Title

A fully human scFv phage display library for rapid antibody fragment reformatting

Permalink

https://escholarship.org/uc/item/3t25g849

Journal

Protein Engineering Design and Selection, 28(10)

ISSN

1741-0126

Authors

Li, Keyu
Zettlitz, Kirstin A
Lipianskaya, Julia
et al.

Publication Date

2015-10-01

DOI

10.1093/protein/gzv024

Peer reviewed

OXFORD

Original Article

A fully human scFv phage display library for rapid antibody fragment reformatting

Keyu Li[1], Kirstin A. Zettlitz[1], Julia Lipianskaya[1], Yu Zhou[2], James D. Marks[2], Parag Mallick[3], Robert E. Reiter[4,5,6], and Anna M. Wu[1,6,*]

[1]Crump Institute for Molecular Imaging, Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at UCLA, 570 Westwood Plaza, Box 951770, Los Angeles, CA 90095, USA, [2]Department of Anesthesia and Perioperative Care, University of California, San Francisco, San Francisco General Hospital, 1001 Potrero Ave, Rm 3C-38, San Francisco, CA 94110, USA, [3]Canary Center for Cancer Early Detection, Stanford University, Palo Alto, CA 94304, USA, [4]Department of Urology, UCLA, Los Angeles, CA 90095, USA, [5]Molecular Biology Institute at UCLA, Los Angeles, CA 90095, USA, and [6]Jonsson Comprehensive Cancer Center at UCLA, Los Angeles, CA 90095, USA

*To whom correspondence should be addressed. E-mail: awu@mednet.ucla.edu
Edited by Andrew Bradbury

Abstract

Phage display libraries of human single-chain variable fragments (scFvs) are a reliable source of fully human antibodies for scientific and clinical applications. Frequently, scFvs form the basis of larger, bivalent formats to increase valency and avidity. A small and versatile bivalent antibody fragment is the diabody, a cross-paired scFv dimer ($\sim$55 kDa). However, generation of diabodies from selected scFvs requires decreasing the length of the interdomain scFv linker, typically by overlap PCR. To simplify this process, we designed two scFv linkers with integrated restriction sites for easy linker length reduction (17-residue to 7-residue or 18-residue to 5-residue, respectively) and generated two fully human scFv phage display libraries. The larger library (9×10^9 functional members) was employed for selection against a model antigen, human N-cadherin, yielding novel scFv clones with low nanomolar monovalent affinities. ScFv clones from both libraries were reformatted into diabodies by restriction enzyme digestion and re-ligation. Size-exclusion chromatography analysis confirmed the proper dimerization of most of the diabodies. In conclusion, these specially designed scFv phage display libraries allow us to rapidly reformat the selected scFvs into diabodies, which can greatly accelerate early stage antibody development when bivalent fragments are needed for candidate screening.

Key words: Antibody fragment, diabody, N-cadherin, phage display, scFv

Introduction

Invented in 1980s, phage display technology has provided a robust approach for generating peptide affinity reagents *in vitro* by mimicking the selection and amplification strategies of the immune system (Smith, 1985; Parmley and Smith, 1988; Cwirla *et al.*, 1990). Shortly after the advent of this technology, a number of laboratories have extended the concept to the display and selection of small antibody fragments such as single-chain variable fragments (scFvs) and fragment antigen-binding (McCafferty *et al.*, 1990; Barbas *et al.*, 1991; Breitling *et al.*, 1991; Garrard *et al.*, 1991; Hoogenboom *et al.*, 1991), leading to a revolutionary new route for antibody discovery and development. Cloning of human antibody repertoires into the phage genome (Marks *et al.*, 1991) has also enabled the *in vitro* selection of fully human antibodies that are preferred for clinical applications. Currently, phage display technology has become a major source of human antibodies and has led to the development of therapeutic antibodies including

adalimumab (Humira™) and belimumab (Benlysta™) (Schirrmann et al., 2011).

In addition to intact full length antibodies composed of separate heavy and light chains, single-chain antibody fragments such as diabodies, minibodies and scFv-Fcs have drawn increasing interest for various diagnostic and therapeutic applications (Holliger and Hudson, 2005; Kenanova et al., 2005; Wu and Senter, 2005; Olafsen et al., 2006; Nimmagadda et al., 2010; Girgis et al., 2013). These fragments are built on the scFv platform: small (25–27 kDa) monovalent fragments composed of antibody V_H and V_L domains linked by a flexible linker (typically 15–20 aa residues). ScFvs typically produce well in bacterial systems and are the preferred format for many antibody phage display libraries (de Kruif et al., 1995; Sheets et al., 1998; Okamoto et al., 2004; Wajanarogana et al., 2006). Larger single-chain fragments add mass and function, including minibodies (dimeric scFv-C_H3 fusions; ~80 kDa) and scFvs fused to full Fc regions (scFv-Fc; ~110 kDa). The smallest bivalent fragment, diabody (50–55 kDa), is created when the linker in an scFv is shortened (3–10 residues) to induce dimerization (Holliger et al., 1993; Kortt et al., 1997; Atwell et al., 1999; Hudson and Kortt, 1999). Depending on goals and applications, researchers need to routinely reformat the selected scFvs into the aforementioned fragments. Using the incorporated restriction sites in most phage display libraries, it is relatively easy to reformat an scFv into a minibody or an scFv-Fc by subcloning. However, reformatting a selected scFv into a diabody requires a reduction in the length of the polypeptide linker, which is usually achieved by time-consuming overlap PCR (Shimazaki et al., 2008) (Fig. 1).

As simple, self-assembling bivalent antibody fragments, diabodies are readily produced in bacterial/microbial systems. Their small size and unique pharmacokinetic properties also make them attractive for applications such as nanoparticle conjugation (Barat et al., 2009; Girgis et al., 2013) and in vivo imaging (Santimaria et al., 2003;

Sundaresan et al., 2003; Robinson et al., 2005; Leyton et al., 2009; Eder et al., 2010; Li et al., 2014). Furthermore, biological effects of antibodies may depend on the cross-linking of targets on the cell surface, thus bivalent fragments are required for certain functional assays. Diabodies may provide a rapid path for evaluating antibody candidates in the early development process even if the final application requires an intact antibody. Given the broad applications of diabodies, a phage display library with a specially designed linker to rapidly convert scFvs into diabodies would accelerate the development process and save resources and time.

Here we describe two large naive human scFv phage display libraries built using different polypeptide linkers containing restriction sites that enable rapid linker length reduction through restriction enzyme digestion and re-ligation. Antibody selection from one of these libraries using N-cadherin (Ncad) as a model antigen has generated multiple positive candidate antibodies with promising binding properties and affinities. Multiple scFv clones from both libraries were reformatted into diabodies using the linker restriction sites, and purified proteins assessed by size-exclusion chromatography to evaluate the dimerization induced by the shorter linkers.

