Flow Cytometric Isolation of Human Antibodies from a Nonimmune *Saccharomyces cerevisiae* Surface Display Library

Michael J. Feldhaus[*†], Robert W. Siegel[*], Lee K. Opresko[*], James R. Coleman[*], Yik A. Yeung, Jennifer R. Cochran, Peter Heinzelman, David Colby, Jeffrey Swers, Christilyn Graff, H. Steven Wiley[*], K. Dane Wittrup[†]

Dept. Chemical Engineering and
Biological Engineering Div.,
Massachusetts Institute of Technology

[*]Pacific Northwest National Laboratories
Biological Sciences Division

[†]Corresponding authors:

902 Battelle Blvd.
P.O. Box 999
Richland, WA 99352
PH 509-376-2592
FAX 509-376-2592
Michael.Feldhaus@pnl.gov

MIT 66-552
25 Ames Street
Cambridge, MA 02139
PH 617-253-4578
FAX 617-258-5766
wittrup@mit.edu

Appeared as an article in Nature Biotechnology 21(2):163-70 January 2003 doi: 10.1038/nbt785

Abstract

A nonimmune library of 10^9 human antibody scFv fragments has been cloned and expressed on the surface of yeast, and nanomolar-affinity scFvs routinely obtained by magnetic bead screening and flow cytometric sorting. The yeast library can be amplified 10^{10}-fold without measurable loss of clonal diversity, enabling effectively indefinite expansion of the library. The expression, stability, and antigen binding properties of more than 50 isolated scFv clones were assessed directly on the yeast cell surface by immunofluorescent labeling and flow cytometry, obviating separate subcloning, expression, and purification steps and thereby expediting the isolation of novel affinity reagents. The ability to use multiplex library screening demonstrates the utility of this approach for high throughput antibody isolation for proteomics applications.

Therapeutic, diagnostic, and analytical applications for antibodies are growing rapidly [1-3]. The identification and development of protein-specific affinity reagents, in a rapid and cost effective manner, can facilitate the deciphering of the function, location, and interactions of the encoded protein products identified through various sequencing projects. Human antibodies are of particular interest for therapeutic applications due to potential immunogenicity issues [4]. Consequently, human antibody repertoires have been cloned and expressed in transgenic mice [5]; however, this in vivo technology is not amenable to high throughput genomic applications. Single chain variable fragments of antibodies (scFvs) displayed on phage provide an alternate means for isolating antibodies in vitro[6]. A number of large ($> 10^9$) scFv libraries have been constructed and displayed using phage display [7-9] or ribosome display[10].

Each scFv display system has particular strengths and weaknesses. For example, antibodies that bind tightly to specific epitopes within a given expression host will be lost in each case. Yeast surface display (Figure 1A) was originally developed to enhance affinity, stability, and expression maturation of proteins[11-13], but has not to date been used to identify new antibodies from nonimmune repertoires. Fluorescence activated cell sorting (FACS) used in the yeast display system allows rapid quantitative isolation of rare clones with defined binding parameters, with high statistical confidence[14-16]. With equilibrium screening the minimum affinity threshold can be dictated, and enrichment monitored throughout the screening process to identify clones with the highest affinity (Figure 1B). The binding properties of multiple individual isolated clones can be rapidly and quantitatively determined directly on the yeast surface, obviating the need for subcloning, expression, and purification. Additionally, screening with monovalent soluble antigen eliminates avidity effects and problems with inaccessible epitopes of immobilized antigens encountered with some other display systems.

We have cloned a library of 10^9 human antibody scFv fragments, expressed the repertoire on the yeast cell surface, and developed flow cytometric and magnetic bead screening methods to isolate scFvs with nanomolar antigen binding affinity. These scFvs

have been utilized without further engineering for microplate assays and immunofluorescent labeling of cells. Novel attributes of this yeast antibody library include essentially indefinite expansion without detectable expression biases, and the capability to perform simultaneous multiplex screens with at least a dozen different protein, peptide, and hapten antigens. This is the first nonimmune antibody library to be cloned, expressed, and screened on the surface of cells. It is also the largest reported yeast expression library of any kind.

Results

Repertoire Cloning. A broadly representative sample of the expressed human antibody V gene repertoire was cloned by PCR from commercially available human spleen and lymph node poly A mRNA pooled from 58 adults (Clontech, Origene, Biochain), by previously described methods. First strand cDNA synthesis was performed with IgG, IgM, kappa, and lambda primers [7], and the cDNA template was amplified by PCR with primers designed and validated to amplify all human germline gene families [17]. The underlying cDNA template diversity was confirmed for each reaction by RT-PCR, and estimated to be greater than 10^6 for each primer pair (data not shown). The primers were modified for cloning into heavy chain and light specific vectors. The base vectors are pBluescript with several modifications needed for subsequent overlap PCR to join the fragments. Separate VH (8×10^7 clones) and VL (1.2×10^6 clones) libraries were constructed, then assembled together in the scFv format by overlap extension PCR as previously described [8]. The scFv library was then cloned into the modified pCT302 vector for expression as an Aga2p fusion on the yeast surface, flanked by an HA epitope tag at the N terminus and a c-myc tag at the C terminus (Figure 1A). 88 scFv clones from the unselected library were sequenced, and representation of the major VH and VL gene families is shown in Figure 2. The yeast library exhibits germline gene family usage that is similar to that reported for human clinical samples. However, significant over-representation of VH6 and Vκ3, and mild under-representation of VH3 and Vκ1 was observed. These differences are most likely due to the normalization of the PCR products before cloning and variation in PCR efficiency. The cloned repertoire samples

the germline gene families most frequently observed in clinical samples (e.g. VH3, Vκ1, Vκ3)[18-20]. The gene family usage of scFvs screened from the library for antigen binding, as described below, is in general similar to the diversity of the unscreened library (Figure 2).

scFv expression on the yeast cell surface and library propagation. 10^9 yeast transformants of the human scFv library were obtained. Presence of the full-length Aga2p-scFv fusion on the cell surface is highly indicative of proper folding, due to the quality control function of the eucaryotic secretory apparatus [21]. ScFv expression in the display format was characterized for over 100 individual yeast transformants with ten full length clones being sequenced and germline gene usage determined (Figure 3A). The surface expression of the N-terminal HA and/or C-terminal c-myc epitope tag was detected by indirect immunofluorescent labeling and flow cytometry (Figure 3A). 61/90 (68%) unsequenced clones expressed c-myc (Figure 3B), indicating that the majority of the library is expressed as a full-length scFv on the yeast surface in a format suitable for screening for antigen-binding activity. Sequencing confirmed the presence of stop codons inadvertently introduced by PCR upstream of the c-myc epitope for the two clones exhibiting no c-myc labeling (Figure 3A). It should be noted that as a general phenomenon with the yeast display system, for a given single scFv clone, only 40-80% of the cells in a culture express the scFv on the surface. The non-expressing cell subpopulation is attributable to a combination of causes: plasmid loss; the polarized growth morphology of yeast, such that mother cells do not express the fusion on their cell walls[22]; and variable expression amongst scFv clones[23]. The overall library labels approximately 40% c-myc positive (Figure 3C), as a product of the average clone's 60% c-myc labeling and 68% of all clones coding for an expressible scFv (i.e. 0.6 * 0.68 = 0.41.)

