecular basis of familial hypercholesterolaemia from structure of ldl receptor module.** *Nature* (1997) **388**(6643):691-693.

46. Bolin KA, Anderson DJ, Trulson JA, Thompson DA, Wilken J, Kent SB, Gantz I, Millhauser GL: **Nmr structure of a minimized human agouti related protein prepared by total chemical synthesis.** *FEBS letters* (1999) **451**(2):125-131.

47. Eigenbrot C, Ultsch M, Dubnovitsky A, Abrahmsen L, Hard T: **Structural basis for high-affinity her2 receptor binding by an engineered protein.** *Proceedings of the National Academy of Sciences of the United States of America* (2010) **107**(34):15039-15044.

48. Bandeiras TM, Hillig RC, Matias PM, Eberspaecher U, Fanghanel J, Thomaz M, Miranda S, Crusius K, Putter V, Amstutz P, Gulotti-Georgieva M *et al*: **Structure of wild-type plk-1 kinase domain in complex with a selective darpin.** *Acta crystallographica Section D, Biological crystallography* (2008) **64**(Pt 4):339-353.

49. Shen BW, Spiegel PC, Chang CH, Huh JW, Lee JS, Kim J, Kim YH, Stoddard BL: **The tertiary structure and domain organization of coagulation factor viii.** *Blood* (2008) **111**(3):1240-1247.

50. Xu L, Aha P, Gu K, Kuimelis RG, Kurz M, Lam T, Lim AC, Liu H, Lohse PA, Sun L, Weng S *et al*: **Directed evolution of high-affinity antibody mimics using mrna display.** *Chemistry & biology* (2002) **9**(8):933-942.

51. Lipovsek D, Pluckthun A: **In-vitro protein evolution by ribosome display and mrna display.** *Journal of immunological methods* (2004) **290**(1-2):51-67.

52. Boder ET, Midelfort KS, Wittrup KD: **Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity.** *Proceedings of the National Academy of Sciences of the United States of America* (2000) **97**(20):10701-10705.

53. Koide A, Abbatiello S, Rothgery L, Koide S: **Probing protein conformational changes in living cells by using designer binding proteins: Application to the estrogen receptor.** *Proceedings of the National Academy of Sciences of the United States of America* (2002) **99**(3):1253-1258.

54. Tolcher AW, Sweeney CJ, Papadopoulos K, Patnaik A, Chiorean EG, Mita AC, Sankhala K, Furfine E, Gokemeijer J, Iacono L, Eaton C *et al*: **Phase i and pharmacokinetic study of ct-322 (bms-844203), a targeted adnectin inhibitor of vegfr-2 based on a domain of human fibronectin.** *Clin Cancer Res* (2011) **17**(2):363-371.

55. Spangler JB, Manzari MT, Rosalia EK, Chen TF, Wittrup KD: **Triepitopic antibody fusions inhibit cetuximab-resistant braf and kras mutant tumors via egfr signal repression.** *Journal of molecular biology* (2012) **422**(4):532-544.

56. Oganesyan V, Ferguson A, Grinberg L, Wang L, Phipps S, Chacko B, Drabic S, Thisted T, Baca M: **Fibronectin type iii domains engineered to bind cd40l: Cloning, expression, purification, crystallization and preliminary x-ray diffraction analysis of two complexes.** *Acta Crystallogr Sect F Struct Biol Cryst Commun* (2013) **69**(Pt 9):1045-1048.

57. Jacobs SA, Diem MD, Luo J, Teplyakov A, Obmolova G, Malia T, Gilliland GL, O'Neil KT: **Design of novel fn3 domains with high stability by a consensus sequence approach.** *Protein engineering, design & selection : PEDS* (2012) **25**(3):107-117.

58. Boersma YL, Pluckthun A: **Darpins and other repeat protein scaffolds: Advances in engineering and applications.** *Curr Opin Biotechnol* (2011) **22**(6):849-857.

59. Schmidt SR: *Fusion protein technologies for biopharmaceuticals : Applications and challenges.* John Wiley & Sons, Hoboken, N.J. (2013).

60. Steiner D, Forrer P, Pluckthun A: **Efficient selection of darpins with sub-nanomolar affinities using srp phage display.** *Journal of molecular biology* (2008) **382**(5):1211-1227.

