Exploiting recombination in single bacteria to make large phage antibody libraries

Daniele Sblattero and Andrew Bradbury†*

*International School for Advanced Studies (SISSA), Biophysics Sector, Via Beirut 2-4, Trieste, 34014, Italy. †Current address: Bioscience Division, MS M888, Los Alamos National Laboratory, Los Alamos, NM 87545. *Corresponding author (amb@telomere.lanl.gov).*

Received 16 August 1999; accepted 10 November 1999

The creation of large phage antibody libraries has become an important goal in selecting antibodies against any antigen. Here we describe a method for making libraries so large that the complete diversity cannot be accessed using traditional phage technology. This involves the creation of a primary phage scFv library in a phagemid vector containing two nonhomologous lox sites. Contrary to the current dogma, we found that infecting Cre recombinase–expressing bacteria by such a primary library at a high multiplicity of infection results in the entry of many different phagemid into the cell. Exchange of VH and VL genes between such phagemids creates many new VH/VL combinations, all of which are functional. On the basis of the observed recombination, the library is calculated to have a diversity of 3 x 10^{11}. A library created using this method was validated by the selection of high affinity antibodies against a large number of different protein antigens.

Keywords: antibody engineering, phage display, recombination, Cre recombinase, filamentous phage, single-chain Fv (scFv)

Phage display was recently introduced as a means of making antibodies in vitro[1–5]. In general, the affinity of the antibodies isolated is proportional to the initial size of the library used for selection[6,7]. Thus large libraries have become important as sources of high-affinity antibodies to virtually any antigen[3,7–9]. Most such libraries have been made using cloning, and when cloning procedures are optimized, can have diversities of $>10^{10}$, from which subnanomolar affinity antibodies have been isolated[7,9].

Recombination has been proposed as an alternative to cloning for creation of large libraries[10–12]. Lambda recombinase has been used to recombine Fabs[10], and Cre recombinase to recombine both Fabs[12] and single-chain Fv's (scFv)[11]. For making large phage antibody libraries, however, there has been only one report describing the use of Cre to recombine Fabs[9]. This library, derived from synthetic libraries of 10^8 VH and 8×10^5 VL, had a final estimated diversity of 6.5×10^{10}. Although antibodies with nanomolar affinities were isolated, the system is difficult to use, since deletion of antibody genes occurs at relatively high frequencies because the library is formed in phage.

In all the recombination systems described above, VH genes are cloned in one vector, VL in another, and recombination is used to create a third vector capable of displaying functional antibodies. When the products of recombination are selected by the use of newly created antibiotic resistances[10,11], all plasmids with the appropriate resistance should be recombined correctly. This reflects an underlying strategy to render recombination irreversible, since in all three cases, a significant proportion of the vectors present after recombination are derived either from nonfunctional parental plasmids or from undesired products of recombination.

Here we describe a method that uses a single vector to exploit the reversibility of Cre-catalyzed recombination. First we created a relatively small primary library (7×10^7 was used here) in a phagemid vector in which the VH and VL genes are separated by two nonhomologous lox sites. The VH and VL genes in this primary library were then recombined by infecting the phagemid into Cre-expressing bacteria at high multiplicity of infection (MOI). Under these conditions, many different phagemids enter a single bacterium, and the VH and VL genes were exchanged between different phagemids, creating many new VH/VL combinations, all of which are functional.

Results

A phagemid antibody display vector with incorporated lox sites. To use Cre recombinase to shuffle V genes in the scFv format, a lox site must be placed between the VH and VL genes, which requires a translated lox site as a protein linker. By examining the six possible frames for the wild-type loxP site and the mutated loxP511 site (which will not recombine with the wild-type loxP (ref. 15)), we identified a translation of loxP511 (ITSYNVYYTKL) that had only a single basic amino acid (to reduce the possibility of proteolysis), lacked stop codons, and was the least hydrophobic. The ability of this sequence to act as an scFv linker was tested following the construction of a new phage display vector, pDAN5 (Fig. 1A), in which the described translation of the loxP511 site was used as a protein linker between VL and VH. Three different monoclonal antibodies (mAbs) (D1.3, antilysozyme[16]; Y13.259, anti-p21ras (ref. 17); and GL30, antigliadin), were cloned into pDAN5. In each case, scFvs with the loxP511 linker were able to recognize the appropriate antigen specifically and gave ELISA signals comparable to those obtained using display vectors with the standard glycine-serine linker (Fig. 1B).