A yeast library of this size requires considerable effort to construct and characterize, and it is therefore highly desirable that such a library can be amplified and propagated without introducing expression biases that would alter clonal representation. However, propagation of expression libraries can introduce substantial biases due to

growth selections[24-26]. To allow expansion of the yeast display library without repertoire bias, scFv expression is under control of the tightly repressible GAL1,10 promoter. When the library or individual scFv clones are grown under non-inducing conditions (i.e. glucose media) scFv expression is not detectable by flow cytometry (Figure 3C). To test the potential expansion capability, an aliquot of the library was mixed with yeast cells expressing a particular anti-fluorescein antibody (4-4-20), such that approximately 5% of the mixture displayed the 4-4-20 scFv. This mixture was then passaged five times, each passage consisting of a 100-fold amplification, and a –80°C freeze/thaw cycle. Following each passage, an aliquot of the mixture was induced for scFv surface expression, and the fluorescein-binding fraction determined by flow cytometry. As a further gross measure of library diversity, the fraction of cells displaying the C-terminal c-myc epitope tag was determined. Remarkably, throughout an overall 10^{10}-fold amplification of the library, both the 4-4-20 fraction and the c-myc immunoreactive fraction were essentially constant, indicating that there was no detectable bias in library diversity introduced during amplification (Figure 3D). Further indirect evidence that library diversity is not biased by propagation is the qualitative agreement in germline gene family usage between the unscreened library clones and clones isolated for specific antigen binding properties (Figure 2, open vs. gray bars).

Library screening by magnetic beads and flow cytometry. A sequential magnetic bead / flow cytometric sorting strategy was developed to isolate antigen-specific scFvs from the full library. High speed flow cytometric sorting at 10^8 cells/hr would require several days to comprehensively oversample (10X) the full library diversity of 10^9 clones, while magnetic bead sorting is amenable to screening 10^{10} cells in under an hour [16]. A validation experiment was performed in which yeast cells expressing a FITC-binding scFv (4-4-20) were mixed with a 10^6-fold excess of cells from the nonimmune scFv library, and then serially passaged five times as described above for the library expansion validation experiments (Figure 3D). The cell populations from the first passage and the fifth passage were labeled with 100 nM FITC-biotin, and one round of enrichment with magnetic beads (either Miltenyi or Dynal) followed by one round of flow cytometric sorting using streptavidin-phycoerythrin labeling was performed.

Remarkably, a 10^5-fold enrichment was obtained for the rare 4-4-20 cells in both the first and the fifth passages, indicating that clones present at 1 in 10^6 cells are retained and can be isolated in two steps, even following a 10^{10}-fold amplification of the library (Table I). The rapid and robust enrichment of a rare clone by two different magnetic bead screening methods suggests that novel clones can be isolated from the full library with similar success. This reproducibility is an important strength of the yeast-displayed library[27].

We screened the yeast-displayed human scFv repertoire against protein, peptide, and hapten antigens using magnetic bead screens, flow cytometry, and a combination of both of the two (Table II). We found a key methodological step is to use either directly fluorophore-labeled antigen, or alternation of capture reagents in subsequent enrichment cycles (e.g. for biotinylated antigens, anti-biotin IgG vs. streptavidin) to avoid isolation of scFvs that bind specifically to the capture reagents (e.g. scFvs that bind to streptavidin-phycoerythrin can be isolated from the library in the absence of such alternation.) The numbers of clones isolated against our antigen panel and reported here are not exhaustive, but instead represent single-pass screens and classification of a dozen isolated colonies by BstNI fingerprinting and/or sequencing (Experimental Protocols). No correlation was seen between gene family usage and the affinity of the scFv for its cognate antigen. Gene family usage of the clones isolated for specific antigen binding properties is similar to that for the unscreened library (Figure 2). Thus, although VH6 is infrequently expressed in clinical samples[20], it is isolated in screens from this library in proportion to its representation in the unscreened library (Figure 2A).

The soluble antigen concentration used for library screening determines the diversity and affinity of the clones isolated: labeling with EGFR-ECD (epidermal growth factor extracellular domain) antigen at a concentration of 100-1,000 nM resulted in the isolation of more than 15 scFv clones with affinity in the same range (Table II). More aggressive labeling strategies at 1-10 nM antigen results in isolation of scFv clones in that affinity range (e.g. fluorescein, HEL (hen egg lysozyme), Table II). For EGF, two parallel selections were performed with antigen concentrations of 100 nM or 1 μM, and the same high affinity scFv was isolated in both screens, while lower affinity scFvs were only isolated in the 1μM screen. The ability to select high affinity clones based on

soluble antigen concentration significantly reduces the time required to classify isolated clones according to desired affinity. ScFv/antigen binding constants can be measured by titration of yeast-displayed scFv with labeled antigen, as described previously [15,28]. Sample titrations of high affinity scFvs isolated by screening at 10 nM (lysozyme, fluorescein) or 100 nM (EGF) soluble antigen concentrations are shown in Figure 4A. It should be noted that although these curves represent clones obtained from three independent library screens against three different antigens, the affinities of the isolated clones cluster near the antigen concentration utilized for each screen, indicative of the quantitative nature of the screening threshold.

Multiplex library screening against multiple antigens. The yeast antibody library can be screened for simultaneous identification of scFvs against multiple antigens. For example, the data presented in Table II for EGF, amphiregulin, betacellulin, HB-EGF, and XPA were obtained from a simultaneous screen against these 5 antigens equally mixed into a single antigen pool. As a stringent test of the ability to simultaneously isolate scFvs with fine differences in specificity, 12 different phospho-peptide fragments of human p53, each twelve amino acids in length, were mixed together, for a total of 12 closely related antigens. The peptides were biotinylated at their N termini, and the library was screened at 1 μM concentration of each peptide in the pool, using magnetic bead screening followed by flow cytometry. The binding affinities and specificities, comparing the phosphorylated to the unphosphorylated peptide, of five different scFvs isolated in this screen are shown in Table III. One scFv, designated 18-36, bound to 3 different peptides, peptide 18T in the phosphorylated and unphosphorylated form and interestingly to 20S which overlaps the 18T peptide at 10 of the amino acids. The highly specific scFv, designated 9a, bound only the 9S peptide in the phosphorylated form, showing quantitative specificity.