61. Steiner D, Forrer P, Stumpp MT, Pluckthun A: **Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display.** *Nature biotechnology* (2006) **24**(7):823-831.

62. Dreier B, Mikheeva G, Belousova N, Parizek P, Boczek E, Jelesarov I, Forrer P, Pluckthun A, Krasnykh V: **Her2-specific multivalent adapters confer designed tropism to adenovirus for gene targeting.** *Journal of molecular biology* (2011) **405**(2):410-426.

63. Munch RC, Janicki H, Volker I, Rasbach A, Hallek M, Buning H, Buchholz CJ: **Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer.** *Molecular therapy : the journal of the American Society of Gene Therapy* (2013) **21**(1):109-118.

64. Silverman J, Liu Q, Bakker A, To W, Duguay A, Alba BM, Smith R, Rivas A, Li P, Le H, Whitehorn E *et al*: **Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains.** *Nature biotechnology* (2005) **23**(12):1556-1561.

65. Gebauer M, Skerra A: **Anticalins small engineered binding proteins based on the lipocalin scaffold.** *Methods Enzymol* (2012) **503**(157-188.

66. Schlehuber S, Skerra A: **Lipocalins in drug discovery: From natural ligand-binding proteins to "anticalins".** *Drug Discov Today* (2005) **10**(1):23-33.

67. Mross K, Richly H, Fischer R, Scharr D, Buchert M, Stern A, Gille H, Audoly LP, Scheulen ME: **First-in-human phase i study of prs-050 (angiocal), an anticalin targeting and antagonizing vegf-a, in patients with advanced solid tumors.** *PLoS One* (2013) **8**(12):e83232.

68. Terwisscha van Scheltinga AG, Lub-de Hooge MN, Hinner MJ, Verheijen RB, Allersdorfer A, Hulsmeyer M, Nagengast WB, Schroder CP, Kosterink JG, de Vries EG, Audoly L *et al*: **In vivo visualization of met tumor expression and anticalin biodistribution with the met-specific anticalin 89zr-prs-110 pet tracer.** *J Nucl Med* (2014) **55**(4):665-671.

69. Lofblom J, Feldwisch J, Tolmachev V, Carlsson J, Stahl S, Frejd FY: **Affibody molecules: Engineered proteins for therapeutic, diagnostic and biotechnological applications.** *FEBS letters* (2010) **584**(12):2670-2680.

70. Kimura RH, Cheng Z, Gambhir SS, Cochran JR: **Engineered knottin peptides: A new class of agents for imaging integrin expression in living subjects.** *Cancer Res* (2009) **69**(6):2435-2442.

71. Silverman AP, Levin AM, Lahti JL, Cochran JR: **Engineered cystine-knot peptides that bind alpha(v)beta(3) integrin with antibody-like affinities.** *Journal of molecular biology* (2009) **385**(4):1064-1075.

72. Liu S, Liu H, Ren G, Kimura RH, Cochran JR, Cheng Z: **Pet imaging of integrin positive tumors using f labeled knottin peptides.** *Theranostics* (2011) **1**(403-412.

73. Moore SJ, Hayden Gephart MG, Bergen JM, Su YS, Rayburn H, Scott MP, Cochran JR: **Engineered knottin peptide enables noninvasive optical imaging of intracranial medulloblastoma.** *Proceedings of the National Academy of Sciences of the United States of America* (2013) **110**(36):14598-14603.

74. McIntosh M, Cruz LJ, Hunkapiller MW, Gray WR, Olivera BM: **Isolation and structure of a peptide toxin from the marine snail conus magus.** *Archives of biochemistry and biophysics* (1982) **218**(1):329-334.

75. Skov MJ, Beck JC, de Kater AW, Shopp GM: **Nonclinical safety of ziconotide: An intrathecal analgesic of a new pharmaceutical class.** *International journal of toxicology* (2007) **26**(5):411-421.

76. Dennis MS, Herzka A, Lazarus RA: **Potent and selective kunitz domain inhibitors of plasma kallikrein designed by phage display.** *J Biol Chem* (1995) **270**(43):25411-25417.

77. Dennis MS, Lazarus RA: **Kunitz domain inhibitors of tissue factor-factor viia. Ii. Potent and specific inhibitors by competitive phage selection.** *J Biol Chem* (1994) **269**(35):22137-22144.

