Published in final edited form as:

Mol Immunol. 2010 January ; 47(4): 912–921. doi:10.1016/j.molimm.2009.09.039.

A large human domain antibody library combining heavy and light chain CDR3 diversity

Weizao Chen^a, Zhongyu Zhu^{a,b}, Yang Feng^a, and Dimiter S. Dimitrov^{a,*}

^aProtein Interactions Group, Center for Cancer Research Nanobiology Program, National Cancer Institute (NCI)-Frederick, National Institutes of Health (NIH), Frederick, MD 21702-1201, USA

^bBasic Research Program, Science Applications International Corporation-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA

Abstract

Domain antibodies (dAbs) are promising candidate therapeutics and diagnostics. Efficient selection of novel potent dAbs with potential for clinical utility is critically dependent on the library diversity and size, and the scaffold stability. We have previously constructed a large (size $\sim 2.5 \times 10^{10}$) dAb library by grafting human antibody heavy chain complementarity determining regions (CDRs) 2 and 3 (H2s, H3s) into their cognate positions in a human heavy chain variable domain (VH) scaffold and mutagenizing the CDR1 (H1). High-affinity binders against some antigens were selected from this library but panning against others was not very successful likely due to limited diversity. We have hypothesized that by grafting highly variable, both in length and composition, human CDRs into non-cognate positions, the dAb library diversity could be significantly increased and the library would allow for more efficient selection of high-affinity antibodies against some targets. To test this hypothesis we designed a novel type of dAb library containing CDRs in non-cognate positions. It is based on our previous library where H1 was replaced by a library of human light chain CDR3s (L3s) thus combining three most diversified fragments (L3, H3 and H2) in one VH scaffold. This large (size $\sim 10^{10}$) phage-displayed library was highly diversified as determined by analyzing the sequences of 126 randomly selected clones. Novel high-affinity dAbs against components of the human insulin-like growth factor (IGF) system were selected from the new library that could not be selected from the previously constructed one. Most of the newly identified dAbs were highly soluble, expressible, monomeric and may have potential as candidate cancer therapeutics. The new library could be used not only for selection of such dAbs thus complementing existing libraries but also as a research tool for exploration of the mechanisms determining folding and stability of human antibody domains.

Keywords

human domain antibody; library; phage display; light chain CDR3; grafting

*Corresponding author. Mailing address: Protein Interactions Group, CCRNP, CCR, NCI-Frederick, NIH, Bldg. 469, Rm. 150B, Frederick, MD 21702-1201, Phone: (301) 846-1352; Fax: (301) 846-5598; dimitr.dimitrov@nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Currently, almost all therapeutic antibodies (except ReoPro, Lucentis and Cimzia which are Fabs) approved by the U.S. Food and Drug Administration and the vast majority of those in clinical trials are full-size antibodies mostly in IgG1 format of about 150 kDa size (Dimitrov and Marks, 2009). A fundamental problem for such large molecules is their poor penetration into tissues (e.g., solid tumors) and poor or absent binding to functionally important regions on the surface of some molecules (e.g., the human immunodeficiency virus envelope glycoprotein) which are accessible by molecules of smaller size. Decreasing the size of the molecule dramatically, non-linearly, increases its penetration in tissues (Yokota *et al.*, 1992; Yokota *et al.*, 1993). Similarly, antibody size dependence of epitope accessibility can be highly nonlinear and some protein surface-exposed structures can be completely obstructed for full size antibodies. Therefore, a large amount of work especially during the last decade has been aimed at developing novel scaffolds of much smaller size (Holt *et al.*, 2003; Nygren and Skerra, 2004; Binz *et al.*, 2005; Hey *et al.*, 2005; Holliger and Hudson, 2005; Skerra, 2007; Kolmar and Skerra, 2008; Saerens *et al.*, 2008; Dimitrov, 2009; Dimitrov and Marks, 2009). Several scaffolds are derived from single antibody domains which are about 10-fold smaller than full size antibodies (Holt *et al.*, 2003; Saerens *et al.*, 2008; Dimitrov, 2009). Such scaffolds are stable, soluble, and easy to format, manufacture and express in microbial cell cultures.

One of the most advanced antibody domain scaffold is based on the single heavy chain variable domain (VH) (Ward *et al.*, 1989; Holt *et al.*, 2003; Chen *et al.*, 2008b). Binders derived from libraries based on mammalian VH or light chain variable domain (VL) scaffolds are called domain antibodies (dAbs). The human dAb, ART621 (targeting TNF α), is now in phase II clinical trials (www.arana.com). The efficient selection of high-affinity binders against various targets is critically dependent on the size and diversity of the antibody library. To minimize immunogenicity it is desirable to use fully human sequences for diversification. We have recently constructed a large (size, $\sim 2.5 \times 10^{10}$) phage-displayed dAb library by grafting naturally occurring human antibody heavy chain complementarity determining regions (CDRs) 2 and 3 (H2s, H3s) into a scaffold based on a newly identified fully human VH and randomly mutating four putative solvent-accessible residues in the CDR1 (H1) (Chen *et al.*, 2008b; Chen *et al.*, 2009). High-affinity dAbs were selected from this library against viral and human cancer-related antigens (Chen *et al.*, 2008a; Chen *et al.*, 2009).