Recreating a functional scFv by shuffling the V genes. The ability to shuffle VH and VL genes in vivo to create functional antibodies was tested using an scFv derived from the antilysozyme mAb D1.3 (ref. 16). Two scFvs that contained either D1.3 VH or D1.3 VL with irrelevant partner chains, X and Y, were created (VL/X-VH/D1.3 and VL/D1.3-VH/Y). Recognition of lysozyme by D1.3 scFv was shown to require the presence of both D1.3 heavy and light chains. Single D1.3 chains associated with irrelevant partner chains were nonfunctional by ELISA. Phagemid containing these scFv genes were allowed to infect *Escherichia coli* expressing Cre recombinase at an MOI of 20:1. If recombination was successful, each bacterium would contain four different scFv genes (VL/D1.3-VH/D1.3; VL/D1.3-VH/Y; VL/X-

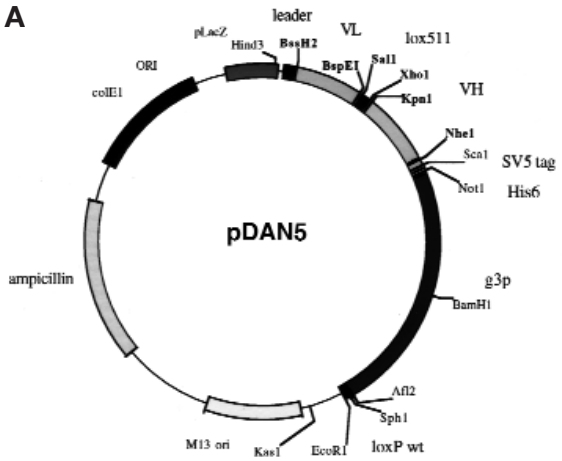
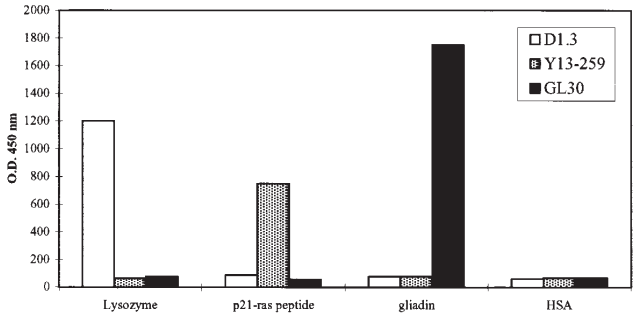


Figure 1. (A) Map of the display vector pDAN5 with an scFv cloned. Sites used for V gene cloning are in bold. *Bss*HII, *Bsp*EI, *Sal*I, *Xho*I, *Kpn*I, *Nhe*I. (B) ELISA signals with scFvs derived from monoclonal antibodies. VH and VL genes cloned into pDAN5 in the scFv format were tested for binding to their own antigens (D1.3: lysozyme; Y13-259: p21ras peptide; GL30: gliadin), as well as to the other tested antigens and a control antigen: human serum albumin (HSA).



VH/D1.3 and VL/X-VH/Y; Fig. 2). Phagemid were rescued from such bacteria and coupled to the appropriate scFv protein by passage through normal *E. coli* infected at MOI < 1 (Fig. 2). In the presence of Cre recombinase, recombination was demonstrated in 25% of phagemid by PCR and 17% by ELISA (Table 1), whereas in wild-type bacteria, no recombination was observed. These results indicate that recombination induced by Cre recombinase can be used to shuffle VH and VL genes between different phagemids, thereby giving rise to new specificities. Furthermore, the recombination reaction appears to go to equilibrium, as VL/D1.3-VH/D1.3 comprises approximately 25% of the total.

Creating a large recombined antibody library in pDAN5. A small scFv phagemid library consisting of 7×10^7 independent clones was created in pDAN5 by cloning PCR-assembled scFv derived from peripheral blood Vμ, Vλ, and Vκ genes (see Experimental Protocol for details). This primary library was used to infect bacteria expressing Cre recombinase following the scheme illustrated in Figure 2, except that instead of the two D1.3 phagemids, the library of 7×10^7 phagemids was used, and the MOI was 200:1 instead of 20:1. This procedure results in bacteria containing multiple phagemids, each of which encodes different VH and VL genes, which can be recombined by the Cre recombinase. Following recombination, phagemid were derived from these bacteria and used to infect bacteria not expressing Cre (DH5αF') at MOI <1 to couple phenotype and genotype, as indicated in Figure 2.

In order to determine the number of different phagemid entering a single bacterium, and the degree of recombination occurring between them, single bacteria that had undergone infection and recombination were plated out to isolate single colonies. Such colonies contain the complement of VH and VL genes present in the single bacteria before plating out. These colonies were individually amplified in liquid culture, used to prepare phagemid, and subsequently rescued as individual colonies. The scFvs present in all the individual colonies represent those found in the original cell giving rise to the single colony that was amplified. The scFvs present in a number of these colonies were then characterized by either sequencing or separately amplifying VH and VL genes and fingerprinting them with the restriction enzyme *Bst*NI. Results obtained with the two methods were essentially identical, with different fingerprint patterns always representing different VH and VL genes.

On the basis of either sequencing or fingerprint patterns, the different VH and VL genes arising from a single cell were numbered for five different individual colonies, with 40–80 V genes analyzed for each. Results for these were very similar, with 12–18 different VH and 12–15 different VL genes identified for each cell (one colony is shown in Table 2). These were present in 19–30 different scFv combinations, with the vast majority of scFvs containing shared V genes. In the cell illustrated, 11 examples of all four possible combinations of VH/VL gene pairs were identified (shown in underlined bold copy in Table 2), indicating that recombination had been extensive. Only one VH/VL pair did not appear to have participated in recombination in this cell.