78. Dennis MS, Lazarus RA: **Kunitz domain inhibitors of tissue factor-factor viia. I. Potent inhibitors selected from libraries by phage display.** *J Biol Chem* (1994) **269**(35):22129-22136.

79. Lehmann A: **Ecallantide (dx-88), a plasma kallikrein inhibitor for the treatment of hereditary angioedema and the prevention of blood loss in on-pump cardiothoracic surgery.** *Expert opinion on biological therapy* (2008) **8**(8):1187-1199.

80. LaFleur DW, Abramyan D, Kanakaraj P, Smith RG, Shah RR, Wang G, Yao XT, Kankanala S, Boyd E, Zaritskaya L, Nam V *et al*: **Monoclonal antibody therapeutics with up to five specificities: Functional enhancement through fusion of target-specific peptides.** *MAbs* (2013) **5**(2):208-218.

81. Heinis C, Rutherford T, Freund S, Winter G: **Phage-encoded combinatorial chemical libraries based on bicyclic peptides.** *Nat Chem Biol* (2009) **5**(7):502-507.

82. Kariolis MS, Kapur S, Cochran JR: **Beyond antibodies: Using biological principles to guide the development of next-generation protein therapeutics.** *Curr Opin Biotechnol* (2013) **24**(6):1072-1077.

83. Gilbert GE, Kaufman RJ, Arena AA, Miao H, Pipe SW: **Four hydrophobic amino acids of the factor viii c2 domain are constituents of both the**

membrane-binding and von willebrand factor-binding motifs. *J Biol Chem* (2002) **277**(8):6374-6381.

84. Pellequer JL, Chen SW, Saboulard D, Delcourt M, Negrier C, Plantier JL: **Functional mapping of factor viii c2 domain.** *Thromb Haemost* (2011) **106**(1):121-131.

85. Lenting PJ, van Mourik JA, Mertens K: **The life cycle of coagulation factor viii in view of its structure and function.** *Blood* (1998) **92**(11):3983-3996.

86. Novakovic VA, Cullinan DB, Wakabayashi H, Fay PJ, Baleja JD, Gilbert GE: **Membrane-binding properties of the factor viii c2 domain.** *Biochem J* (2011) **435**(1):187-196.

87. Pratt KP, Shen BW, Takeshima K, Davie EW, Fujikawa K, Stoddard BL: **Structure of the c2 domain of human factor viii at 1.5 a resolution.** *Nature* (1999) **402**(6760):439-442.

88. Liu Z, Lin L, Yuan C, Nicolaes GA, Chen L, Meehan EJ, Furie B, Furie B, Huang M: **Trp2313-his2315 of factor viii c2 domain is involved in membrane binding: Structure of a complex between the c2 domain and an inhibitor of membrane binding.** *J Biol Chem* (2010) **285**(12):8824-8829.

89. Franchini M, Lippi G: **Acquired factor viii inhibitors.** *Blood* (2008) **112**(2):250-255.

90. Lusher JM, Scharrer I: **Evolution of recombinant factor viii safety: Kogenate and kogenate fs/bayer.** *Int J Hematol* (2009) **90**(4):446-454.

91. Bray GL, Gomperts ED, Courter S, Gruppo R, Gordon EM, Manco-Johnson M, Shapiro A, Scheibel E, White G, 3rd, Lee M: **A multicenter study of recombinant factor viii (recombinate): Safety, efficacy, and inhibitor risk in previously untreated patients with hemophilia a. The recombinate study group.** *Blood* (1994) **83**(9):2428-2435.

92. Powell JS, Josephson NC, Quon D, Ragni MV, Cheng G, Li E, Jiang H, Li L, Dumont JA, Goyal J, Zhang X *et al*: **Safety and prolonged activity of recombinant factor viii fc fusion protein in hemophilia a patients.** *Blood* (2012) **119**(13):3031-3037.

93. Mei B, Pan C, Jiang H, Tjandra H, Strauss J, Chen Y, Liu T, Zhang X, Severs J, Newgren J, Chen J *et al*: **Rational design of a fully active, long-acting pegylated factor viii for hemophilia a treatment.** *Blood* (2010) **116**(2):270-279.