In the absence of the VH-VL combinatorial diversity, the importance of constructing highly diversified libraries increases. The diversity of dAbs, however, is inherently limited by using only three CDRs compared to six CDRs of a conventional antibody. Remarkably, camelidae (and other species) naturally produce functional antibodies which are composed solely of heavy chains, designated heavy-chain antibodies or HCAs (Hamers-Casterman *et al.*, 1993). The antigen-binding site of the HCAs contains a single variable domain (referred to as V_HH). Compared to human VHs, V_HHs underwent remarkable changes in sequence and structure during evolution (Nguyen *et al.*, 2000). Most strikingly, an extra hypervariable region is present exclusively in the H1s of V_HHs and their H3s are, on average, longer than those of human VHs (17 versus 12 residues). These changes, together with others, dramatically increase the diversity of V_HH repertoire and enlarge the surface area interacting with antigens resulting in novel paratopes that are different from those of conventional antibodies (Nguyen *et al.*, 2000). In our previously constructed library (Chen *et al.*, 2008b) H1 was mutagenized but its length remained constant. To increase diversity, both in length and composition, we have hypothesized that H1s, which are relatively weakly diversified, could be replaced by more diverse non-cognate CDR3s, specifically those of the light chain (L3s), without significantly affecting the structural integrity of the scaffold. Here, we describe the generation and characterization of a fully human large (size, $\sim 10^{10}$) phage-displayed dAb library, which was constructed by combining naturally occurring human antibody H2s and H3s with L3s on the same scaffold. We identified novel

dAbs against human cancer-related proteins, including components of the human insulin-like growth factor (IGF) system, that exhibit high solubility, affinity, specificity, and were not selected by panning of our previously constructed library based on mutated H1s.

2. Materials and Methods

2.1. Amplification of L3 repertoire

Primers used for PCR amplification of gene fragments are described in Table 1. L3 repertoire was harvested from five different sources. Two of them were plasmid DNAs of a phage-displayed naive human Fab library (5×10^9 members) constructed from peripheral blood B cells of 10 healthy donors (Zhu *et al.*, 2006) and a phage-displayed immune human Fab library ($\sim 10^9$ members) constructed from bone marrow of 3 long-term nonprogressors whose sera exhibited the broadest and most potent human immunodeficiency virus 1 (HIV-1) neutralization among 37 HIV-infected individuals (Zhang *et al.*, 2003), respectively. To reduce PCR amplification bias, cDNAs, which were commercially purchased or produced in our group, were directly used as the other three sources. These include (a) a cDNA mixture from bone marrow of 10 healthy donors and fetal spleen of 24 spontaneously aborted male/female Caucasian fetuses; (b) a cDNA mixture from cord blood of two healthy babies, from which two naive Fab libraries (6×10^8 and 7.2×10^8 members, respectively) have been recently constructed; and (c) a cDNA mixture from peripheral blood B cells of 22 healthy donors, spleens of 3 healthy donors, and lymph nodes of 34 healthy donors which has also been used for construction of a large nonimmune Fab library (1.5×10^{10} members). To increase the efficiency for amplification of L3, full-length kappa (KL) and lambda (LL) light chains were first amplified from the five sources, respectively, as described previously (Zhu and Dimitrov, 2009) (Fig. 1A). The products from the same source were pre-mixed, gel-purified and then used as a template for L3 amplification. To maintain maximal diversity, we carried out five and six PCRs for each template separately to obtain kappa (KL3) and lambda (LL3) L3 repertoires, respectively (Fig. 1B), by using different primer combinations (sense primer L3F1 paired with different L3 antisense primers, respectively) (Table 1). PCR was performed in a volume of 50 μ l using High Fidelity PCR Master (Roche, Indianapolis, IN), 500 pM concentration of each primer, and 0.2 μ g of templates (purified full-length light chains) for 30 cycles (45 s at 94°C, 45 s at 50°C, and 1 min at 72°C). KL3 and LL3 products from different primer combinations for each template were first pooled, respectively, purified from 2% agarose gel using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA), quantified by reading the optical density (OD) at 260 nm (1 OD unit=50 μ g/ml) and then mixed at a molarity ratio of 3:1 for kappa and lambda. Finally, all products were mixed at a molarity ratio of 2:2:4:1:4 for the five templates in the order described above.

2.2. L3 mutagenesis

The purified L3 repertoire was subjected to the first-round mutagenesis by PCR (primer: L3F2 and L3R) for 30 cycles (45 s at 94°C, 45 s at 40°C, and 1 min at 72°C) (Fig. 1C). The product was purified and then re-amplified (primer: L3F3 and L3R) for the second-round mutagenesis under the same conditions (Fig. 1D).

2.3. Assembly of VHs

For assembly of chimeric VHs, H2 and H3 repertoires (primer: FR2F and HISR) as well as FR1 fragment (primer: FR1F and FR1R) were PCR amplified from m81 plasmid DNA (Chen *et al.*, 2008b), respectively (Fig. 1E). The L3 repertoire was then joined to FR1 fragment by overlapping PCR performed in a volume of 100 μ l by using both templates (in the same molarities) for 7 cycles (45 s at 94°C, 45 s at 55°C, and 1 min at 72°C) in the absence of primers and 15 additional cycles in the presence of primers (FR1F and L3R) (Fig. 1F). The entire VHs were formed by further annealing the products to H2-H3 fragments using overlapping PCR

under the same conditions with the extension primers FR1F and HISR appended with SfiI restriction sites (Fig. 1G).