ScFv clones isolated from the yeast library can be used directly in a number of immunoassay applications. For example, a number of scFvs against the EGFR-ECD were subcloned into yeast secretion vectors and expressed in soluble form at 1-10 mg/L with an N-terminal FLAG epitope tags and a C-terminal c-myc tag. The purified scFvs were used in microplate assays to quantify soluble EGFR ECD, and can also be used to

immunofluorescently label fibroblasts expressing human EGFR (data not shown). Soluble anti-EGF scFvs can be used to immunofluorescently label CHO cells expressing surface-bound EGF (data not shown.)

Discussion

Several alternative systems have been reported for isolation of human antibodies from large nonimmune human antibody repertoires: transgenic mice, phage display, ribosome display, periplasmic expression in *E. coli*[29], and yeast display as reported here. Key attributes of these alternative systems are: repertoire size and expression bias; screening efficiency; and mutation frequency and control. Large antibody repertoires (> 10^9 clones) have been constructed several times previously in phage display formats, with nonimmune diversity [8,9,30-32] or synthetic diversity incorporated into CDRs [10,33,34]. Such libraries have ranged from 10^7-10^{11} clones. In libraries of 10^7-10^8, typical affinities of screened antibodies were generally in the 100-1,000 nM range, while repertoires from 10^9-10^{11} produced binders in the 1-100 nM affinity range. Antibodies isolated against haptens generally have higher affinity than those against protein antigens, and anti-peptide antibodies are generally of lower affinity than anti-protein antigens. The yeast-displayed antibody library described here produces a distribution of affinities comparable to those reported for phage-displayed repertoires of similar or larger size (Figure 4B). For applications demanding picomolar affinity (e.g. tumor targeting, immunoassays for dilute analytes), yeast display is well suited for subsequent rapid affinity maturation [12,13,35].

Library size is a critical variable for isolation of novel binders. However, the true expressed diversity of a library is not experimentally determinable, and is distinct from the raw number of independent transformants. The ability to monitor both the starting nonimmune library and the selection process by FACS allows a direct estimate of functional diversity at each step in the process. For example, polyreactive scFvs that bind secondary detection reagents are present in most nonimmune antibody libraries[36], and these clones can be detected and then eliminated from a yeast display screen by alternating secondary reagents (e.g. streptavidin vs. anti-biotin IgG) or antigen-deleted depletion sorting (data not shown). Furthermore, as a single-celled eucaryote, yeast is

well suited to the expression of eucaryotic secretory proteins such as antibodies. For example, the 4-4-20 anti-fluorescein scFv is poorly expressed in E. coli [37] and poorly displayed on phage [38], but is well displayed on the yeast cell wall and secreted at approximately 1 mg/L from yeast [13].

Expression biases and growth selections can rapidly skew library diversity[26,39], so much so that it is common practice not to serially propagate phage display libraries. The problem of toxic clone expression in phage display has been partially addressed by the use of regulated promoters in *E. coli*[24-26], however some level of fusion protein expression is a necessary component of each screening step, introducing potential selection biases. One approach to minimizing the loss of library diversity due to growth selection during phage propagation is to array antibodies and screen for binders in a single step[40]; however, the number of clones addressable in this format is several orders of magnitude smaller than in solution phage screening or the yeast display library described here. In this regard, the stable propagation of the full 10^9 clone yeast library following a 10^{10}-fold amplification (Figure 2 and Table I) demonstrates an ability to greatly expand this library.

The starting antibody repertoire[41] of a mouse is no larger than 10^8, and initial binders from that population are iteratively improved by somatic mutagenesis and clonal selection. However, the affinity maturation process in vivo is constrained by B cell physiology [42], and neither the biophysical endpoints or the timing of the process is under the investigator's control. It is interesting to note that antibodies with single-digit nanomolar affinity are generally present in a repertoire of 10^9 that can be quantitatively screened in vitro as described here, indicating that such high affinity clones may in many cases be present in the pre-immunized repertoire of an animal[43] or created by novel V_H and V_L pairing during library construction. To the extent that variable domain immunogenicity is problematic [4], it is desirable to develop "fully human" antibodies with framework residues identical to germline sequences. Somatic mutation in transgenic mice targets both framework and CDR residues, with an average of 5 framework

mutations per antibody [5]. However, since human somatic mutagenesis also alters V region framework residues [44], it is not clear how significant a problem this will be. An advantage of both yeast display and phage display is that mutagenesis during affinity maturation can be targeted specifically to CDR residues.

The high rate of protein discovery and the need to evaluate biological processes at the organismal level requires high-throughput generation of affinity reagents, and antibodies are particularly useful in this regard[1,2]. Multiplex antigen screening will enable parallel isolation of antibodies against multiple antigens in a rapid (~2-3 weeks) and robust manner. For example, a bank of expressed protein for each ORF in the human genome is being accumulated [45], which could be labeled and pooled in batches of at least ten, or potentially hundreds, of ORF gene products for screening against this library. In addition, panels of antibodies against multiple isoforms of a given protein can be identified as shown here for p53 phosphopeptides. This capability will be important for functional genomics studies that utilize immunoprecipitation and mass spectrometry in organisms not well suited to large scale epitope tag expression [46].

The yeast antibody library described here is a robust, propagatable source of human antibodies with high affinity and specificity. Properties of lead antibodies can be rapidly screened without subcloning, and if necessary lead optimization can be performed in the same display format with high efficiency. These capabilities are complementary to those of the existing suite of antibody engineering methodologies, and extend the potential reach of immunotechnology.

Acknowledgments. This work was funded in part by NSF BPEC, the Hereditary Disease Foundation, and by the Laboratory Directed Research and Development program of the DOE. The p53 phosphopeptides were kindly provided by Ettore Appella, Shari Mazur and Yuichiro Higashimoto (NCI). The authors are grateful for comments on the manuscript by David Kranz, Michael Roguska, and Eric Shusta.

Experimental Protocols

cDNA cloning by PCR and construction of scFv in yeast display vector. Commercial sources of RNA (Clontech Human Normal lymph node 5μg PolyA+ RNA pooled from 29 females/males, Clontech Human Normal Spleen 5μg PolyA+ RNA pooled from 14 females/males, Origene Human Normal Spleen 5μg PolyA+ RNA pooled from 7 females/males, Biochain Inst. Human normal Spleen 5μg PolyA+ RNA pooled from 8 females/males) was utilized as template. Superscript II first strand synthesis reactions were set up to generate cDNA from RNA obtained from four sources listed above. 16 reactions total (4 reactions for each RNA source) were set up using gene specific primers to amplify the IgG, IgM heavy chains, and the λ and κ light chains. The HuIgG , HuIgM , Huκ For and Huλ For: primers used for the gene specific cDNA synthesis are from Marks *et al* [7]. 16 PCR reaction were set up on the Roche light cycler and followed in real time, for the 4 samples x 4 first strand primers. The first strand synthesis reactions from the four different sources were pooled as the real time PCR verified amplification from the first strand material was about equal. The 4 first strand reactions that were synthesized using the HuIgM primer were pooled, as were the 4 for HuIgG, 3 for HuGk, and 4 for HuGL.