94. Jiang H, Lillicrap D, Patarroyo-White S, Liu T, Qian X, Scallan CD, Powell S, Keller T, McMurray M, Labelle A, Nagy D *et al*: **Multiyear therapeutic benefit of aav serotypes 2, 6, and 8 delivering factor viii to hemophilia a mice and dogs.** *Blood* (2006) **108**(1):107-115.

95. Kitazawa T, Igawa T, Sampei Z, Muto A, Kojima T, Soeda T, Yoshihashi K, Okuyama-Nishida Y, Saito H, Tsunoda H, Suzuki T *et al*: **A bispecific antibody to factors ixa and x restores factor viii hemostatic activity in a hemophilia a model.** *Nature medicine* (2012) **18**(10):1570-1574.

96. Kemball-Cook G, Tuddenham EG, Wacey AI: **The factor viii structure and mutation resource site: Hamsters version 4.** *Nucleic acids research* (1998) **26**(1):216-219.

97. Nguyen PC, Lewis KB, Ettinger RA, Schuman JT, Lin JC, Healey JF, Meeks SL, Lollar P, Pratt KP: **High-resolution mapping of epitopes on the c2 domain of factor viii by analysis of point mutants using surface plasmon resonance.** *Blood* (2014) **123**(17):2732-2739.

98. Baumgartner S, Hofmann K, Chiquet-Ehrismann R, Bucher P: **The discoidin domain family revisited: New members from prokaryotes and a homology-based fold prediction.** *Protein Sci* (1998) **7**(7):1626-1631.

99. Vogel W: **Discoidin domain receptors: Structural relations and functional implications.** *FASEB J* (1999) **13 Suppl**(S77-82.

100. Vogel WF, Abdulhussein R, Ford CE: **Sensing extracellular matrix: An update on discoidin domain receptor function.** *Cell Signal* (2006) **18**(8):1108-1116.

101. Kiedzierska A, Smietana K, Czepczynska H, Otlewski J: **Structural similarities and functional diversity of eukaryotic discoidin-like domains.** *Biochim Biophys Acta* (2007) **1774**(9):1069-1078.

102. Liu Z, Wang F, Chen X: **Integrin alpha(v)beta(3)-targeted cancer therapy.** *Drug Dev Res* (2008) **69**(6):329-339.

103. Marelli UK, Rechenmacher F, Sobahi TR, Mas-Moruno C, Kessler H: **Tumor targeting via integrin ligands.** *Front Oncol* (2013) **3**(222.

104. Agnew HD, Rohde RD, Millward SW, Nag A, Yeo WS, Hein JE, Pitram SM, Tariq AA, Burns VM, Krom RJ, Fokin VV *et al*: **Iterative in situ click chemistry creates antibody-like protein-capture agents.** *Angew Chem Int Ed Engl* (2009) **48**(27):4944-4948.

105. Hey T, Fiedler E, Rudolph R, Fiedler M: **Artificial, non-antibody binding proteins for pharmaceutical and industrial applications.** *Trends Biotechnol* (2005) **23**(10):514-522.

106. De Groot AS, Scott DW: **Immunogenicity of protein therapeutics.** *Trends Immunol* (2007) **28**(11):482-490.

107. Gonzales NR, De Pascalis R, Schlom J, Kashmiri SV: **Minimizing the immunogenicity of antibodies for clinical application.** *Tumour Biol* (2005) **26**(1):31-43.

108. Binz HK, Amstutz P, Kohl A, Stumpp MT, Briand C, Forrer P, Grutter MG, Pluckthun A: **High-affinity binders selected from designed ankyrin repeat protein libraries.** *Nature biotechnology* (2004) **22**(5):575-582.

109. Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S, Conrath K: **General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold.** *J Biol Chem* (2009) **284**(5):3273-3284.

110. Wurch T, Lowe P, Caussanel V, Bes C, Beck A, Corvaia N: **Development of novel protein scaffolds as alternatives to whole antibodies for imaging and therapy: Status on discovery research and clinical validation.** *Curr Pharm Biotechnol* (2008) **9**(6):502-509.

111. Roovers RC, Vosjan MJ, Laeremans T, el Khoulati R, de Bruin RC, Ferguson KM, Verkleij AJ, van Dongen GA, van Bergen en Henegouwen PM: **A biparatopic anti-egfr nanobody efficiently inhibits solid tumour growth.** *Int J Cancer* (2011) **129**(8):2013-2024.

