

Antibody affinity maturation using bacterial surface display

Patrick S. Daugherty, Gang Chen¹, Mark J. Olsen¹, Brent L. Iverson¹ and George Georgiou²

Departments of Chemical Engineering and ¹Chemistry and Biochemistry, University of Texas, Austin, TX 78712, USA

²To whom correspondence should be addressed

A quantitative system for screening combinatorial single-chain Fv (scFv) antibody libraries was developed utilizing surface display on *Escherichia coli* and fluorescence-activated cell sorting (FACS). This system was employed to isolate clones with high-affinity to a fluorescently-labeled hapten from libraries constructed by randomizing heavy and light-chain residues in the anti-digoxin 26-10 derived antibody, scFv(dig). The use of flow cytometry enabled the detection of rare library members directly in heterogeneous populations and the optimization of selection conditions prior to sorting. A heavy-chain mutant having wild-type affinity ($K_D = 0.91 \pm 0.22$ nM) and an expected representation frequency of less than 1×10^6 , was selected to homogeneity after three rounds utilizing increasingly stringent selection conditions. The isolated clone possessed two distinct point mutations relative to the wild-type DNA sequence, yet still coded for the wild-type amino acid sequence, suggesting that the wild-type residues may be optimal at the randomized positions. An affinity improved clone ($K_D = 0.30 \pm 0.05$ nM), having a dissociation constant approximately threefold lower than the wild-type antibody, was isolated from a smaller light-chain library in a single sorting step. Flow cytometry was shown to be a simple and rapid method for the determination of the relative hapten dissociation rate constants of selected clones without requiring subcloning. The relative rate constants estimated by FACS were confirmed by producing the scFv antibodies in soluble form and measuring hapten binding kinetics by surface plasmon resonance (SPR). These results demonstrate that *E. coli* surface display, coupled with quantitative selection and analysis using FACS, has the potential to become a powerful tool for rapid isolation and characterization of desirable mutants from large polypeptide libraries.

Keywords: affinity maturation/*E. coli*/FACS/polypeptide library/surface display

Introduction

Polypeptide library screening methodologies have proven extremely valuable for basic and applied biological research (Marks *et al.*, 1991; Barbas, 1993). Although various techniques for screening libraries of short peptides have been developed, thus far, the screening of protein libraries has relied almost exclusively on phage display. Phage display has been used to probe polypeptide–ligand interactions and extensively for the isolation of high-affinity proteins, including antibodies (Winter and Milstein, 1991; Marks *et al.*, 1991; Short *et al.*,

1995; Vaughan *et al.*, 1996). Phage libraries are typically screened using repeated cycles of phage capture and elution from an immobilized ligand followed by phage amplification in bacteria. Another display technology wherein polypeptides are physically linked *in vitro* to their coding RNA through a ribosome was recently described (Hanes and Pluckthun, 1997), though its utility for affinity-based protein library screening has not yet been demonstrated.

As an alternative to screening by solid support capture and elution, fluorescence-activated cell sorting (FACS) is a mature technology with significant unexplored potential for molecular evolution. Despite the extensive application of FACS instrumentation for rare cell detection and isolation in clinical medicine, reports describing the use of FACS for library screening have been limited. A library of green fluorescent protein (*gfp*) mutants generated in *Escherichia coli* was screened using FACS allowing the isolation of red-shifted excitation mutants of *gfp* with a 20- to 35-fold increase in fluorescence intensity when excited at 488 nm (Cormack *et al.*, 1996). FACS has also been used to screen libraries of short synthetic peptides immobilized on beads for high-affinity peptide ligands (Needels *et al.*, 1993). By incubating the library population with various concentrations of fluorescently-tagged ligand and examining the resultant fluorescence distribution, optimal screening conditions could be easily identified. Recently, a method for screening protein libraries on the basis of affinity has been developed which utilizes display on the surface of microorganisms and FACS for quantitative isolation of desired clones (Francisco *et al.*, 1993; Boder and Wittrup, 1997; Georgiou *et al.*, 1997). High-affinity library members are enriched using FACS, after incubating cells with a fluorescently-tagged antigen in solution, and the resulting population is amplified by growth and resorted as necessary.

Technologies for the display of heterologous polypeptides on the surface of bacteria have been developed only recently (Francisco *et al.*, 1993; Gunneriusson *et al.*, 1996; Rode *et al.*, 1996; Georgiou *et al.*, 1997; Maurer *et al.*, 1997; Piard *et al.*, 1997). In particular, *E. coli* is well suited for library screening applications because it possesses a high transformation efficiency and is relatively simple to manipulate in the laboratory. A number of heterologous proteins have been displayed on the external surface of *E. coli* by fusing the *N*-terminus of the desired polypeptide to a chimera termed Lpp-OmpA' that consists of *E. coli* lipoprotein (aa 1–9) followed by aa 46–159 of the major outer membrane protein OmpA (Georgiou *et al.*, 1997). As a model protein for the development of library screening methodologies, we used a variant of a high-affinity scFv specific to digoxin (Huston *et al.*, 1988), derived from the well-characterized monoclonal antibody 26-10 (Huston *et al.*, 1988; Jeffrey *et al.*, 1993; Schildbach *et al.*, 1993). The crystal structure of the 26-10 Fab, complexed with digoxin, has been solved to 2.5 Å resolution (Jeffrey *et al.*, 1993). All residues that form important contacts with digoxin have been subjected to saturation mutagenesis and binding data for the

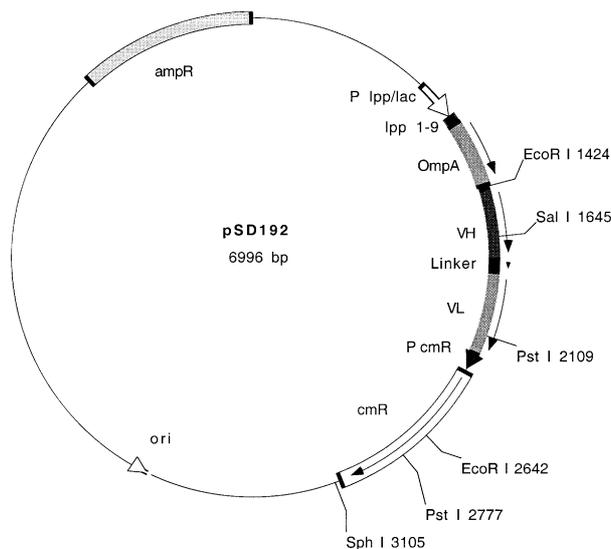


Fig 1. The surface display vector pSD192. The chloramphenicol acetyl transferase gene was introduced just downstream of scFv(dig), as described in the Materials and methods, enabling direct selection of library clones having scFv inserts.

resulting 190 mutant antibodies are available (Burks *et al.*, 1997) (unpublished data). Furthermore, phage display of the 26-10 Fab has been used previously to analyze the contribution of the heavy-chain CDR1 to digoxin affinity, resulting in the isolation of two mutants with 2–4-fold improved affinities (Short *et al.*, 1995). In this report, the use of FACS to screen *E. coli* displayed scFv libraries enabled quantitative detection and isolation of rare high-affinity clones.