Materials and methods

Construction of the V_H and V_L libraries

The antibody variable genes were cloned from the preexisting fully human V_H and V_L libraries via PCR and incorporated into pHEN1 phage display vectors (Sheets et al., 1998) that have been modified to contain the customized 17aa-SSA linker or the 18aa-SX linker to build the V_H and V_L libraries. Figure 2 shows the pHEN1 map and the sequences of the traditional (G4S)3 linker (Huston et al., 1988; Sheets et al., 1998), the 17aa-SSA linker (GGSTGGGSGGGGST
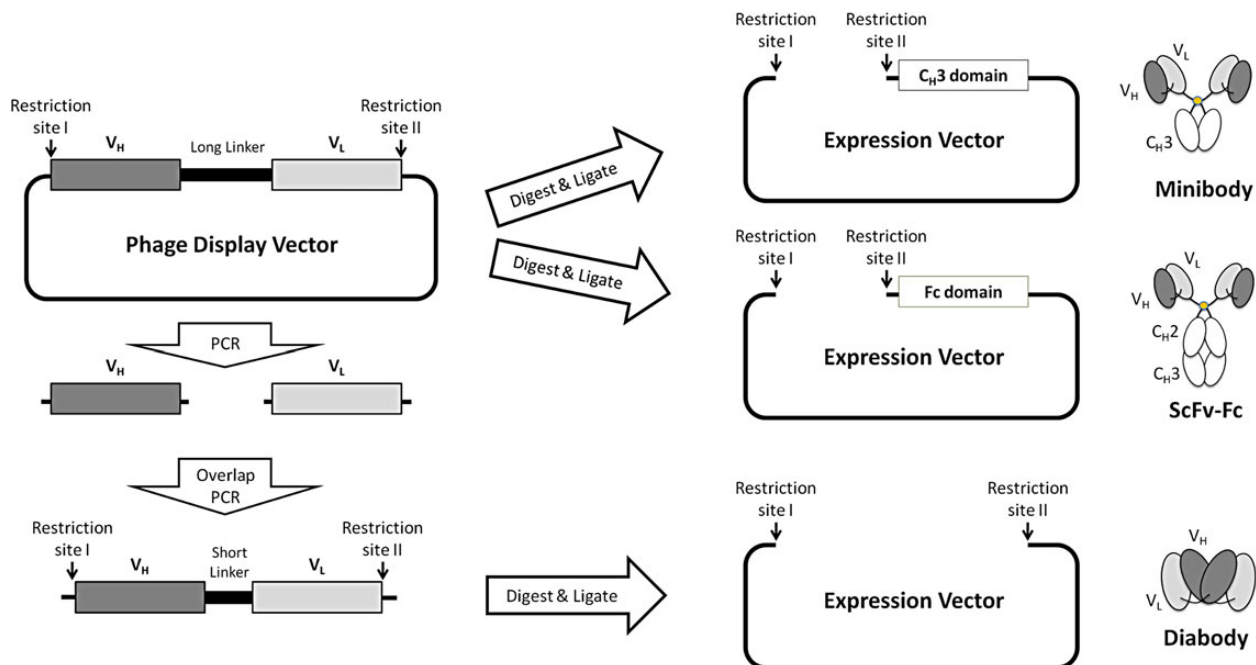
Fig. 1 Reformatting selected scFvs from common phage libraries. In most conventional scFv phage display libraries, the flanking restriction sites (I and II as shown here) can be utilized to rapidly make minibody and scFv-Fc constructs. However, to reformat an scFv into a diabody, the long linker in an scFv has to be shortened in order to induce dimerization. This is usually accomplished by a series of PCRs, which is far more complicated and time consuming, requiring careful design of multiple sets of primers.
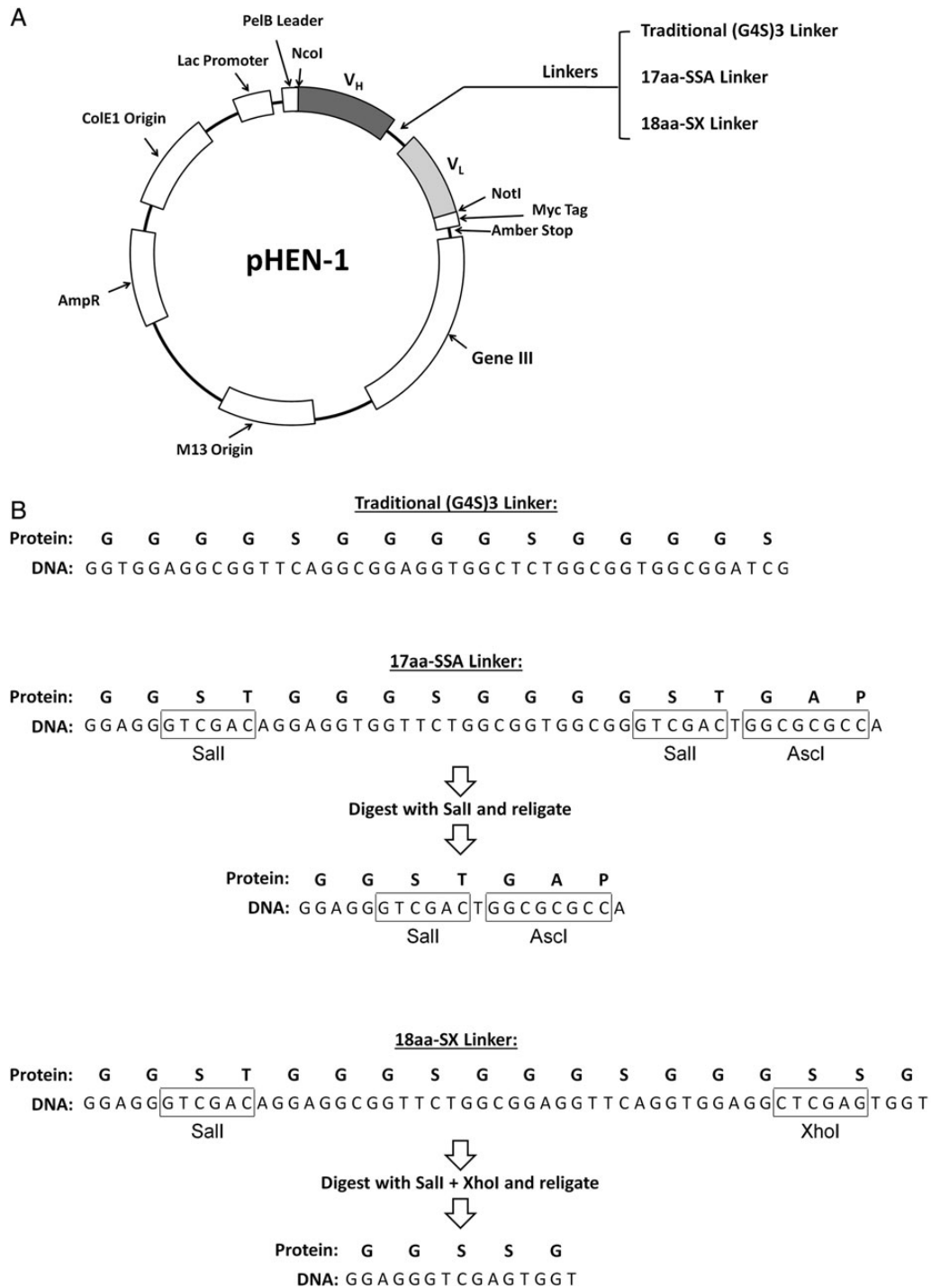
Fig. 2 (**A**) The map of major components of the pHEN1 phagemid vector. V_H and V_L genes are in the same reading frame with the PelB leader, the linker, the Myc tag, the amber stop codon and the phage minor coat protein gene III. Restriction sites NcoI and NotI can be used for subcloning of the scFv gene. (**B**) DNA and protein sequences of the linkers. Additional restriction sites were engineered into the 17aa-SSA linker and 18aa-SX linker for easier linker length reduction and library construction.