One concern with Cre/lox systems is their perceived instability. We analyzed full-length scFvs by PCR (data not shown) and found that 4 (4%) of 96 from the primary library and 6 (6%) of 96 from the recombined library contained genes that had either the VH or VL gene deleted. The lox site was preserved in all sequenced scFvs, including those that did not appear to participate in recombination. Therefore, these deletions are likely to have occurred during library construction, probably as a result of spurious priming during PCR assembly.

Testing the large phage antibody library. The library was tested by selection on a number of different protein antigens (Table 3A, and data not shown). Antibodies were obtained against all 15 antigens attempted, with a range of 3 to 11 antibodies per antigen and a mean of over 6 antibodies per antigen. Antibodies specifically recognized their antigens and were not polyreactive. Eight of these scFvs were purified by affinity chromatography using the His tag and were characterized by gel filtration. All scFvs were monomeric and did not show the 'diabody' peak associated with scFv containing shorter linkers[18]. Four of the monomeric scFv peaks were purified, and their affinities were assessed by surface plasmon resonance. All had affinities <90 nM, with the best having an affinity of 15 nM (Table 3B).

Variable gene sequence analysis. The different V genes found in the primary library, the recombined library and a number of selected scFvs were analyzed by sequencing 24–40 V genes for each category. After sequencing, the origin of the V gene and complementarity determining region 3 (CDR3) length were determined using the

Table 1. Analysis of D1.3 recombination by PCR and ELISA.

ELISA testing		
Cre recombinase bacteria	16/96	17%
Wild-type bacteria (DH5αF')	0/96	0%
PCR testing using D1.3 specific primers		
Cre recombinase bacteria	12/48	25%
Wild-type bacteria (DH5αF')	0/48	0%

The ELISA signals observed for 96 different colonies, and PCR signals for 48 different colonies, derived after growth in Cre-expressing bacteria or wild-type bacteria are shown. ELISAs were performed using phagemid binding to lysozyme adsorbed to 96-well plates. PCR was performed using a 5′ primer specific for the CDR3 of VL and a 3′ primer specific for the CDR3 of VL. This combination is able to give a PCR product only when D1.3 VH and VL are present in the same scFv.

V BASE immunoglobulin V gene database[19]. The primary and recombined libraries were diverse, with almost all families represented (Fig. 3A). These were not evenly distributed, with VH genes having a predominance of VH4 genes, and VL genes having more Vκ1 and Vκ3 genes. There was no great difference between the primary and recombined libraries, indicating that diversity does not appear to be compromised by recombination.

Upon selection, the V gene distribution changed somewhat, as shown previously[9]. Whereas the primary and recombined libraries had an excess of VH4 genes, the V genes found in antibodies recognizing selected antigens appeared to be more widely distributed, with VH1, VH3, VH5, and VH6 being the VH gene classes found most frequently. Among the VL genes, 38% of all selected antibodies contained a gene of the VL3 family, even though this accounts for <10% of the primary and recombined libraries. The CDR3 length, on the other hand (Fig. 3B), does not appear to change very much between the three populations of antibodies, with a wide distribution of all lengths found, ranging between six and 22 amino acids for the VH genes and seven and 13 for the VL genes.

Discussion

Although phage antibody libraries have been used to isolate antibodies against a large number of antigens[3,5,7–9,20,21], the procedure is still not widely used, possibly because of the difficulty of making large phage antibody libraries, which require that at least 10^9 independent clones be derived from a good source of diverse VH and VL genes. In general, such libraries are made by carrying out a large number of ligations and transfections, and once made, become a limited resource as amplification cannot be carried out without a potential reduction in diversity. The method described here overcomes this problem, as diversity is regenerated each time recombination is used to create each new secondary library. In fact, given sufficient resources, library adequate for 10^7 selections can be made from an initial primary library without losing diversity. This is likely to be an important factor in the application of phage antibodies to functional genomics.

Other recombination methods used or proposed for making antibody libraries[3,10–12] use two plasmids to generate the library. The VH genes are cloned on one plasmid, and the VL genes on another. After recombination, four plasmid populations are generated, only one of which is correct (containing heavy and light library chains). The others contain copies of either single library chains in combination with dummy chains, or dummy chains alone. Although these extra products should not be packaged into phage, we have found that plasmids that lack an Ff origin of replication may still be packaged, and as a result will contaminate the library. The use of a single plasmid to generate diversity ensures that all recombination products are functional. It is generally believed that one bacterium cannot be infected with more than one phage or phagemid, because p3 expression is believed to inhibit pilus synthesis. We have overcome this problem by inhibiting p3 synthesis with glucose. This does not, however, explain the maintenance of so many different phagemids with the same origin of replication within a single cell. The mechanisms of this are presently under investigation.

