

Efficient construction of a large nonimmune phage antibody library: The production of high-affinity human single-chain antibodies to protein antigens

(single-chain Fv/phage display/antibody libraries/human mAbs)

MICHAEL D. SHEETS*^{†‡}, PETER AMERSDORFER[§], RICARDA FINNERN[§], PETER SARGENT[¶], ERICKA LINDQVIST[¶], ROBERT SCHIER[§], GRETE HEMINGSEN[§], CINDY WONG[§], JOHN C. GERHART*, AND JAMES D. MARKS^{‡§}

Departments of *Molecular Cell Biology and [¶]Public Health, University of California, Berkeley, CA 94720; [§]Department of Anesthesia and Pharmaceutical Chemistry, University of California, Room 3C-38, San Francisco General Hospital, 1001 Potrero, San Francisco, CA 94110; [†]Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI 53706; and [‡]Department of Stomatology, University of California, San Francisco, CA 94143

Contributed by John C. Gerhart, March 16, 1998

ABSTRACT A large library of phage-displayed human single-chain Fv antibodies (scFv), containing 6.7×10^9 members, was generated by improving the steps of library construction. Fourteen different protein antigens were used to affinity select antibodies from this library. A panel of specific antibodies was isolated with each antigen, and each panel contained an average of 8.7 different scFv. Measurements of antibody–antigen interactions revealed several affinities below 1 nM, comparable to affinities observed during the secondary murine immune response. In particular, four different scFv recognizing the ErbB2 protein had affinities ranging from 220 pM to 4 nM. Antibodies derived from the library proved to be useful reagents for immunoassays. For example, antibodies generated to the *Chlamydia trachomatis* elementary bodies stained *Chlamydia*-infected cells, but not uninfected cells. These results demonstrate that phage antibody libraries are ideally suited for the rapid production of panels of high-affinity mAbs to a wide variety of protein antigens. Such libraries should prove especially useful for generating reagents to study the function of gene products identified by genome projects.

Antibodies that bind with high specificity and high affinity to a target molecule are essential tools for biological research. These reagents have proven invaluable for: (i) detecting and quantitating levels of gene expression; (ii) determining the subcellular, cellular, and tissue location of gene expression; and (iii) identifying the molecules interacting with a gene product, for example by immunoprecipitation.

Numerous new applications for basic research, as well as clinical use, have resulted from the development of recombinant antibodies constructed from Ig variable (V) region genes (1–3). Single-chain Fv antibodies (scFv) have proven particularly useful. scFv consist of the antigen-binding domains of Ig heavy (V_H) and light (V_L) chain regions connected by a flexible peptide linker (4), all encoded by a single gene. The single gene design of scFv simplifies the construction of fusion proteins such as cancer immunotoxins (5) and facilitates intracellular expression in eukaryotic cells to achieve phenotypic knockout of antigen function (6–8). The intracellular expression of antibodies is proving to be an effective new strategy for studying the function of specific proteins *in vivo* where conventional genetic approaches are not feasible.

Genome projects have led to an increasing rate of gene discovery and an accelerating need for antibodies to study gene

expression and function. Until recently, hybridoma technology, a slow and cumbersome process, was used to produce mAbs for such applications. Separate immunizations are required for each antigen, and the cell fusion process required to generate hybridomas is laborious and inefficient. In addition, production of antibodies to antigens conserved between species is difficult and antibodies from hybridomas are murine and hence immunogenic if used therapeutically.

Recent advances using antibody phage display now make it possible to overcome these limitations and generate human mAbs that recognize any desired antigen (1–3, 9). For phage display, the antigen-binding regions of V_H and V_L genes are cloned and used to construct scFv (or Fab) gene repertoires. A phage antibody library is created by cloning these repertoires as fusion proteins with a minor coat protein of bacteriophage (the gene 3 protein) (10–12). Each resulting phage has a functional antibody protein on its surface and contains the gene encoding the antibody incorporated into the phage genome. Particular phage antibodies that specifically bind to proteins and small molecules can be separated from nonbinding phage antibodies with affinity chromatography techniques (12–15). This strategy requires no immunization, the antibody genes are cloned, and generally the antibody fragments express well in *Escherichia coli*. The number and affinity of the antibodies generated to a particular antigen is a function of library size and diversity, with larger libraries yielding a greater number of high-affinity antibodies (14, 15). Unfortunately, the construction of large phage-displayed antibody libraries has remained difficult. If such libraries are to be a common tool of life scientists the efficient production of these reagents must become routine, especially because library diversity and utility are lost on library reamplification.

In this paper, we describe a strategy to optimize the construction of phage-display antibody libraries. By using this strategy, a very large phage-displayed single-chain antibody library consisting of 6.7×10^9 members was produced. This library then was used to isolate panels of antibodies to 14 different protein antigens. Analysis of antibody–antigen interactions revealed high-affinity binding with K_d s for the ErbB2 protein ranging between 220 pM and 4 nM.