The primers used are from Sblattero &Bradbury[17] with a slight modification to allow for cloning into our heavy and light chain vectors. The modifications consist of the addition of the following sequence to the 5'end of the "Forward" primers and 5'end of the "Back" primers (GCCACCTAGAATTCC for VH front primers), (TGGCACTAGTGCTAGC for all VH back primers), (TGTTCTAGAATTCCGGA for all forward Vλ and Vκ primers), and (TGGTTCCGGAGGCGGCGGTTCT for all back Vλ and Vκ primers). Appropriate volumes of the PCR reactions were mixed together to normalize the amount of material being pooled from each PCR reaction. This created 3 pools: lambda + Kappa light chain variable region mix, IgG heavy chain variable region, and IgM heavy chain variable region. The material for the light chain library is digested with EcoR1/BspE1 and cloned into pVariable-light. The library is 1.2 x 10^6 clones. Ninety six colonies were screened by colony PCR to determine presence of insert and then sequenced. 90% of the clones contained an insert of the appropriate size. The IgG and IgM Heavy chain material was use to create an IgG and a separate IgM heavy chain library by the parallel processing of the material. The material was digested with the restriction enzymes SpeI and EcoRI and run on an agarose gel and band purified before being cloned into a modified pBluescript plasmid. Two ligations with 100 ng of cut vector were set up, cleaned up and electroporated into electro competent *E. coli*. This created two libraries, IgM variable region heavy chain (2.9 x 10^7 clones) and an IgG variable heavy chain library (4.9 x 10^7 clones). PFU polymerase was used for PCR amplification of λκ light chain library, IgG, and IgM.

ScFvs were constructed by overlap extension PCR essentially as described previously[8] Ten 50μL amplifications were set up for each. Approximately 3 x 10^8 templates were used in each reaction (1 ng of super coiled plasmid containing the heavy chain library or the light chain library). 1 PCR reaction was set up to join IgG to λκ and another to join IgM to λκ. 50 ng of PFU amplified IgG template (2 x 10^{11} molecules) is mixed with 50 ng of PFU amplified λκ template. The material is joined by overlap extension using 10 cycles of the following protocol in 100 μL reaction. 5 μL of the joined heavy and light chain material (representing approximately 10^{10} molecules) is added to a 50 μL PCR reaction containing the outside primers of the constructs. The material from both sets of reaction were pooled in equal amounts and digested with the restriction enzymes Not1 and Nhe1.

pCTCON, a derivative of pCT302 was digested with Xho1 and the following oligo (TCGATGCGGCCGCA) was annealed to itself to create a double stranded fragment that could be cloned, to create pPNL6. This oligo destroys the Xho1 site and replaces it with a Not1 site. This was done since the restriction enzyme Not1 does not cut within any of the V, D, or J genes of either heavy or light chains. pPNL6 was prepared to accept the inserts by double digesting the vector with Not1 and Nhe1 and then band isolating the material and treating the plasmid with Shrimp alkaline phosphatase. The Nhe1/Not1 digested band isolated vector and insert were ligated together is 10 separate ligations. All of the cleaned and pooled ligation was transformed into ElectroMax competent *E. coli* in 25 electroporations. The dilution plates demonstrated that 1.53 x 10^9 total transformants were generated for the scFv library. The library consisted of 90% insert of the correct size and 10% insert of a smaller fragment representing heavy chain only clones. The library was then transformed into EBY100 yeast using high efficiency lithium acetate transformation [47]. ScFv clones were identified by 4-base cutter fingerprinting essentially as described previously.[7] Colony PCR was used to amplify the scFv out of the expression vector and was subsequently

restriction digested with either BstN1 or Sau3a. The digestion products were electrophoresed on a 2-3% NuSieve gel and visualized.

Magnetic bead enrichment.

The library is resuspended in a total volume of 5 mL and the antigen (one or more) to be screened is added at the desired concentration (typically 100 nM). The library is incubated at room temp for 30 min. followed by a 5 min. incubation on ice. 45 mL of ice cold PBS (phosphate buffered saline) is added and the cells are then pelleted by centrifugation at 4 ^{o}C. The cells are washed two more times with 50 mL and then resuspended in a total volume of 5 mL PBS with 500 µL Miltenyi Macs streptavidin magnetic particles. The cells are incubated on ice for 20 min. with mixing every few minutes. 45 mL of ice cold wash buffer is added and the cells are then pelleted by centrifugation at 4 ^{o}C. The cells are washed once more with 50 mL buffer and then resuspended in 10 mL of buffer, passed through a 20 µm mesh screen and then loaded on the LS Mac column. The column is washed 1x with 3 mL and 2x with 7 mL wash buffer and gravity flow. The bound cells are eluted by removing the column from the magnetic apparatus and adding 7 mL of buffer which is forced though the column with a plunger. These cells are pelleted and grown overnight in 10-100 mL of SD+CAA[28]. The culture will generally be at or near 5OD/mL. The cells are then pelleted and induced as previously described[28]. The selection process is repeated.

Cell staining for flow cytometry.

Typically 1-5 x 10^7 cells are eluted from the column. Cells are resuspended in 200 µL wash buffer to which 1µg of anti-c-myc (9E10, Covance, BAbco) mAb is added. The cells are incubated on ice for 30 min. The cells are washed 3x with 1 mL of wash buffer and then the secondary detection reagents are added together. For the antigen we prefer streptavidin Alexa 633 and to alternate in subsequent sorts with Neutravidin FITC, anti-bio Alexa 488, or tetrameric-streptavidin phycoerythrin, (Molecular Probes.) Goat anti-mouse Alexa 488 or phycoerythrin conjugated antibodies are used to detect the bound anti-c-myc antibody. The cells are incubated on ice for 30 min. and washed twice before being sorted on the flow cytometer. Typically the top 0.01-0.1% of the c-myc$^+$ Ag$^+$ cells are sorted. Therefore, approximately 50,000-200,000 cells are sorted into a tube for further expansion and screening.

K_d determination.

Quantitative equilibrium binding was determined as described previously [15,28] To verify the protein-ligand dissociation constant K_d within the surface display context we performed flow cytometric analysis of the c-myc normalized antigen binding. In general we used 10 different antigen concentrations that covered 1nM-1μM total antigen concentration.