Although assessing the degree of diversity created in a single cell is difficult, we found similar results for five colonies analyzed. In the colony shown, 13 different VH genes and 12 different VL genes were identified. These were recombined in a total of 30 different combinations, with all except one scFv containing VH or VL genes also found in other scFvs. One VL was found with six dif-

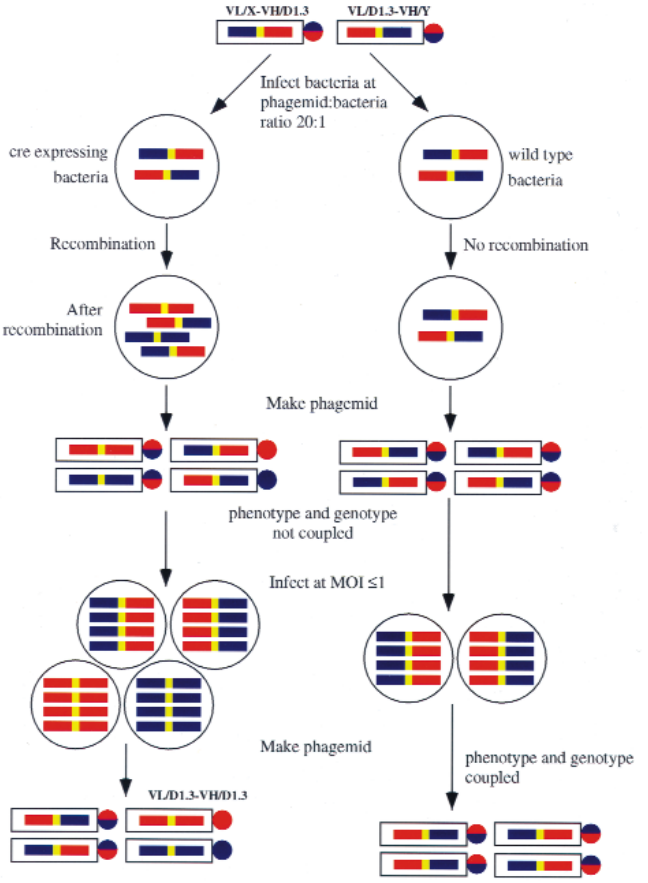


Figure 2. scheme of D1.3 recombination experiment. Two phagemid containing VL/X-VH/D1.3 and VL/D1.3-VH/Y (where X and Y represent irrelevant V genes) are added to either Cre-expressing bacteria or wild-type DH5αF at an MOI of 20:1. After overnight growth, phagemid are made, reinfected into DH5αF' at a phagemid:bacteria ratio of <1 (to couple genotype and phenotype), and tested by PCR and ELISA for the presence of functional VL/D1.3-VH/D1.3.

Table 2. Assessing the diversity of Cre bacteria infected by multiple phagemid.

	VH genes												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	1									1	1		
2		1	**1**	**1**	1	**3**						**5**	
3			**1**	**1**									
4			**1**						**1**				1
5	**1**			**2**		**1**			1			**1**	
6							1						
7			**1**						**2**				
8									1				
9	**3**			**2**				1					
10			**1**	**1**									
11				1									
12						1							

Single scFv genes from individual colonies of Cre-expressing bacteria containing multiple recombined V genes were isolated by phagemid preparation and reinfection at low MOI into DH5αF'. The individual VH and VL chains present in each phagemid were identified by sequencing. The different sequences found for VH and VL were numbered, with the VH genes listed horizontally and the VL fingerprints vertically. The number of scFvs found with a particular VH/VL combination is indicated in the corresponding cell. V genes that are derived from two VH/VL combinations in which all four possible recombination products have been identified are indicated by bold, underlined characters.

Table 3A. Antigens against which antibodies have been selected.

Immunotube selection

Antigen	Source	Tag	Different scFv
Rad52	Human	His	11
PIGS 12B	Pyrobaculum	His	4
Cyclin D	Human	GST	4
cdk2	Human	GST	8
cdc25A	Human	GST	3
HSA	Human		7
loop2 gp120	HIV		6

HAS, human serum albumin; glutathione–S–transferase (GST),
A list of all the antigens against which the library has been tested is given. For those selected in immunotubes, this includes the number of different scFvs derived.

Table 3B. Affinities of selected scFvs.

Antigen	Clone	K_{on} $M^{-1}s^{-1}$	K_{off} s^{-1}	K_d M
rad52	rad52-1	1.26e5	1.96e-3	1.56e-8
cdk2	cdk2-1	5.19e4	4.30e-3	8.29e-8
	cdk2-9	1.47e5	4.38e-3	2.97e-8
	cdk2-3	8.90e4	5.33e-3	5.98e-8

Monomeric scFvs were purified by IMAC and gel filtration as described in the Experimental Protocol. On and off rates were determined by surface plasmon resonance in BIAcore.

ferent VH's, one VH with six different VL's, and 11 overlapping cases of all four combinations of a VH/VL pair were identified (Table 2), indicating that recombination had been extensive.