2.4. Library preparation and analysis

Library preparation and analysis of library performance in terms of antibody diversity, biophysical properties and panning for antigen-specific binders were performed mainly as described previously (Chen et al., 2008b). For selection of antibodies specific to the ligand-binding site of the receptor (IGF-1R) of the human IGF system, bound phages were eluted with a mixture of the ligands (IGF-1 and IGF-2) each at 200 µg/ml concentration in 50 µl PBS after incubation of the libraries with IGF-1R. After 30 min of incubation with rotation at 50 rpm at room temperature, the suspension was centrifuged at 25,000g for 5min at 4°C and the eluted phages (supernatant) were rescued by infection of *Escherichia coli* (E. coli) TG1 cells for 45 min at 37°C. A phage library was prepared for the next round of panning.

2.5. Polyclonal phage-based ELISA (ppELISA)

PpELISA was performed by using Corning high-binding 96-well plates coated with 1 µg/ml of antigens and blocked with 3% non-fat dry milk in PBS (MPBS). Briefly, the microplate wells were inoculated with 50 µl each well of MPBS containing about 10¹⁰ PFU of pooled phage purified from each round of panning for 2 h at room temperature. After 4 washes with PBS containing 0.05% Tween 20 (PBST), bound phage was detected by adding 50 µl of 1:5000 diluted HRP-conjugated mouse anti-M13 antibody (Sigma, St. Louis, MO) to each well. Following incubation for 1 h at room temperature, the plates were washed 4 times with PBST and the assay was developed at 37°C with ABST substrate (Roche, Indianapolis, IN) and monitored at 405 nm.

2.6. Soluble expression-based monoclonal ELISA (semELISA)

SemELISA was performed by using Corning high-binding 96-well plates coated with 1 µg/ml of antigens and blocked with MPBS. Clones were randomly picked from the infected TG1 cells after several rounds of panning and each inoculated into 100 µl 2YT medium containing 100 µg/ml ampicillin and 0.2% glucose in 96-well plates. After the bacterial cultures reached an optical density at 600 nm (OD₆₀₀) of 0.5, isopropyl-β-D-thiogalactopyranoside (IPTG) at 0.5 mM (final concentration) was added to the medium, and the plates were further incubated at 30°C overnight in a shaker at 250 rpm. The microplate wells coated with antigens and blocked with milk were then inoculated with 50 µl each well of the supernatant of TG1 culture for 2 h at room temperature. After 4 washes with PBST, FLAG-tagged soluble VHs were detected by adding 50 µl of 1:5000 diluted HRP-conjugated anti-FLAG antibody (Sigma, St. Louis, MO) to each well and the assay was developed as described above.

2.7. Expression, purification and binding of soluble VHs

Soluble VHs were expressed, purified and their binding activity was measured as described previously (Chen et al., 2008b).

3. Results

3.1. Design and construction of a large human VH-based library with grafted human libraries of L3, H2 and H3

Previously (Chen et al., 2008b), we identified a completely natural human antibody VH domain (designated m0), which differs from the closest human antibody germline sequence VH3-23 by mutations in all framework regions (FRs) as well as in the CDRs. M0 was stable, highly soluble, monomeric, expressed at high levels in bacteria as an isolated single domain and therefore, was used as a framework scaffold to construct a large phage-displayed human dAB

library (designated m8l) by grafting in vivo-formed H2 and H3, and randomizing four putative solvent accessible residues in the H1 of m0 to alanine, aspartic acid, serine and tyrosine, which are residues most widely used in CDRs of human antibodies. However, the H1 diversity of m8l is relatively low due to the limited number (n=256) of theoretically possible sequences. In addition, the H1 length remains constant. To increase diversity we grafted human L3 repertoire into the H1 of m8l generating a combination of the most diversified CDRs of an antibody, L3, H2 and H3 in the same scaffold.

To amplify efficiently highly diverse L3 repertoires, we designed a new set of primers (Table 1) based on human germline VL sequences from the ImMunoGeneTics (IMGT) database (http://imgt.cines.fr/textes/IMGTrepertoire/Proteins/alleles/human/HuA1_list.html). The sense primers target the last three residues of the light chain FR3 (LFR3) that are highly conserved among different families of KL and LL, and the antisense primers – the first three residues of the J gene product (Fig. 2). Therefore, the PCR products contain L3 plus three additional residues from LFR3. In a separate experiment, we found that the amplification of L3 was inefficient when cDNA or plasmid DNA was directly used as a template. Therefore, full-length KL and LL were first amplified under standard conditions by using primers as described previously (Zhu and Dimitrov, 2009) (Fig. 1A), gel-purified and then used as templates for L3 amplification (Fig. 1B). The last residue (position 104, IMGT annotation) of LFR3, however, is cysteine in all germline sequences which could affect protein folding leading to low yield of properly folded antibodies. To address this potential problem, this cysteine was mutagenised to serine or glycine following the amplification and pooling of the KL3 and LL3 repertoires (Fig. 1C and 2). The other two residues left from LFR3 (position 102 and 103, IMGT annotation) are most frequently tyrosine which is highly hydrophobic. Based on the observation that the first (position 27, IMGT annotation) residue of germline H1s is always glycine which is neutral, we speculate that the hydrophobic nature of the first tyrosine from LFR3 could affect VH stability after grafting of L3 fragments containing the tyrosine from LFR3 into H1. To reduce the possibility of antibody aggregation and for convenience in mutagenesis via PCR, we mutated the tyrosine at position 102 into aspartic acid, asparagine and histidine (Fig. 1D and 2). Low annealing temperature (40°C) was used for the PCR mutagenesis to allow for efficient secondary amplification of the L3 fragments containing the desired mutation.