Materials and methods

Strains and reagents

Escherichia coli strain JM109 {e14⁻(McrA⁻) *recA1 endA1 gyrA96 thi-1 hsdR17*(r_K⁺m_K⁺) *supE44 relA1Δ* (*lac-proAB*) [F' *traD36 proAB lacIqZΔM15*]} was used for all cloning steps and library experiments. For scFv expression, pET25b from Novagen (Madison, WI) was used with strain BL21(DE3) [F' *ompT hsd S_B gal dcm* (DE3)]. The scFv(dig) gene was a gift from Becton Dickinson Research Center (Research Triangle Park, NC), and the DNA sequence has been reported previously (Francisco *et al.*, 1993). Restriction enzymes were from Promega (Madison, WI), and *Pfu* polymerase from Stratagene (La Jolla, CA). Oligonucleotide primers were from Genosys (The Woodlands, TX). Fluorescein–digoxin was purchased from The Binding Site (Eugene, OR). BODIPY–digoxin was prepared via oxidation of the terminal sugar residue of digoxin (Sigma, St Louis, MO) with NaIO₄ followed by covalent attachment to BODIPY-FL-EDA, from Molecular Probes (Eugene, OR), through reductive amination in the presence of NaCNBH₃ (Smith *et al.*, 1970). The mixture was centrifuged for 15 min at 8000 g, and the supernatant was purified using a PepRPC 15 μm FPLC column from Pharmacia (Piscataway, NJ).

Construction of surface display vector pSD192

The surface display vector pSD192 (Figure 1), which carries the gene encoding Lpp-OmpA'-scFv(dig) transcribed from an *lpp-lac* promoter and a gene encoding a C-terminal fragment of chloramphenicol acetyltransferase (*cat*), was derived from pTX152 (Francisco *et al.*, 1993) as follows: the gene encoding

Lpp-OmpA'-scFv(dig) was removed from pTX152 by digestion with *PvuI* and *Bam*HI, and ligated into similarly digested pET-22b to yield pGC183. A gene encoding a C-terminal fragment of chloramphenicol acetyl transferase (*cat*) was amplified from pBR325 with primers #1 (5'-AGGGCAGCATGCCTAA-3') and #2 (5'-GACCCCGAGGACTACGTCTTCGAATAAATAC-3'), digested with *Eco*RI and *Sph*I and ligated into similarly digested pGC183 to yield pGC185. The full length scFv(dig)-*cat* gene was constructed by overlap PCR (Ho *et al.*, 1989). The scFv and *cat* fragments were separately amplified from pTX152 [primers #3 (5'-TGGACCAACAACATCGGTGACGCACAC-3') and #4 (5'-TATTCGAAGACGTTAGCTCTCGGGGTC-3')] and pBR325 (primers #1 and #2), annealed and reamplified with primers #3 and #2. The final PCR product was digested with *Eco*RI and ligated into similarly digested pGC185 to yield pGC182. The existing *Pst*I site was removed from pGC182 by replacing the *Alw*NI-*Pvu*I fragment with that from pUC18 giving rise to pSD182. pSD192 was obtained by introducing *Pst*I sites upstream of the light-chain CDR3 and within the *cat* gene by overlap PCR. Mutagenic fragments were amplified with primers #5 (5'-TTGGCTGCAGTAATATATTGCAGCAT-3') and #6 (5'-TGGACCAACAACATCGGTGACGCACAC-3'), #7 (5'-TGCAATATATTACTGCAGCCAAACTACGCAT-3') and #6 (5'-CTTGC GAATATTCTGCAGAAACTGCCG-3'), and #8 (5'-CGGCAGTTTCTGCAGATATATTTCGCAAGAT-3') and #9 (5'-CGAGGGCATGCAAGGGCACCAATAACTGCTTA-3'). Fragments were gel purified, annealed, and amplified with primers #6 and #8. The final PCR product was digested with *Eco*RI and *Sph*I and ligated into similarly digested pSD182 to yield pSD192. Removal of the small *Pst*I fragment from pSD192 and religation yielded pSD195.

Affinity discrimination of scFv point mutants

Heavy-chain residue Y33 was mutated to N, S and G by overlap extension PCR (Ho *et al.*, 1989). PCR products were digested with *Eco*RI, and ligated into similarly digested pSD195, yielding single point mutants Y33N, Y33S and Y33G. After transformation into JM109, overnight cultures were grown at 37°C and subcultured 1:100 at 25°C for 20 h. Cells (200 μl) were pelleted by centrifugation at 3000 g for 4 min, washed with 1 ml PBS, pH 7.1, and resuspended in 1.5 ml PBS. Cells were aliquoted into 1.5 ml tubes and BODIPY–digoxin or fluorescein–digoxin was added to a final concentration of 10⁻⁷, 10⁻⁸ or 10⁻⁹ M. Cells were incubated at 24°C for 1 h with gentle shaking, pelleted and resuspended in 1 ml PBS for analysis by flow cytometry. At least 10 000 events were acquired on a Becton Dickinson (San Jose, CA) FACSsort. Parameters were set in LOG mode as follows: forward scatter (FSC) threshold: 50, FSC pre-amp: E01, FL1: 800, FL2: 800, side scatter (SSC): 400. Debris and instrument noise were excluded from analysis by defining a FSC-SSC region around the cell population. For equilibrium affinity analysis, 20 μl of culture were added directly from a 24 h, 25°C culture to 180 μl BODIPY–digoxin at 10, 5, 2.5, 1.25, 0.62 and 0 nM control in a 96-well plate, and allowed to incubate 1 h with gentle shaking. Cell samples were diluted 5-fold in PBS and analyzed immediately by flow cytometry.

ScFv library construction, analysis and screening

ScFv(dig) libraries were constructed by randomizing heavy-chain CDR3 residues K99, W100, A100a and M100b, and separately light-chain CDR3 residues T91, V94 and P96, by overlap extension PCR using NNS (S = G,C) or NNB

(B = G,C,T) randomization schemes, respectively. PCR products were digested with *SalI* and *PstI* (heavy-chain) or *PstI* (light-chain), purified by electroelution, and ligated into similarly treated pSD195. Ligation products were ethanol precipitated, resuspended in 15 μ l of 10 mM Tris-HCl (pH 8.5), and transformed into 300 μ l JM109 cells using electroporation. Electroporation cuvettes were washed with 3 ml SOC and cells were incubated for 1 h at 37°C in a total volume of 30 ml. Plating revealed the heavy and light-chain libraries to contain 4×10^6 and 3×10^5 individual transformants, respectively. Prewarmed LB media (500 ml) supplemented with 0.4 % glucose, and 75 μ g/ml ampicillin was added and plasmid library DNA was prepared after growth for 8 h at 37°C.