GAP) and the 18aa-SX linker (GGSTGGGSGGGSGGGSSG). The high-fidelity, proofreading DNA polymerase KOD Xtreme (Novagen) was used for the PCR amplifications, and restriction enzymes were purchased from New England Biolabs. The V gene PCR primers were designed based on previously published primer sequences (Marks *et al.*, 1991; Sheets *et al.*, 1998), with minor

modifications according to the IMGT database (http://www.imgt. org). Restriction sites were incorporated where needed. All the primers used for the library construction are summarized in Supplementary Table SI.

The 17aa-SSA linker DNA was produced by direct annealing and amplification using the RJH12-LSSA primer and the RVλ2-LSSA

primer. The V_H and V_L genes of one phage clone from the Sheets library (Sheets *et al.*, 1998) were separately PCR-amplified. The three fragments (V_H, V_L and linker) were joined together with overlap extension PCR using the V gene PCR primers, and then cloned into the pHEN1 vector using the NcoI and NotI sites. The 18aa-SX-linker was produced by direct annealing and amplification using the primer SXlinker-up and the primer SXlinker-down. It was then fused to a random V_L gene via overlap PCR, and then cloned into the 17aa-SSA-linker-pHEN1 clone via SalI and NotI sites to build the 18aa-SX-linker-pHEN1 vector.

The V_H and V_L libraries are then constructed using the scheme outlined in Fig. 3. A preexisting pCITE-V_H library with 2×10^8 members (Sheets *et al.*, 1998) was used as the V_H gene source. The V_H gene repertoire was PCR-amplified with the human V_H back primers and J_H forward primers. The amplified V_H DNA fragments were amplified again with HuVHBackNco primers and LSSA-JH primers to add the NcoI restriction site and the sequence overlapping the linker region. The 17aa-SSA linker was amplified from the 17aa-SSA-scFv-pHEN1 vector using LSSA-up primer and the RVλ2-LSSA primer. The gel-purified V gene repertoire was joined to the 17aa-SSA

linker via overlap extension PCR using the HuVHBackNco primers and the RVλ2-LSSA primer. The PCR product was then purified and digested with NcoI and AscI restriction enzymes overnight at 37°C. The 17aa-SSA-linker-pHEN1 vector DNA was extracted from the bacteria with the QIAprep Spin Miniprep Kit (Qiagen, cat#: 27106), and digested with of NcoI and AscI for 6 h before another hour of calf intestinal alkaline phosphatase (New England Biolabs) treatment at 37°C. The gel-purified V_H and vector DNA was ligated with T4 ligase (New England Biolabs) at 16°C overnight, and then gel purified and electroporated into electrocompetent TG1 bacteria (Lucigen). The V_L gene repertoire was cloned into the 17aa-SSA-linker-pHEN1 vector and the 18aa-SX-linker-pHEN1 vector to build the 17aa-SSA-V_L and the 18aa-SX-V_L libraries. V_L genes were PCR-amplified from a small previous scFv library with 3×10^7 members (Sheets *et al.*, 1998) using human V_κ, V_λ back primers and J_κ, J_λ forward primers. The restriction sites were then added to V_L genes by PCR with HuJκForNot and HuJλForNot primers in combination with LSSA-Vκ and LSSA-Vλ primers for the 17aa-SSA-V_L library construction, or in combination with SXLinker-Vκ and SXLinker-Vλ for the 18aa-SX-V_L library construction. The V_L genes and vectors were
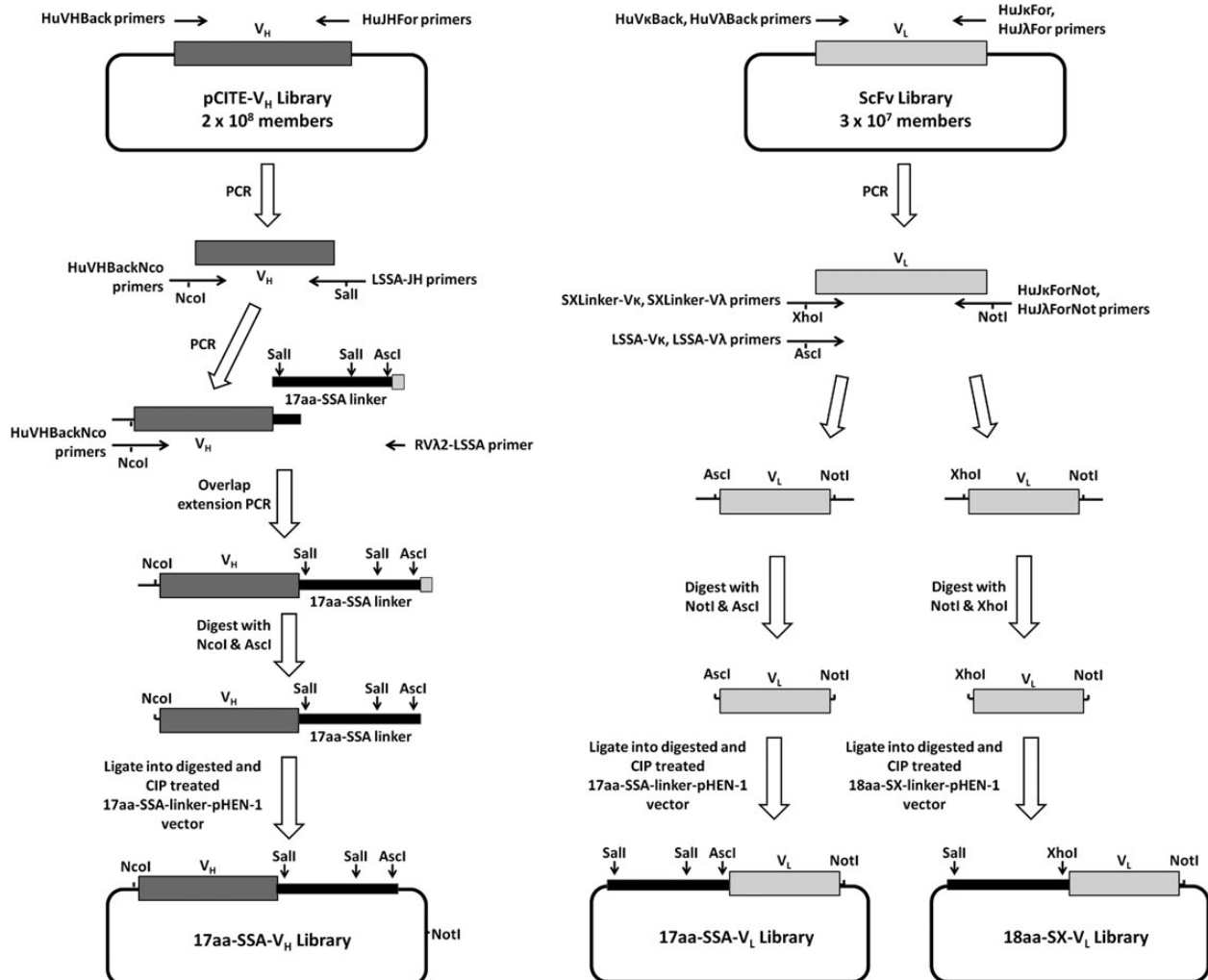
Fig. 3 Schematic outline of the construction of the V_H and V_L libraries. The library construction begins with two existing V gene libraries (Sheets *et al.*, 1998). The V genes were PCR-amplified and cloned into 17aa-SSA-linker-pHEN1 and 18aa-SX-linker-pHEN1 vectors to build the V_H and V_L libraries. CIP stands for calf intestinal alkaline phosphatase.

then digested, ligated and electroporated into TG1 bacteria to build the two V$_L$ libraries.