For assembly of full-length VHs, human H2 and H3 repertoires, and the FR2, 3 and 4 of m0 were amplified as a whole from m8l (Chen *et al.*, 2008b) (Fig. 1E). The FR1 of m0 was also PCR-amplified from the m8l plasmid DNA and joined to the L3 repertoire by overlapping PCR (Fig. 1F). The entire chimeric VHs were assembled by further joining the L3 repertoire to the H2 and H3 repertoires (Fig. 1G). The products were cloned into phagemid pComb3X, and a large (size ~ 10¹⁰) library (designated m9l) was obtained by performing 100 electroporations as described in Materials and Methods.

3.2. Sequence diversity of the library - randomly selected dAbs

To estimate the degree of the new library diversity we analyzed 126 randomly selected dAb clones for their gene usage, somatic mutations, and CDR length, and how different CDRs are combined. There were no identical L3 and H3 sequences found. H2s were less diverse - three groups (19, 4 and 3 members, respectively) of sequences each contained identical H2s from VH families 7, 4 and 4, respectively. Of the 126 L3 sequences, 88 were derived from all families of KL except family 7; about 75% of the sequences were from family 1 (Fig. 3A). They had lengths from 10 to 14 residues but more than 60% of them were 10 residues long (data not shown). The other 38 L3 sequences were from only two (of 11) LL families; more than 80% were from family 1 (Fig. 3B). Their lengths ranged from 9 to 12 residues (data not shown). Most of these L3s (82.5%) were mutated in their V genes compared to the closest corresponding

germline genes; 57% contained two or more mutations in their amino acid sequences (data not shown). The distribution and somatic mutations of H2s in m9l were consistent with those in m8l (Chen *et al.*, 2008b) except that in m9l, H2 gene usage was biased slightly toward VH1 and 7 while VH2- and 6-derived sequences were absent (Fig. 3C). The H3 lengths of dAbs from m9l ranged from 5 to 23 residues; the distribution was in agreement with the reported frequency in m8l (Chen *et al.*, 2008b) but there was an increased number of H3s with length shorter than 9 residues and longer than 18 residues (Fig. 3D). To determine the combinatorial diversity of the library, we plotted the pairing between each other of L3 origin, H2 origin and H3 length (Fig. 4). Regardless of the preferential amplification of gene fragments from certain families, the CDRs were paired randomly. These results suggest high degree of diversity of the new library.

3.3. Selection and initial characterization of high-affinity antigen-specific antibodies

To evaluate the usefulness of the new library, m9l, for selection of potent dAbs against therapeutically relevant proteins compared to the previously constructed one, m8l, we panned both libraries side by side against components (the ligand IGF-2 and the receptor IGF-1R) of the human IGF system which is of considerable interest as a target for cancer therapy. After several rounds of selection, enrichment and binding specificity of pooled phage were determined by ppELISA. To identify individual antibody that specifically bound to the antigens, clones were selected from the two libraries after panning and subjected to semELISA. Positive (with reading three times higher than that of the negative control) clones were sequenced and analyzed for CDR sequence diversity. Unique clones were expressed, purified and tested for binding activity and specificity.

Much higher and more rapid enrichment was observed when the new library, m9l, was panned against IGF-2 (Fig. 5A). We found that 25 of 190 clones selected from m9l after four rounds of panning showed significant binding to IGF-2 but not to the irrelevant antigen, bovine serum albumin (BSA). The previously constructed library, m8l, was panned five rounds because of lower enrichment. 18 of 190 clones were positive against IGF-2 although with lower binding compared to those selected from m9l. Four clones selected from m9l were unique and there was a repeated use of sequences in both L3 and H3 (Table 2). Three unique clones were identified from m8l, they had H2 of different origin and H3 of varying length. All the clones were highly expressible in *E. coli* HB2151 strain - yield of soluble dAbs from m8l of about 15 mg l⁻¹ and of those from m9l – about 2.5–20 mg l⁻¹. However, only one of the three dAbs obtained from m8l bound with reasonable activity (EC₅₀, ~ 50 nM) to IGF-2 as measured by ELISA; in contrast, three of the four dAbs selected from m9l bound with higher strength (EC₅₀, ~ 5–20 nM) (Table 2). One of these dAbs was also cross-reactive for IGF-1. This antibody and another one selected from m9l significantly inhibited IGF-2-induced IGF-1R phosphorylation in the human cancer cell line MCF-7; the only binder from m8l did not exhibit measurable inhibitory activity (Chen, Feng and Dimitrov, unpublished work), suggesting that these dAbs target different epitopes on IGF-2.

In the IGF-1R panning, specific and comparable enrichment was achieved with both libraries even after the first round of panning (Fig. 5B); three and two unique clones were selected from m8l and m9l, respectively that all exhibited high solubility, yield and comparable EC₅₀s ranging from 10 to 30 nM (Chen, Feng and Dimitrov, unpublished work). All selected IGF-1R antibodies, however, did not significantly inhibit IGF-2-induced IGF-1R phosphorylation in MCF-7 cells suggesting they did not precisely target the ligand-binding site on IGF-1R. They could be directed against different surface areas of the human IGF-1R which is a large protein. This prompted us to further compare both libraries with respect to selection of antibodies against a specific area – the ligand-binding site of IGF-1R. In a new panning, bound phage was eluted with a mixture of IGF-1 and IGF-2 after incubation of the libraries with IGF-1R. The

ppELISA showed that a significant specific enrichment was obtained from m91 but not from m81 (Fig. 5C). Two novel dAbs with relatively high affinity (EC_{50} , 30–50 nM) were identified from the new panning that significantly competed with IGF-1 in binding to IGF-1R (Chen, Feng and Dimitrov, unpublished work).