For heavy-chain library screening, 3 μ l of library DNA was transformed into 80 μ l JM109. Cells were grown at 37°C for 8 h and subcultured 1:100 for 14 h at 25°C. Cells were diluted 1:40 into PBS with BODIPY-digoxin at a final concentration of 15 nM, and incubated with gentle shaking at 24°C for 45 min. Propidium iodide was added to a final concentration of 5 μ g/ml and cells were incubated an additional 15 min. Prior to sorting, cells were filtered through a 5 μ m syringe filter (Gelman Sciences), to remove any cell aggregates and diluted in a volume of PBS to yield an event rate of approximately 2000 s^{-1} . The FACSsort was sterilized with 70% ethanol and a total of $>10^7$ cells were sorted in recovery mode into 150 ml LB media supplemented with antibiotics as above. The cells were regrown as above, and labeled with 10 nM BODIPY-digoxin, washed 2 times, resuspended in 1 μ M digoxin and allowed to incubate for 30 min at room temperature prior to sorting in exclusion mode. A third sort was performed as in round two, except that the cells were labeled with 0.2 nM BODIPY-digoxin. After each round of sorting and regrowth, cells were plated to isolate clones. Colonies from the first round of sorting were picked at random and grown in 200 μ l LB in a 96-well plate overnight with shaking at 37°C. Cells were subcultured 1:100 and grown at 24°C for 20 h. In a fresh 96-well plate, 10 μ l each of cell culture was transferred to 90 μ l of 10 nM BODIPY-digoxin. Cells were incubated for 1 h with gentle shaking at 25°C, and analyzed by flow cytometry.

The light-chain library was screened essentially as described above. For two-step selection, cells were labeled first with 10 nM BODIPY-digoxin, washed once in PBS and sorted in recovery mode. A second round of sorting was performed after labeling with 5 nM BODIPY-digoxin, using exclusion mode. Aliquots were plated and positive clones were confirmed using flow cytometry. For single-step selection, cells were labeled with 5 nM BODIPY-digoxin, washed two times in PBS and incubated with 1 μ M digoxin in PBS, for 30 min at room temperature. The cells were sorted in exclusion mode directly into SOC media supplemented with antibiotics.

Flow cytometric kinetic analysis

Overnight cultures were grown at 25°C and aliquots were frozen in 10% glycerol. Aliquots were thawed on ice and diluted 1:20 in PBS, pH 7.0. BODIPY-digoxin was added to a final concentration of 10 nM. Samples were incubated for 1 h at 25°C with gentle shaking and the cells were pelleted by centrifugation at 6000 r.p.m. for 4 min. The cell pellet was resuspended in 1 μ M digoxin in PBS and analyzed immediately by FACS. The mean fluorescence intensity from 5000 events was recorded at 1 or 2 min intervals and plotted as a first-order exponential decay.

Expression and purification of soluble scFvs

ScFv genes were subcloned into pET-25b, and transformed into BL21(DE3) for expression. The respective proteins were refolded from inclusion bodies as described (Burks and Iverson, 1995), and purified using Ni-NTA spin columns (Qiagen, Santa Clarita, CA). Glycerol was added to the eluate at a final concentration of 10% and the material was dialyzed (4×500 ml) at 4°C over a 20 h period into 10% (w/v) glycerol, 0.05 M Tris-HCl, 0.05 M KCl (pH 7.4). The dialyzed material was verified to contain greater than 95% monomeric scFv by size-exclusion chromatography, and was stored at 4°C at less than 50 μ g/ml. Protein concentrations were determined from the A_{280} .

Kinetic analysis of purified scFvs

ScFv binding kinetics were determined using surface plasmon resonance with the BIAcore instrument (Pharmacia). Digoxin-conjugated BSA (Burks *et al.*, 1997) in 5 mM acetate buffer (pH 5.5) was immobilized on a carboxy-dextran CM5 sensor chip, using EDC/NHS amine coupling chemistry, at a level equivalent to 150–300 RU (Karlsson and Falt, 1997). The reaction was quenched with ethanolamine to eliminate unreacted, activated ester groups. Association and dissociation rate constants were measured at 25°C with a continuous flow of 60 μ l/min (Karlsson and Falt, 1997) using HBS (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.005% P20). Rate constants were calculated using the BIAevaluation software. Association rate constants (k_{assoc}) were determined from a plot of $\ln(dR/dt)/t$ versus concentration over a range of 10–150 nM scFv. The k_{diss} values were determined within the first 2 min of dissociation from a plot of $\ln(R_0/R)$ versus time from the highest concentration of scFv analyzed. Unlabeled digoxin (1 μ M) was added to the running buffer to minimize rebinding effects (Yang *et al.*, 1995; Schier *et al.*, 1996).

Results

Discrimination of surface displayed antibody affinities by flow cytometry

Escherichia coli clones displaying mutants of scFv(dig) that differ in hapten binding affinity were analyzed using flow cytometry to determine if their respective fluorescence signals could be resolved. Since heavy-chain residue Y33 of scFv(dig) is critical for digoxin binding (Short *et al.*, 1995; Burks *et al.*, 1997), single-point mutants Y33N, Y33S and Y33G were constructed by site-directed mutagenesis and the respective scFv(dig) mutants were displayed on the cell surface as Lpp-OmpA' fusions using the surface display vector pSD192 (Figure 1). To obtain optimal protein display on the *E. coli* surface, the cells were first grown at 37°C, and subsequently for 12 h at 25°C, prior to labeling. Harvesting the cells in exponential phase resulted in decreased signals whereas prolonged growth at 25°C was found to adversely affect culture viability and the recovery of sorted clones (data not shown). In initial experiments, cells were labeled with a fluorescein-digoxin conjugate. Though the fluorescence distribution of cells expressing each of the scFv(dig) mutants could be resolved, the low quantum yield of fluorescein in the physiological pH range resulted in suboptimal population resolution. To enhance the fluorescence signal, and thus improve the resolution of signals, we synthesized a pH-insensitive BODIPY-digoxin conjugate. Labeling of the cells with 100 nM BODIPY-digoxin, a value well above the digoxin equilibrium