Interestingly, in all five cells analyzed, no single scFv dominated the analysis, with the most abundant being present in only five copies (14%), indicating that all scFv appear to have similar probabilities of remaining within the cell. The single sequenced scFv that did not appear to have participated in recombination was normal, with a functional lox site present. Thus, its apparent nonparticipation was probably due to not having analyzed enough clones rather than an intrinsic problem of the scFv itself. The minimum diversity generated by a single cell can be estimated to be the number of recombined scFv actually observed: 26–30, which gives a minimum library size estimate of 3×10^{11}. However, if the 13–18 different VH genes rescued from a single colony are matched by 13–18 different VL genes, the potential diversity identified in this small sample is 169–324 (132–182), giving a maximum estimate that approaches the 500–700 copy number of pUC-based plasmids[22], and the potential for a library at least 10-fold larger.

Previous libraries created using the Cre/lox system[9] have suffered from instability as a result of deletion of antibody V genes. This has been attributed to the fact that the library was constructed in a filamentous phage vector, although the possibility that the lox site in some way confers instability cannot be excluded. In this study, we found that the scFv genes containing the lox site are remarkably stable, with the percentage of full-length scFv genes remaining essentially identical in the passage from primary (4%) to recombined library (6%). These figures probably represent the growth advantage of preexisting deleted clones[23] during the growth cycles required for the recombination process. Furthermore, in all full-length scFvs, the lox site was always found to be present by sequencing.

The possibility that the recombination process itself may induce bias in the library was examined by sequencing V genes from the primary and recombined libraries. An increase in the representation of Vκ3 and Vκ2 genes, and a reduction in VH1 and Vκ4 genes in the recombined library compared with the primary was found, but in general the ratio of the different V genes was remarkably well preserved. Interestingly, neither the primary nor the recombined library was dominated by single V gene families, as has been found in other naive libraries, and representative genes from almost all VH and VL gene families were found. This may be because of the V gene primer set used[24], which was specifically designed to be able to amplify all known germline V genes, or because amplification of V genes was carried out with individual primers, rather than mixtures. The diversity of V gene families in the primary and recombined libraries was also reflected in the V genes of scFvs selected for antigen binding. Although Vλ3 genes are frequently found, as shown for scFvs selected from other published naive libraries[7,9], there is no predominance of other V gene families, and members of almost all V gene families can be found. Of 22 different selected scFvs sequenced, 14 different VH genes, and 13 different VL genes were found, with no example of identical V genes being found in two scFvs recognizing different antigens. By comparing V gene distribution in the recombined library to that found in selected scFvs, we found that some V genes (e.g., VH4, Vκ1, and Vκ3) were recovered more frequently in the library than in selected scFvs, whereas others (e.g., VH1, VH3, Vλ1, and Vλ3) were recovered more often in selected scFvs than in the recombined library. Whether this represents V genes that, in general, are more likely to recognize antigens of interest, or reflects the fact that relatively few antigen-binding scFvs have been analyzed, awaits further analysis.

The affinities of the antibodies isolated were all better than 90nM, with the best having an affinity of 15 nM. This is lower than the best affinities reported for the larger libraries, which in some cases were subnanomolar[7,9]. However, the highest affinities reported in these papers were all obtained by selection in the soluble phase using biotinylated antigen and magnetic streptavidin beads, or alternatively by selecting on haptens. When selections were performed as described here (protein antigens coupled to immunotubes), the affinities obtained were similar (see Sheets et al. for a full discussion of this point). Furthermore, the bacterial elution method used here has recently been shown to be far less efficient than more stringent methods (e.g., 100 mM HCl, pH 1.1, or 100 mM triethylamine) at eluting high-affinity antibodies[25], suggesting

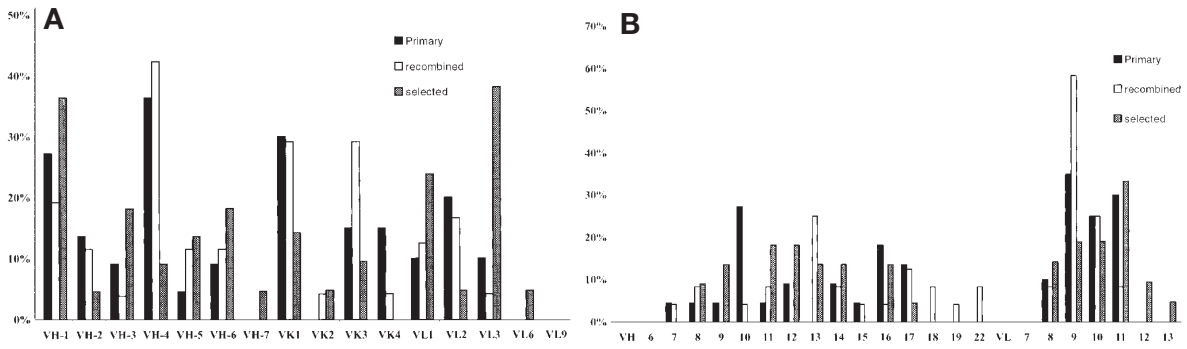


Figure 3. V gene use and CDR3 length in primary, recombined, and selected scFvs. V gene family (A) and CDR3 length (B) were determined using V BASE[19], and the percentage of V genes deriving from a particular V gene family and the CDR3 length are indicated for scFvs derived from the primary or recombined library, and for selected scFvs recognizing specific antigens.

that the antibodies with the highest affinities may well have been left on the immunotube.