These results suggest that the new library contains potentially useful candidate therapeutic dAbs that cannot be selected from the previously constructed library, and that those antibodies selected from the two libraries against the same antigen target different epitopes. The newly identified dAbs from m91 are currently being additionally characterized and further improved for testing in animal models of cancer.

3.4. Antibody oligomerization and folding

To analyze the dAbs from this library for certain biophysical properties such as oligomerization, aggregation and degradation four dAbs were randomly selected and more extensively characterized. After purification on Ni-nitrilotriacetic acid resin, these proteins were dialyzed against PBS, pH 7.4, and concentrated to 12, 7.8, 3 and 5 mg/ml, respectively. No precipitation was observed with these four protein solutions immediately after purification and concentration. After they were stored for about one year at 4°C, one of them showed obvious precipitation (pellet) after centrifugation. Supernatant fractions were collected and measured for optical density at 280 nm (OD₂₈₀). No significant degradation was observed; they all ran on reducing SDS-PAGE with apparent molecular weight (aMW) of about 16 kDa, similar to the calculated MW (cMW) of 15–17 kDa (Fig. 6A). The oligomerization of these long-term stored dAbs was measured by size-exclusion chromatography on Superdex-75 column. They were eluted as a mono-disperse symmetric peaks indicating that they did not stick to the column matrix (Fig. 6B). Only one of them eluted at the expected size of a monomer while variation in the aMW was observed with the other dAbs which eluted more rapidly or slowly. Interestingly, the elution of two of these dAbs as well as m36 (Chen *et al.*, 2008a), which is a well-characterized monomeric dAb (cMW, ~15 kDa) from m81, was further delayed in the presence of 300 mM NaCl suggesting an increased overall hydrophobicity of these antibodies. Three freshly prepared dAbs selected by panning against IGF-2 (Fig. 5A; Table 2) were further tested for oligomerization in an additional experiment. The result showed that two of them were monomers while another one appeared to be a dimer (Fig. 6C). In agreement with our previous finding, variation in the aMW was observed with these dAbs which eluted more rapidly or slowly. These data suggest that randomly or antigen selected dAbs from the new library are in general stable against aggregation but some may exhibit variations in their aMWs.

To further investigate possible conformational changes in the scaffold caused by the grafted L3s we used the observation that the staphylococcal protein A (SPA) binds to the VH domain containing the VH3 gene products. Thus the library was panned against SPA, and 46 and 43 clones were randomly picked from the third and fourth round of panning, respectively, sequenced and analyzed for L3 and H2 gene usage, and H3 length. Compared with the original library, there was an increased number (70%, 89% and 91% for the original library, the third and fourth round, respectively) of antibodies with KL3s (Fig. 7A). However, the frequencies of V (Fig. 7B) and J (data not shown) gene usages in the KL3s selected after panning were comparable with those for the original library. The frequency of antibodies composed of VH3-derived H2 was also dramatically increased (four to fivefold) and their H3s were diverse with length from 6 to 20 residues (Fig. 7C), in agreement with our previous study (Chen *et al.*, 2008b). Eight clones were randomly selected from the fourth round of panning for expression; they all were expressible and purified with high solubility and yield from the soluble fraction of *E. coli* periplasm (Fig. 7D). These results suggest that the VH3-based scaffold used in the library, m0, preserves its conformational integrity after grafting of KL3s from almost all families as evaluated by the SPA binding activity.

4. Discussion

A major finding of this study is that combining CDR3s from heavy and light chains in a VH-based scaffold results in a highly diverse library that can be a source for high-affinity novel binders that could not be selected from the libraries constructed based on previous designs. Antibody diversity is generated through a complex series of events in which pairing between heavy and light chains resulting in tremendous amount of combinations among six CDRs is one of the most important contributing factors. Due to lack of the VH-VL combinatorial diversity, the importance of constructing highly diverse dAb libraries increases and in addition, there could be a need to compensate the loss of antigen-interacting surface contributed by the hypervariable loops of the VL.

In general L3 is significantly more diversified than H1. Combinations between V and J genes of light chains, junctional insertions and deletions as well as extensive somatic hypermutations result in generation of L3s with sequence and structural diversity much higher than that of H1. In addition L3s are on average longer than H1s (10 versus 8 residues, IMGT annotation) and in contrast to H1, vary in length although not very much. To amplify and graft L3, three additional residues from LFR3 were also added and mutated that further increased the length and diversity of the insert. The longer loops could contribute to an enlarged surface area of interaction with antigens potentially resulting in antibodies with higher affinity. Finally, grafting of L3s into the H1 position provides a further mechanism for introducing the diversity into the antibody tertiary structure. After grafting, the orientation and location of L3 with respect to H2 and H3 is altered compared to those in conventional antibodies. Taken together, we speculate that these changes would lead to novel paratopes and therefore, the library could be more suited for selection of antibodies to some antigens or epitopes on the same antigen than the existing libraries based on conventional designs. In line with this speculation, our results showed efficient selection of different dAbs from m91 against the human cancer-related antigens, IGF-2 and IGF-1R, compared to dAbs selected from our previously constructed library m81.