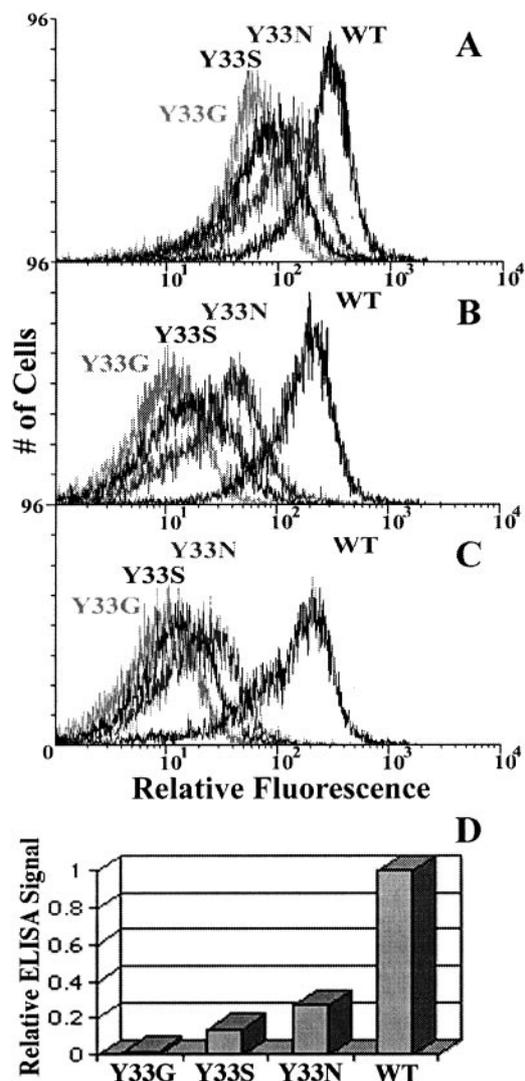


Fig. 2. Flow cytometric analysis of BODIPY–digoxin binding to *E.coli* cells expressing scFv(dig) and mutants Y33N, Y33S and Y33G. (a) 100 nM; (b) 10 nM; (c) 1 nM. (d) Relative ELISA signals for scFv(dig), and Y33N, Y33S and Y33G mutants *in vitro* (Burks *et al.*, 1997).

constant of the wild-type scFv(dig) antibody, resulted in significant overlap of the fluorescence distributions of high and low affinity clones (Figure 2A). However, at 1 nM BODIPY–digoxin, a concentration equal to the equilibrium dissociation constant of the wild-type scFv(dig) for digoxin, the mean fluorescence intensity of cells expressing antibodies with wild-type affinity (MFLI = 200) could be readily distinguished from that of the low affinity mutants (Figure 2C). The fluorescence distributions of cultures displaying mutant scFv(dig) antibodies correlated with the relative affinities of the corresponding *in vitro* generated antibodies as measured by ELISA (Figure 2D). The fluorescence signal of the Y33G clone (MFLI \approx 15) coincides with background cellular autofluorescence. For cells displaying wild-type scFv(dig), the BODIPY–digoxin concentration required to achieve 50% of maximum fluorescence intensity was approximately 2 nM, in agreement with the equilibrium dissociation constant ($K_D = 0.91 \pm 0.22$ nM) for the purified scFv determined by SPR.

Library construction, analysis and screening by FACS

Preliminary library screening experiments revealed that selection efficiency is reduced by cloning artifacts resulting from

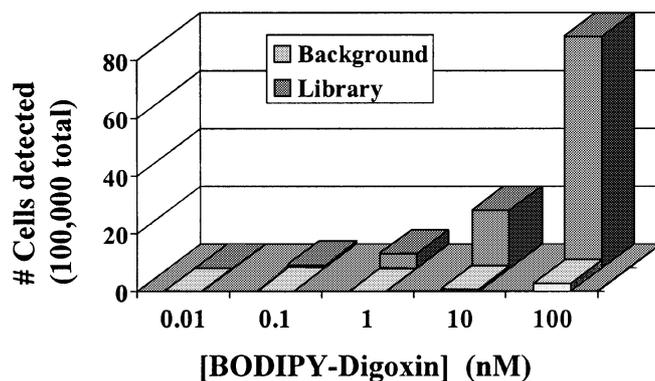


Fig. 3. Frequency of rare high-fluorescence cells detected in a heavy-chain CDR3 library per 100 000 cells using flow cytometry. Cells were considered to be target cells if they possessed FSC and FL1 signals falling into the high-fluorescence region (R1) in Figure 4. The background control cell population consisted of a pool of 20 random clones characterized by DNA sequencing and flow cytometry, having K_D values above 100 nM.

the presence of a small number of clones that do not contain scFv(dig) gene inserts and thus exhibit improved growth. A simple cloning strategy was designed to ensure that all transformants growing under selective conditions contain full-size inserts. Briefly, a 3' fragment of the chloramphenicol acetyltransferase gene (*cat*) was cloned downstream of the *Lpp-OmpA'* gene in the expression vector pSD195. Antibody genes were amplified by PCR together with the 5' *cat* fragment, such that ligation into pSD195 restored *cat* function (Figure 1). Only clones that contained the full-length PCR insert, in the correct orientation, were able to grow in the presence of chloramphenicol.

A scFv(dig) library was created by PCR randomization of heavy-chain CDR3 residues K99, W100, A100a and M100b using an NNS (S = G or C) randomization scheme. To eliminate the possibility that the wild-type DNA sequence (AAATGGGCCATG) will be selected upon screening, the mutagenized DNA was treated with *NcoI*, which is unique to the wild-type DNA sequence, such that the wild-type is eliminated without affecting the total amino acid diversity encoded by the library. In the resulting library, the wild-type amino acid sequence was encoded by a single, non-wild-type DNA sequence (AAGTGGGCGATG). Sequence analysis of the library DNA in aggregate and of 20 randomly selected transformants did not reveal any codon bias, and all clones contained full-length scFv inserts. Given the number of transformants, and the equal distribution of bases at randomized positions, the probability that all possible sequences are represented is greater than 95%, assuming that a Poisson distribution applies (Lowman *et al.*, 1991). The 20 randomly selected transformants were cultured in liquid media and labeled with 100 nM BODIPY–digoxin. The fluorescence of 17 of 20 clones was indistinguishable from background cellular autofluorescence (MFLI \approx 15), while three clones displayed weak fluorescence (MFLI < 60).