METHODS

Construction of the V_H Library. Total RNA was prepared from three different samples of human spleen cells and two

Abbreviations: BMP, bone morphogenetic protein; BoNT, botulinum neurotoxin; ECD, extracellular domain; CDR, complementarity determining region; EB, elementary body; scFv, single-chain Fv fragment; V_K , Ig kappa light chain variable region; V_λ , Ig lambda light chain variable region; V_L , Ig light chain variable region; V_H , Ig heavy chain variable region.

[‡]To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/956157-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

different samples of human peripheral blood lymphocytes. cDNA was synthesized from total RNA primed with the HuIgMFOR primer (12). V_H gene repertoires were amplified from the cDNA by using Vent DNA polymerase (New England Biolabs) in combination with the HuIgMFOR primer and an equimolar mixture of HuVHBACK primers (12). PCR products were agarose gel-purified and reamplified to append *NcoI* and *NotI* restriction sites by using Tth DNA polymerase (Epicentre Technologies, Madison, WI) and an equimolar mixture of the HuVHBACKSfi primers (that contain a *NcoI* site for cloning) and the HuCMForNot primer (5'-GAGTC-ATTCTCGACTTGC GGCCGCTGGAAGAGGCACGTTCTTTTCTTT-3') (12). The PCR products were cut with restriction enzymes *NcoI* and *NotI* and agarose gel-purified. The resulting DNA fragments were ligated into the plasmid pCITE3A (Novagen) cut with restriction enzymes *NcoI* and *NotI* and the ligated DNA was electroporated into the *E. coli* strain TG1. A library of V_H genes containing 2.3×10^8 members was generated from the products of seven ligation reactions and 15 electroporations. The resulting library was termed pCITE- V_H . Cloning efficiency and library diversity was determined by PCR screening (12, 16). The pCITE3A plasmid was used to create the V_H gene repertoire because of the presence of unique sequences for PCR amplification that surround the *NcoI* and *NotI* cloning sites. These sequences allow the specific amplification of the V_H genes for scFv assembly. This strategy is advantageous for amplification of the V_H genes and also the subsequent amplification of scFv genes assembled from the V_H genes. Although we chose the pCITE3A plasmid for production of our V_H gene repertoire, any plasmid that contains the proper restriction sites for cloning and unique sequences for specific PCR amplification would have been suitable.

Construction of the scFv Library. The V_H gene repertoire was PCR-amplified from the pCITE- V_H library by using 300 ng of library plasmid DNA as a template, Vent DNA polymerase, the CITE3 primer (5'-GATCTGATCTGGGGCCTCGGTGC-3'), and an equimolar mixture of Hu J_H primers (12). The V_L genes for scFv assembly were obtained from a previously constructed scFv phage antibody library (12). The V_L gene repertoire, including DNA encoding the scFv peptide linker (G_4S)₃ (4), was amplified from 300 ng of library plasmid DNA by using Vent DNA polymerase, the Gene3 primer (5'-GC-AAGCCCAATAGGAACCCATGTACCG-3'), and an equimolar mixture of RHu J_H primers (12). The amplified V_H and V_L genes were agarose gel-purified and spliced together with overlap extension PCR to create a scFv gene repertoire (11). To accurately join V_H and V_L gene repertoires with overlap extension PCR, the input DNA fragments must have blunt ends. Therefore, the proofreading DNA polymerase Vent was used to generate the V_H and V_L DNA fragments for scFv assembly. For all subsequent PCR steps of library construction Tth DNA polymerase was found to be the optimal enzyme. The V_H and V_L gene repertoires were spliced together in 100- μ l PCRs containing 100 ng of the V_H and V_L DNA fragments and Tth DNA polymerase. The reactions were cycled eight times (95°C 2 min, 55°C 1 min, and 72°C 3 min) to join the fragments. Then the CITE3 and Gene3 primers were added and the reaction was cycled 30 times (94°C 1 min, 55°C 1 min, and 72°C 3 min) to amplify the assembled scFv genes. The scFv genes were cut with restriction enzymes *NcoI* and *NotI*, agarose gel-purified, and ligated into the plasmid pHEN-1 (17) cut with *NcoI* and *NotI*. The ligated DNA was electroporated into *E. coli* TG1 cells.

Proteins. The extracellular domains of the *Xenopus* activin receptor type I (A. Suzuki and N. Ueno, personal communication), activin receptor type II (18), bone morphogenetic protein (BMP) receptor type I (19, 20), and fibroblast growth factor receptor (21) were cloned into pMAL expression plasmids as fusions with the gene encoding maltose binding protein

expressed and purified from *E. coli*. (New England Biolabs). Neuronal bungarotoxin was purchased from Biotoxins. *Clostridium botulinum* neurotoxin type A (BoNT/A) was provided by Ray Stevens (Univ. of California, Berkeley), and BoNT/B, C, and E were provided by Theresa Smith (United States Army Medical Research Institute of Infectious Disease). BoNT/A C-fragment was purchased from Ophidian (Madison, WI). Human ErbB-2 extracellular domain (ECD) was provided by James Huston (Creative Biomolecules) (22), human cytochrome b5 was provided by Lucy Waskell (Univ. of California, San Francisco), and human vascular endothelial growth factor was provided by James Hoeffler (Invitrogen).