The relatively high level of diversity generated by the non-cognate grafting appears to come with a price - the expression of soluble dAbs from randomly selected clones was somewhat lower than that from the previously constructed library m81 (Chen *et al.*, 2008b) (data not shown). It is to be expected that the use of non-cognate grafting could result in misfolding and aggregation of some dAbs. Of note is that most of the antigen selected dAbs from this library were soluble and expressed at high levels. They contain L3s and H2s from different germ lines and H3s of varying length (Table 2). Human antibodies binding SPA in the Fab regions are encoded by gene segments belonging to the VH3 group; a previous study (Potter *et al.*, 1996) shows that SPA simultaneously interacts with FR1, CDR2, and FR3; thus, SPA binding has been used as a marker for proper folding of human VH3. To see whether the recombinant VHs with a fixed VH3 scaffold can retain the activity, we cycled the library through four rounds of selection against SPA. Gene usage and length distribution of CDRs of clones randomly picked from the third and fourth round of selection show that VH3-derived H2 are enriched (Fig. 7A), in agreement with our previous study (Chen *et al.*, 2008b). Interestingly, there is also a selective use of KL3 (Fig. 7A) but not a particular family (Fig. 7B) suggesting that KL3 sequences provide better structural compatibility than LL3 sequences. Moreover, the H3s are highly diverse after selection (Fig. 7C). The gene usage and CDR length are not significantly related to each other suggesting the ability of the scaffold, m0, to simultaneously stably hold three randomly combined diverse CDRs of different origin and length. A size exclusion chromatography analysis of seven dAbs showed that they exhibit different oligomeric states including monomeric and dimeric (Fig. 6). Interestingly, there is a substantial variation in the aMW of some of these dAbs which differs from that of a well-defined monomer (15–17 kDa). This phenomenon was observed by other groups who suggested as a possible explanation the

weak transient interactions with the column matrix (Ewert *et al.*, 2003; Jespers *et al.*, 2004). In our study, the elution of two randomly selected antibodies was further delayed in the presence of 300 mM NaCl (Fig. 6B) suggesting possible involvement of L3 that could result in a higher level of hydrophobicity of the proteins and although very unlikely, a significant alteration in density of protein folding.

The new library can be a useful source of high-affinity antibodies with potential applications for development of therapeutics. It can be also used as a research tool to study mechanisms of folding and aggregation of engineered antibody domains. The selected high-affinity dAbs against components of the human IGF system and other antigens could be further developed as potential therapeutics, diagnostics and research reagents.

Abbreviations

VH	heavy chain variable domain
VL	light chain variable domain
dAb	domain antibody
Fab	antigen-binding fragment
CDR	complementarity-determining region
L3	light chain CDR3
H2	heavy chain CDR2
H3	heavy chain CDR3
FR	framework region
KL	kappa light chain
LL	lambda light chain
IGF-2	human insulin-like growth factor 2
IGF-1R	IGF-1 receptor
HIV	human immunodeficiency virus
SPA	staphylococcal protein A
MW	molecular weight
PBS	phosphate-buffered saline
PBST	PBS containing 0.05% Tween-20

Acknowledgments

This work was supported by the NIH Intramural AIDS Targeted Antiviral Program, by the NIH National Institute of Allergy and Infectious Diseases Intramural Biodefense Program (DSD), by the NIH NCI Center for Cancer Research Intramural Research Program, and by federal funds from the NIH NCI under contract N01-CO-12400.

Reference List

- Binz HK, Amstutz P, Pluckthun A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol* 2005;23:1257–1268. [PubMed: 16211069]
- Chen W, Zhu Z, Feng Y, Dimitrov DS. Human domain antibodies to conserved sterically restricted regions on gp120 as exceptionally potent cross-reactive HIV-1 neutralizers. *Proc. Natl. Acad. Sci. USA* 2008a;105:17121–17126. [PubMed: 18957538]