The presence of rare clones expressing high-affinity scFv(dig) antibodies in the heavy-chain library was evaluated by labeling the cells with 100, 10, 1, 0.1 or 0.01 nM BODIPY–digoxin (Figure 3). High-fluorescence events, defined as those occurring in region 1 (R1; Figure 4), were detected at a frequency of about 0.01% using 1 nM BODIPY–digoxin and about 0.001% with 0.1 nM BODIPY–digoxin. For comparison, the 20 randomly selected transformants that had been

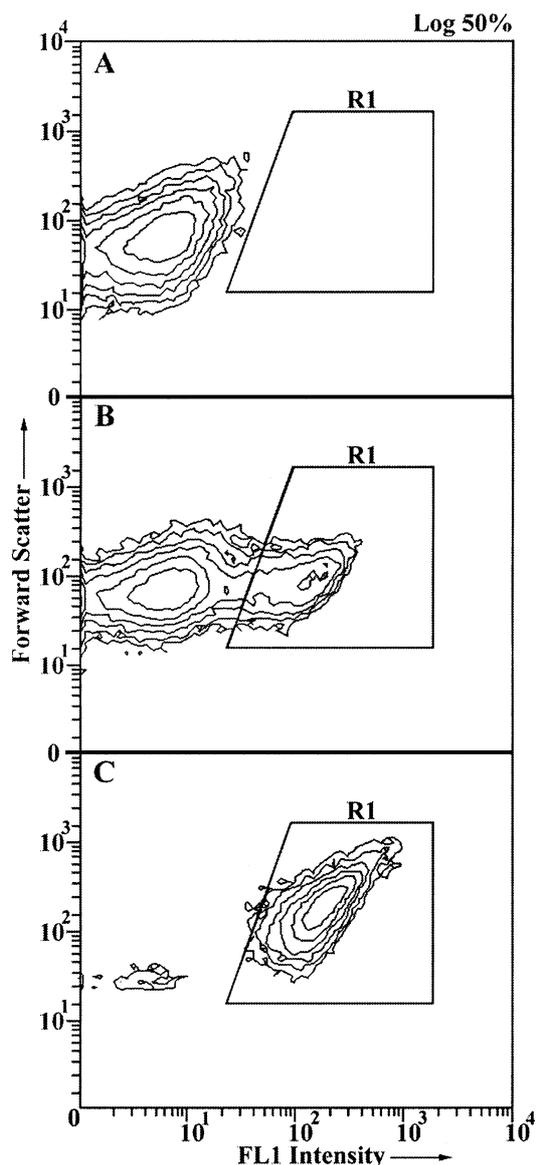


Fig. 4. Enrichment of highly fluorescent bacteria using FACS. The heavy-chain CDR3 library cell population was labeled with BODIPY-digoxin and analyzed by flow cytometry. (a) Presort (15 nM BODIPY-digoxin); (b) post-round 1 (10 nM BODIPY-digoxin); (c) post-round 2 (10 nM BODIPY-digoxin).

sequenced to assess the diversity of the library were pooled and analyzed in the same manner. In this case, high fluorescence clones were not detected for antigen concentrations below 10 nM.

Prior to sorting, the library cell population was also labeled with propidium iodide to preferentially label non-viable cells (Lopez-Amoros *et al.*, 1995). The use of propidium iodide, as both a viability and non-specific control stain (Leary, 1994), was found to enhance the enrichment of the target cells. The library was first screened after labeling with 10 nM BODIPY-digoxin and washing once in PBS (Figure 4). A sort window was defined such that cells were recovered only if they exhibited high BODIPY fluorescence (FL1 channel), low propidium iodide fluorescence (FL2 channel) and an appropriate range of forward scatter values to account for cell size. More than 10^7 cells were sorted in recovery mode in less than 2 h. Recovery mode enables the collection of all clones in the

target window regardless of coincident non-target cells. Highly-fluorescent clones were enriched from a frequency of 0.01% in the pre-sort population to about 20% following sorting (Figure 4A and B, respectively), an enrichment factor of approximately 2000-fold. The sorted population was regrown in LB media and plated on selective agar plates. 96 colonies were picked at random, grown in a 96-well microtiter plate and subsequently analyzed by flow cytometry. Approximately 20% of clones (18/96) were highly fluorescent at 10 nM BODIPY-digoxin exactly as expected from the fluorescence distribution of the sorted population. DNA sequencing revealed a clear consensus at H(99)- (100b) for (K/R)RAL (Table I). All HC1.1 clones analyzed (14/14) expressed scFv(dig) antibodies with hapten affinities less than 3 nM.

The entire collection of cells selected in the first round were regrown in liquid culture and a second round of enrichment was performed using conditions favoring clones with slower dissociation kinetics. Specifically, the cells were incubated with 10 nM BODIPY-digoxin, washed two times with PBS, and resuspended in 1 μ M unlabelled digoxin in PBS. The number of washing steps and the time of incubation in the presence of unlabelled digoxin (30 min) were optimized by monitoring the fluorescence distribution of the cells after each step. Cells were sorted in exclusion mode, which rejects coincident events, to achieve high purity (Figure 4C). The sorted cells were regrown and subjected to a final round of selection with 0.2 nM BODIPY-digoxin, a concentration equal to 20% of the equilibrium dissociation constant of the purified, wild-type scFv(dig) in solution. During the third round, a significant percentage of the population remained fluorescent and was sorted in exclusion mode. After each round, selected clones were analyzed by flow cytometry (Figure 5), and their amino acid sequences were determined by DNA sequencing (Table I). Clones isolated in the first round displayed a consensus sequence of (K/R)RAL, which evolved as the selection stringency was increased in rounds two and three. W100 emerged in round two, changing the consensus to (K/R) (R/W)AL. As might be expected, clones with high codon-based representation frequencies (HC1.2-1, HC1.1-2 and 1.2-1) displayed variation at the DNA level. In round three, a single mutant with a protein sequence identical to the wild-type (KWAM) was isolated as five out of five clones examined. However, all round three clones analyzed were encoded by the non-wild-type DNA sequence and the original wild-type DNA sequence was not found.

The selected clones were conveniently ranked in terms of relative affinity by measuring the hapten dissociation rate from whole cells using flow cytometry (Figure 5). Whole cells were labeled with an excess of BODIPY-digoxin, pelleted by centrifugation and resuspended into buffer containing large excess of unlabelled digoxin at which time the MFLI was measured at 1 or 2 min intervals. As shown in Figure 5, a plot of $\log(\text{MFLI}/\text{MFLI}_0)$ as a function of time is linear, with the slope giving the apparent dissociation rate constant. Replicate measurements for a given clone were reproducible, varying by less than 15%. Clones selected in the first round exhibited fast apparent dissociation rates. The average dissociation rate decreased substantially between rounds one and two, and further decreased by about two-fold after the third round.