Despite these caveats, the affinities of antibodies selected from the recombined library still exceed those selected from libraries with diversities similar to the primary library used (e.g., the 3×10^7 library described by Marks et al.[21]), indicating that antibodies of high affinity can be selected from initial small libraries when recombination is carried out. This is not surprising, given that the large libraries created using traditional methods do not have greater numbers of different VH and VL genes, but a greater proportion of the possible combinations of these genes. This is reflected in the recent trend to first make small VH and VL libraries and then combine these by cloning[7–9], a procedure that is carried out far more efficiently by in vivo recombination.

While there have been practical demonstrations that there is an affinity advantage to creating larger libraries, the search for larger and larger libraries reaches a practical limit in the volume required to perform a selection. This is fixed at approximately 10^{13} phage in a volume of 1 ml, indicating that diversity $>10^{12}$ probably remains untapped (assuming a display level of 10%[26]). It is unlikely that standard methods of library creation could reach this level of diversity, given that libraries 100 times smaller already require hundreds of transfections to create. However, it should be possible to tap this level of diversity using the method we describe here by increasing the volume of bacteria in the preparation of the final library to 3–5 L. With still higher volumes, one can contemplate the creation of far larger libraries, which could be used at the industrial level.

By including cycles of recombination between rounds of selection, the effective diversity accessed would be expected to approach the theoretical maximum diversity of the library. Binding V regions could be shuffled either against one another or against the starting library. In both cases, this would be the equivalent of shuffling the chains of all binding antibodies in parallel, in a fashion similar to that which has been carried out in series for single antibodies with notable increases in affinity[4,27]. We expect the affinities of antibodies obtained in this way to be higher than those reported here. This approach could also be applied to affinity maturation, with mutations introduced simultaneously in both heavy and light chains, giving the advantage that all potential combinations of mutations can be sampled simultaneously.

Experimental protocol

Bacterial strains. DH5αF' (Gibco BRL, Rockville, MD): F'/*end*A1 *hsdR17* (r_K- mK^+) *supE44 thi-1 recA1 gyrA (Nal^r) relA1 Δ (lacZYA-argF)U169 deoR* (ϕ80*dlacΔ(lacZ)M15)*

BS1365: BS591 F' kan (BS591: recA1 endA1 gyrA96 thi-1 Δ lacU169 supE44 hsdR17 [lambda imm434 nin5 X1-cre][28])

Creating pDAN5 (scFv display vector) and pDAN5-scFv derivatives. To make pDAN5, a new polylinker was cloned into pUC119 by overlap PCR of two long oligonucleotides. This introduced a bacterial leader sequence, a polycloning cassette, the SV5 tag[29], a His_6 tag, and an amber stop codon (Fig. 1). Gene 3 was subsequently cloned by PCR amplification from fdtet, with the 5′ end downstream of the amber stop codon, and the 3′ end upstream of a wild-type lox sequence. The Y13-259 scFv was assembled (in the order VL-VH) from Y13-259 scFv (VH-VL order)30 using murine V region–specific primers based on those previously described31 in which the standard $(gly4ser)_3$ scFv linker was replaced with the loxP511 site. This was cloned into pDAN5, to create pDAN5-Y13-259, using *Bss*HII and *Nhe*I. Other pDAN-scFvs were cloned, either as assembled scFv PCR fragments, or as individual VH and VL genes amplified to incorporate *Bss*HII and *Sal*I for VL and *Xho*I and *Nhe*I for VH.

Primary library construction. Total RNA was prepared from 40 different samples of human peripheral blood lymphocytes purified by Ficoll Hypaque (Amersham Pharmacia Biotech, UK). cDNA was synthesized using random hexamers and reverse transcriptase following standard protocols. IgM V regions were amplified using an IgM 3′ primer (GGA AAA GGG TTG GGG CGG AT) and the 5′ VH primers and methods as previously described[24]. The primers described in this paper were also used to amplify VL and Vκ genes from random primed cDNA, and VH genes from gel-purified IgM V genes. VH and VL genes were reamplified to add a region of overlap in the scFv linker as well as long tails to facilitate restriction enzyme digestion. The scFv were assembled by mixing equimolar amounts of VH and VL genes and performing assembly as previously described[31]. The scFv obtained were cloned into *Bss*HII/*Nhe*I-cut pDAN5–Y13-259 to obtain a primary library of 7×10^7 clones.

Recombination and the creation of the secondary library. To induce recombination, BS1365 (which expresses Cre-recombinase constitutively) was grown in 20 ml 2 × TY, 100 mg/ml kanamycin, 1% glucose at 37°C to OD_{550} 0.5. Phagemid were added at MOI 20:1 for the D1.3 shuffling experiment and 200:1 when the library was created. Phagemid were left for 1 h without shaking at 37°C to allow infection, and ampicillin was then added to 100 mg/ml. Recombination was allowed to continue by shaking overnight growth at 30°C.