- Chen W, Zhu Z, Feng Y, Xiao X, Dimitrov DS. Construction of a large phage-displayed human antibody domain library with a scaffold based on a newly identified highly soluble, stable heavy chain variable domain. *J. Mol. Biol* 2008b;382:779–789. [PubMed: 18687338]
- Chen W, Zhu Z, Xiao X, Dimitrov DS. Construction of a human antibody domain (VH) library. *Methods Mol. Biol* 2009;525:81–99. [PubMed: 19252835]
- Dimitrov DS. Engineered CH2 domains (nanoantibodies). *mAbs* 2009;1:26–28. [PubMed: 20046570]
- Dimitrov DS, Marks JD. Therapeutic antibodies: Current state and future trends - Is a paradigm change coming soon? *Methods Mol. Biol* 2009;525:1–27. [PubMed: 19252861]
- Ewert S, Huber T, Honegger A, Pluckthun A. Biophysical properties of human antibody variable domains. *J. Mol. Biol* 2003;325:531–553. [PubMed: 12498801]
- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R. Naturally occurring antibodies devoid of light chains. *Nature* 1993;363:446–448. [PubMed: 8502296]
- Hey T, Fiedler E, Rudolph R, Fiedler M. Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends Biotechnol* 2005;23:514–522. [PubMed: 16054718]
- Holliger P, Hudson PJ. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol* 2005;23:1126–1136. [PubMed: 16151406]
- Holt LJ, Herring C, Jespers LS, Woolven BP, Tomlinson IM. Domain antibodies: proteins for therapy. *Trends Biotechnol* 2003;21:484–490. [PubMed: 14573361]
- Jespers L, Schon O, Famm K, Winter G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat. Biotechnol* 2004;22:1161–1165. [PubMed: 15300256]
- Kolmar H, Skerra A. Alternative binding proteins get mature: rivalling antibodies. *FEBS J* 2008;275:2667. [PubMed: 18435760]
- Nguyen VK, Hamers R, Wyns L, Muyldermans S. Camel heavy-chain antibodies: diverse germline V (H)H and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J* 2000;19:921–930. [PubMed: 10698934]
- Nygren PA, Skerra A. Binding proteins from alternative scaffolds. *J. Immunol. Methods* 2004;290:3–28. [PubMed: 15261569]
- Potter KN, Li Y, Capra JD. Staphylococcal protein A simultaneously interacts with framework region 1, complementarity-determining region 2, and framework region 3 on human VH3-encoded Igs. *J. Immunol* 1996;157:2982–2988. [PubMed: 8816406]
- Saerens D, Ghassabeh GH, Muyldermans S. Single-domain antibodies as building blocks for novel therapeutics. *Curr. Opin. Pharmacol* 2008;8:600–608. [PubMed: 18691671]
- Skerra A. Alternative non-antibody scaffolds for molecular recognition. *Curr. Opin. Biotechnol* 2007;18:295–304. [PubMed: 17643280]
- Ward ES, Gussow D, Griffiths AD, Jones PT, Winter G. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* 1989;341:544–546. [PubMed: 2677748]
- Yokota T, Milenic DE, Whitlow M, Schlom J. Rapid tumor penetration of a single-chain Fv and comparison with other immunoglobulin forms. *Cancer Res* 1992;52:3402–3408. [PubMed: 1596900]
- Yokota T, Milenic DE, Whitlow M, Wood JF, Hubert SL, Schlom J. Microautoradiographic analysis of the normal organ distribution of radioiodinated single-chain Fv and other immunoglobulin forms. *Cancer Res* 1993;53:3776–3783. [PubMed: 8339291]
- Zhang MY, Shu Y, Phogat S, Xiao X, Cham F, Bouma P, Choudhary A, Feng YR, Sanz I, Rybak S, Broder CC, Quinnan GV, Evans T, Dimitrov DS. Broadly cross-reactive HIV neutralizing human monoclonal antibody Fab selected by sequential antigen panning of a phage display library. *J. Immunol. Methods* 2003;283:17–25. [PubMed: 14659896]
- Zhu Z, Dimitrov DS. Construction of a large naive human phage-displayed Fab library through one-step cloning. *Methods Mol. Biol* 2009;525:129–142. [PubMed: 19252833]
- Zhu Z, Dimitrov AS, Bossart KN, Crameri G, Bishop KA, Choudhry V, Mungall BA, Feng YR, Choudhary A, Zhang MY, Feng Y, Wang LF, Xiao X, Eaton BT, Broder CC, Dimitrov DS. Potent neutralization of Hendra and Nipah viruses by human monoclonal antibodies. *J. Virol* 2006;80:891–899. [PubMed: 16378991]

Fig. 1.

Library construction procedure. (A) KL and LL were PCR-amplified respectively by using cDNA or library plasmid DNA as templates and purified. (B) L3 repertoire was harvested from the purified KL and LL by PCR with a set of primer combinations. (C) The last residue that the L3 sense primers target was mutated from cysteine to serine and glycine. (D) The first residue that the L3 sense primers target was mutated from tyrosine to aspartic acid, asparagine and histidine. (E) H2 and H3 repertoires and FR fragments were amplified from m81 plasmid DNA. (F) L3 repertoire was joined to FR1 by overlapping PCR. (G) Entire VHs were assembled by overlapping PCR, combining L3 repertoire with H2 and H3 repertoires.

caggtgcagctggtgcagctctgggggaggcttggtacagcctggaggggtccctgagactc
 Q V Q L V Q S G G G L V Q P G G S L R L

L3 sense → L3

tcctgtgcagcctctgactacagccaacagacttacagtaccccatcacttttatgagc
 S C A A S **D** Y **S** Q Q T Y S T P **F** T **F** M S
 (102Y) (104C)

L3 antisense ←

tgggtccgccaggctccagggaagggactagagtggatcggggaatatcttttataatgga
 W V R Q A P G K G L E W I G N I F Y N G

aacaccttctactccccgtccctcaagagtcgagtcaccatctccagagacaattccaag
 N T F Y S P S L K S R V T I S R D N S K

aacacgctgtatctgcaaatgaacaccctgagagccgaggacacagccacgtattactgt
 N T L Y L Q M N T L R A E D T A T Y Y C

gcgaaagatcgtcccaaaatagcagtggtgccttgactactggggccagggcacccctg
 A K D R P K I A V A A L D Y W G Q G T L

gtcaccgtctcctca
 V T V S S

Fig. 2.

Cloning of L3 repertoire. To detail our design for amplification of L3 repertoire, the sequence of an antibody randomly selected from the library is shown. The regions on LFR3 and J gene that the L3 sense and antisense primers target, respectively, are indicated by arrows. The two positions on LFR3 which are mutated after L3 amplification are highlighted with bracket and the residues in the two positions for this particular sequence - with gray background.

Fig. 3. Analysis of library sequence diversity based on the 126 randomly selected clones. (A) Gene usage of KL3. (B) Gene usage of LL3. (C) Gene usage of H2. (D) H3 length distribution.