Selection of Phage Antibodies. Phagemid particles were rescued from the library, as described (23) except that the procedure was scaled up to 2 liters of culture media. Specific phage-displayed scFv were affinity-selected by using proteins adsorbed to Immunotubes (Nunc) (12). For selections with maltose binding protein (MBP) fusion proteins, phage were preincubated with 50 μ g of purified MBP to deplete the library of MBP antibodies. For selection of scFv to the ErbB-2 ECD, Immunotube selection was alternated with selection using decreasing concentrations of biotinylated ErbB-2 ECD and

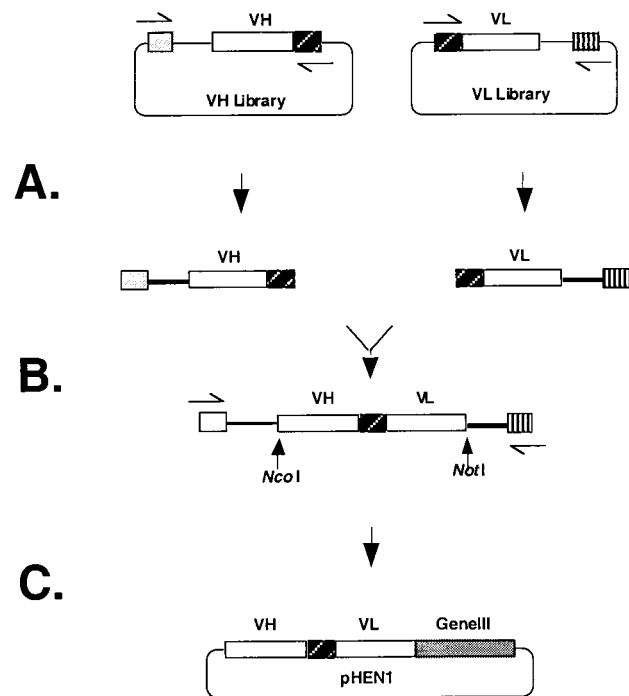


FIG. 1. Schematic outline of the approach used for library construction. A library of V_H and linker- V_L gene repertoires were generated from the plasmid DNA of the separate libraries. The V_H genes were amplified by using a plasmid specific primer (□) and an equimolar mixture of Hu J_H primers (■). The linker DNA and V_L genes were amplified by using a plasmid specific primer (□) and an equimolar mixture of RHu J_H primers (■). The RHu J_H primers are complementary to the Hu J_H primers. (B) The V_H and linker DNA- V_L gene repertoires were PCR assembled into a scFv gene repertoire. (C) The assembled scFv gene repertoire was cut with the restriction enzymes *NcoI* and *NotI* and cloned into the plasmid pHEN1 (17) for phage display.

capture of bound phage using streptavidin paramagnetic beads (23). For selection of scFv that bind *Chlamydia* antigens, immunotubes were coated overnight at room temperature with 1 ml of *C. trachomatis* strain L2/434/Bu elementary bodies (EBs) at a concentration of 0.1 mg/ml (in PBS) purified from a suspension culture of L929 cells (24). Phage eluted from each selection were used to infect *E. coli* TG1 cells. Phage particles were rescued from the cells and used for the subsequent round of antigen selection. The rescue-selection-plating cycle was repeated 3–4 times, after which individual clones were analyzed for specific antigen binding by ELISA.

Antibody Binding Specificity. The binding specificity of all scFv was determined by ELISA using the target antigen and at least nine other proteins as substrates (12). The number of unique scFv was estimated by PCR fingerprinting of the scFv genes with the restriction enzyme *Bst*NI and confirmed by DNA sequencing (12, 16). Putative V_H and V_L germ-line gene segment derivation was determined with the VBASE sequence directory (25).

scFv Purification and Affinity Measurements. For purification, scFv genes were subcloned, expressed, and purified to homogeneity (26). scFv dissociation equilibrium constants (K_d) were calculated from the association (k_{on}) and dissociation (k_{off}) rate constants determined by using surface plasmon resonance in a BIAcore (23, 27).

Fluorescent Cell Staining. Monolayers of HeLa 229 cells were grown on coverslips in 24-well cell culture plates. Two hundred microliters of *C. trachomatis* EBs at 8×10^6 inclusion forming units/ μ l were used to infect the monolayers (28). The infected cells were incubated for 48 hr at 37°C, washed with PBS, and fixed with 100% methanol for 10 min. Purified scFv (50 μ g/ml) was incubated with fixed cells for 1 hr at room temperature. scFv binding was detected with the 9E10 mAb that recognizes the c-myc epitope present in the scFv (29) (1 μ g/ml) followed by fluorescein isothiocyanate-conjugated anti-mouse Fc (Zymed). Cells were counterstained with Evans blue and visualized with fluorescence microscopy.