References

1. Schweitzer, B., *et al*. Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nature Biotechnology* **20**, 359-365 (2002).

2. MacBeath, G. & Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760-1763. (2000).

3. Gura, T. Therapeutic antibodies: Magic bullets hit the target. *Nature* **417**, 584-586. (2002).

4. Clark, M. Antibody humanization: a case of the 'Emperor's new clothes'? *Immunol Today* **21**, 397-402. (2000).

5. Mendez, M. J. *et al*. Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice. *Nat Genet* **15**, 146-156. (1997).

6. Winter, G., Griffiths, A. D., Hawkins, R. E. & Hoogenboom, H. R. Making antibodies by phage display technology. *Annu Rev Immunol* **12**, 433-455 (1994).

7. Marks, J. D. *et al*. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol* **222**, 581-597. (1991).

8. Sheets, M. D. *et al*. Efficient construction of a large nonimmune phage antibody library: the production of high-affinity human single-chain antibodies to protein antigens. *Proc Natl Acad Sci U S A* **95**, 6157-6162. (1998).

9. Vaughan, T. J. *et al*. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat Biotechnol* **14**, 309-314. (1996).

10. Knappik, A. *et al*. Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* **296**, 57-86. (2000).

11. Shusta, E. V., Holler, P. D., Kieke, M. C., Kranz, D. M. & Wittrup, K. D. Directed evolution of a stable scaffold for T-cell receptor engineering. *Nat Biotechnol* **18**, 754-759. (2000).

12. Holler, P. D. *et al*. In vitro evolution of a T cell receptor with high affinity for peptide/MHC. *Proc Natl Acad Sci U S A* **97**, 5387-5392. (2000).

13. Boder, E. T., Midelfort, K. S. & Wittrup, K. D. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc Natl Acad Sci U S A* **97**, 10701-10705. (2000).

14. Daugherty, P. S., Iverson, B. L. & Georgiou, G. Flow cytometric screening of cell-based libraries. *J Immunol Methods* **243**, 211-227. (2000).

15. VanAntwerp, J. J. & Wittrup, K. D. Fine affinity discrimination by yeast surface display and flow cytometry. *Biotechnol Prog* **16**, 31-37. (2000).

16. Yeung, Y. A. & Wittrup, K. D. Quantitative screening of yeast surface-displayed polypeptide libraries by magnetic bead capture. *Biotechnol Prog* **18**, 212-220. (2002).

17. Sblattero, D. & Bradbury, A. A definitive set of oligonucleotide primers for amplifying human V regions. *Immunotechnology* **3**, 271-278. (1998).

18. Ignatovich, O., Tomlinson, I. M., Jones, P. T. & Winter, G. The creation of diversity in the human immunoglobulin V(lambda) repertoire. *J Mol Biol* **268**, 69-77. (1997).

19. Foster, S. J., Brezinschek, H. P., Brezinschek, R. I. & Lipsky, P. E. Molecular mechanisms and selective influences that shape the kappa gene repertoire of IgM+ B cells. *J Clin Invest* **99**, 1614-1627. (1997).

20. Brezinschek, H. P. *et al*. Analysis of the human VH gene repertoire. Differential effects of selection and somatic hypermutation on human peripheral CD5(+)/IgM+ and CD5(-)/IgM+ B cells. *J Clin Invest* **99**, 2488-2501. (1997).

21. Ellgaard, L., Molinari, M. & Helenius, A. Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882-1888. (1999).

22. Field, C. & Schekman, R. Localized secretion of acid phosphatase reflects the pattern of cell surface growth in Saccharomyces cerevisiae. *J Cell Biol* **86**, 123-128. (1980).

23. Shusta, E. V., Raines, R. T., Pluckthun, A. & Wittrup, K. D. Increasing the secretory capacity of Saccharomyces cerevisiae for production of single-chain antibody fragments. *Nat Biotechnol* **16**, 773-777. (1998).

24. Huang, W., McKevitt, M. & Palzkill, T. Use of the arabinose p(bad) promoter for tightly regulated display of proteins on bacteriophage. *Gene* **251**, 187-197. (2000).

25. Krebber, A., Burmester, J. & Pluckthun, A. Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p. *Gene* **178**, 71-74. (1996).

26. Daugherty, P. S., Olsen, M. J., Iverson, B. L. & Georgiou, G. Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. *Protein Eng* **12**, 613-621. (1999).

27. Lou, J. *et al*. Antibodies in haystacks: how selection strategy influences the outcome of selection from molecular diversity libraries. *J Immunol Methods* **253**, 233-242. (2001).

28. Boder, E. T. & Wittrup, K. D. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol* **328**, 430-444 (2000).

29. Chen, G. *et al*. Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol* **19**, 537-542. (2001).

30. Sblattero, D. & Bradbury, A. Exploiting recombination in single bacteria to make large phage antibody libraries. *Nat Biotechnol* **18**, 75-80. (2000).

31. de Haard, H. J. *et al*. A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J Biol Chem* **274**, 18218-18230. (1999).

32. Griffiths, A. D. *et al*. Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J* **13**, 3245-3260. (1994).

33. Soderlind, E. *et al*. Recombining germline-derived CDR sequences for creating diverse single- framework antibody libraries. *Nat Biotechnol* **18**, 852-856. (2000).

34. Pini, A. *et al*. Design and use of a phage display library. Human antibodies with subnanomolar affinity against a marker of angiogenesis eluted from a two-dimensional gel. *J Biol Chem* **273**, 21769-21776. (1998).

35. Kieke, M. C. *et al*. High affinity T cell receptors from yeast display libraries block T cell activation by superantigens. *J Mol Biol* **307**, 1305-1315. (2001).

36. Llorente, M. *et al*. Natural human antibodies retrieved by phage display libraries from healthy donors: polyreactivity and recognition of human immunodeficiency virus type 1gp120 epitopes. *Scand J Immunol* **50**, 270-279. (1999).

37. Nieba, L., Honegger, A., Krebber, C. & Pluckthun, A. Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng* **10**, 435-444. (1997).

38. Bothmann, H. & Pluckthun, A. Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat Biotechnol* **16**, 376-380. (1998).

39. de Bruin, R., Spelt, K., Mol, J., Koes, R. & Quattrocchio, F. Selection of high-affinity phage antibodies from phage display libraries. *Nat Biotechnol* **17**, 397-399. (1999).

40. de Wildt, R. M., Mundy, C. R., Gorick, B. D. & Tomlinson, I. M. Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nat Biotechnol* **18**, 989-994. (2000).