After recombination, bacteria were diluted 1/20 in the same growth medium, and grown to OD_{550} 0.5 at 37°C. M13K07 helper phage were added at an MOI of 20:1, and left without shaking for 1 h at 37°C before further growth (6 h to overnight). Additional kanamycin was not added because BS1365 expresses kanamycin resistance. Phagemid were prepared by centrifugation and polyethylene glycol precipitation as previously described[4]. As these phagemid arise from bacteria containing many different scFvs, there is no coupling between phenotype and genotype. This is overcome by using the isolated phagemid to infect DH5αF' grown to OD_{550} 0.5 at MOI ≤ 1. Phagemid used for selection were prepared from overnight cultures of bacteria derived at this low MOI using standard techniques. When making the library, 20 ml of Cre-expressing bacteria were used. These were diluted into 400 ml to prepare phagemid, and 5×10^{11} of these phagemid were used to infect 1 L of DH5αF' (5×10^{11} bacteria). When recombining D1.3, all volumes used were 1 ml with appropriate dilutions.

Assessment of diversity in individual cells. After overnight growth in Cre-expressing bacteria, the culture was plated out to isolate individual colonies. These contain multiple recombined V genes. A number of these colonies were grown in 10 ml 2 × TY, 1% glucose, and 100 μg/ml ampicillin at 37°C to OD_{550} 0.5. M13K07 helper phage were added at an MOI of 20:1, the culture was left for 1 h at 37°C without shaking, and then left to grow 2–4 h at 30°C shaking (250 r.p.m.). Phagemids were prepared from these cultures as described above. All phagemids present in each individual culture represent all the different scFv combinations that have arisen within the original starting cell. Colonies were derived from these individual phagemid by infection at MOI ≤ 1 into DH5αF', and individual VH and VL chains present in each phagemid were identified by sequencing or *Bst*NI fingerprinting.

Proteins. Proteins used for selection were kindly provided by Min Park (Rad52), Scott Peterson (cyclin D, cdk2, and cdc25A), Tom Peat (phosphoglycerate dehydrogenase). Human serum albumin was purchased from Sigma (St. Louis, MO).

Selection of phage antibodies. Phage antibody selection was performed essentially as previously described[4], with proteins coupled to immunotubes (Nunc, Rochester, NY) at 10 mg/ml overnight, blocked in 2% nonfat milk–phosphate buffered saline (MPBS) and incubated with the phage antibody library (also blocked in MPBS) for 1–2 h. Washing after the first cycle involved five PBS and five PBS–0.1% Tween-20 washes. Phage were eluted by the addition of 1 ml DH5αF at OD_{550} 0.5. Following elution, bacteria were amplified and phage prepared for further cycles of selection. Subsequent washes were more stringent, and phage antibodies were tested for reactivity by ELISA after the second or third cycle.

Phage ELISA testing of antibody binding specificity. Phage ELISA was used to identify positive lysozyme binding D1.3 scFvs created by recombination. Phage ELISAs were also performed on individual phage clones isolated following selection on target proteins. Positive clones gave signals at least three times the background signal of 0.08. These were fingerprinted using V gene primers[23] for amplification and *Bst*NI to identify the number of different positive clones.

Sequencing. Sequencing was carried out using the Epicentre Sequitherm Excel II kit (Alsbyte, Mill Valley, CA) and analyzed using specific labeled primers annealing within the SV5 tag region and the leader sequence. Sequences were analyzed on a Li-Cor 4000L automatic sequencer (Lincoln, NE). The identity of the different V genes was analyzed by submitting the sequence to V BASE[19]. Twenty-two scFvs were sequenced for the primary and 26 for the recombined libraries; 22 different selected scFvs were sequenced; and in the single cell analysis, 35 different scFvs were sequenced.

scFv purification. Periplasmic extracts were made from 500 ml volumes of bacteria induced with 0.5 mM isopropylthiogalactoside at an OD_{550} of 0.6. The scFv were purified using the Ni-NTA kit (Qiagen, Valencia, CA) and checked by 12% PAGE. Gel filtration was carried out on a Superdex 75 column (Amersham Pharmacia Biotech, UK) to analyze and isolate the monomeric scFv fraction.

Affinity determination. We calculated scFv dissociation equilibrium constants (K_d) from the association (k_{on}) and dissociation (k_{off}) rate constants determined using surface plasmon resonance in a BIACORE 2000 instrument (Biacore AB, Uppsala, Sweden). Calculation of K_d values was performed by fitting the data according to a single-site model, using the BIAevaluation 3 software (Biacore AB). Approximately 200 RU of each antigen were immobilized on CM5 sensor chips (Biacore AB). The scFv were purified by immobilized metal-ion affinity chromatography (IMAC) and gel filtration as described above and used at 50–300 nM.