Construction of the scFv libraries

As outlined in Fig. 4, the 17aa-SSA-scFv phage display library and the 18aa-SX-scFv phage display library were constructed via V gene shuffling of the V$_H$ and V$_L$ libraries. The plasmids of the V$_H$ and V$_L$ libraries were extracted from the bacteria stock with the QIAprep Spin Miniprep Kit, and then digested with the enzymes as described in Fig. 4. The V$_L$ gene repertoire was gel purified and ligated into the digested pHEN1-V$_H$ vector. The ligated DNA is then gel purified and electroporated into the TG1 bacteria to build the two scFv libraries. Random scFv clones from the final transformation were sequenced with M13R primer (5′-CAGGAAACAGCTATGACCATG-3′, priming to the upstream of the scFv gene) and fdseq primer (5′-GAAT TTTCTGTATGAGG-3′, priming to the gene III region) from both directions to determine the percentage of functional clones. Sequences of random functional clones from each library were analyzed using the abYsis system (http://www.bioinf.org.uk/abysis/) to determine the germline V gene usage.

Phage library selection

The 17aa-SSA-scFv library was used for the selections against the human Ncad protein (extracellular domain fused to C-terminal polyhistidine tag, Sino Biological, cat#: 11039-H08H) using previously published methods (Marks and Bradbury, 2004). Specific phage displayed scFvs were selected by Ncad proteins absorbed to immunotubes (Nunc). After three rounds of selections, random clones were picked and sequenced to identify distinct clones, and their binding to the recombinant Ncad proteins were evaluated by phage ELISA

according to previously published methods (Marks and Bradbury, 2004).

Scfv subcloning, reformatting and production

Four anti-Ncad 17aa-SSA-scFvs and four random 18aa-SX-scFvs were subcloned into the expression vector pSYN1 (Schier *et al.*, 1995; Bai *et al.*, 2003) via NcoI and NotI site to fuse to a C-terminal sequence containing a Myc tag and a 6-his tag (AAAEQKLISEED LNGAAHHHHHH, 'AAA' is the NotI site), and then transformed into chemically competent TG1 bacteria for protein production. To reformat the scFvs into diabodies, the pSYN1-17aa-scFv and pSYN1-18aa-scFv plasmids were digested either with SalI alone (for 17aa-SSA linker) or with SalI and XhoI (for 18aa-SX linker). The digested plasmids were spin column purified and then re-ligated with T4 ligase. After the digestion and re-ligation, the 17-residue or 18-residue linkers were shortened to 7-residue or 5-residue linkers, respectively. These pSYN1-7aa-diabody and pSYN1-8aa-diabody plasmids were also transformed into TG1 bacteria. The expression of scFvs and diabodies was induced by 0.5 mM IPTG overnight at room temperature. The expressed proteins were then released from bacteria with BugBuster Master Mix (Novagen), and purified by immobilized metal affinity chromatography.

Affinity determination by ELISA

The affinities of the anti-Ncad 17aa-SSA-scFvs were determined by ELISA. The human Ncad protein (Sino Biological, cat#: 11039-H08H) was coated to a 96-well ELISA plate (3 μg/mL, 100 μl/well, overnight incubation at 4°C). After washing and blocking with 2% milk-PBS, the plate was incubated with serial dilutions of scFvs (100 μl/well; the scFvs were diluted in 2% milk-PBS). After
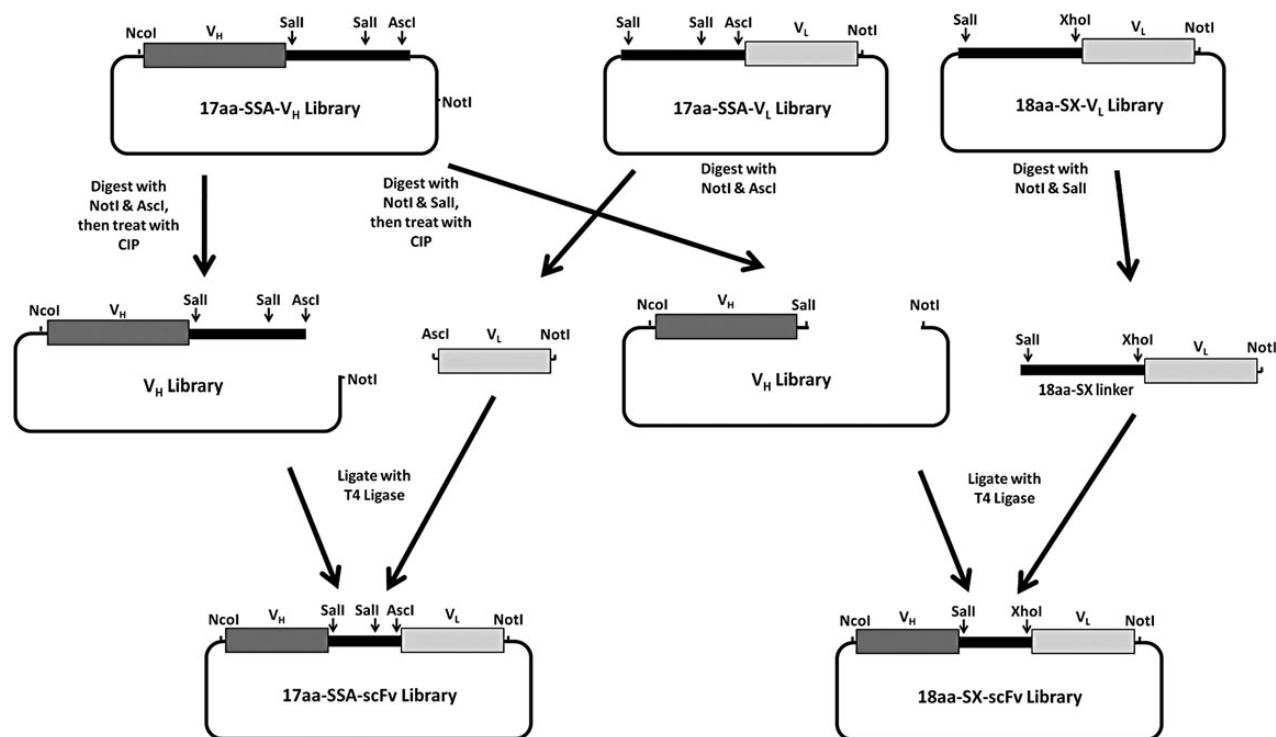
Fig. 4 Upon the completion of the V$_H$ and V$_L$ libraries, the scFv libraries were constructed via V gene shuffling. The V$_L$ gene repertoire was cut out from the V$_L$ libraries using NotI enzyme in combination with either AscI or SalI enzyme, and then cloned into the V$_H$ library to build the scFv libraries with the two different linkers. CIP stands for calf intestinal alkaline phosphatase.