RESULTS

Library Construction. A very large phage antibody library was created for the routine isolation of high-affinity scFv

antibodies to any target protein. This library was generated by optimizing the individual steps of library construction to increase the efficiency of scFv gene assembly and increase the efficiency of cloning scFv genes (Fig. 1). First, scFv antibodies were assembled from cloned V_H and V_L gene repertoires contained in separate plasmid vectors. A library of V_H genes, containing 2.3×10^8 members, was specifically created for generating an additional scFv repertoire. The V_L genes for scFv assembly were derived from an existing scFv repertoire containing 3.0×10^7 members (Fig. 1A). The use of cloned libraries as a source of V-genes provided a stable and limitless supply of material for scFv assembly. For the construction of previous antibody libraries, scFv gene repertoires were directly assembled from V_H and V_L reverse transcription-PCR (RT-PCR) products (12). With this previous approach, RNA availability and the efficiency of RT-PCR limited the quantity of V-genes available for scFv construction. Second, the efficiency of scFv assembly was increased by exploiting the presence of the DNA encoding the peptide (G₄S)₃ linker located at the 5' end of the V_L library (Fig. 1B). Using V_L genes already fused to the peptide linker allowed us to construct scFv from only two DNA fragments. Previously, scFv gene repertoires were inefficiently assembled from three separate DNA fragments consisting of V_H and V_L gene repertoires and linker DNA (12). Third, the V_H and V_L gene repertoires and the scFv genes assembled from these repertoires were amplified with primers that annealed to sequences approximately 200 bp 5' of the V_H genes and to sequences approximately 200 bp 3' of the V_L genes. This strategy generated long sequence extensions at the ends of the individual V_H , V_L gene segments, and the assembled scFv. These sequence extensions ensured efficient cutting with the restriction enzymes *Nco*I and *Not*I that were used for scFv cloning and facilitated the identification of the correctly assembled scFv (Fig. 1C).

By using these three modifications a repertoire of scFv genes was efficiently assembled and cloned to create a phage antibody library containing 6.7×10^9 members. This library was generated from the products of only 12 ligation reactions and 36 electroporations. DNA sequencing of the V-genes from 36

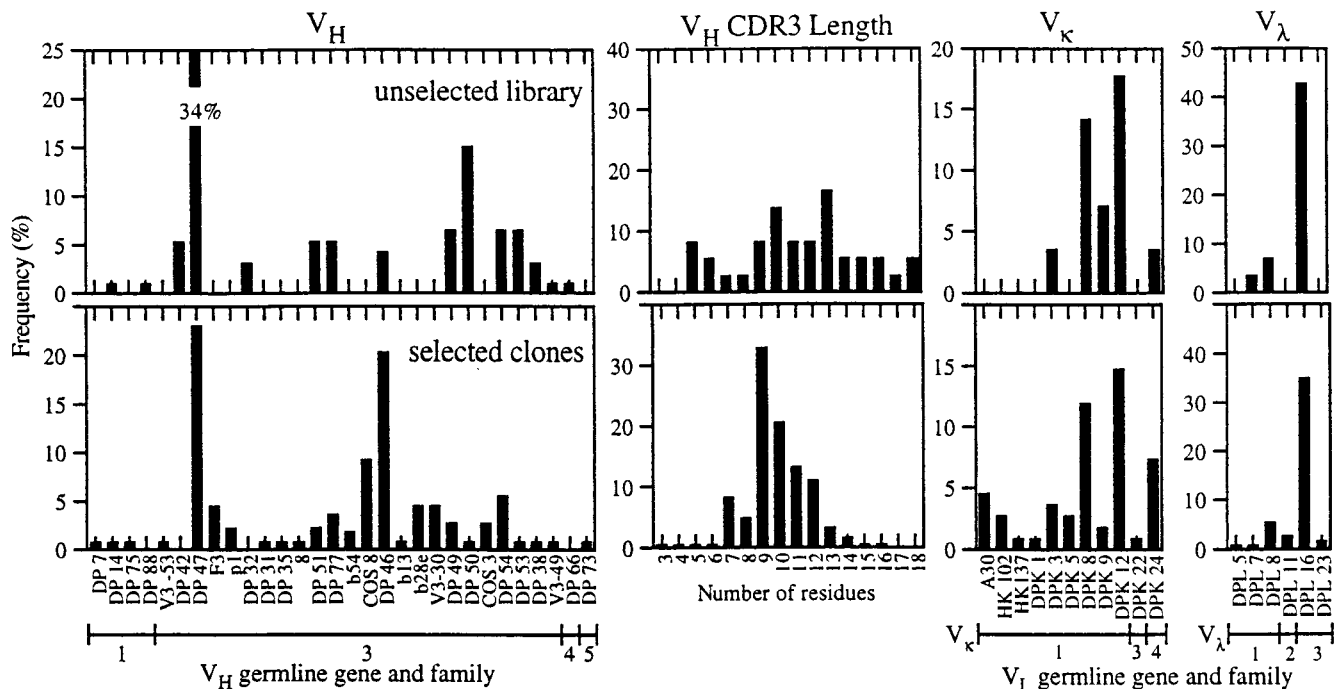


FIG. 2. V-gene usage and V_H CDR3 length of unselected and antigen-specific scFv. The V_H and V_L genes were sequenced and the germ-line gene was assigned based on homology to a database (VBASE) of germ-line V-genes compiled by Tomlinson *et al.* (25). Specific V_H , V_K , or V_L germ-line genes are listed on the ordinate, with the V_H , V_K , or V_L germ-line gene family indicated below. Only V-genes in unselected or selected clones are listed.