Acknowledgments

We are grateful to Roberto Marzari, Jianlong Lou, and Chonglin Yang for helpful discussions; to Francesco Tedesco for RNA from human lymphocytes; to Gabriella Rossi and Jessica Franzot for excellent technical help; and to the following for the proteins they kindly provided for selection: Min Park (Rad52), Scott Peterson (cyclin D, cdk2, cdc25A), Tom Peat (phosphoglycerate dehydrogenase). We would also like to thank Brian Sauer for providing BS1365 as well as many useful discussions on Cre recombinase. This work was partly financed by Fondo Trieste.

1. Barbas, C.F., Kang, A.S., Lerner, R.A. & Benkovic, S.J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl. Acad. Sci. USA* **88**, 7978–7982 (1991).
2. Griffiths, A.D. et al. Human anti-self antibodies with high specificity from phage display libraries. *EMBO J*. **12**, 725–734 (1993).
3. Griffiths, A.D. et al. Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* **13**, 3245–3260 (1994).
4. Marks, J.D. et al. By-passing immunization: building high affinity human antibodies by chain shuffling. *Bio/Technology*. **10**, 779–783 (1992).
5. Nissim, A. et al. Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *EMBO J*. **13**, 692–698 (1994).
6. Perelson, A.S. & Oster, G.F. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self non-self discrimination. *J. Theor. Biol.* **81**, 645–670 (1979).
7. 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).
8. 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).
9. Sheets, M.D. et al. Efficient construction of a large nonimmune phage antibody library; the production of panels of high affinity human single-chain antibodies to protein antigens. *Proc. Natl. Acad. Sci. USA* **95**, 6157–6162 (1998).
10. Geoffroy, F., Sodoyer, R. & Aujame, L. A new phage display system to construct multicombinatorial libraries of very large antibody repertoires. *Gene* **151**, 109–113 (1994).
11. Tsurushita, N., Fu, H. & Warren, C. Phage display vectors for in vivo recombination if immunoglobulin heavy and light chain genes to make large combinatorial libraries. *Gene* **172**, 59–63 (1996).
12. Waterhouse, P., Griffiths, A.D., Johnson, K.S. & Winter, G. Combinatorial infection and in vivo recombination: a strategy for making large phage antibody repertoires. *Nucleic Acids Res*. **21**, 2265–2266. (1993).
15. Hoess, R., Wierzbicki, A. & Abremiski, K. The role of the loxP spacer region in P1 site-specific recombination. *Nucleic Acids Res*. **14**, 2287–2300 (1986).
16. Mariuzza, R.A. et al. Preliminary crystallographic study of the complex between the Fab fragment of a monoclonal anti-lysozyme antibody and its antigen. *J. Mol. Biol*. **170**, 1055–1058 (1983).
17. Furth, M.E., Davis, L.J., Fleurdelys, B. & Scolnic, E.M. Monoclonal antibodies to the p21 products of the transforming gene of Harvey murine sarcoma virus and of the cellular *ras* gene family. *J. Virol.* **43**, 294–304 (1982).
18. Holliger, P., Prospero, T. & Winter, G. "Diabodies": small bivalent and bispecific antibody fragments. *Proc. Natl. Acad. Sci. USA*. **90**, 6444–6448 (1993).
19. Tomlinson, I.M., Williams, S.C., Corbett, S.J., Cox, J.P.L. & Winter, G. *VBASE sequence directory.* (MRC Centre for Protein Engineering, Cambridge, UK; 1996) http://www.mrc-cpe.cam.ac.uk/imt-doc/vbase-home-page.html.
20. de Kruif, J., Boel, E. & Logtenberg, T. Selection and application of human single chain Fv antibody fragments from a semi-synthetic phage antibody display library with designed CDR3 regions. *J. Mol. Biol.* **248**, 97–105 (1995).
21. Marks, J.D. et al. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* **222**, 581–597 (1991).
22. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1989).
23. 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).
24. Sblattero, D. & Bradbury, A. A definitive set of oligonucleotide primers for amplifying human V regions. *Immunotechnology* **3**, 271–278 (1998).
25. Schier, R. & Marks, J.D. Efficient in vitro affinity maturation of phage antibodies using BIAcore guided selections. *Hum. Antib. Hybrid*. **7**, 97–105. (1996).
26. Clackson, T. & Wells, J.A. In vitro selection from protein and peptide libraries. *Trends Biotechnol.* **12**, 173–184. (1994).
27. Schier, R. et al. Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection. *J. Mol. Biol.* **255**, 28–43 (1996).
28. Sauer, B. & Henderson, N. The cyclization of linear DNA in *Escherichia coli* by site-specific recombination. *Gene*. **70**, 331–341 (1988).
29. Hanke, T., Szawlowski, P. & Randall, R.E. Construction of solid matrix-antibody-antigen complexes containing simian immunodeficiency virus p27 using tag-specific monoclonal antibody and tag-linked antigen. *J. Gen. Virol.* **73**, 653–660 (1992).
30. Werge, T.M., Bradbury, A., Di Luzio, A. & Cattaneo, A. A recombinant cell line, expressing a form of the anti-p21ras antibody, Y13-259, which binds protein A and may be produced as ascites. *Oncogene* **7**, 1033–1035 (1992).
31. Krebber, A. et al. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods* **201**, 35–55 (1997).