41. Cohn, M. & Langman, R. E. The immune system: A look from a distance. *Front Biosci* **1**, d318-323. (1996).

42. Foote, J. & Eisen, H. N. Kinetic and affinity limits on antibodies produced during immune responses. *Proc Natl Acad Sci U S A* **92**, 1254-1256. (1995).

43. Newman, M. A., Mainhart, C. R., Mallett, C. P., Lavoie, T. B. & Smith-Gill, S. J. Patterns of antibody specificity during the BALB/c immune response to hen eggwhite lysozyme. *J Immunol* **149**, 3260-3272. (1992).

44. Zeng, X. *et al*. DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol* **2**, 537-541. (2001).

45. Braun, P. *et al*. Proteome-scale purification of human proteins from bacteria. *Proc Natl Acad Sci U S A* **99**, 2654-2659. (2002).

46. Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147. (2002).

47. Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. Studies on the transformation of intact yeast cells by the LiAc/SS- DNA/PEG procedure. *Yeast* **11**, 355-360. (1995).

48. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* **15**, 553-557. (1997).

49. Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M. & Wittrup, K. D. Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J Mol Biol* **292**, 949-956. (1999).

50. Krebs, B. *et al*. High-throughput generation and engineering of recombinant human antibodies. *J Immunol Methods* **254**, 67-84. (2001).

Figure 1. A) scFv expressed on the surface of *S. cerevisiae* as a tightly regulated fusion protein to Aga2p.[48] B) Bivariate flow cytometric analysis of a greatly enriched antigen binding population shows the discrimination of clones with slight differences in affinities and non-binder populations. Cells are double labeled with biotinylated antigen/streptavidin-phycoerythrin, and anti-c-myc/anti-mouse FITC labels. Subpopulation 1 expresses high affinity scFvs, and labels more brightly with antigen at a given scFv expression level. Subpopulation 2 expresses somewhat lower affinity scFvs, and subpopulation 3 does not detectably bind antigen. Subpopulation 4 is not expressing scFv on the cell surface.

Figure 2. Germline gene family usage for heavy (Panel A) and light chains (Panel B) in the scFv library. Unscreened yeast scFv library clones (open bars, 88 clones), reported human gene usage (dark bars[18-20]), and yeast scFv library clones obtained by antigen binding screens (gray bars, 40 clones.)

Figure 3. A) Yeast surface expression of 10 randomly selected individual scFv clones from the library, assessed by HA and c-myc epitope tag labeling and flow cytometry. B) Expression levels for 90 individual scFvs on the yeast cell surface. The C-terminal c-myc epitope tag was detected by indirect immunofluorescence and flow cytometry. Fluorescence intensity was normalized to 1 for a nondisplaying population, and a separate positive peak is distinguishable for any value over two. C) Histogram of c-myc surface expression level for scFv library as a whole, assessed by indirect immunofluorescent labeling and flow cytometry. The peak for uninduced cells indicates background levels for the measurement. D) Stability of the expressed scFv repertoire following multiple library passages. Each passage consists of a 100-fold amplification of the library and a –80°C freeze/thaw step. Cells expressing an anti-fluorescein scFv were mixed in at a concentration so as to produce 5% fluorescein-binding cells in the mixture at the beginning of the experiment.

Figure 4. A) Measurement of binding constant K_d for scFvs specific for HEL (4 distinct clones), FITC, and EGF on the yeast cell surface. The mean fluorescence intensity, determined by flow cytometry, is plotted against varying concentrations of antigen for several antigen specific scFv clones B) Comparison of protein antigen-binding affinities of antibodies cloned from large antibody repertoires in phage display (dark bars) or yeast display (this work). The phage data represent 36 antibodies reported from phage libraries of the following sizes: 6.5×10^{10} [32]; 6.7×10^{9} [8]; 1.4×10^{10} [9]; 3×10^{11} [30]; 3.7×10^{10} [31]; 2×10^{9} [10,50];3×10^{8} [34]; 2×10^{9} [33]. The yeast display data represent 11 antibodies selected against the proteins hen egg lysozyme, XPA, HBEGF, Amphiregulin, and EGF.

Table I. Rapid enrichment of rare clones following library propagation

	Passage 1	Passage 5
3 μm magnetic beads (Dynal)	58,000 cells retained/ 1.5 x 10^8 cells screened	200,000 cells retained/ 1.5 x 10^8 cells screened
Cell sorter	313 cells sorted/ 2,500,000 cells analyzed	500 cells sorted/ 5,000,000 cells analyzed
Individual clones	6/10 FITC positive	4/10 FITC positive
Overall enrichment	**6 x 10^5**	**4 x 10^5**
50 nm magnetic beads (Miltenyi)	68,000 cells retained/ 1.5 x 10^8 cells screened	26,000 cells retained/ 1.5x10^8 cells screened
Cell sorter	528 cells sorted/ 1,500,000 cells analyzed	101 cells sorted/ 1,500,000 cells analyzed
Individual clones	6/10 FITC positive	9/10 FITC positive
Overall enrichment	**6 x 10^5**	**9 x 10^5**

Cells displaying a FITC-binding scFv were mixed with the library at a frequency of 1 cell in 10^6, and the library was subjected to five 100-fold passages as described for Figure 3D. The FITC-binding cells were isolated sequentially by one magnetic bead screen and one flow cytometric sorting step.

Table II scFvs isolated against protein, peptide, and hapten antigens

Antigen	# clones	K_d
Hen egg lysozyme	4	6-9 nM
EGF	3	3nM, >1μM
Betacellulin	2	140 nM
Amphiregulin	12	$\geq$120 nM
HB-EGF	3	7.5 nM, 61 nM
XPA	2	82 nM, 460 nM
EGFR ECD 273-621	15	100-1,000 nM
P53 peptides	5	15-800 nM
Fluorescein	2	5 nM, 150 nM

Table III. Multiplex screening against p53 peptides

```
9S:        PQSDPSVEPPLS
15S:             VEPPLSQETFSD
18T:                PLSQETFSDLWK
20S:                  SQETFSDLWKLL

376S:       SKKGQSTSRHKK
378S:         KGQSTSRHKKLM
```

peptide / scFv	9S	15S	18T	20S	376S	378S
18-36	nt/nt	NB/NB	**121/412**	NB/>**1,000**	nt/nt	NB/nt
18a	nt/nt	NB/NB	**310/361**	NB/NB	nt/nt	NB/nt
20a	nt/nt	NB/NB	NB/NB	**641/148**	nt/nt	nt/nt
9a	**16**/NB	NB/nt	NB/nt	NB/nt	nt/nt	nt/nt
378a	nt/nt	nt/nt	NB/-	nt/nt	NB/NB	**147/1,006**

Numbers given are K_d in nM units, phosphorylated/nonphosphorylated peptide
NB = No Binding detected at 1 μM
nt = Not tested