Table 1. Results of phage antibody library selections

Protein antigen used for selection	Percentage (number) of ELISA positive clones	Number of different antibodies isolated
FGF receptor ECD	69 (18/26)	15
BMP receptor type I ECD	50 (12/24)	12
Activin receptor type I ECD	66 (16/24)	7
Activin receptor type II ECD	66 (16/24)	4
Erb-B2 ECD	91 (31/34)	14
VEGF	50 (48/96)	6
BoNT/A	28 (26/92)	14
BoNT-A C-fragment	95 (87/92)	10
BoNT/B	10 (9/92)	5
BoNT/C	12 (11/92)	5
BoNT/E	9 (8/92)	3
Bungarotoxin	67 (64/96)	15
Cytochrome b5	55 (53/96)	5
<i>C. trachomatis</i> EB	66 (63/96)	7

For each antigen (column 1), the number and the percentage of positive clones selected (column 2) and the number of different antibodies isolated (column 3) is indicated. FGF, fibroblast growth factor; VEGF, vascular endothelial growth factor.

randomly chosen scFv revealed 36 unique sequences and a relatively random distribution of V_H complementarity determining region (CDR) 3 length of between 5 and 18 residues (Fig. 2). There was, however, bias in V-gene usage, with both over-representation of specific V-gene families (V_{H3} , $V_{\kappa 1}$, and $V_{\lambda 3}$) and V-genes (DP-47, DPL 16) (Fig. 2). This bias partially reflects differential V-gene usage observed in the human B-cell repertoire (30–33) but also may be caused by differences in PCR primer annealing to the different V-genes. Previous work indicates that more diverse repertoires could be created by using V_H and V_{κ} gene family specific primers individually rather than pooled for construction of the V-gene repertoires (34).

Selection and Characterization of Antigen-Specific scFv. Antibodies from the phage antibody library were affinity-selected by using 13 different purified protein antigens from a variety of species, including human and EBs from *C. trachomatis* (Table 1). Given our interest in developmental biology, four of these proteins were the extracellular domains of different *Xenopus* growth factor receptors: the activin receptor types I and II, the BMP receptor type I, and the fibroblast growth factor receptor (19–21). After at least three rounds of selection with a particular antigen, the binding specificity of individual scFv was determined by ELISA. A high percentage of the clones analyzed specifically bound the antigen used for selection (Table 1, second column). To determine the number of different scFv that recognized each antigen, ELISA-positive clones first were characterized by DNA fingerprinting (12, 16) and then DNA sequencing (23). This analysis revealed an average of 8.7 different antibodies were generated to each protein antigen, with the number of scFv ranging from 3 to 15 (Table 1). Because only a small number of clones from each selection were analyzed, it is likely that screening of more clones would yield additional antibodies.

The binding of scFv to antigens was highly specific. For example, serotype specific scFv were isolated against each of the four different types of BoNT, despite 32–59% sequence homology between the toxins (Fig. 3). Another example of scFv specificity is shown in Fig. 4, where a *C. trachomatis*-specific scFv stains *C. trachomatis* elementary bodies within infected cells while neighboring uninfected cells remain unstained.

V-gene derivation of scFv antibodies that bound to the different antigens was diverse (Fig. 2). V_H genes were derived from three of the six V_H gene families (nos. 1, 3, and 5) and

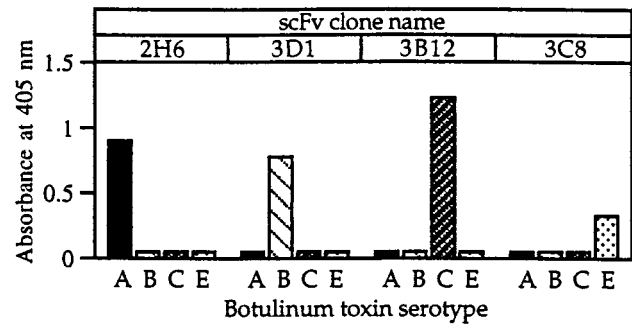


FIG. 3. Specificity of anti-Botulinum neurotoxin scFv. Representative scFv (2H6, 3D1, 3B12, and 3C8) isolated respectively from selections on BoNT serotypes A, B, C, and E were studied. Specificity was determined by ELISA.

from 26 different germ-line genes. V_L genes were derived from three of the six V_{κ} gene families (nos. 1, 3, and 4) and 11 different V_{κ} germ-line genes, from three of the nine V_{λ} gene families (nos. 1–3), and nine different V_{λ} germ-line genes. Despite the diversity, there was a bias seen in the V-gene usage. V_H genes largely were derived from the V_{H3} family, particularly DP46 and DP47. V_{κ} genes most frequently were derived from the $V_{\kappa 1}$ family while V_{λ} genes most frequently were derived from the $V_{\lambda 3}$ family, especially DPL-16. This bias partially reflects the greater frequency of certain V-genes in the B-cell repertoire (30–33) and also in the unselected library (for example DP-47 and DPL-16). Differential V-gene usage also may reflect expression biases of *E. coli*. The number of sequenced V-genes from previous nonimmune phage antibody libraries is small (approximately 30) but a similar bias in V-gene usage is observed (12, 35, 36). V_H CDR length of selected clones was not as evenly distributed as in the unselected clones (Fig. 2) with the majority of lengths between 7 and 15 amino acids. A similar peak is seen in V_H CDR3 length of antibodies generated *in vivo* (37).