equilibrium, the unbound scFvs were washed off with PBS and the bound scFvs were detected using a 1:1000 diluted rabbit anti-Myc antibody incubation followed by incubation with 1:1000 diluted alkaline phosphatase conjugated goat anti-rabbit antibody. The final development was performed using the PNPP substrate (Thermo Scientific) according to the recommended protocol. Each concentration was performed in duplicate. A_{405} of each well was read and the binding curves were analyzed using Graphpad Prism to calculate dissociation constants using the 'one site-specific binding' model.

Size-exclusion chromatography

In order to evaluate the linker length-dependent dimerization, the native structural sizes of the antibody fragments were evaluated by size-exclusion chromatography. Approximately 20 µg of the purified antibody fragments were applied to a Superdex 75 column (GE Healthcare) using PBS as the running buffer at the flow rate of 0.5 ml/min. The curves were exported and analyzed using Graphpad Prism.

Results

Library construction

The V_H and V_L gene repertoires were first cloned from preexisting V_H library (2×10^8 members) and V_L library (3×10^7 members) (Sheets *et al.,* 1998) to build new V_H and V_L libraries with customized linkers. The 17aa-SSA-V_H library was constructed with multiple electroporations of a total of 77 µg of ligated DNA. The final 17aa-SSA-V_H library has 5.3×10^8 transformants, which represents more than 2-fold coverage of the original V_H library. The 17aa-SSA-V_L and the 18aa-SX-V_L libraries were built with electroporations of 4.9 and 10.0 µg of ligated DNA, respectively, and the total numbers of transformants were 9.4×10^7 and 7.1×10^7, also covering more than two times the original V_L library.

By V gene shuffling, two large fully human scFv phage display libraries with customized linkers were successfully constructed. The numbers of transformants in the final 17aa-SSA-scFv library and 18aa-SX-scFv library were 1.8×10^{10} and 4.1×10^9, using 65 and 83 µg of ligated DNA, respectively. However, sequencing results of the random clones from the 17aa-SSA-scFv library and 18aa-SX-scFv library show that in both libraries, about half of the clones contain functional scFv sequences, with the other half showing either no sequencing signal, no scFv insert or non-functional sequences with premature stop codons. Therefore, the effective library sizes should be ~9×10^9 and 2×10^9, respectively.

Thirty functional clones from each scFv library were analyzed with the abYsis on line tool (http://www.bioinf.org.uk/abysis/), and the germline genes and subgroups were assigned according to the IMGT database (http://www.imgt.org) (Fig. 5). The V_H gene usage showed an over-representation of the IGHV3 subgroup (IMGT classification) and the IGHV3-23 germline gene (DP-47), which is very similar to the Sheets library (Sheets *et al.,* 1998) using the same V_H gene source library, probably reflecting the bias in the original pCITE-V_H library. The V_L gene usage seemed to be more variable. IGKV4-01 and IGLV1-40 were over-represented in the 17aa-SSA-scFv library while IGLV3-19 was the preferred germline gene in the 18aa-SX-scFv library. This is probably because the V_L gene libraries were created with fewer ligations and electroporations, and the variations in the reaction conditions resulted in different V_L gene usage in the two V_L libraries. However, such variation probably has a limited effect on the

final scFv libraries because the V_H genes contribute the primary diversity in an scFv library.

Anti-Ncad scFvs

The 17aa-SSA-scFv library was used for a selection against the Ncad extracellular domain. After three rounds of selection, 96 random clones were sequenced. The sequencing result revealed 35 distinct functional sequences, 15 of which have two or more duplicates. Twenty-eight of the distinct functional phage clones showed at least twice the background signal in the phage ELISA assay. Four of the scFvs have been successfully expressed and purified (A1, A6, D6 and G6). The affinities of these four purified scFvs were estimated by ELISA (Fig. 6). A1, A6 and D6 showed low nanomolar affinities ranging from 1.5 to 28 nM, confirming the diversity of the 17aa-SSA-scFv library. The G6 scFv failed to bind to the recombinant Ncad protein in the ELISA test. Its binding in the phage selection and phage ELISA could be related to the stickiness of the phage.

Rapid reformatting of scFvs to diabodies

scFv clones from the two phage libraries can be conveniently reformatted into diabodies by shortening the linker length using the incorporated restriction sites. The aforementioned four anti-Ncad 17aa-linker scFvs (A1, A6, D6, G6) were reformatted to 7aa-linker diabodies through SalI digestion and re-ligation. Four random 18aa-linker scFvs from the 18aa-SX-scFv library were reformatted to 5aa-linker diabodies through SalI–XhoI double digestion and re-ligation. The scFvs and diabodies were then produced and analyzed by size-exclusion chromatography on a Superdex 75 column (Fig. 7). Most of the diabodies with 7aa- or 5aa-linker formed only dimers except the 7aa-linker A6 diabody, which formed a mixture of monomer and dimer. Despite having a long linker, the 17aa-linker D6 scFv also formed a mixture of monomer and dimer. Other than these two exceptions, the customized linkers performed well in most cases: the 17aa- and 18aa-linker can provide enough flexibility to allow the formation of self-pairing scFv monomers, while the 7aa- and 5aa-linker usually induce the formation of dimers.