To confirm the flow cytometric dissociation rate measurements, His-tagged scFv(dig) mutants were expressed without the Lpp-OmpA' sequence and purified to >95% by refolding from inclusion bodies (Burks and Iverson, 1995). Analytical

Table I. Summary of scFv(dig) HCDR3 library screening results

Clone ^a	Isolated	Frequency	HCDR3	k_{diss} (s ⁻¹)	k_{diss} (s ⁻¹)	k_{assoc} (M ⁻¹ s ⁻¹)	K_{D} (M)
				(FCM)	(SPR)	(SPR)	(SPR)
				($\times 10^3$) ($\pm 15\%$)	($\times 10^3$) ($\pm 15\%$)	($\times 10^{-6}$) ($\pm 10\%$)	($\times 10^9$)
WT	–	–	SSGNKWAMDY	0.23	0.83	0.91	0.91
HC1.3-1 ^b	5	1	SSGNKWAMDY	0.23	0.83	0.91	0.91
HC1.2-1	5	54	SSGNRRALDY	0.68	1.2	0.86	1.4
HC1.2-2	7	6	SSGNKWALDY	0.32	0.99	0.85	1.2
HC1.1-1	1	18	SSGNKRALDY	1.5	3.6	1.4	2.6
HC1.1-2	3	54	SSGNRRSLDY	1.3	2.7	1.2	2.3
HC1.1-3	2	3	SSGNKRAMDY	1.3	2.7	1.1	2.5
HC1.1-4	6	54	SSGNRRALDY	0.68	1.2	0.86	1.4
HC1.1-5	1	24	SSGNGRALDY	1.7	2.6	1.1	2.4
HC1.1-6	1	18	SSGNRYALDY	1.7	–	–	–

^aClone designations are (H)heavy(C)hain.EXP#.ROUND# - CLONE#. *Isolated* refers the total number of clones isolated possessing the given sequence, and *frequency* is the codon-based representation in the library. ScFv heavy-chain CDR3 sequences are shown with randomized residues in boldface type. Kinetic constants were determined by flow cytometry (FCM) or surface plasmon resonance (SPR).

^bThe sequence of HC1.3-1 is unique from the wild-type at the DNA level.

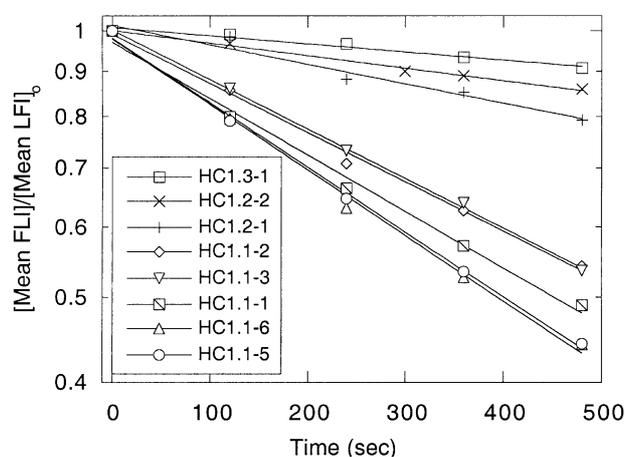


Fig. 5. The decay in mean fluorescence as BODIPY–digoxin dissociates from whole cells displaying mutant scFv selected from a heavy-chain CDR3 library. At least 4000 events were acquired at 1–2 min time intervals and the mean fluorescence is plotted as an exponential decay.

gel filtration confirmed that greater than 95% of the purified protein was present as monomeric scFv. The binding kinetics of soluble purified scFv were analyzed by SPR using a BIAcore instrument. A high flow rate and low coupling density were used to minimize mass transport and rebinding effects (Karlsson and Falt, 1997). Importantly, the K_{D} values for the purified antibodies were all at, or below, the antigen concentrations originally used to isolate the corresponding clones by FACS. Overall, association rate constants, k_{assoc} , for all isolated clones varied by 1.6-fold, whereas the dissociation rate constants, k_{diss} , improved by 3–4-fold between the first and last round of selection (from $2.6\text{--}3.6 \times 10^{-3} \text{ s}^{-1}$ to $0.83 \times 10^{-3} \text{ s}^{-1}$). The dissociation rate constants of the purified scFvs correlated with the values obtained by flow cytometric analysis of the corresponding antibodies displayed on *E. coli* (Table I).

A second scFv(dig) library was created by randomizing light-chain residues T91, V94, and P96. As with the heavy-chain library, DNA sequencing analysis did not reveal any codon bias. Given the number of transformants obtained, the Poisson probability that all sequences are represented in the library is greater than 85%. The light-chain library was screened using two increasingly stringent sorting steps resulting in the isolation of clones designated LC2.1 and LC2.2. The

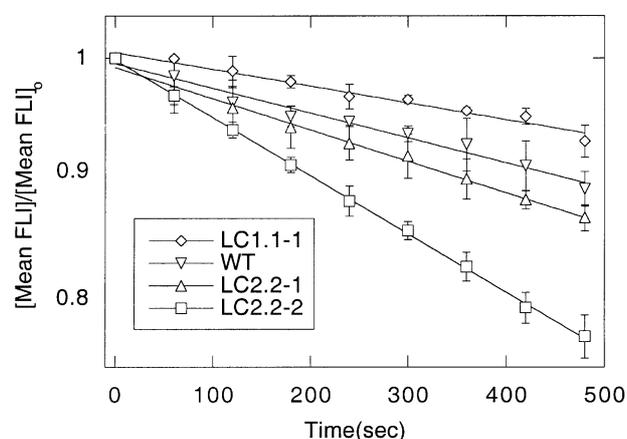


Fig. 6. The decay in mean fluorescence as BODIPY–digoxin dissociates from whole cells displaying mutant scFv selected from a light-chain CDR3 library. Selection conditions are given in Table II. At least 4000 events were acquired at 1–2 min time intervals and the mean fluorescence is plotted as an exponential decay.

dissociation rates were again measured by flow cytometry (Figure 6, Table II). The average apparent dissociation rates decreased by 3.5-fold between rounds 1 and 2. T91 was conserved among the highest affinity clones isolated. The substitution P96A occurred in 6/8 isolated LC2 clones. V94 displayed a large tolerance to substitution but with a preference for moderately sized hydrophobic residues: V, I, and L.

The light-chain library was also screened using conditions to favor the isolation of only the highest affinity clones using just a single round of sorting. Cells were labeled with 5 nM BODIPY–digoxin, washed twice in PBS, and then incubated with 1 μM digoxin in PBS for 30 min prior to sorting. Highly-fluorescent clones were enriched from 0.01% to greater than 10% of the population. Three clones having high MFLI signals at 5 nM BODIPY–digoxin were selected at random and their dissociation rate constants were measured by flow cytometry. All three clones had an apparent dissociation rate constant of $1.5 (\pm 0.2) \times 10^{-4} \text{ s}^{-1}$, a value reproducibly lower than that of the wild-type antibody [$k_{\text{diss}} = 2.3 (\pm 0.3) \times 10^{-4} \text{ s}^{-1}$] (Figure 6). DNA sequencing revealed that the scFvs expressed by all three clones possessed the same amino acid substitutions, V94I and P96A, while the wild-type residue T91 was conserved. The

Table II. Characterization of scFv(dig) mutants selected from a light-chain library by FACS