Affinity of Selected Antibodies. The antibody-antigen binding affinities were measured for several of the anti-ErbB-2 and anti-BoNT/A scFv. The genes of four anti-ErbB-2 scFv and four anti-BoNT/A scFv were subcloned into a plasmid to add a hexahistidine tag, then expressed and purified from *E. coli*. The dissociation equilibrium constants (K_d) of purified soluble anti-ErbB-2 and anti-BoNT/A scFv were calculated from association and dissociation rate constants measured by using surface plasmon resonance (Table 2) (23, 27). The K_d of the antibodies ranged from 220 pM to 4 nM for anti-ErbB-2 scFv and 38 nM to 71 nM for anti-BoNT/A scFv. The affinity of the

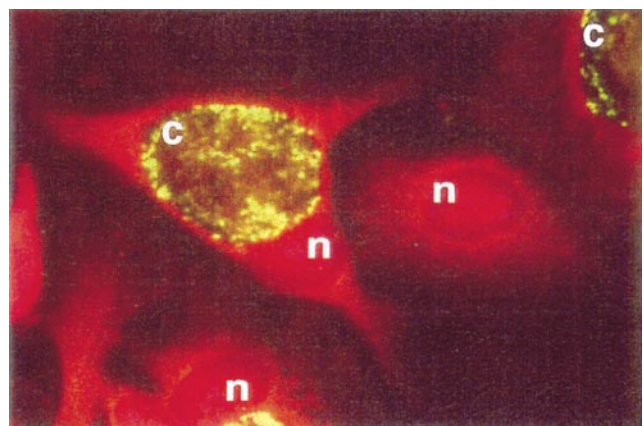


FIG. 4. Staining of HeLa cells infected with *C. trachomatis* with the scFv 2A10. The scFv specifically stains *C. trachomatis* elementary bodies (c) within infected HeLa cells but does not stain uninfected cells. n = nucleus.

Table 2. Affinities and binding kinetics of anti-BoNT A C-fragment and anti-Erb-B2 scFv

Specificity and clone	K_d ($\times 10^{-9}$ M)	k_{on} ($\times 10^5$ M $^{-1}$ s $^{-1}$)	k_{off} ($\times 10^{-3}$ s $^{-1}$)
ErbB-2 B7A	0.22	4.42	0.1
ErbB-2 G11D	0.48	2.19	0.11
ErbB-2 A11A	0.49	3.69	0.18
ErbB-2 F5A	4.03	1.62	0.65
BoNT-A 2A9	26.1	0.25	0.66
BoNT-A 2H6	38.6	2.2	8.5
BoNT-A 3F6	66.0	4.7	30.9
BoNT-A 2B6	71.5	1.1	7.8

Association (k_{on}) and dissociation (k_{off}) rate constants for purified scFv were measured by using surface plasmon resonance (BIAcore) and K_d was calculated as (k_{off}/k_{on}).

anti-ErbB-2 scFv B7A is the highest observed for any antibodies isolated from nonimmune phage antibody libraries (14, 15). The affinities of the isolated scFv are also comparable to affinities of mAbs derived from the secondary immune response (38).

The different K_d s observed for scFv that bind ErbB-2 and BoNT/A are probably a consequence of the different selection conditions used to isolate each panel of antibodies. ErbB-2 antibodies were selected with decreasing concentrations of soluble antigen captured with magnetic beads alternated with selections using immobilized antigen. The use of soluble antigen is a more efficient method for controlling the concentration of antigen used for selection and isolating scFv with higher affinity (23, 39). Therefore, reselection of antibodies using decreasing concentrations of BoNT/A would likely lead to the isolation of antibodies with higher binding affinity.

DISCUSSION

A very large scFv phage antibody library was efficiently generated and its use as a resource for the production of antibodies was extensively evaluated. By using a number of different criteria, the results validate our methods for constructing large libraries of this type and validate the use of these libraries as a resource for the rapid production of antibodies (Table 1). First, by using 14 different proteins for affinity selection, specific antibodies were successfully generated to each of these antigens (Table 1). Second, a high percentage of the antibodies that resulted from affinity selection specifically recognize antigen (Table 1). Third, multiple different antibodies were produced to each antigen (Table 1). Fourth, the binding affinities of the antibodies isolated were comparable to those of mAbs from the secondary murine immune response (Table 2). In addition, these antibody antigen binding affinities are the highest reported for antibodies from nonimmune phage antibody libraries (12, 14, 15). Fifth, isolated scFv served as functional reagents in a number of different immunoassays including ELISA, immunofluores-

cence (Figs. 3 and 4), Western blotting, epitope mapping, and immunoprecipitation (data not shown).