Figure 1

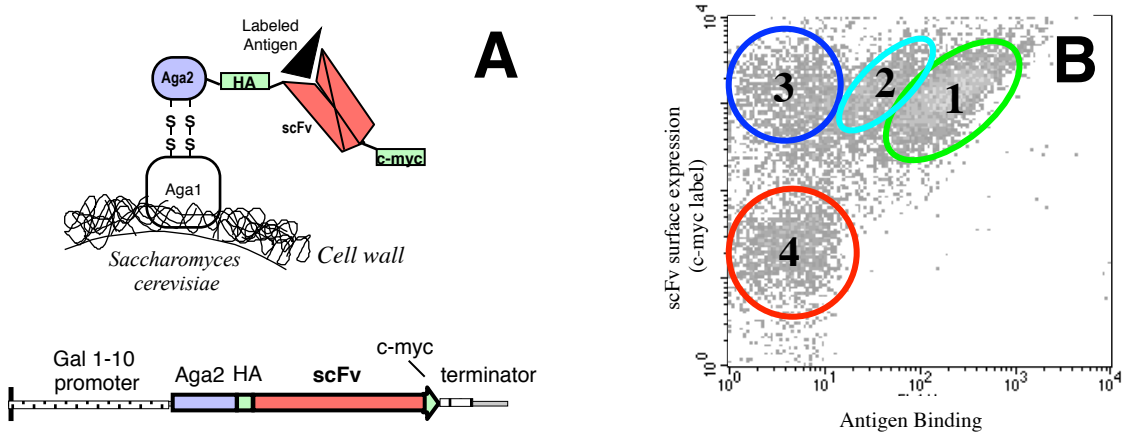

A
Labeled Antigen
Aga2
HA
scFv
c-myc
Aga1
Saccharomyces cerevisiae
Cell wall
Gal 1-10 promoter
Aga2 HA
scFv
c-myc
terminator
B
3
2
1
4
scFv surface expression (c-myc label)
Antigen Binding

Figure 2

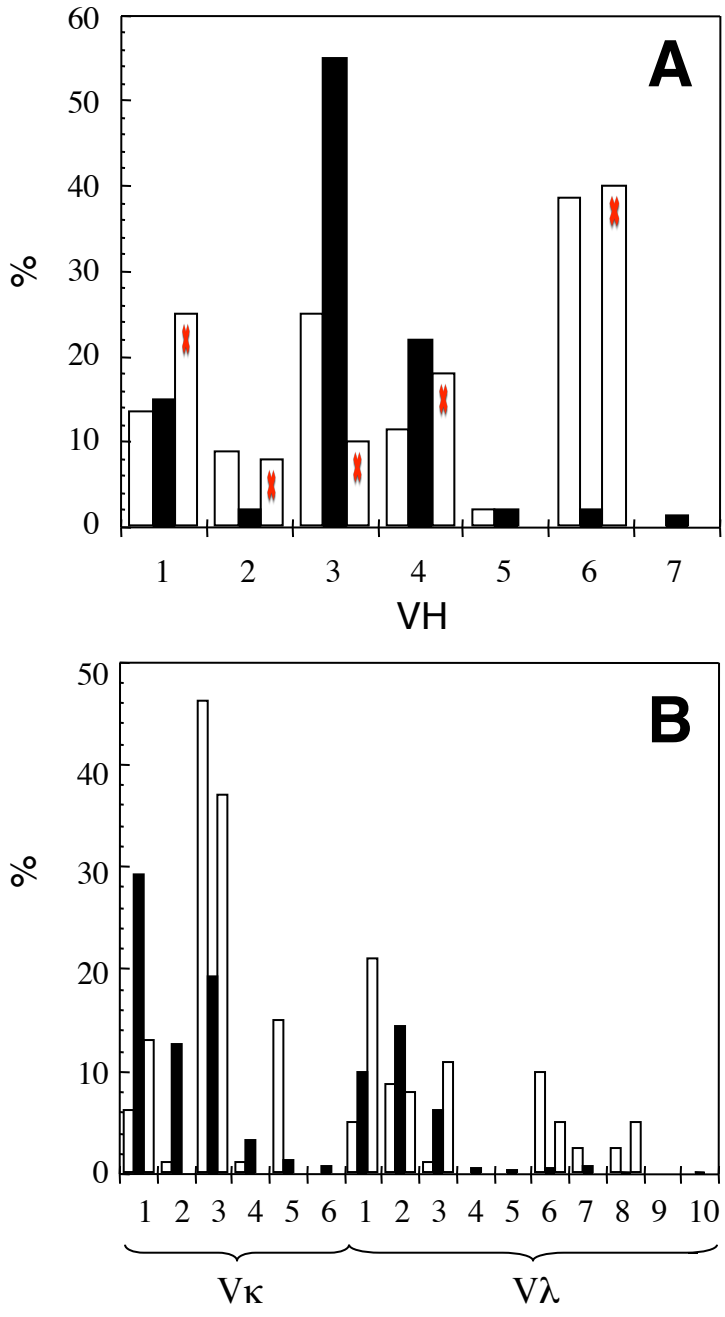

A
%
60
50
40
30
20
10
0
1
2
3
4
5
6
7
VH
B
%
50
40
30
20
10
0
1 2 3 4 5 6 1 2 3 4 5 6 7 8 9 10
Vκ
Vλ