Clone ^a	Selection conditions ^b	LCDR3	k_{diss} (s ⁻¹) (Flow cytometry) ($\times 10^3$) ($\pm 15\%$)
WT		SQTTHVPPT	0.23
LC1.1-1	5 nM B-d/2 wash/ORS ^b	SQTTHIPAT	0.15
LC2.2-1	5 nM B-d/1 wash	SQTTHTPPT	0.29
LC2.2-2	5 nM B-d/1 wash	SQTTHFPAT	0.54
LC2.1-1	10 nM B-d/1 wash	SQATHLPAT	1.7
LC2.1-2	10 nM B-d/1 wash	SQSTHRPPT	2.3
LC2.1-3	10 nM B-d/1 wash	SQPTHLPAT	–
LC2.1-4	10 nM B-d/1 wash	SQSTHFPAT	1.6
LC2.1-5	10 nM B-d/1 wash	SQSTHVPAT	0.85
LC2.1-6	10 nM B-d/1 wash	SQSTHIPAT	0.72

^aClone designations are: (L)ight(C)hain.EXP#.ROUND# - CLONE#.

Selection conditions indicate the concentration of fluorescent probe utilized, the number of wash steps performed, and ^bwhether off-rate selection (ORS) was performed by incubating washed cells in buffer containing an excess of unlabeled digoxin for 30 min before sorting. Randomized residues are shown in bold-face. Apparent dissociation rate constants were measured by flow cytometry and % error values were determined from triplicate measurements.

^bB-d, BODIPY-digoxin.

affinity-improved mutant (LC1.1-1), was produced in soluble form as above and its hapten binding kinetics were measured by SPR in triplicate. The k_{assoc} and k_{diss} rate constants were $1.0 (\pm 0.1) 10^6 \text{ M}^{-1} \text{ s}^{-1}$ and $0.30 (\pm 0.04) 10^{-3} \text{ s}^{-1}$, respectively. The value of k_{diss} for LC1.1-1 was threefold lower than that of the wild-type antibody (Table II) resulting in an equilibrium dissociation constant of $K_D = 0.30 \pm 0.05 \text{ nM}$.

Discussion

Escherichia coli display coupled with flow cytometry was shown to be an effective approach for rapid isolation and analysis of high-affinity scFv antibodies. A heavy-chain library was constructed in which the four scFv(dig) binding site residues H(99)–H(100b) were fully randomized. ScFv selected in rounds 1, 2 and 3 possessed a mean k_{diss} of $2.110^{-3} \text{ s}^{-1}$, $1.110^{-3} \text{ s}^{-1}$ and $0.8310^{-3} \text{ s}^{-1}$, respectively, indicating that subtle differences in k_{diss} and hapten affinity could be selected for in a single round using FACS. Each round of FACS identified changes in the consensus sequence which yielded corresponding improvements in affinity. The consensus identified from round one, where 15 nM antigen was used for selection, was (K/R)RAL. K99 was either conserved or conservatively substituted with K99R. A100a was almost completely conserved, occurring in 27/30 selected clones. Despite the strong consensus for the substitutions W100R and M100bL in round 1, increasing the stringency of selection in rounds two and three demonstrated both wild-type residues W100 and M100b to result in improved affinity. The apparent evolution of the consensus sequence is likely due to the expected high occurrence of R100 and L100b which are both three times more likely to occur than the wild-type residues W100 and M100b, given the randomization scheme.

Several independent observations suggest that the sequence KWAM observed in 5/5 clones in the third round of FACS might yield the highest affinity scFv(dig) in the randomized heavy-chain library. (i) The murine immune system that produced the original 26-10 monoclonal antibody selected an

identical amino acid sequence at these positions, an otherwise unlikely coincidence. (ii) Saturation mutagenesis and SPR kinetic analysis of the wild-type HCDR3 residues W100 and M100b show that each residue confers the highest affinity (Burks *et al.*, 1997) (G.Chen, G.Georgiou and B.L.Iverson, unpublished data). The substitutions W100R and M100bL, which predominated in the selected population in rounds one and two, were also found to confer nearly wild-type affinity in saturation mutagenesis experiments. (iii) Due to the method used for library construction, the amino acid sequence KWAM was encoded by only one non-wild-type DNA sequence out of over 10^6 sequences that can arise from the randomization of the codons for H(99)–H(100b). (iv) The isolation of the KWAM mutant was not due to a growth advantage (data not shown). (v) The KWAM mutant exhibited a subtly improved affinity relative to sequences having only the single amino acid substitutions W100R or M100bL which were isolated in rounds one and two respectively (Figure 5, Table I). While it is not possible to determine with certainty whether other higher affinity mutants might exist within a library population, the above results demonstrate an effective selection and strongly suggest that the sequence KWAM at H(99)–(100b) confers the highest digoxin affinity in the library.

An affinity-improved scFv(dig) mutant with a subnanomolar dissociation constant, threefold lower than that of the wild-type, was isolated from a smaller light-chain library in a single step, simply by including a dissociation phase prior to sorting. Though the light-chain library encompassed a relatively small protein sequence space, coding for 8000 unique mutants, the isolation of an improved antibody with subnanomolar dissociation constant in a single step is remarkable. T91 was conserved in LC1.1-1, as well as in clones LC2.2-1 and LC2.2-2, the only other isolated clones with affinities within 2-fold of the wild-type. The conservative substitution V94I is presumably preferred due to the slightly larger volume of isoleucine, since V94 helps to define the interior of the binding pocket yet does not make close contact with digoxin (Jeffrey *et al.*, 1993). Though, this interpretation assumes that the binding pocket of scFv(dig), which differs only with respect to light-chain framework residues, closely mimics that of the 26-10 Fab. Though the structural basis of the substitution P96A is unclear, it might confer favorable flexibility to the light-chain CDR3.

The isolation of the high-affinity heavy and light-chain mutants HC1.3-1 and LC1.1-1 demonstrates an effective affinity-based selection. Fusions to Lpp-OmpA' are displayed on the surface of the *E.coli* at a level of $3\text{--}5 \times 10^4$ copies per cell (Francisco *et al.*, 1993; Chen *et al.*, 1996). The number of displayed polypeptides is sufficiently high to enable selection on the basis of affinity to monovalent antigen in solution. Consequently, selections may be accomplished without using polyvalent antigen, thus precluding avidity dominated selections. The isolation of high-affinity clones in other polypeptide screening methodologies can be complicated by avidity effects (Cramer *et al.*, 1996; Boder and Wittrup, 1997).