A

B

C

Fig. 4.

Combinatorial diversity of the library. 126 randomly selected clones were sequenced and analyzed for L3 and H2 gene usage, and length of H3s. The pairings between L3 and H2 gene usage (A), L3 gene usage and H3 length (B), and H2 gene usage and H3 length (C) were plotted for each sequence, respectively. The number of individual sequences found for each pairing is shown on the top right corner.

Fig. 5. Selection of antigen-specific antibodies. PpELISA for panning of both libraries side by side against IGF-2 (A), IGF-1R (B) and IGF-1R with elution by the IGF ligands (C).

A

B

C

Fig. 6.

Antibody oligomerization. (A) Reducing SDS-PAGE of four and three dAbs randomly and antigen selected from the library, respectively. (B) Size-exclusion chromatography with Seperdex-75 column for four dAbs randomly selected from m91. (C) Size-exclusion chromatography with Seperdex-75 column for three dAbs selected by panning m91 against IGF-2. The proteins were eluted at 0.5 ml/min with PBS or PBS containing 300 mM NaCl. Arrows indicate the elution volumes of molecular mass standards in PBS: ribonuclease A (14 kDa), chymotrypsin (25 kDa) and ovalbumin (44 kDa).

Fig. 7.

Selection of phage-displayed VHs to SPA. Library (phage) was cycled through four rounds of selection against SPA. 46 and 43 clones were randomly picked from the third and fourth round of selection, respectively, sequenced and analyzed for gene usage of L3 and H2 (A and B), and length distribution of H3 (C). Eight antibodies (clone numbers 101–107) were randomly selected from the fourth round of panning and expressed in *E. coli* HB2151 (D).

Table 1

Primers used for library construction

Primer description	Name	Sequence	Target
L3 sense	L3F1	CTC TCC TGT GCA GCC TCT <u>TAY TAC TGY</u>	L3
	L3F2	CTC TCC TGT GCA GCC TCT <u>TAY TAC RGY</u>	L3
	L3F3	CTC TCC TGT GCA GCC TCT <u>NAY TAC RGY</u>	L3
KL3 antisense	KL3R1	CTG GCG GAC CCA GCT CAT <u>GAA CGT CCA</u>	KJ1 ^a
	KL3R2	CTG GCG GAC CCA GCT CAT <u>AAA AST GYA</u>	KJ2
	KL3R3	CTG GCG GAC CCA GCT CAT <u>GAA AGT GAA</u>	KJ3
	KL3R4	CTG GCG GAC CCA GCT CAT <u>GAA MGT GAG</u>	KJ4
	KL3R5	CTG GCG GAC CCA GCT CAT <u>GAA GGT GAT</u>	KJ5
LL3 antisense	LL3R1	CTG GCG GAC CCA GCT CAT <u>GAA GAC ATA</u>	LJ1 ^b
	LL3R2	CTG GCG GAC CCA GCT CAT <u>GAA TAC CAC</u>	LJ2, 3
	LL3R3	CTG GCG GAC CCA GCT CAT <u>AAA TAC AAA</u>	LJ4
	LL3R4	CTG GCG GAC CCA GCT CAT <u>AAA CAC CCA</u>	LJ5
	LL3R5	CTG GCG GAC CCA GCT CAT <u>GAA CAC ATT</u>	LJ6
	LL3R6	CTG GCG GAC CCA GCT CAT <u>GAA CAC AGC</u>	LJ7
FR1 sense	FR1F	TGG TTT CGC TAC CGT GGC CCA GGC GGC CCA GGT GCA GCT GGT G	FR1
FR1 antisense	FR1R	AGA GGC TGC ACA GGA GAG	FR1
FR2 sense	FR2F	ATG AGC TGG GTC CGC CAG	FR2
Extension 1	L3R	CTG GCG GAC CCA GCT CAT	5 end of L3 antisense primer
Extension 2	HISR	GTC GCC GTG GTG GTG GTG GTG GTG GCC GGC CTG GCC ACT TG	pComb3X vector

The sequences underlined enable amplification of human antibody L3 repertoire.

^aFamily of human antibody KL J gene.

^bFamily of human antibody LL J gene.

Table 2
CDR diversity of unique phage-displayed VHs selected after panning with human IGF-2

Antigen	Library	Clone	H1 or L3		H2		H3
			Sequence	Gene usage	Sequence	Gene usage	
IGF-2	m81	1G8	DFAFSYYE	NA	IKQDGSVT	VH3-7 ^c	RELGV DY
		2A1	DFDFYSYE	NA	VYSSGST	VH4-39, et al.	RDLGRDDY
		2E6	SFDFDY YE	NA	ISKSGG TK	VH3-30-3 et al.	KDRATCSGGSCYSFY YGMDV
	m91	1A1	QQYKTYPLTF	KV1-8 ^a et al., κJ4 ^b	ISGSGGTT	VH3-23	ARVASRDYFDY
		1E4	LQHDNFPYTF	KV5-2, κJ2	INHS GST	VH4-34	ARDYGYAFDI
		2D6	QQNKTYPLTF	KV1-8 et al., κJ4	INDSGGTT	VH3-20 et al.	ARVASRDYFDY
		2G2	LQHDNFPYTF	KV5-2, κJ2	ISGSGGTT	VH3-23	ARIRWLQDL DY

^aSub-family of human antibody light chain V gene.

^bFamily of human antibody light chain J gene.

^cSub-family of human antibody heavy chain V gene.

NA, not applied.