112. Boersma YL, Chao G, Steiner D, Wittrup KD, Pluckthun A: **Bispecific designed ankyrin repeat proteins (darpins) targeting epidermal growth factor receptor inhibit a431 cell proliferation and receptor recycling.** *J Biol Chem* (2011) **286**(48):41273-41285.

113. Eggel A, Baumann MJ, Amstutz P, Stadler BM, Vogel M: **Darpins as bispecific receptor antagonists analyzed for immunoglobulin e receptor blockage.** *Journal of molecular biology* (2009) **393**(3):598-607.

114. Lusher JM, Arkin S, Abildgaard CF, Schwartz RS: **Recombinant factor viii for the treatment of previously untreated patients with hemophilia a. Safety, efficacy, and development of inhibitors. Kogenate previously untreated patient study group.** *N Engl J Med* (1993) **328**(7):453-459.

115. Pfaff M, Tangemann K, Muller B, Gurrath M, Muller G, Kessler H, Timpl R, Engel J: **Selective recognition of cyclic rgd peptides of nmr defined conformation by alpha iib beta 3, alpha v beta 3, and alpha 5 beta 1 integrins.** *J Biol Chem* (1994) **269**(32):20233-20238.

116. Maheshwari G, Brown G, Lauffenburger DA, Wells A, Griffith LG: **Cell adhesion and motility depend on nanoscale rgd clustering.** *J Cell Sci* (2000) **113 (Pt 10)**(1677-1686.

117. Hoover DM, Lubkowski J: **Dnaworks: An automated method for designing oligonucleotides for pcr-based gene synthesis.** *Nucleic acids research* (2002) **30**(10):e43.

118. Gibson DG: **Enzymatic assembly of overlapping DNA fragments.** *Methods Enzymol* (2011) **498**(349-361.

119. Crooks GE, Hon G, Chandonia JM, Brenner SE: **Weblogo: A sequence logo generator.** *Genome Res* (2004) **14**(6):1188-1190.

120. Kabsch W: **Integration, scaling, space-group assignment and post-refinement.** *Acta crystallographica Section D, Biological crystallography* (2010) **66**(Pt 2):133-144.

121. McCoy AJ: **Solving structures of protein complexes by molecular replacement with phaser.** *Acta crystallographica Section D, Biological crystallography* (2007) **63**(Pt 1):32-41.

122. Emsley P, Cowtan K: **Coot: Model-building tools for molecular graphics.** *Acta crystallographica Section D, Biological crystallography* (2004) **60**(Pt 12 Pt 1):2126-2132.

123. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW, McCoy AJ *et al*: **The phenix software for automated determination of macromolecular structures.** *Methods* (2011) **55**(1):94-106.

124. Collaborative Computational Project N: **The ccp4 suite: Programs for protein crystallography.** *Acta crystallographica Section D, Biological crystallography* (1994) **50**(Pt 5):760-763.

125. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr., Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M: **The protein data bank: A computer-based archival file for macromolecular structures.** *Journal of molecular biology* (1977) **112**(3):535-542.

126. Novotny J, Bruccoleri R, Newell J, Murphy D, Haber E, Karplus M: **Molecular anatomy of the antibody binding site.** *J Biol Chem* (1983) **258**(23):14433-14437.

127. Liu Y, Pan Y, Xu Y: **Binding investigation of integrin alphavbeta3 with its inhibitors by spr technology and molecular docking simulation.** *J Biomol Screen* (2010) **15**(2):131-137.

128. Mena MA, Daugherty PS: **Automated design of degenerate codon libraries.** *Protein engineering, design & selection : PEDS* (2005) **18**(12):559-561.

129. Shimaoka M, Springer TA: **Therapeutic antagonists and conformational regulation of integrin function.** *Nat Rev Drug Discov* (2003) **2**(9):703-716.

130. Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaout MA: **Crystal structure of the extracellular segment of integrin alpha vbeta3 in complex with an arg-gly-asp ligand.** *Science* (2002) **296**(5565):151-155.

131. Takagi J, Petre BM, Walz T, Springer TA: **Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling.** *Cell* (2002) **110**(5):599-511.

132. 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.

133. Besenmatter W, Kast P, Hilvert D: **Relative tolerance of mesostable and thermostable protein homologs to extensive mutation.** *Proteins* (2007) **66**(2):500-506.