While little information is available describing the detection and isolation of rare bacterial cells using flow cytometry (Cormack *et al.*, 1996; Valdivia and Falkow, 1996), it was relatively straightforward to develop protocols that routinely achieve enrichment factors of 2000-fold and as high as 10 000-fold per round (unpublished data) under optimal conditions. Higher enrichment factors should be obtainable by minimizing non-specific staining and using multi-parameter sorting (Gross

et al., 1995). In the present work, the ability to determine the frequency of rare high fluorescence clones in a library illustrates the utility of extending rare cell analysis to cell surface libraries. Large *E. coli* displayed libraries can be analyzed directly for the presence of desired clones prior to sorting. Using FACS instrumentation optimized for high-throughput sorting (Leary, 1994), the screening of libraries as large as 10^9 members could be accomplished rapidly in a single step.

Measurements of k_{diss} obtained by flow cytometry were a reliable predictor of relative affinity and provided a convenient method of affinity ranking selected clones. The ability to analyze directly, by flow cytometry, the dissociation kinetics of selected clones prevents unnecessary, time-consuming subcloning and purification steps. Despite the agreement between relative affinities determined by flow cytometry and SPR, the absolute values of k_{diss} measured by flow cytometry were 2–3-fold slower than those determined for purified antibodies by SPR. The absence of flow conditions providing buffer exchange in the FACS sample tube may play a significant role in this discrepancy, since such conditions give rise to rebinding effects which slow the observed dissociation rate (Karlsson and Falt, 1997). Modules for kinetic measurements, enabling constant sample mixing and temperature control, are available for many flow cytometers and may help overcome this problem.

The combination of protein display on bacteria and FACS instrumentation should provide a generally applicable strategy for screening protein libraries more quantitatively than present methods allow. The ability to analyze protein libraries directly for the presence of rare, high-affinity clones may prove beneficial for rapidly probing numerous distinct library populations. Flow cytometry also provides a convenient measure of selection success and allows for rapid and reliable affinity ranking of the selected clones without any subcloning or protein purification steps. This system brings refinement to the library screening process and should enable the isolation and analysis of high-affinity polypeptides in a shorter time span than currently achievable.

Acknowledgements

We would like to thank Carlos Barbas for advice regarding library construction and Michael Cieslewicz for critically reading the manuscript. P.S.D. was partially supported through a NIH Biotechnology Training Fellowship. This work was supported by grants from the National Science Foundation (BES-9412502), DOE (DE-FG07-96ER 62322) and through a contract from DARPA.

References

- Barbas, C.F. (1993) *Curr. Opin. Biotechnol.*, **4**, 526–530.
 Boder, E.T. and Wittrup, K.D. (1997) *Nature Biotechnol.*, **15**, 553–557.
 Burks, E.A., Chen, G., Georgiou, G. and Iverson, B.L. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 412–417.
 Burks, E.A. and Iverson, B.L. (1995) *Biotechnol. Prog.*, **11**, 112–114.
 Chen, G., Cloud, J., Georgiou, G. and Iverson, B.L. (1996) *Biotechnol. Prog.*, **12**, 572–574.
 Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) *Gene*, **173**, 33–38.
 Cramer, A., Cwirla, S. and Stemmer, W.P. (1996) *Nature Med.*, **2**, 100–102.
 Francisco, J.A., Campbell, R., Iverson, B.L. and Georgiou, G., (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10444–10448.
 Georgiou, G., Stathopoulos, C., Daugherty, P.S., Nayak, A.R., Iverson, B.L. and Curtiss, R. (1997) *Nature Biotechnol.*, **15**, 29–34.
 Gross, H.J., Verwer, B., Houck, D., Hoffman, R.A. and Recktenwald, D. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 537–541.
 Gunneriusson, E., Samuelson, P., Uhlen, M., Nygren, P.A. and Stahl, S. (1996) *J. Bacteriol.*, **178**, 1341–1346.
 Hanes, J. and Pluckthun, A. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 4937–4942.
 Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) *Gene*, **77**, 51–59.

- Huston, J.S., Levinson, D., Mudgett-Hunter, M. *et al.* (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5879–5883.
 Jeffrey, P.D., Strong, R.K., Sieker, L.C., Chang, C.Y., Campbell, R.L., Petsko, G.A., Haber, E., Margolies, M.N. and Sheriff, S. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10310–10314.
 Karlsson, R. and Falt, A. (1997) *J. Immunol. Methods*, **200**, 121–133.
 Leary, J.F. (1994) *Methods Cell. Biol.*, **42**, 331–358.
 Lopez-Amoros, R., Comas, J. and Vives-Rego, J. (1995) *Appl. Environ. Microbiol.*, **61**, 2521–2526.
 Lowman, H.B., Bass, S.H., Simpson, N. and Wells, J.A. (1991) *Biochemistry*, **30**, 10832–10838.
 Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D. and Winter, G. (1991) *J. Mol. Biol.*, **222**, 581–597.
 Maurer, J., Jose, J. and Meyer, T.F. (1997) *J. Bacteriol.*, **179**, 794–804.
 Needels, M.C., Jones, D.G., Tate, E.H., Heinkel, G.L., Kochersperger, L.M., Dower, W.J., Barrett, R.W. and Gallop, M.A. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 10700–10704.
 Piard, J.C., Hautefort, I., Fischetti, V.A., Ehrlich, S.D., Fons, M. and Gruss, A., (1997) *J. Bacteriol.*, **179**, 3068–3072.
 Rode, H.J., Little, M., Fuchs, P., Dorsam, H., Schooltink, H., de Ines, C., Dubel, S. and Breitling, F. (1996) *Biotechniques*, **21**, 650, 652–653, 655–656, 658.
 Schier, R., Bye, J., Apell, G., McCall, A., Adams, G.P., Malmqvist, M., Weiner, L.M. and Marks, J.D. (1996) *J. Mol. Biol.*, **255**, 28–43.
 Schildbach, J.F., Near, R.I., Brucoleri, R.E., Haber, E., Jeffrey, P.D., Ng, S.C., Novotny, J., Sheriff, S. and Margolies, M.N. (1993) *J. Biol. Chem.*, **268**, 21739–21747.
 Short, M.K., Jeffrey, P.D., Kwong, R.-F. and Margolies, M. (1995) *J. Biol. Chem.*, **270**, 28541–28550.
 Smith, T.W., Butler, V.P., Jr. and Haber, E. (1970) *Biochemistry*, **9**, 331–337.
 Valdivia, R.H. and Falkow, S. (1996) *Mol. Microbiol.*, **22**, 367–378.
 Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J. and Johnson, K.S. (1996) *Nature Biotechnol.*, **14**, 309–314.
 Winter, G. and Milstein, C. (1991) *Nature*, **349**, 293–299.
 Yang, W.P., Green, K., Pinz-Sweeney, S., Briones, A.T., Burton, D.R. and Barbas, C.F. (1995) *J. Mol. Biol.*, **254**, 392–403.

Received February 15, 1998; revised April 24, 1998; accepted May 1, 1998