Nonimmune phage antibody libraries can be constructed as either scFv or Fab antibody fragments and from either V-genes rearranged *in vivo* or synthesized *in vitro*. The scFv format was chosen for this library as the expression levels in *E. coli* are typically higher than Fab. This results in more efficient antibody display on phage and more efficient production of native antibody fragments for use. V-genes rearranged *in vivo* were used for library construction to eliminate the need for cloning the individual gene segments necessary for *in vitro* V-gene synthesis. In addition, use of Ig mRNA as the source of V-genes ensures that close to 100% of the gene sequences will be functionally rearranged with ORFs (results from this work and ref. 34). Fewer V-genes will have an ORF when constructed from synthetic oligonucleotides. Furthermore, V-genes rearranged *in vivo* have V_H CDR3s largely derived from the D-gene segments. These genes are not of random sequence but encode amino acids with a propensity for loop formation (40). In contrast, synthetic CDR3s consist of random sequence and thus may be less likely to fold properly or produce usefully shaped binding pockets.

The number and affinities of antibodies produced from this library compare favorably to results from the limited number of phage antibody libraries previously described (Table 3). A comparison of nonimmune libraries illustrates the importance of library size and also suggests that to date, the most useful libraries are those in the scFv format constructed from V-genes rearranged *in vivo*.

Nonimmune phage antibody libraries already are being used as a source of diagnostic and therapeutic antibodies. It is likely that their greatest utility, however, may lie in the laboratory. New genes are rapidly being identified by the genome projects, and the next generation of experiments will shift to elucidating the function of the protein products encoded by these newly identified genes (41). The production of antibodies with phage-displayed libraries is ideally suited for the large-scale determination of protein function. For example, once a gene has been sequenced, the protein(s) that it encodes can be overexpressed and then used to rapidly select phage-displayed antibodies. The resulting antibodies would provide immunological reagents for protein characterization. In addition, the production of antibodies with phage display also provides access to the genes that encode specific antibodies. These antibody genes can be used to express antibody proteins within cells to block and elucidate the function of specific molecules *in vivo* (6–8).

In summary, the steps of phage antibody library construction have been optimized to facilitate the rapid and efficient construction of large phage antibody libraries. With this current library we obtain panels of high-affinity antibodies to a wide array of antigens. The approach used puts this technique within reach of laboratories skilled in molecular biology. Subsequent uses for these libraries will be limited only by the investigator's imagination.

Table 3. Comparison of protein binding antibodies selected from nonimmune phage-display antibody libraries

Library	Library size and type*	Number of protein antigens studied	Average number of antibodies per protein antigen	Number of affinities measured	Range of affinities for protein antigens K_d ($\times 10^{-9}$ M)
Marks <i>et al.</i> (12)	3.0×10^7 (scFv, N)	2	2.5	1	100–2000
Nissim <i>et al.</i> (13)	1.0×10^8 (scFv,SS)	15	2.6	ND	ND
deKruif <i>et al.</i> (42)	3.6×10^8 (scFv,SS)	12	1.9	3	100–2,500
Griffiths <i>et al.</i> (14)	6.5×10^{10} (Fab, SS)	30	4.8	3	7.0–58
Vaughan <i>et al.</i> (15)	1.4×10^{10} (scFv, N)	3	7.0	3	4.2–8.0
Sheets <i>et al.</i> (this work)	6.7×10^9 (scFv, N)	14	8.7	8	0.22–71.5

*For library type, N = V-gene repertoires obtained from V-genes rearranged *in vivo*; SS = semisynthetic V-genes constructed from cloned V-gene segments and synthetic oligonucleotides encoding V_H CDR3. ND, not determined.

We thank A. Suzuki and N. Ueno for the BMP and activin receptor cDNAs, E. Amaya for the fibroblast growth factor receptor cDNA, and Catherine Fox for valuable comments on the manuscript. This research was supported by National Institutes of Health Postdoctoral Grant GM15203-02 (M.D.S.), National Institutes of Health Grant R01 GM 19363 (J.C.G.), U.S. Army Medical Research Grants DAMD17-94-C-4034, and DAMD17-94-J-4433 (J.D.M.), and the CaPCURE Foundation (J.D.M.).