Figure 3

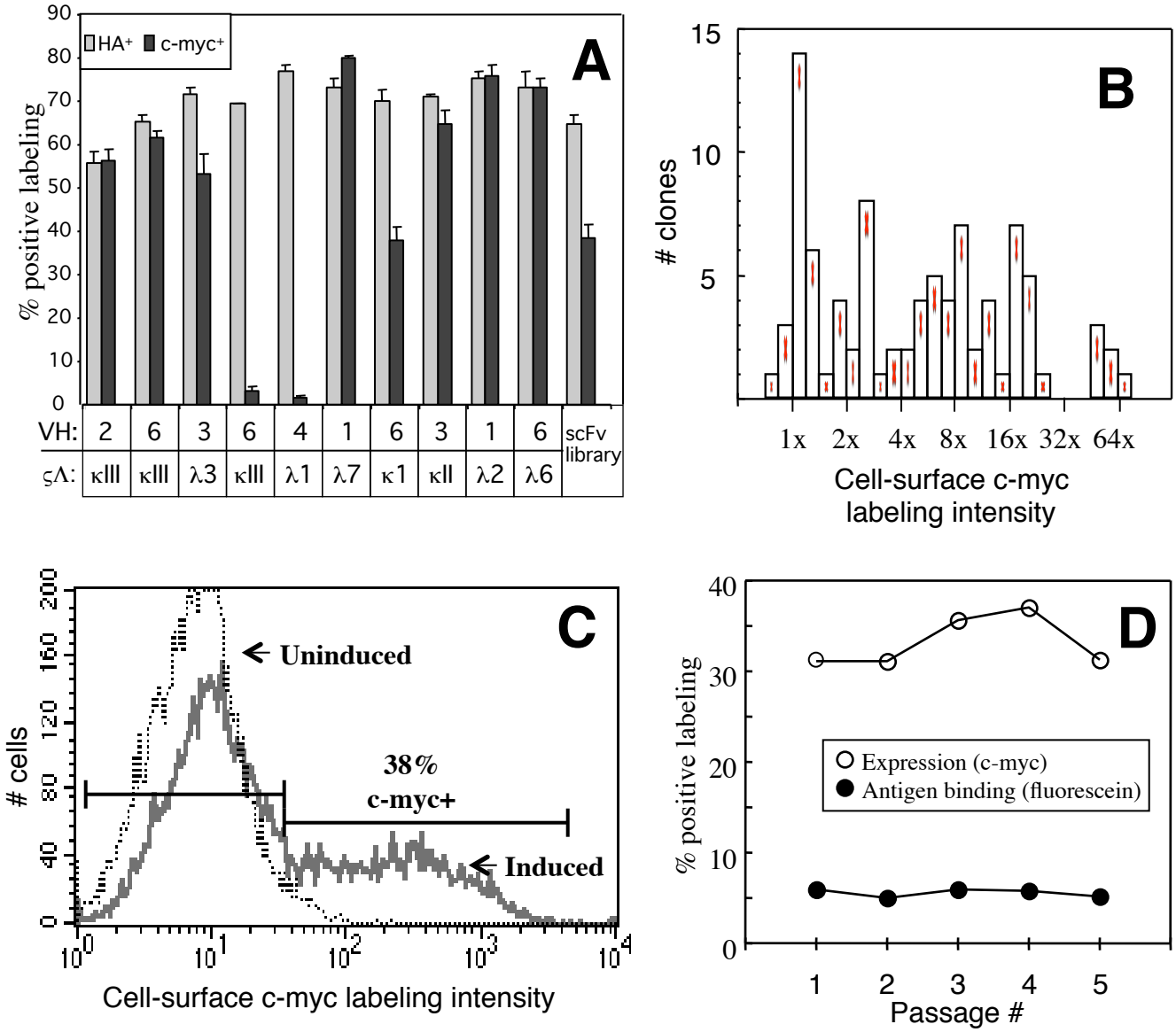
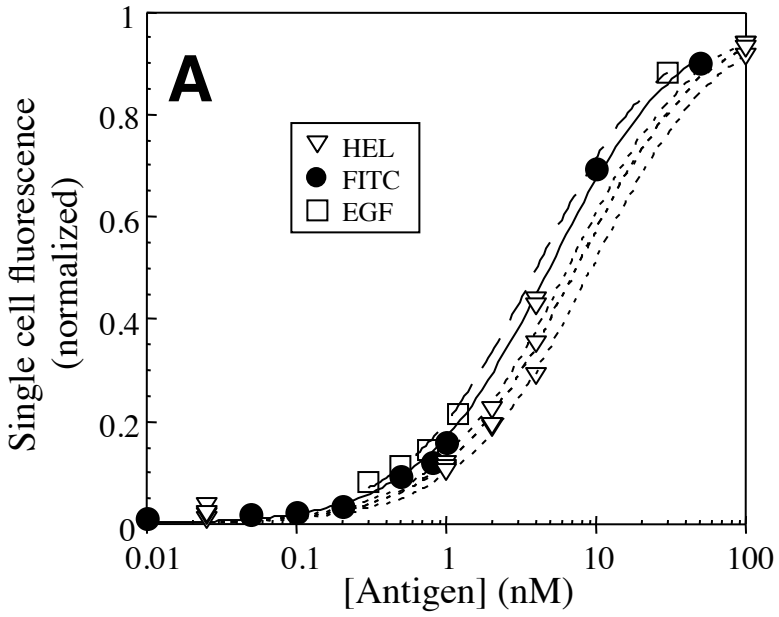
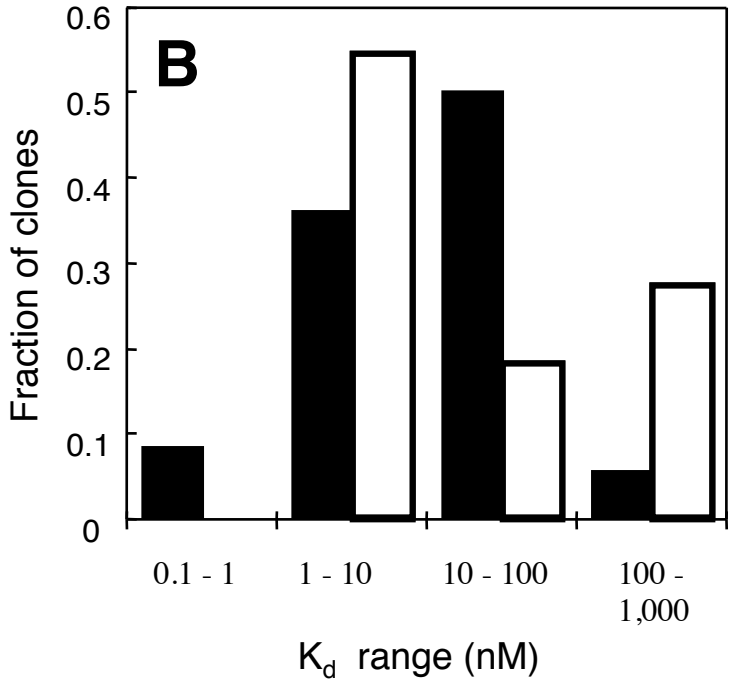

A
HA+ c-myc+
% positive labeling
VH: 2 6 3 6 4 1 6 3 1 6 scFv library
ςΛ: κIII κIII λ3 κIII λ1 λ7 κ1 κII λ2 λ6
B
clones
1x 2x 4x 8x 16x 32x 64x
Cell-surface c-myc labeling intensity
C
cells
Uninduced
38% c-myc+
Induced
Cell-surface c-myc labeling intensity
D
% positive labeling
Expression (c-myc)
Antigen binding (fluorescein)
Passage #

Figure 4


A
HEL
FITC
EGF
Single cell fluorescence (normalized)
[Antigen] (nM)
0
0.2
0.4
0.6
0.8
1
0.01
0.1
1
10
100


B
Fraction of clones
0
0.1
0.2
0.3
0.4
0.5
0.6
0.1 - 1
1 - 10
10 - 100
100 - 1,000
K_d range (nM)