Discussion

In this work, we successfully constructed two novel fully human scFv libraries with two different customized linkers. Unlike the traditional 15-residue (G4S)3 scFv linker introduced by Huston *et al.* (1988), which is used in many scFv phage display libraries (de Kruif *et al.,* 1995; Sheets *et al.,* 1998; Okamoto *et al.,* 2004; Wajanarogana *et al.,* 2006), these customized linkers have various restriction sites incorporated into the DNA sequences to be employed for rapid linker length reduction via restriction digest and ligation. The multimerization of scFvs depends on the different lengths of the linkers. Generally, long flexible linkers allow self-pairing to form scFv monomers, and shorter linkers will induce the formation of multimers. In one example, Kortt, Atwell and their colleagues studied the effect of the linker length on the anti-neuraminidase antibody NC10 scFv. While NC10 scFv with 15-residue linker forms a monomer, the 3–10-residue linkers induced dimerization, and the 0–2-residue linkers resulted in trimers (Kortt *et al.,* 1997; Atwell *et al.,* 1999). Whitlow *et al.* also reported an scFv dimer with a 12-residue linker (Whitlow *et al.,* 1994). However, such requirement on linker length is not universal. The multimerization is also dependent on the different structures of individual scFv clones and the properties of different
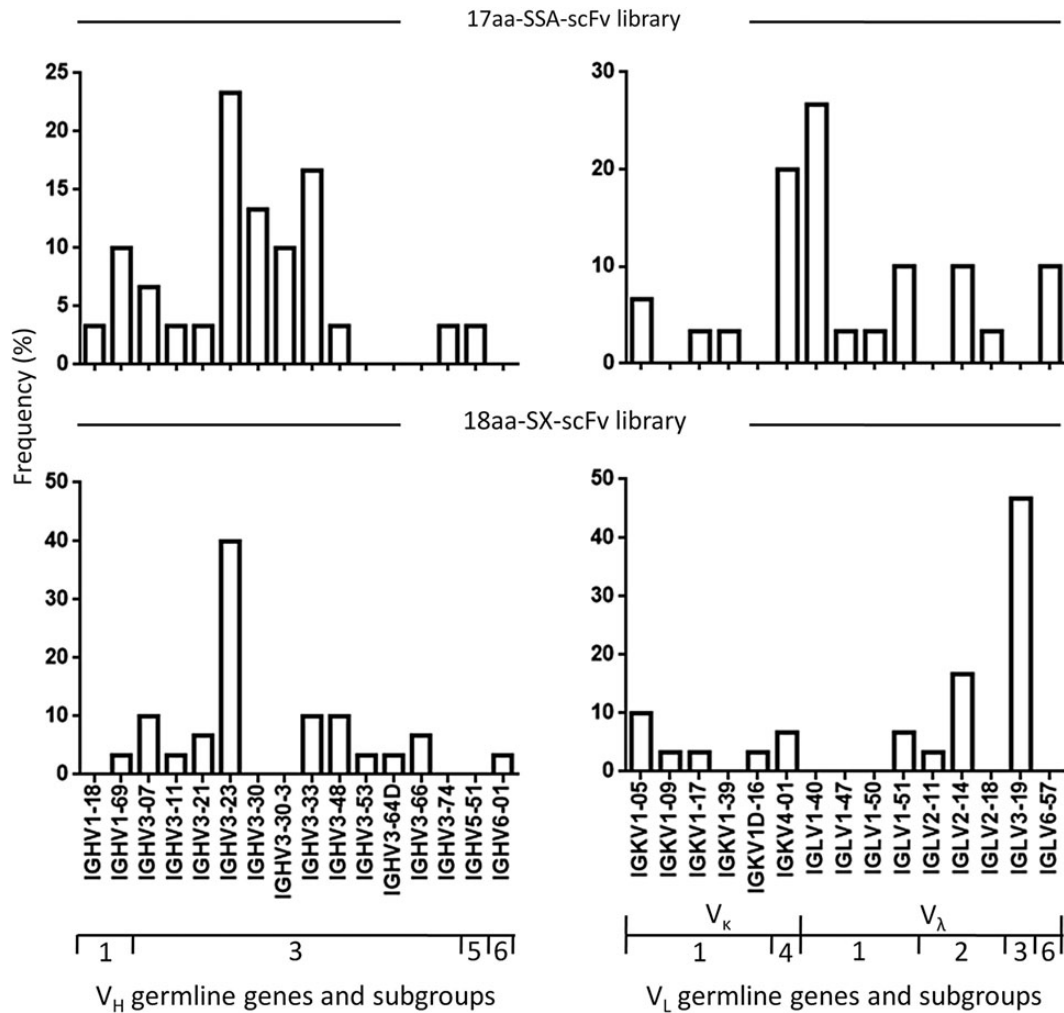
Fig. 5 V gene usage of the 17aa-SSA-scFv library and the 18aa-SX-scFv library. Thirty functional clones from each scFv library were analyzed, and the germline genes and subgroups were assigned. The frequency of specific V_H and V_L genes is listed. The V genes are classified into subgroups based on the IMGT (http://www.imgt.org) classification. (V_H subgroups: IGHV1, IGHV3, IGHV5, IGHV6; V_κ subgroups: IGKV1, IGKV4; V_λ subgroups: IGLV1, IGLV2, IGLV3, IGLV6.)
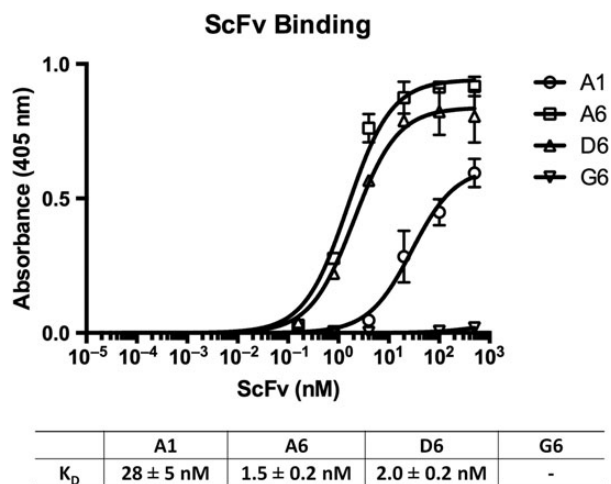

Fig. 6 The affinities of the anti-Ncad scFvs were estimated by ELISA on immobilized Ncad recombinant protein. Each concentration was performed in duplicate. The binding curves were analyzed in Graphpad Prism to calculate dissociation constants using the 'one site-specific binding' model.

linkers. In one study by Le Gall et al., an anti-CD19 scFv, HD37, formed a mixture of monomer, dimer and tetramer, even with an 18-residue linker (Le Gall et al., 1999). The linker length requirement also changes when the orientation of the V domains changes from V_H-V_L to V_L-V_H, because the distance between the C-terminus of V_L and the N-terminus of V_H is larger than the one between the C-terminus of V_H and the N-terminus of V_L (Huston et al., 1991; Arndt et al., 1998; Dolezal et al., 2000; Kortt et al., 2001). According to experience in our laboratory, 15–18-residue linkers with small, hydrophilic amino acids such as glycine, serine and threonine work well for scFv monomers in most cases, and 5–8-residue linker can usually induce successful dimerization (Sundaresan et al., 2003; Shimazaki et al., 2008; Sirk et al., 2008; Leyton et al., 2009; Girgis et al., 2011). We expect an scFv phage display library with a flexible linker that can facilitate easy linker length change in this range can accommodate the needs of most of the monomers and dimers. However, considering the enormous number of different scFv clones in one library, it will be difficult to say such linker length change will be perfect for every clone.