1. Marks, J. D., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1992) *J. Biol. Chem.* **267**, 16007–16010.
2. Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. (1994) *Annu. Rev. Immunol.* **12**, 433–455.
3. Marks, C. & Marks, J. D. (1996) *N. Eng. J. Med.* **335**, 730–733.
4. Huston, J. S., Levinson, D., Mudgett, H. M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R., *et al.* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5879–5883.
5. Pastan, I., Pai, L. H., Brinkmann, U. & Fitzgerald, D. (1996) *Breast Cancer Res. Treat.* **38**, 3–9.
6. Marasco, W. A., Haseltine, W. A. & Chen, S. Y. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 7889–7893.
7. Biocca, S., Pierandrei-Amaldi, P., Campioni, N. & Cattaneo, A. (1994) *Bio/Technology* **12**, 396–399.
8. Marasco, W. A. (1997) *Gene Therapy* **4**, 11–15.
9. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990) *Nature (London)* **348**, 552–554.
10. Parmley, S. F. & Smith, G. P. (1988) *Gene* **73**, 305–318.
11. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991) *Nature (London)* **352**, 624–628.
12. Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. & Winter, G. (1991) *J. Mol. Biol.* **222**, 581–597.
13. Nissim, A., Hoogenboom, H. R., Tomlinson, I. M., Flynn, G., Midgley, C., Lane, D. & Winter, G. (1994) *EMBO J.* **13**, 692–698.
14. Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L., Kontermann, R. E., Jones, P. T., Low, N. M., Allison, T. J., *et al.* (1994) *EMBO J.* **13**, 3245–3260.
15. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. & Johnson, K. S. (1996) *Nat. Biotechnol.* **14**, 309–314.
16. Gussow, D. & Clackson, T. (1989) *Nucleic Acids Res.* **17**, 4000.
17. Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Hudson, P. & Winter, G. (1991) *Nucleic Acids Res.* **19**, 4133–4137.
18. Nishimatsu, S., Iwao, M., Nagai, T., Oda, S., Suzuki, A., Asashima, M., Murakami, K. & Ueno, N. (1992) *FEBS Lett.* **312**, 169–173.
19. Maeno, M., Ong, R. C., Suzuki, A., Ueno, N. & Kung, H. F. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10260–10264.
20. Suzuki, A., Thies, R. S., Yamaji, N., Song, J. J., Wozney, J. M., Murakami, K. & Ueno, N. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10255–10259.
21. Musci, T. J., Amaya, E. & Kirschner, M. W. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8365–8369.
22. McCartney, J. E., Tai, M.-S., Hudziak, R. M., Adams, G. P., Weiner, L. M., Jin, D., Stafford III, W. F., Liu, S., Bookman, M. A., Laminet, A., *et al.* (1995) *Protein Eng.* **8**, 301–314.
23. Schier, R., Bye, J. M., Apell, G., McCall, A., Adams, G. P., Malmqvist, M., Weiner, L. M. & Marks, J. D. (1996) *J. Mol. Biol.* **255**, 28–43.
24. Koehler, J. E., Burgess, R. B., Thompson, N. E. & Stephens, R. S. (1990) *J. Biol. Chem.* **265**, 13206–13214.
25. Tomlinson, I. M., Williams, S. C., Corbett, S. J., Cox, J. P. L. & Winter, G. (1995) *vBASE Sequence Directory* (MRC Centre for Protein Engineering, Cambridge, U.K.).
26. Schier, R., Marks, J. D., Wolf, E. J., Apell, G., Huston, J. S., Weiner, L. M. & Adams, G. P. (1995) *Immunotechnology* **1**, 63–71.
27. Johnsson, B., Löfås, S. & Lindqvist, G. (1991) *Anal. Biochem.* **198**, 268–277.
28. Zhang, J. P. & Stephens, R. S. (1992) *Cell* **69**, 861–869.
29. Evan, G. I., Lewis, G. K., Ramsay, G. & Bishop, J. M. (1985) *Mol. Cell. Biol.* **5**, 3610–3616.
30. Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. & Winter, G. (1992) *J. Mol. Biol.* **227**, 776–798.
31. Cox, J. P. L., Tomlinson, I. M. & Winter, G. (1994) *Eur. J. Immunol.* **24**, 827–836.
32. Davidkova, G., Pettersson, S., Holmberg, D. & Lundkvist, I. (1997) *Scand. J. Immunol.* **45**, 62–73.
33. Williams, S. C., Frippiat, J.-P., Tomlinson, I. M., Ignatovich, O., Lefranc, M.-P. & Winter, G. (1996) *J. Mol. Biol.* **264**, 220–232.
34. Marks, J. D., Tristrem, M., Karpas, A. & Winter, G. (1991) *Eur. J. Immunol.* **21**, 985–991.
35. Marks, J. D., Ouwehand, W. H., Bye, J. M., Finnern, R., Gorick, B. D., Voak, D., Thorpe, S. J., Hughes-Jones, N. C. & Winter, G. (1993) *Bio/Technology* **11**, 1145–1149.
36. Griffiths, A. D. & Malmqvist, M. (1993) *EMBO J.* **12**, 725–734.
37. Wu, T. T., Johnson, G. & Kabat, E. A. (1993) *Proteins* **16**, 1–7.
38. Foote, J. & Milstein, C. (1991) *Nature (London)* **352**, 530–532.
39. Hawkins, R. E., Russell, S. J. & Winter, G. (1992) *J. Mol. Biol.* **226**, 889–896.
40. Abergel, C. & Claverie, J.-M. (1991) *Eur. J. Immunol.* **21**, 3021–3025.
41. Lander, E. S. (1996) *Science* **274**, 536–539.
42. deKruif, J., Boel, E. & Logtenberg, T. (1995) *J. Mol. Biol.* **248**, 97–105.