In order to keep the maximal diversity during V gene cloning and shuffling, a recognition site of the rare-cutter enzyme AscI was
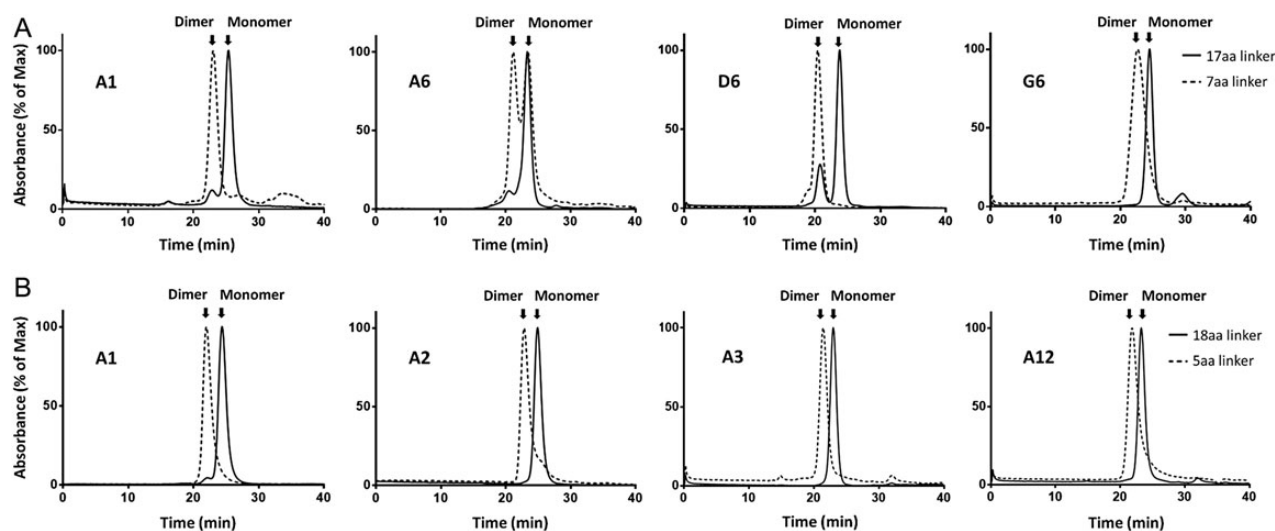
Fig. 7 (**A**) Four pairs of anti-Ncad antibody fragments were analyzed by Superdex 75 size-exclusion chromatography. The 17aa-linker and the 7aa-linker worked best for the A1 and G6 clones. For the A6 clone, the 7aa-linker diabody formed a mixture of monomer and dimer. For the D6 clone, the 17aa-linker scFv formed a mixture of monomer and dimer. (**B**) Four random clones from the 18aa-SX-scFv library were reformatted, expressed and purified. The 5aa-linker successfully induced dimerization for all these clones.

incorporated into the 17aa-SSA linker for cloning purpose. AscI has no known cuts in the germline V genes, and its 8-bp recognition site is less likely to be found in the mature antibody DNA. However, this leads to some less common residues in the protein sequence (GGSTGGGSGGGGSTGAP). In order to provide enough flexibility to allow self-pairing, people usually choose small, hydrophilic amino acids such as glycine, serine and threonine for scFv linkers. (G4S)3 linker is an example of such principle. On the other hand, due to its unique structure, prolines are usually used in more rigid linkers where the inter domain interactions are to be avoided (Zhao *et al.*, 2008). Proline-rich sequences such as $(XP)_n$ are commonly seen in natural peptides displaying relatively elongated and rigid conformations (Chen *et al.*, 2013). Considering the possible stiffness caused by the proline, the linker length was set to the longer side, using a 17-residue linker (GGSTGGGSGGGGSTGAP) for monomers and a 7-residue linker (GGSTGAP) for dimers.

The completed 17aa-SSA-scFv library was used for selection on Ncad, and successfully generated multiple positive hits. Some of the selected scFvs showed high affinities with dissociation constant as low as 1.5 nM, consistent with the estimated diversity of the library. These 17aa-linker scFvs were quickly reformatted to 7aa-linker diabodies, and the purified proteins were analyzed by size-exclusion chromatography. The 7aa-linker-induced successful dimerization of three of four clones, but one of the 7aa-linker diabody (A6) formed a mixture of monomers and dimers, probably partially contributed to its high apparent affinity ($K_D = 1.5 \pm 0.2$ nM). Although we expected certain specific clones may have special requirements for linker length and flexibility, it was quite surprising to see the formation of monomers with a linker length as short as seven amino acid residues. Previously our laboratory created a diabody construct with an 8-residue linker (GGGGSGGG) that formed monomers instead of dimers (McCabe *et al.*, 2012; Tavare *et al.*, 2014). It was still unexpected that a 7aa-linker can provide enough bridging to allow self-pairing, particularly since this 7aa-linker contains a proline residue. It remains to be investigated whether these monomers are functional, self-pairing

scFvs or just non-functional monomers that fail to form dimers due to certain unique conformations induced by the relatively rigid proline residue. On the other hand, the 17aa-linker D6 scFv also formed a monomer–dimer mixture. This demonstrated the complicated nature of linker length-dependent multimerizaiton. As mentioned before, the specific requirement on the linker length is also highly dependent on the nature of the antibody clones. There are cases that scFv clones with 12-residue or even 18-residue linkers still formed dimers (Whitlow *et al.*, 1994; Le Gall *et al.*, 1999). Multimerization can be induced by a shorter linker, but long linkers cannot forbid scFvs from multimerization.

Given these complications, we then constructed the second library with the 18aa-SX linker. For the 18aa-SX linker, we adopted a more traditional and flexible design, using only glycines, serines and threonines. It also incorporates a larger length difference between monomers (18-residue: GGSTGGGSGGGSGGGSSG) and dimers (5-residue: GGSSG). We assessed four random clones from this new library, and all proteins behaved as expected. However, considering the limited number of tested clones, it is difficult to tell whether the better performance is caused by better linker design or is simply due to the innate properties of the individual antibodies. Exactly how well these linker designs will work for diverse antibody sequences remains to be verified in the future applications of these libraries.

The purpose of building these libraries with novel linkers was to promote convenient linker length reduction to rapidly generate diabody constructs from scFvs. The linker designs worked well for this purpose. The 17aa-SSA linker or the 18aa-SX linker can be shortened to 7 residues or 5 residues by simple digestion, spin column purification and re-ligation. There is no need for PCR, gel purification or primer designing. Such specialized linker design may be less important for occasional applications when many laboratories routinely order synthetic genes for desired proteins, given the falling cost of DNA synthesis. However, in a situation where a large number of lead candidates need to be screened in the diabody format due to its size, avidity or pharmacokinetic properties, these novel libraries present a unique approach which can save

time and effort and greatly accelerate the development process of diabodies and other engineered antibody fragments.

Supplementary data

Supplementary data are available at *PEDS* online.

Funding

References

Arndt,K. M., Muller,K. M. and Pluckthun,A. (1998) *Biochemistry,* 37 (37), 12918–26.

Atwell,J. L., Breheney,K. A., Lawrence,L. J., McCoy,A. J., Kortt,A. A. and Hudson,P. J. (1999) *Protein Eng,* 12 (7), 597–604.

Bai,J., Sui,J., Zhu,R. Y., Tallarico,A. S., Gennari,F., Zhang,D. and Marasco,W. A. (2003) *J Biol Chem,* 278 (3), 1433–42.

Barat,B., Sirk,S. J., McCabe,K. E., *et al.* (2009) *Bioconjug Chem,* 20 (8), 1474–81.

Barbas,C. F., 3rd, Kang,A. S., Lerner,R. A. and Benkovic,S. J. (1991) *Proc Natl Acad Sci U S A,* 88 (18), 7978–82.

Breitling,F., Dubel,S., Seehaus,T., Klewinghaus,I. and Little,M. (1991) *Gene,* 104 (2), 147–53.

Chen,X., Zaro,J. L. and Shen,W. C. (2013) *Adv Drug Deliv Rev,* 65 (10), 1357–69.

Cwirla,S. E., Peters,E. A., Barrett,R. W. and Dower,W. J. (1990) *Proc Natl Acad Sci U S A,* 87 (16), 6378–82.

de Kruif,J., Boel,E. and Logtenberg,T. (1995) *J Mol Biol,* 248 (1), 97–105.

Dolezal,O., Pearce,L. A., Lawrence,L. J., McCoy,A. J., Hudson,P. J. and Kortt, A. A. (2000) *Protein Eng,* 13 (8), 565–74.

Eder,M., Knackmuss,S., Le Gall,F., Reusch,U., Rybin,V., Little,M., Haberkorn, U., Mier,W. and Eisenhut,M. (2010) *Eur J Nucl Med Mol Imaging,* 37 (7), 1397–407.

Garrard,L. J., Yang,M., O'Connell,M. P., Kelley,R. F. and Henner,D. J. (1991) *Biotechnology (N Y),* 9 (12), 1373–7.

Girgis,M. D., Kenanova,V., Olafsen,T., McCabe,K. E., Wu,A. M. and Tomlinson,J. S. (2011) *J Surg Res,* 170 (2), 169–78.

Girgis,M. D., Federman,N., Rochefort,M. M., McCabe,K. E., Wu,A. M., Nagy, J. O., Denny,C. and Tomlinson,J. S. (2013) *J Surg Res,* 185 (1), 45–55.

Holliger,P. and Hudson,P. J. (2005) *Nat Biotechnol,* 23 (9), 1126–36.

Holliger,P., Prospero,T. and Winter,G. (1993) *Proc Natl Acad Sci U S A,* 90 (14), 6444–8.

Hoogenboom,H. R., Griffiths,A. D., Johnson,K. S., Chiswell,D. J., Hudson,P. and Winter,G. (1991) *Nucleic Acids Res,* 19 (15), 4133–7.

Hudson,P. J. and Kortt,A. A. (1999) *J Immunol Methods,* 231 (1–2), 177–89.

Huston,J. S., Mudgett-Hunter,M., Tai,M. S., McCartney,J., Warren,F., Haber, E. and Oppermann,H. (1991) *Methods Enzymol,* 203, 46–88.

Huston,J. S., Levinson,D., Mudgett-Hunter,M., *et al.* (1988) *Proc Natl Acad Sci U S A,* 85 (16), 5879–83.

Kenanova,V., Olafsen,T., Crow,D. M., *et al.* (2005) *Cancer Res,* 65 (2), 622–31.

Kortt,A. A., Dolezal,O., Power,B. E. and Hudson,P. J. (2001) *Biomol Eng,* 18 (3), 95–108.

Kortt,A. A., Lah,M., Oddie,G. W., *et al.* (1997) *Protein Eng,* 10 (4), 423–33.

Le Gall,F., Kipriyanov,S. M., Moldenhauer,G. and Little,M. (1999) *FEBS Lett,* 453 (1–2), 164–8.

Leyton,J. V., Olafsen,T., Sherman,M. A., Bauer,K. B., Aghajanian,P., Reiter,R. E. and Wu,A. M. (2009) *Protein Eng Des Sel,* 22 (3), 209–16.

Li,K., Tavare,R., Zettlitz,K. A., Mumenthaler,S. M., Mallick,P., Zhou,Y., Marks,J. D. and Wu,A. M. (2014) *Mol Cancer Ther,* 13 (11), 2607–17.

Marks,J. D. and Bradbury,A. (2004) in Benny,K. C. Lo (ed.), *Antibody engineering : methods and protocols* (Totowa, N.J.: Humana Press), 161 p.

Marks,J. D., Hoogenboom,H. R., Bonnert,T. P., McCafferty,J., Griffiths,A. D. and Winter,G. (1991) *J Mol Biol,* 222 (3), 581–97.

McCabe,K. E., Liu,B., Marks,J. D., Tomlinson,J. S., Wu,H. and Wu,A. M. (2012) *Mol Imaging Biol,* 14 (3), 336–47.

McCafferty,J., Griffiths,A. D., Winter,G. and Chiswell,D. J. (1990) *Nature,* 348 (6301), 552–4.

Nimmagadda,S. V., Aavula,S. M., Biradhar,N., Rao,V. S., Shanmugham,R., Chandran,D., Thirumeni,N., Singanallur,N. B. and Villuppanoor,S. A. (2010) *Clin Vaccine Immunol,* 17 (8), 1261–8.

Okamoto,T., Mukai,Y., Yoshioka,Y., *et al.* (2004) *Biochem Biophys Res Commun,* 323 (2), 583–91.

Olafsen,T., Kenanova,V. E. and Wu,A. M. (2006) *Nat Protoc,* 1 (4), 2048–60.

Parmley,S. F. and Smith,G. P. (1988) *Gene,* 73 (2), 305–18.

Robinson,M. K., Doss,M., Shaller,C., Narayanan,D., Marks,J. D., Adler,L. P., Gonzalez Trotter,D. E. and Adams,G. P. (2005) *Cancer Res,* 65 (4), 1471–8.

Santimaria,M., Moscatelli,G., Viale,G. L., *et al.* (2003) *Clin Cancer Res,* 9 (2), 571–9.

Schier,R., Marks,J. D., Wolf,E. J., *et al.* (1995) *Immunotechnology,* 1 (1), 73–81.

Schirrmann,T., Meyer,T., Schutte,M., Frenzel,A. and Hust,M. (2011) *Molecules,* 16 (1), 412–26.

Sheets,M. D., Amersdorfer,P., Finnern,R., *et al.* (1998) *Proc Natl Acad Sci U S A,* 95 (11), 6157–62.

Shimazaki,K., Lepin,E. J., Wei,B., *et al.* (2008) *Clin Cancer Res,* 14 (22), 7367–77.

Sirk,S. J., Olafsen,T., Barat,B., Bauer,K. B. and Wu,A. M. (2008) *Bioconjug Chem,* 19 (12), 2527–34.

Smith,G. P. (1985) *Science,* 228 (4705), 1315–7.

Sundaresan,G., Yazaki,P. J., Shively,J. E., *et al.* (2003) *J Nucl Med,* 44 (12), 1962–9.

Tavare,R., Wu,W. H., Zettlitz,K. A., Salazar,F. B., McCabe,K. E., Marks,J. D. and Wu,A. M. (2014) *Protein Eng Des Sel,* 27 (10), 317–24.

Wajanarogana,S., Prasomrothanakul,T., Udomsangpetch,R. and Tungpradabkul,S. (2006) *Biotechnol Appl Biochem,* 44 (Pt 1), 55–61.

Whitlow,M., Filpula,D., Rollence,M. L., Feng,S. L. and Wood,J. F. (1994) *Protein Eng,* 7 (8), 1017–26.

Wu,A. M. and Senter,P. D. (2005) *Nat Biotechnol,* 23 (9), 1137–46.

Zhao,H. L., Yao,X. Q., Xue,C., Wang,Y., Xiong,X. H. and Liu,Z. M. (2008) *Protein Expr Purif,* 61 (1), 73–7.